МРНТИ 06.71.07
JEL: H56, L51, O33

Artificial Intelligence as a Tool of National Security Resilience: Evidence from Singapore

Aigerim Azatbekova[1]*, Zere Serikbayeva[1], Nurbolat Nyshanbayev[1]

[1] *Turan University, Almaty, Kazakhstan*

For citation: Azatbekova, A., Serikbayeva, Z. & Nyshanbayev, N. (2025). Artificial Intelligence as a Tool of National Security Resilience: Evidence from Singapore. Qainar Journal of Social Science, 4(4), 64-81, https://doi.org/10.58732/2958-7212-2025-4-64-81

Abstract

In the context of accelerated digitalization and the expansion of the use of artificial intelligence (hereinafter – AI) in critical sectors, the importance of forming effective AI management models focused on ensuring the sustainability of national security is increasing. The purpose of this study is to analyze the risk-based approach to artificial intelligence management in Singapore and assess its contribution to strengthening national security resilience in the period 2020-2025. The methodological basis of the study was a qualitative analysis of regulatory and strategic documents, comparative institutional analysis, as well as thematic coding of AI management tools from the perspective of risk-based regulation theory. The results of the study show that Singapore's AI management model is based on a combination of "soft" regulation, technical verification, and intersectoral collaboration, which minimizes the risks associated with cyber threats, vulnerability of critical infrastructure, and reduced public trust, without limiting innovation activity. In 2023-2024, the level of AI adoption among small and medium—sized enterprises increased by more than three times, and among large companies - by more than 18 percentage points. The share of employees using AI tools in their professional activities has reached almost 74%, which indicates the deep integration of AI into socio-economic processes. The practical significance of the work lies in the possibility of adapting the Singapore model in the development of national AI management systems in countries with a high degree of digitalization, including Kazakhstan..

Keywords: Artificial Intelligence, National Security, Regulation, Smart Technology, Social Sustainability, Social Trust, Singapore

Жасанды интеллект ұлттық қауіпсіздік орнықтылығын арттыру құралы ретінде: Сингапур тәжірибесі

Азатбекова Ә.[1]*, Серікбаева З.[1], Нышанбаев Н.[1]

[1]Тұран университеті, Алматы, Қазақстан

Дәйексөз үшін: Азатбекова Ә., Серікбаева З., Нышанбаев Н. (2025). Жасанды интеллект ұлттық қауіпсіздік орнықтылығын арттыру құралы ретінде: Сингапур тәжірибесі. Қайнар әлеуметтік ғылымдар журналы, 4(4), 64-81, https://doi.org/10.58732/2958-7212-2025-4-64-81

Түйін

Жедел цифрландыру және жасанды интеллекттің (бұдан әрі – ЖИ) критикалық маңызы бар секторларда кеңінен қолданылуы жағдайында ұлттық қауіпсіздіктің орнықтылығын қамтамасыз етуге бағытталған ЖИ тиімді басқару модельдерін қалыптастырудың маңыздылығы арта түсуде. Осы зерттеудің мақсаты — Сингапурда жасанды интеллектті басқарудың тәуекелге бағдарланған тәсілін талдау және оның 2020–2025 жж. аралығында ұлттық қауіпсіздік орнықтылығын нығайтуға қосқан үлесін бағалау. Зерттеудің әдіснамалық негізін нормативтік-стратегиялық құжаттарды сапалық талдау, салыстырмалы институционалдық талдау, сондай-ақ тәуекелге бағдарланған реттеу теориясы тұрғысынан жасанды интеллектті басқару құралдарын тақырыптық кодтау құрады. Зерттеу нәтижелері Сингапурдағы жасанды интеллектті басқару моделі «жұмсақ» реттеуді, техникалық верификацияны және секторлар арасындағы өзара іс-қимылды ұштастыруға негізделгенін көрсетеді. Бұл тәсіл инновациялық белсенділікті шектемей, киберқауіптермен, критикалық инфрақұрылымның осалдықтарымен және қоғамдық сенімнің төмендеуімен байланысты тәуекелдерді барынша азайтуға мүмкіндік береді. 2023–2024 жж. шағын және орта кәсіпорындар арасында жасанды интеллектті енгізу деңгейі үш еседен астам өсті, ал ірі компаниялар арасында бұл көрсеткіш 18 пайыздық пункттен астам артты. Кәсіби қызметінде жасанды интеллект құралдарын пайдаланатын қызметкерлердің үлесі шамамен 74%-ға жетіп, ЖИ-дің әлеуметтік-экономикалық үдерістерге терең интеграцияланғанын дәлелдейді. Жұмыстың практикалық маңыздылығы жоғары цифрландыру деңгейіне ие елдерде, соның ішінде Қазақстанда, ұлттық жасанды интеллектті басқару жүйелерін әзірлеу барысында сингапурлық модельді бейімдеу мүмкіндігінде көрініс табады..

Түйін сөздер: жасанды интеллект, ұлттық қауіпсіздік, реттеу, ақылды технологиялар, әлеуметтік тұрақтылық, әлеуметтік сенім, Сингапур

Искусственный интеллект как фактор укрепления устойчивости национальной безопасности: опыт Сингаппура

Азатбекова А.[1]*, Серикбаева З.[1], Нышанбаев Н.[1]

[1]*Университет Туран, Алматы, Казахстан*

Для цитирования: Азатбекова А., Серикбаева З., Нышанбаев Н.(2025). Искусственный интеллект как фактор укрепления устойчивости национальной безопасности: опыт Сингаппура. Кайнар журнал социальных наук, 4(4), 64Р-81, https://doi.org/10.58732/2958-7212-2025-4-64-81

Аннотация

В условиях ускоренной цифровизации и расширения применения искусственного интеллекта (ИИ) в критически важных секторах возрастает значимость формирования эффективных моделей управления искусственным интеллектом (далее – ИИ), ориентированных на обеспечение устойчивости национальной безопасности. Целью данного исследования является анализ риск-ориентированного подхода к управлению искусственным интеллектом в Сингапуре и оценка его вклада в укрепление устойчивости национальной безопасности в период 2020–2025 гг. Методологической основой исследования послужили качественный анализ нормативно-стратегических документов, сравнительный институциональный анализ, а также тематическое кодирование инструментов управления ИИ с позиций теории риск-ориентированного регулирования. Результаты исследования показывают, что сингапурская модель управления ИИ основана на сочетании «мягкого» регулирования, технической верификации и межсекторального взаимодействия, что позволяет минимизировать риски, связанные с киберугрозами, уязвимостью критической инфраструктуры и снижением общественного доверия, без ограничения инновационной активности. В 2023–2024 гг. уровень внедрения ИИ среди малых и средних предприятий увеличился более чем в три раза, а среди крупных компаний — более чем на 18 п.п. Доля работников, использующих ИИ-инструменты в профессиональной деятельности, достигла почти 74%, что свидетельствует о глубокой интеграции ИИ в социально-экономические процессы. Практическая значимость работы состоит в возможности адаптации сингапурской модели при разработке национальных систем управления ИИ в странах с высокой степенью цифровизации, включая Казахстан.

Ключевые слова: искусственный интеллект, национальная безопасность, регулирование, умные технологии, социальная устойчивость, социальный доверие, Сингапур

Introduction

Modern national security architectures are changing due to the quick global spread of artificial intelligence (hereinafter – AI) in public administration, cybersecurity, and defense. States are increasingly integrating algorithmic systems into fundamental security and governance functions to use AI for threat detection, intelligence analysis, cyber defense operations, and data-driven public service delivery. The way governments now present AI as a strategic asset for security and geopolitical competitiveness is highlighted by comparative studies of national AI strategies (Radu, 2021; Yerlikaya & Erzurumlu, 2021). AI is also becoming a more significant source of vulnerability. Research on AI for national security (e.g., financial networks, transportation, and energy grids) highlights the "predictability problem" and the emergence of new risk vectors, such as AI-driven cyberattacks, automation-enhanced disinformation, and cascading failures in critical infrastructure systems (Taddeo et al., 2022). AI's dual role as a security enabler and a source of intricate, systemic risks complicates conventional risk-management strategies and calls for more precise, risk-sensitive governance tools (Al-Hawawreh et al., 2024). Between 2020 and 2025, national strategies and governance frameworks for AI in security and defense have become a significant topic of discussion in academic and policy circles. Because it has created a risk-based, innovation-friendly vision of AI governance and begun to apply this logic to security-related fields, Singapore stands out in this broader global discourse. Thus, analyzing Singapore's approach from 2020 to 2025 offers a targeted lens through which small, highly digitalized states try to strike a balance between national security objectives and responsible AI governance.

A growing body of scholarship on AI governance and ethics, risk-based regulation, and AI and national security demonstrates the relevance of several strands to current debates about AI and security (Lütge & Uhl, 2021; Wirtz et al., 2022). More practitioner-oriented, Floridi et al. (2022) published a procedure called cap AI for assessing an AI system's conformity, which aims to serve companies as a governance tool to assess technologies in terms of legal compliance, ethical soundness, and technical robustness. Scholars even remark that AI ethics researchers have previously placed too much focus on the 'what' instead of the 'how' (Morley et al., 2021). However, most analyses focus on the EU, the United States, and China, and provide little examination of Singapore as a state that uses risk-based AI governance as an explicit national-security tool (Bernd et al., 2020).

The accelerating adoption of artificial intelligence in critical sectors has prompted growing scholarly and policy interest in governance frameworks that can mitigate AI-related risks while preserving the benefits of innovation (Bartneck et al., 2021; Kriebitz et al., 2022). In this context, risk-based governance combining ethical principles, institutional oversight, technical assurance, and sectoral adaptation is increasingly seen as essential for national security and systemic resilience. This literature review surveys foundational and recent works on AI governance globally and in Singapore, analyzes their contributions, and identifies research gaps that this article aims to address.

The broad debate around AI governance has generated a variety of frameworks and models aimed at embedding ethics, accountability, and risk management into AI deployment (Allahrakha, 2024). A recent systematic review of the AI governance literature highlights that governance efforts typically address who governs, what elements are governed, when in the AI lifecycle governance occurs, and how it is implemented (Cheng & Zeng, 2022; Cohen & Suzor, 2024).

More technical and standards-oriented perspectives argue that governance must not remain at the level of principles alone. For example, a proposed "roadmap to society's trust" suggests that responsible AI systems should be anchored in four interlinked dimensions: regulatory context; trustworthy AI technologies and standardization; auditability and accountability; and governance processes, thereby ensuring holistic oversight across technical, social, and institutional domains. Moreover, recent scholarship calls for harmonization between international standards and national/regional regulatory contexts. For instance, a 2025 study proposes a "Comparative Risk-Impact Framework" that aligns ISO AI standards with diverse regulatory environments, underscoring the importance of context-aware standardization and risk management.

These global and conceptual works provide a theoretical and normative foundation: robust AI governance should combine ethical principles, enforceable standards, technical assurance, and institutional accountability. However, they also underscore a challenge: governance frameworks that are too abstract risk remaining symbolic, while overly technical ones may lack social or institutional legitimacy. This tension motivates empirical and context-specific studies, such as governance implementation in particular states. This article assesses how Singapore's risk-based approach to AI governance contributed to national security resilience between 2020 and 2025. The object of the study is Singapore's national security strategy in the context of accelerated digitalisation and the expanding role of artificial intelligence. The subject is the set of Singapore's risk-based AI governance instruments and their function in reinforcing national security resilience.

The study pursues four objectives: (1) to conceptualise the linkages between AI governance, risk-based regulation, and national security; (2) to map Singapore's key AI governance initiatives from 2020 to 2025; (3) to analyse how these instruments operationalise a risk-based regulatory logic; and (4) to examine their alignment with Singapore's broader security and defence strategies.

Methodologically, the study employs qualitative document analysis, comparative assessment, and analytical synthesis, drawing on governance theory, risk-regulation concepts, and national-security resilience frameworks. The working hypothesis is that Singapore's risk-based AI governance contributes to national security primarily by strengthening digital resilience, institutional trust, and public–private coordination, rather than through rigid regulatory control.

A pioneering example of national-level AI governance is provided by the Model AI Governance Framework, developed by the IMDA and the Personal Data Protection Commission (hereinafter – PDPC). The first edition was released in January 2019; its second edition followed on 21 January 2020. The 2020 revision refined the 2019 edition by strengthening the implementability of the guidelines: it emphasizes internal governance structures, clear accountability, data governance, risk-based operational measures (such as bias mitigation, robustness, reproducibility), and stakeholder communication. Importantly, the Model Framework is technology and sector-agnostic, allowing it to complement but not replace sectoral/regulatory requirements when appropriate.

To help organizations operationalise the Model Framework, the PDPC/IMDA published complementary documents: an Implementation and Self-Assessment Guide for Organisations (ISAGO), a Compendium of Use Cases, and a "Guide to Job Redesign in the Age of AI." These support practical alignment with the Framework across different contexts. Through these measures, Singapore moved from abstract AI ethics toward a living governance ecosystem, one that offers concrete, actionable practices for organisations deploying AI.

Recognising that principles and guidelines alone may be insufficient, especially for high-risk or sensitive AI applications, Singapore's governance architecture includes a technical verification layer: AI Verify. First released in May 2022, AI Verify is described as "the world's first AI governance testing framework and toolkit", combining technical tests and process-based checks to assess compliance with 11 governance principles: transparency, explainability, reproducibility/robustness, safety, security, fairness, data governance, accountability, human agency and oversight, inclusive growth, and societal/environmental well-being.

By providing objective, reproducible testing and governance reports, AI Verify helps organisations demonstrate that their AI deployments align with their claimed principles, enabling transparency and accountability in practice. As AI technology evolved, especially with the rise of generative models, Singapore updated its governance approach. In January 2024, the AI Verify Foundation, together with IMDA, proposed a Model AI Governance Framework for Generative AI (MGF-GenAI) to extend governance coverage to the unique risks posed by generative AI. This evolution illustrates Singapore's adaptive, risk-based approach: governance is not static but dynamically updated to respond to technological developments and emerging risks.

Despite these advancements, academic and policy literature cautions against over-reliance on principles or voluntary frameworks. The 2025 systematic review of AI governance frameworks finds that while many frameworks exist, they vary widely in scope, coverage, and enforceability; there is no consensus on which combination of principles, tools, and processes constitutes "good governance". Moreover, the "roadmap to trustworthy AI" approach argues that technical and institutional mechanisms of regulation, standardization, and auditability must operate together, but notes that many existing practices remain fragmented or voluntary.

In a global comparison, a 2025 study that aligns ISO AI standards with multiple national regulatory frameworks finds that voluntary standards often lack enforcement mechanisms and may fail to address region-specific risks, such as data privacy, social context, or national security. This underscores a core critique: voluntary or principle-only governance may be insufficient for high-stakes AI applications, especially where national security or critical infrastructure is involved. For Singapore specifically, publicly available reports and documentation do not appear to offer systematic, empirical evaluations of how widely and effectively organisations adopt the Model Framework or AI Verify. There is limited academic research assessing whether the use of AI Verify correlates with a lower incidence of bias, security breaches, or other AI-related harms in critical sectors. This gap complicates assertions about the real-world effectiveness of

Singapore’s governance approach. Finally, while frameworks like MGF-GenAI attempt to anticipate new risks, the dynamic, evolving nature of AI rapid innovation, cross-border deployment, and supply-chain dependencies may outpace governance updates, leaving unforeseen vulnerabilities, especially in national-security relevant domains.

Materials and Methods

This study investigates how Singapore’s national security is reinforced through a risk-based model of artificial intelligence governance between 2020 and 2025. The methodological design was developed to capture both the structural components of Singapore’s governance system and the mechanisms through which AI-related risks are classified, mitigated, and incorporated into broader national security strategies. The research is grounded in the assumption that the systematic deployment of regulatory and technical assurance instruments contributes to reducing technological vulnerabilities, strengthening critical infrastructure resilience, and reinforcing institutional accountability. At the same time, the study explicitly acknowledges several challenges, including regulatory gaps associated with rapidly evolving AI capabilities, partial reliance on private-sector compliance, and persistent tensions among innovation, ethical governance, and security imperatives. To address the complexity of these dynamics, the research was structured into several sequential stages. Each stage applied distinct analytical procedures and methodological approaches, allowing for a comprehensive and multi-layered examination of Singapore’s AI governance model and its security implications.

Stage 1. Systematic literature review and conceptual grounding

The first stage consisted of a structured and systematic examination of academic and policy literature related to AI governance, risk management, and national security. The primary objective of this stage was to identify existing conceptual frameworks, establish the theoretical foundations of the study, and assess the relevance of international research to Singapore’s governance model.

The literature search focused on peer-reviewed journal articles, government publications, and policy analyses that addressed four core thematic areas: (1) risk-based approaches to AI management; (2) governance mechanisms for high-risk AI applications; (3) the interaction between AI systems and national security; (4) models of digital and institutional resilience in technologically advanced states.

Sources were identified using major academic databases, including Scopus, Web of Science, and SSRN. These databases were selected for their high coverage of peer-reviewed research in technology governance, security studies, and public policy. In addition, materials from authoritative international organizations such as the Organization for Economic Cooperation and Development (hereinafter – OECD) and institutions of the European Union were included to ensure alignment with global regulatory standards.

To ensure both relevance and currency, the review covered publications issued between 2019 and 2024, a period that corresponds to the consolidation of risk-based regulatory models and the global acceleration of AI deployment. Keywords used for the literature search included combinations of the following terms: AI governance, risk-based regulation, national security, digital resilience, AI assurance, algorithmic accountability, and critical infrastructure protection. This stage also included a comparative review of international regulatory frameworks, most notably the OECD AI Principles and the European Union's risk-tiered AI regulatory approach. These frameworks were used as reference points to assess the normative positioning of Singapore's governance model within global discourse. Through this comparison, the study identified key governance dimensions that informed subsequent analysis.

Based on the insights extracted from the literature, three analytical dimensions were formulated and used consistently throughout the study: (1) governance tools and regulatory instruments; (2) risk classification logic and mitigation strategies; (3) the institutional and strategic links between AI governance and national security outcomes. These dimensions enabled the development of a coherent conceptual framework for analyzing how Singapore integrates security considerations into civil sector AI regulation.

Stage 2. Collection and selection of primary and secondary data

The second stage involved the systematic collection of documentary materials that are central to understanding Singapore's AI governance architecture. This stage relied primarily on document-based qualitative research, which is well-suited for policy-oriented and institutional analyses of governance systems. The primary sources comprised official national strategies, regulatory frameworks, and technical assurance documents issued by Singapore's public authorities. These included: The National AI Strategy (2019); The National AI Strategy 2.0 (2023); The Model AI Governance Framework (Second Edition, 2020); The Model Framework for Generative AI (2024); Official AI Verify documentation and technical guidelines; Sector-specific regulatory guidelines issued by the Monetary Authority of Singapore (MAS), including the FEAT and FAIR principles and the Veritas toolkit.

These materials were selected because they collectively define Singapore's national approach to AI governance, its classification of AI-related risks, procedural safeguards, and the allocation of institutional responsibilities. The documents also provide empirical evidence of how security-relevant principles such as robustness, accountability, and operational reliability are embedded into regulatory practice. Secondary sources included analytical publications by international organisations, policy think tanks, and academic research centres that examine Singapore's digital governance model in the broader context of global regulatory trends. These sources were used to contextualise national policy decisions within comparative governance debates and to identify areas of convergence or divergence between Singapore and other jurisdictions.

All sources were screened using three key criteria: (1) direct relevance to AI risk governance or national security; (2) institutional credibility of the issuing body; (3) analytical depth and empirical grounding. This rigorous screening ensured that the dataset remained focused, reliable, and aligned with the study's objectives.

Stage 3. Qualitative document analysis and thematic coding

The third stage consisted of a detailed qualitative analysis of the collected documents using a structured thematic coding approach. This stage was designed to extract governance-relevant content, identify regulatory patterns, and assess how security principles are operationalised across policy instruments. The analysis proceeded in several sequential steps. First, all policy documents and regulatory frameworks underwent close reading, during which key governance elements, including regulatory mandates, oversight mechanisms, enforcement tools, and security-related provisions, were manually identified and extracted. Second, a manual coding scheme was developed based on the three analytical dimensions identified in Stage 1. This coding scheme was structured around the following thematic categories: AI-related risks identified by policymakers; sector-specific and cross-sectoral risk mitigation instruments; institutional responsibilities and interagency coordination mechanisms; explicit references linking AI governance to digital defence and national security objectives; and procedural frameworks for testing, verification, and technical assurance. Third, an iterative coding process was applied to the documents. This involved multiple rounds of coding to refine analytical categories, identify patterns of consistency or divergence, and detect gaps between formal regulatory intentions and practical enforcement mechanisms.Through this process, the study reconstructed how Singapore operationalises principles of responsible, safe, and secure AI across different regulatory layers. The thematic analysis also made it possible to trace how technical assurance mechanisms such as AI Verify function as bridges between high-level governance principles and real-world system deployment.

Stage 4. Comparative and integrative analysis

The fourth stage employed comparative analysis to evaluate Singapore's AI governance model against both internal and international reference points. This dual comparative design was essential for identifying the internal coherence of Singapore's regulatory system as well as its positioning within global governance trends. Internally, the study compared high-level strategic documents such as the National AI Strategy 2.0 with sector-specific regulatory frameworks and technical assurance mechanisms. This comparison examined how abstract governance principles are translated into operational tools within different sectors, especially finance and public services. Special attention was given to how MAS guidelines interact with national-level strategies and how AI Verify complements sectoral regulation.

This internal comparison revealed the structure of Singapore's so-called "risk to resilience" governance pipeline, demonstrating how risk identification, regulatory classification, assurance testing, and institutional oversight function as mutually reinforcing components. Externally, Singapore's model was compared with international frameworks, particularly the OECD AI Principles and the EU's risk-based regulatory approach. This international comparison allowed the identification of key areas of regulatory convergence, such as transparency and accountability requirements, as well as areas of divergence, particularly in Singapore's stronger emphasis on technical testing infrastructure and whole-of-government coordination. This comparative component enabled the identification of unique features of Singapore's governance architecture, especially its integration of national-security considerations into predominantly civil and commercial regulatory domains.

Stage 5. Synthesis, validation, and assessment of limitations

The final stage consisted of synthesising findings across all previous stages and validating analytical interpretations through cross-source triangulation. Data extracted from academic literature, official policy documents, and international analytical reports were jointly reviewed to ensure internal consistency and methodological robustness. This synthesis stage enabled the construction of a cohesive analytical narrative linking governance mechanisms with national security outcomes. It allowed the study to move beyond descriptive analysis and toward an integrated assessment of how Singapore's AI risk governance contributes to technological resilience, institutional accountability, and security assurance.

The study explicitly acknowledges several methodological limitations. First, the analysis is based exclusively on open-source materials, without access to classified defence documents. Second, rapid advancements in AI technologies generate temporal constraints on the durability of the findings. Third, the absence of expert interviews limits insight into internal policymaking processes and strategic deliberations. Despite these limitations, the multi-stage design ensures a methodologically robust foundation for assessing Singapore's evolving AI governance ecosystem.

In accordance with the journal's ethical standards, artificial intelligence (AI) tools were used strictly within acceptable limits. Artificial Intelligence Technology (ChatGPT, version 5.1) it was used exclusively to improve the language, correct errors and increase clarity, as well as to organize references and verify the consistency of citations. Artificial intelligence tools were not used to interpret data, develop scientific arguments, draw conclusions, or create new research content. The author bears full responsibility for the accuracy, integrity and originality of the research manuscript.

Results

The empirical findings reveal that Singapore's AI governance architecture between 2020 and 2025 is built upon four interdependent pillars: the National AI Strategy 2.0, the Model AI Governance Framework, the AI Verify assurance ecosystem, and a set of sectoral and security-oriented initiatives. Together, these instruments create a multilayered, risk-based governance model that operationalizes both technological innovation and national-security imperatives.

First, The National AI Strategy 2.0 (NAIS 2.0) articulates Singapore's strategic vision of AI as a "force for good," grounding its approach in three overarching systems, ten institutional enablers, and fifteen national-level actions. This structure reflects Singapore's intentional shift from a sector-specific AI policy to a comprehensive governance model that addresses societal, economic, and security dimensions simultaneously.

NAIS 2.0 highlights several national-security concerns:

(a) increasing exposure to algorithmic vulnerabilities,

(b) deepfake-driven disinformation,

(c) cross-border cyberattacks targeting critical infrastructure,

(d) geopolitical risks related to computing dependency and AI supply chains.

As emphasized in official sources (Smart Nation Singapore, 2023), the strategy embeds resilience as a foundational principle. This positions AI not only as a driver of economic transformation but also as a domain requiring defense-oriented safeguards, including secure infrastructure, trusted data flows, and unified public trust.

Second, the Model AI Governance Framework (2nd edition) serves as the ethical and procedural backbone of Singapore's AI governance ecosystem. It operationalizes human-centric values-transparency, fairness, robustness, and accountability - into institutional practices widely adopted across both public and private sectors (PDPC, 2020). Although voluntary, the Framework exerts strong normative influence due to its practical applicability and alignment with international standards. Its expansion in 2023-2024 to address generative AI technologies demonstrates regulatory adaptability in response to emerging risks (Kumar & Narayanan, 2021). This positions Singapore at the forefront of global conversations on AI governance, especially for small states requiring flexible yet credible regulatory systems.

Third, a distinctive feature of Singapore's governance model is its emphasis on assurance and validation, realized through the AI Verify testing infrastructure. As articulated in IMDA documentation (2023), AI Verify enables organizations to: assess explainability and robustness, identify vulnerabilities in AI systems, implement accountability safeguards, and benchmark practices against international standards. The establishment of the AI Verify Foundation and its rapid growth, now comprising more than 60 organizations, indicates strong domestic and international recognition of Singapore's role in shaping AI-assurance norms. For a small state reliant on global digital flows, assurance-driven governance enhances both national trust and external interoperability.

Fourth, MAS's risk-management approach operationalizes responsible AI through concrete tools designed to minimize algorithmic bias, systemic instability, and security vulnerabilities. These instruments serve national-security functions by safeguarding sensitive financial systems and preventing cascading failures (MAS, 2022). Singapore integrates AI directly into its Digital Defense pillar under Total Defense and through the Digital and Intelligence Service (DIS) established in 2022. These initiatives strengthen cyber-intelligence capabilities and support early identification of adversarial AI threats (MINDEF, 2022). Workforce development initiatives, such as Work-Learn programmes, are further enhancing resilience by building AI literacy and defensive capabilities at the societal level.

The integration of these instruments demonstrates a coherent risk-based governance approach grounded in broader theoretical frameworks of technological governance, resilience studies, and national-security strategy. Rather than adopting rigid or punitive regulation, Singapore implements a layered model: high-level ethical guidelines, technical assurance toolkits, sectoral risk-based adaptations, and national-security integration. This aligns with international scholarship emphasizing adaptive, context-sensitive AI governance models.

Figure 1 provides a consolidated representation of Singapore's risk-based AI governance architecture and illustrates how the different components identified in the empirical analysis function as an integrated system that supports national security.

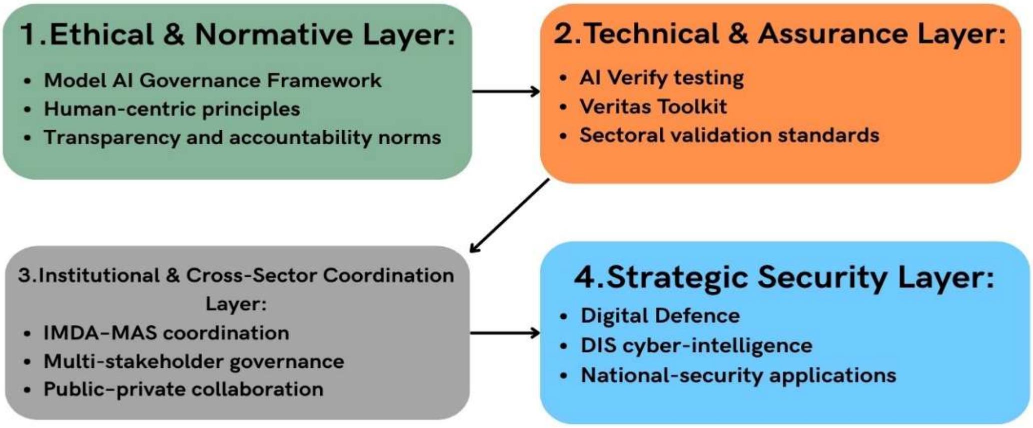
1.Ethical & Normative Layer:
• Model AI Governance Framework
• Human-centric principles
• Transparency and accountability norms
2.Technical & Assurance Layer:
• AI Verify testing
• Veritas Toolkit
• Sectoral validation standards
3.Institutional & Cross-Sector Coordination Layer:
• IMDA-MAS coordination
• Multi-stakeholder governance
• Public-private collaboration
4.Strategic Security Layer:
• Digital Defence
• DIS cyber-intelligence
• National-security applications

Figure 1. Singapore's risk-to-resilience governance pipeline

The figure visually organizes Singapore's governance instruments into four mutually reinforcing layers, demonstrating that the country does not rely on a single regulatory framework, but instead operationalizes AI oversight through a coordinated, multi-level governance pipeline.

The first layer, the Ethical and Normative Layer, reflects Singapore's emphasis on human-centric governance principles articulated in the Model AI Governance Framework. As the article's findings show, these principles form the foundation of all subsequent governance measures by embedding fairness, transparency, and accountability into AI use across sectors. This normative base ensures that risk mitigation begins at the design stage and provides the ethical anchor that guides both public and private actors.

The second layer, the Technical and Assurance Layer, corresponds to the empirical results that highlight the central role of AI Verify and the Veritas Toolkit in Singapore's governance ecosystem. These tools translate high-level principles into concrete procedures for testing, validation, and risk classification. Figure 1 makes clear that assurance mechanisms function as the operational core of Singapore's risk-based approach: they identify vulnerabilities, measure model performance, and ensure compliance with safety and robustness requirements in high-risk domains. This directly connects to the article's conclusion that Singapore prioritizes verifiable risk reduction rather than prescriptive regulation.

The third layer, Institutional and Cross-Sector Coordination, captures the interaction between IMDA, MAS, private firms, and industry consortia that the study identifies as essential for governance adaptability. The empirical analysis demonstrates that Singapore's success relies on its whole-of-government model and constant engagement with industry stakeholders. Figure 1 illustrates this coordination as a distinct functional layer, highlighting its role in aligning technical safeguards with sector-specific needs and enabling rapid policy updates as AI technologies evolve.

The fourth layer, the Strategic Security Layer, visualizes how AI governance is embedded in national-security initiatives such as Digital Defence and DIS cyber-intelligence operations. This reflects one of the study's central findings: Singapore treats AI governance not merely as a regulatory or economic matter, but as a pillar of national resilience. The figure shows how security institutions rely on the upstream layers - ethical standards, technical assurance, and institutional coordination - to support threat detection, infrastructure protection, and defense readiness. This layered integration explains why Singapore's model is particularly effective for a small, highly digitalized state facing complex cyber and geopolitical risks.

By combining these four layers into a single governance pipeline, Figure 1 enhances the interpretive clarity of the article's results. It shows that Singapore's risk-based governance is not an isolated collection of policies, but a structured system in which ethical norms, assurance mechanisms, institutional collaboration, and security strategy work together to produce national-security outcomes. The visual thus reinforces the study's central argument: Singapore's ability to align innovation, risk management, and strategic resilience arises from the coordinated interaction among governance layers rather than from traditional regulatory control.

The novelty of Singapore's model lies in its alignment of AI governance with national resilience, rather than with solely economic or innovation objectives. Unlike the EU (rights-based regulation) or the US and China (market-driven or state-security-driven models), Singapore blends flexible regulation with operational assurance, grounded in the institutional logic of Total Defense. The linkage between public trust, digital resilience, and national security is more explicit in Singapore's model than in most other small states. This makes the Singaporean case analytically significant: it demonstrates how a small, highly digitalized state uses governance capacity, rather than coercive regulation, to secure AI ecosystems and national security simultaneously.

Table 1 presents a consolidated set of indicators that illustrate how Singapore's AI governance ecosystem has expanded and matured between 2019 and 2024.

Table 1. Key quantitative indicators of Singapore's AI governance ecosystem

Governance element	Index	Year
AI adoption SMEs	From 4.2% to 14.5%	2023-2024
AI adoption non-SMEs	From 44.0% to 62.5%	2023-2024
Workers using AI	73.8%	2024
AI Verify Foundation members	>60 general	2023
Veritas consortium	From 17 to 25 organizations	2019-2020

Note: compiled by the authors based on the source (IMDA, 2022)

The most striking finding is the rapid acceleration of AI adoption across segments of the economy, demonstrating that Singapore's risk-based regulatory approach is not merely theoretical but actively shaping behavior among both large enterprises and SMEs. The growth from 4.2% to 14.5% AI adoption among SMEs, alongside a sharp rise from 44.0% to 62.5% among non-SMEs, reflects a national environment in which firms perceive AI as both accessible and strategically necessary. This pattern aligns with Singapore's broader objective under NAIS 2.0 to normalize AI across all sectors, not only among digitally intensive industries.

Equally important is the finding that 73.8% of workers already use AI tools, signaling that AI is no longer limited to specialized technical staff but has diffused into daily professional routines. This mass-level integration strengthens the country's Digital Defense posture by building a workforce capable of recognizing, managing, and responding to digital risks. It is definitely an essential component of Singapore's resilience-oriented national security strategy.

The expansion of the AI Verify Foundation, with more than 60 general members, underscores Singapore's ambition to shape global assurance standards. Its growing network mirrors how small advanced states amplify influence through niche leadership rather than military or economic scale. Similarly, the Veritas consortium's growth from 17 to 25 organizations demonstrates that risk-based governance is gaining traction within finance, a sector central to national security due to data-sensitivity and systemic exposure.

Taken together, the indicators describe a governance ecosystem that is not only expanding in scale but deepening in sophistication. Singapore's model demonstrates an ability to combine voluntary frameworks, technical assurance, and sector-specific initiatives into a coherent architecture that strengthens national security while sustaining innovation - an approach increasingly relevant to other small, highly digitalized states. A pressing question shaping contemporary national security debates is whether AI governance should rely on strict regulatory frameworks, as seen in the EU, or on a risk-based, assurance-driven approach, like Singapore's, to safeguard security while enabling innovation. For small, globally connected states such as Singapore, the challenge has been to identify a governance model that ensures digital resilience, trust, and strategic autonomy, without inhibiting economic competitiveness or technological progress (IMDA, 2023).

Singapore's risk-based model relies on a combination of soft law, public–private collaboration, and verification mechanisms that promote security through shared responsibility rather than government dominance. The development of AI Verify, internationally recognized as the first testing framework combining technical and governance assessments, demonstrates Singapore's emphasis on assurance over coercion (GovTech, 2024). This contrasts with the EU's command-and-control logic under the EU AI Act, which mandates risk classification and compliance obligations enforced through legal penalties (EU Commission, 2023). Singapore's approach tries to avoid the rigidity and compliance burden associated with such regulatory systems.

A central element of Singapore's strategy is its integration of AI governance into its broader security doctrine, particularly Total Defense, which frames national security as a whole-of-society effort. By embedding AI governance into cybersecurity, economic security, and psychological resilience strategies, Singapore positions AI as both an opportunity and a vulnerability requiring systematic management (MHA, 2022). This holistic integration marks a distinctive divergence from many Western models, which treat AI primarily as a regulatory or ethical issue rather than a fundamental aspect of national security.

The concept of the "governance pipeline" in Singapore consists of several interlinked components. First, a technical pipeline built through testing, verification, and risk-classification mechanisms, such as AI Verify, Model AI Governance Frameworks, FEAT/FAIR guidelines for financial institutions, and technical sandboxes. These systems enhance trust and allow critical sectors (finance, healthcare, transport) to adopt AI safely. Second, an institutional pipeline that involves continuous cooperation among state agencies, private-sector firms, multinational technology companies, and academic institutions. This pipeline ensures that governance adapts quickly to technological change and aligns with industry needs (IMDA-MAS Joint Report, 2024). Third, a strategic security pipeline, which integrates AI governance with counter-disinformation strategies, cyber defense operations, and critical infrastructure protection.

The core strategic question "Should we regulate AI more heavily?" may be reframed as: "Does Singapore's risk-based, soft-law pipeline provide greater national security and resilience than adopting a rigid regulatory system like the EU?" Balancing these competing models remains one of Singapore's most significant governance challenges (Tan, 2025).

Strengthening Singapore's risk-based pipeline offers several advantages:

(1) It enhances international credibility by aligning with OECD transparency and accountability principles.

(2) It maintains innovation agility in a fast-changing technological environment.

(3) It reinforces digital trust, a cornerstone of Singapore's national security strategy.

Adopting a strictly regulated system similar to the EU AI Act would strengthen Singapore's international alignment with major economic blocs and offer more enforceability. However, it may significantly hinder innovation, reduce investment attractiveness, and undermine Singapore's competitiveness in global AI development. The environment in which Singapore operates, marked by geopolitical rivalry, rapid technological shifts, cybersecurity threats, and economic vulnerability to global supply chains, requires a governance model that is flexible, resilient, and strategically integrated with defence planning. Singapore's risk-based approach, supported by its strong cybersecurity institutions and its whole-of-government coordination, positions the country as a leading example of how small states can maintain autonomy amid global technological competition (Liew, 2025).

The "pipeline paradigm" in Singapore is therefore more than a regulatory model: it represents a flow of standards, institutions, verification mechanisms, and security practices connecting industry, state, and society. Singapore's decision to deepen this risk-based pipeline while also selectively aligning with international regulatory blocs provides the most realistic strategy for maintaining both national security and technological leadership. A hybrid approach allows Singapore to remain interoperable with the EU and OECD ecosystems while preserving the flexibility, speed, and innovation advantages of its soft-law governance. This hybrid strategy enhances autonomy, reduces vulnerability, and positions Singapore as a regional leader in AI governance capable of shaping emerging security norms across Asia.

Conclusion

This study examined Singapore's risk-based approach to AI governance from 2020 to 2025 and evaluated its contribution to national security. The findings confirm the hypothesis that flexible, assurance-driven AI governance can strengthen national security not by enforcing strict legal controls, but by building digital resilience, institutional trust, and whole-of-society preparedness. Singapore's model illustrates how small, technologically ambitious states can safeguard critical systems while simultaneously promoting innovation. The analysis identified several features that make Singapore's model distinct in the global governance landscape. Its reliance on collaborative, consensus-based mechanisms, the use of soft-law instruments combined with technical assurance tools, and the integration of AI oversight within broader doctrines such as Total Defense position AI governance as an essential pillar of national resilience. Unlike more prescriptive frameworks such as the EU AI Act, Singapore's system prioritizes adaptability, sectoral partnerships, and continuous testing rather than strict regulatory enforcement.

However, the study also revealed structural limitations. The voluntary nature of Singapore's assurance ecosystem may create enforcement gaps, especially for actors operating outside the cooperative environment. Additionally, the innovation security trade-off remains difficult to manage, as overly flexible frameworks risk underestimating high-risk AI misuse. Measuring the direct security impact of these policies is another persistent challenge. Despite these constraints, Singapore's approach offers meaningful lessons for other states. For countries like Kazakhstan, balancing digital modernization, security concerns, and the need for multi-vector cooperation, Singapore provides a viable template. A risk-based, partnership-oriented model allows states to enhance cybersecurity, strengthen public trust, and promote innovation without adopting overly rigid regulatory systems or compromising strategic autonomy.

Overall, Singapore demonstrates that AI governance can function as a security multiplier when embedded within a broader national-resilience strategy. By leveraging transparency, assurance, and cross-sector collaboration, states can create a stable and trustworthy digital environment capable of withstanding emerging AI-related threats.

Author Contributions

Conceptualisation and theoretical framework: AB and ZS; research design and methodology: NN; data collection and processing: AB, ZS and NN; bibliometric analysis and interpretation: AB, ZS and NN; case study analysis and visualisation: AB, ZS and NN; draft writing and manuscript structure: AB; editing and critical revision: ZS and NN; final review and approval: AB, ZS and NN. All authors have read and approved the final version of the manuscript and agreed to its publication.

Received: October 05, 2025
Revised: December 03, 2025
Accepted: December 22, 2025
Published: December 30, 2025

REFERENCES

Al-Hawawreh, M., Baig, Z.A., & Zeadally, S. (2024). AI for Critical Infrastructure Security: Concepts, Challenges, and Future Directions. IEEE Internet of Things Magazine, 7, 136-142. https://doi.org/10.1109/IOTM.001.2300181

Allahrakha, N. (2024). UNESCO's AI Ethics Principles: Challenges and Opportunities. *International Journal of Law and Policy*, *2*(9), 24-36. https://doi.org/10.59022/ijlp.225

Bartneck, C., Lütge, C., Wagner, A., & Welsh, S. (2021). *An introduction to ethics in robotics and AI* (p.117). Springer Nature.

Bernd, W. W., Weyerer, J. C., & Sturm, B. J. (2020). The dark sides of artificial intelligence: An integrated AI governance framework for public administration. *International Journal of Public Administration, 43*(9), 818–829. https://doi.org/10.1080/01900692.2020.1749851

Cheng, J., & Zeng, J. (2022). Shaping AI's Future? China in Global AI Governance. *Journal of Contemporary China, 32*, 794 - 810. https://doi.org/10.1080/10670564.2022.2107391

Cohen, T., & Suzor, N.P. (2024). Contesting the public interest in AI governance. *Internet Policy Review, 13*(3). https://doi.org/10.14763/2024.3.1794

Floridi, L., Holweg, M., Taddeo, M., Amaya Silva, J., Mökander, J., & Wen, Y. (2022). capAI-A Procedure for Conducting Conformity Assessment of AI Systems in Line with the EU Artificial Intelligence Act. Available at SSRN 4064091. https://doi.org/10.2139/ssrn.4064091

Infocomm Media Development Authority (IMDA). (2022). *Annex B: Background on Singapore's AI governance work.*

Infocomm Media Development Authority (IMDA). (2023). *AI Verify Foundation: Technical overview and assurance framework.*

Kriebitz, A., Max, R., & Lütge, C. (2022). The German Act on Autonomous Driving: why ethics still matters. *Philosophy & Technology, 35*(2), 1-13. https://doi.org/10.1007/s13347-022-00526-2

Kumar, S., & Narayanan, A. (2021). Human-centric AI governance in Asia. *AI & Society.*

Lütge, C., & Uhl, M. (2021). *Business Ethics: An Economically Informed Perspective*. Oxford University Press, USA

Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2019). From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices. *Science and Engineering Ethics, 26*, 2141 - 2168. https://doi.org/10.1007/s11948-019-00165-5

OECD. (2020). *Singapore's model framework to balance innovation and trust in AI*. OECD.AI Policy Observatory. Retrieved September 22, 2025, from https://oecd.ai/en/wonk/singapores-model-framework-to-balance-innovation-and-trust-in-a

Personal Data Protection Commission (PDPC). (2020). *Model AI governance framework* (2nd ed.).

Personal Data Protection Commission (PDPC). (2020). *Singapore's approach to AI governance*.

Radu, R. (2021). Steering the governance of artificial intelligence: National strategies in perspective. *Policy and Society, 40*(2), 178–193. https://doi.org/10.1080/14494035.2021.1929728

Smart Nation Singapore. (2023). *National AI Strategy 2.0: AI for the public good, for Singapore and the world*. Government of Singapore. https://file.go.gov.sg/nais2023.pdf

Taddeo, M., Ziosi, M., Tsamados, A., Gilli, L., & Kurapati, S. (2022). Artificial intelligence for national security: The predictability problem. *Centre for Digital Ethics (CEDE) Research Paper No. Forthcoming.*

Wirtz, B. W., Weyerer, J. C., & Kehl, I. (2022). Governance of artificial intelligence: A risk and guideline-based integrative framework. *Government Information Quarterly, 39*, 101685. https://doi.org/10.1016/j.giq.2022.101685

Yerlikaya, S., & Erzurumlu, Y.Ö. (2021). Artificial Intelligence in Public Sector: A Framework to Address Opportunities and Challenges. *Studies in computational intelligence, 935*, 201-216. https://doi.org/10.1007/978-3-030-62796-6_11

Information about the authors

***Aigerim Azatbekova** – Bachelor, Turan University, Almaty, Kazakhstan. Email: a.azatbekova@gmail.com, ORCID ID: https://orcid.org/0009-0000-6921-9359
Zere Serikbayeva – Bachelor, Turan University, Almaty, Kazakhstan. Email: zereserikbaeva1@gmail.com, ORCID ID: https://orcid.org/0009-0003-5639-9073
Nurbolat Nyshanbayev – PhD, Associate Professor, Turan University, Almaty, Kazakhstan. Email: n.nyshanbayev@turan-edu.kz, ORCID ID: https://orcid.org/0000-0001-7563-2254

Авторлар туралы мәліметтер

***Азатбекова Ә.** – бакалавр, Тұран университеті, Алматы, Қазақстан. Email: a.azatbekova@gmail.com, ORCID ID: https://orcid.org/0009-0000-6921-9359
Серікбаева З. – бакалавр, Тұран университеті, Алматы, Қазақстан. Email: zereserikbaeva1@gmail.com, ORCID ID: https://orcid.org/0009-0003-5639-9073
Нышанбаев Н. – PhD, қауымдастырылған профессор, Тұран университеті, Алматы, Қазақстан. Email: n.nyshanbayev@turan-edu.kz, ORCID ID: https://orcid.org/0000-0001-7563-2254

Сведения об авторах

***Азатбекова А.** – бакалавр, Университет Туран, Алматы, Казахстан. Email: a.azatbekova@gmail.com, ORCID ID: https://orcid.org/0009-0000-6921-9359
Серикбаева З. – бакалавр, Университет Туран, Алматы, казахстан. Email: zereserikbaeva1@gmail.com, ORCID ID: https://orcid.org/0009-0003-5639-9073
Нышанбаев Н. – PhD, доцент, Университет Туран, Алматы, Казахстан. Email: n.nyshanbayev@turan-edu.kz, ORCID ID: https://orcid.org/0000-0001-7563-2254