The agentic AI security adoption matrix: Understanding readiness and resistance across domains

Guy WAIZEL,
Cato Networks, Israel
guy.waizel@gmail.com

Abstract
The objective of this paper is to examine how agentic artificial intelligence (AI) is being adopted in practice, with particular attention to its security implications, and to identify domains where adoption is progressing more slowly due to regulatory, ethical, and oversight requirements. This research is important because agentic AI offers new efficiencies and autonomy but simultaneously introduces risks related to trust, accountability, and adversarial exploitation. Prior work on autonomous systems, AI governance, and security automation provides a foundation, and recent studies highlight a contrast between rapid adoption in digitally bounded, low-stakes environments and slower uptake in safety-critical contexts. Building on this foundation, the paper adopts a systematic literature review combined with comparative domain analysis to classify adoption trends. The approach draws on academic publications, industry reports, and regulatory frameworks such as the NIST AI Risk Management Framework and the EU AI Act, which explicitly designate domains like healthcare, defense, and social care as high-risk. The results indicate that agentic AI has advanced most quickly in areas such as customer service, software engineering assistance, and cybersecurity operations centers, where regulatory barriers are minimal and oversight is straightforward. By contrast, adoption remains constrained in healthcare, defense, and social care, where pilot projects exist but mainstream deployment is slowed by requirements for explainability, liability, and human-in-the-loop controls. The implications of these findings are significant for academics extending adoption models, researchers designing security frameworks, and practitioners balancing innovation with governance. The value of this paper lies in presenting a conceptual "Agentic AI Security Adoption Matrix," which offers an original perspective on how adoption speed and security sensitivity interact, and provides guidance on where agentic AI may thrive versus where its deployment will require cautious, regulated progression.

Keywords: adoption readiness, security risks, autonomous systems, trust governance.

1. Introduction
Artificial intelligence (AI) has evolved rapidly from predictive analytics to autonomous decision-making systems. Li [1] and Acha [2] investigated multi-agent learning and secure coordination, respectively, demonstrating how decentralized intelligence can improve decision resilience when supported by robust communication protocols. From a theoretical standpoint, Alanen's Creation–Disruption Theory [3] describes how innovation both transforms and destabilizes existing markets. When applied to agentic AI, this framework explains the simultaneous efficiency gains and governance challenges accompanying autonomous adoption. Gupta [4] advanced large language model (LLM) optimization methods that enhance reasoning and adaptability as technical foundations for agentic systems. Sowmyanarayanan [5] offered a maturity model that charts enterprise readiness for agentic AI, emphasizing governance, interoperability, and ethical oversight as critical prerequisites.

Recent developments in LLMs and multi-agent frameworks have accelerated the transformation toward what researchers term *agentic AI*—AI systems capable of pursuing goals, adapting dynamically, and coordinating autonomously with other agents [6], [7], [8]. These developments have generated both technological excitement and regulatory concern

as organizations strive to balance innovation with oversight [9, 10, 11, 12, 13, 14]. Sapkota *et al.* [6] distinguished between simple AI agents and agentic ecosystems, collaborative, memory-augmented systems characterized by self-directed reasoning. Schneider [9] mapped the evolution from generative to agentic AI, highlighting how these systems expand autonomy and collaborative potential beyond traditional outputs. Hosseini and Seilani [15] and Pati [8] further explored how these agents enhance decision efficiency while introducing explainability and accountability challenges.

At the regulatory level, frameworks such as the NIST AI Risk Management Framework [9], NIST Cybersecurity Framework 2.0 [10], the EU AI Act [12], ISO/IEC 42001 [13], and ISO/IEC 23894 [14] have become cornerstones for trustworthy AI deployment. These frameworks classify AI systems by risk, establish management system requirements, and outline governance procedures. Tam [16] and the Cloud Security Alliance [17] emphasized how the NIST AI RMF is being operationalized within industry to manage AI risk, while ENISA [18] and OECD [19] stress cross-sector cooperation, resilience, and human-centric design.

Enterprise adoption further illustrates this integration of policy and practice. For instance, Waizel and Fried [20] proposed the Cato MCPSaaS framework, which embeds agentic AI within secure, network-native supervision environments to ensure scalable and compliant automation. Waizel et al. [21] demonstrated how adversaries could exploit model context protocols, offering mitigation strategies for "living off AI" threats through access control and anomaly monitoring. Waizel and Fershtman [22] presented the Cato API Assistant, showcasing safe, contextual automation in API management. Finally, Waizel et al. [23] described the Cato MCP Server as a unifying control layer for AI-driven orchestration and telemetry, complementing governance frameworks through centralized visibility and enforcement.

Taken together, these studies indicate that agentic AI adoption is advancing rapidly across enterprise and digital operations contexts while remaining cautious in safety-critical sectors. The convergence of academic theory, regulatory frameworks, and enterprise practice underscores the need for a structured model, leading to this paper's proposed Agentic AI Security Adoption Matrix.

2. Methods

2.1 Research design

A systematic literature review (SLR) and comparative domain analysis were used to identify adoption patterns, readiness levels, and governance maturity. The approach combined peer-reviewed academic publications, doctoral dissertations, regulatory documents, and enterprise whitepapers and technical blog posts to capture both theoretical and applied perspectives. The methodology followed a structured coding and synthesis process designed to ensure coverage, reliability, and traceability across diverse data types.

2.2 Data collection and scope

Sources were identified through databases including IEEE Xplore, Scopus, arXiv, and ResearchGate, alongside regulatory repositories such as NIST, ENISA, ISO, and the

Official Journal of the European Union. Enterprise and analysts publications were also incorporated. The search was restricted to materials published between 2021 and 2025, focusing on themes such as:

Agentic AI systems and multi-agent collaboration; Governance, ethics, and security frameworks; Domain-specific adoption (e.g., healthcare, defense, cybersecurity, enterprise IT).

51 distinct sources were selected, encompassing 27 academic publications, 9 regulatory or standardization frameworks, and 15 enterprise/industry reports. Each source was reviewed to extract insights relevant to readiness, risk, and governance characteristics.

2.3 Coding and analysis

Each source was coded according to thematic categories, including: (1) autonomy and agentic structure, (2) governance and control mechanisms, (3) adoption barriers, (4) regulatory frameworks, and (5) implementation case studies. Coding was conducted manually using structured matrices that aligned each source with one or more domain categories (e.g., healthcare, cybersecurity, public administration).

The process was iterative and comparative, sources were cross-analyzed to identify convergence or divergence in findings. Attention was paid to the interplay between technological readiness and regulatory intensity, allowing the classification of each domain into readiness and security-sensitivity quadrants.

The coding framework was developed based on the full corpus of abstracts and studies listed in the references [24, 1, 2, 25, 26, 4, 5, 6, 7, 15], [27, 28, 29, 30, 31, 32, 33, 34], [35, 36, 37, 38, 39, 40, 41, 42, 16], [17, 43, 9, 11, 10, 12, 14, 13], [18, 19, 9, 44, 45, 46, 47, 48, 49],[50, 23, 20, 22, 21, 51]. The approach ensured that all sources were considered consistently, with representative insights extracted for domain-level synthesis. Each coding decision was verified through iterative review to maintain objectivity and alignment with the paper's research objectives.

3. Results

The synthesis of findings revealed distinct adoption patterns of agentic AI across domains, shaped by regulatory intensity, operational criticality, and organizational readiness. Analysis of the coded corpus showed a clear bifurcation between low-risk digital service sectors, where deployment has accelerated, and high-risk socio-technical sectors, where adoption remains cautious.

3.1 Cross-domain adoption trends

Academic, regulatory, and enterprise data indicated that agentic AI integration follows a predictable maturity gradient. Enterprise environments with established automation practices, such as software engineering, cybersecurity operations, and customer service, display rapid uptake because these contexts permit bounded autonomy and low regulatory friction [37, 38, 39, 40, 41, 42]. In contrast, high-risk sectors such as healthcare, defense, and social care remain governed by strict oversight requirements related to explainability,

liability, and human-in-the-loop controls [16, 17, 43, 9]. The results of this analysis are synthesized and illustrated in the proposed Agentic AI Security Adoption Matrix, which maps adoption maturity against regulatory exposure and operational risk (Fig. 1).

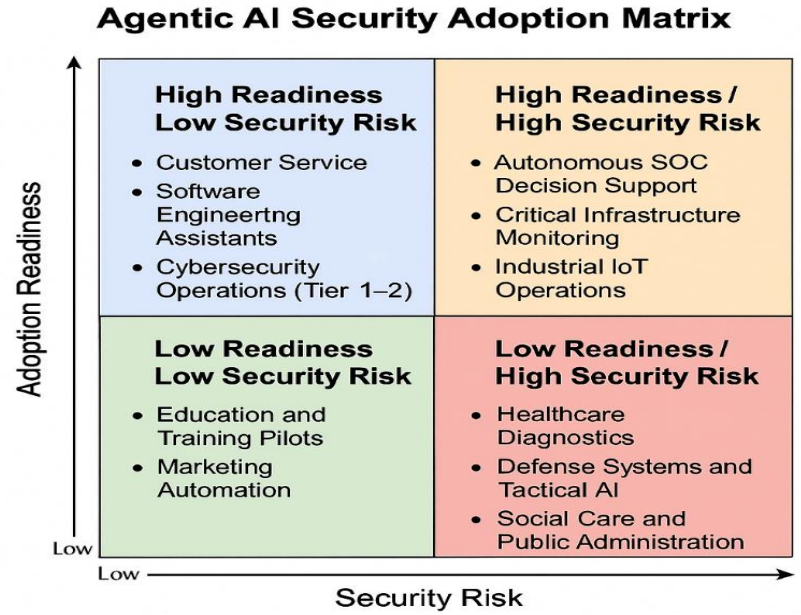

Fig. 1. Agentic AI Security Adoption Matrix
Source: Synthesis based on coded literature [1] - [51]

3.2 Governance and security alignment

Governance analysis revealed that most frameworks converge on layered control structures that span policy, organizational, technical, operational, and infrastructural domains [8, 27, 28, 29, 30, 31, 32, 33, 34, 35], [10, 11, 12, 14, 13]. The Layered Governance Model (Fig. 2) was created based on regulatory and enterprise frameworks and displays this vertical integration, showing how each layer anchors compliance and assurance mechanisms within enterprise agentic-AI systems.

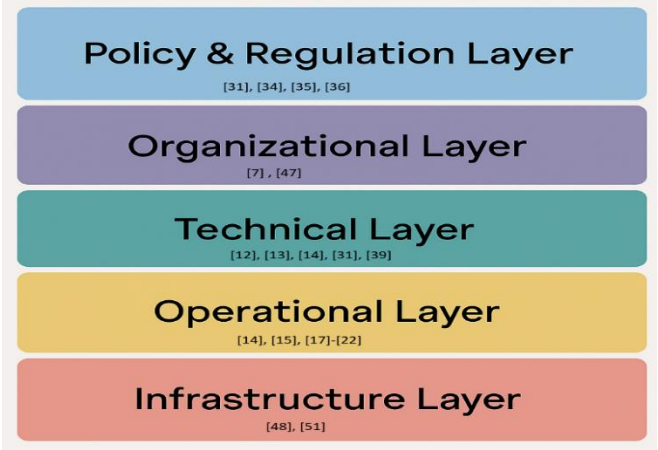

Fig. 2. Layered Governance Model
Source: Based on regulatory and enterprise frameworks

3.3 Domain-wise summary of agentic AI adoption readiness

To compare adoption dynamics across key sectors, the study synthesized domain-level evidence to capture readiness patterns, governance characteristics, and observed use cases of agentic AI, as summarized in Table 1.

Table 1. Domain-Wise Summary of Agentic AI Adoption Readiness

Domain	Key Findings	References
Customer Service and Chat Automation	Demonstrates rapid adoption with high autonomy and minimal oversight requirements.	[1], [8], [9], [11], [17]
Software Development Assistants	LLM-based tools enhance productivity and consistency, showing high readiness.	[6], [16], [41], [42], [43]
Cybersecurity Operations Centers	Adoption focused on automated triage, SOC optimization, and adaptive governance.	[14], [15], [17], [18], [22]
Healthcare Applications	Adoption constrained by regulatory demands for explainability and human oversight.	[27], [28], [31], [34]
Defense and Aerospace	High strategic potential but subject to restricted testing and oversight.	[28], [34], [45]
Social Care and Public Administration	Slow uptake due to ethical sensitivity and institutional inertia.	[23], [25], [26]
Education and Training Systems	Moderate readiness with pilot deployments emphasizing safety and traceability.	[24], [38]
Marketing and Analytics	Increasing experimentation under low governance intensity.	[24], [25]

Source: Aggregated results from coded dataset [1] - [51]

3.4 Aggregate domain readiness and security sensitivity

To consolidate these findings, adoption readiness was evaluated alongside each domain's security sensitivity and governance intensity, providing a comparative view of systemic maturity and control rigor, as displayed in Table 2.

Table 2. Aggregate Domain Readiness and Security Sensitivity

Domain	Adoption Readiness	Security Sensitivity	Governance Intensity
Customer Service & Chat Automation	High	Low	Moderate
Software Development Assistants	High	Low	Low
Cybersecurity Operations Centers	High	High	High
Healthcare Applications	Low	High	Very High
Defense and Aerospace	Low	High	Very High
Social Care and Public Administration	Low	High	High
Education and Training Systems	Moderate	Low	Moderate
Marketing and Analytics	Moderate	Low	Low

Source: Aggregated results from coded dataset [1] - [51]

3.5 Framework intersections and application contexts

The analysis further mapped how major regulatory and governance frameworks intersect with domain-specific applications, illustrating the alignment between policy intent and practical implementation contexts, as shown in Table 3.

Table 3. Framework Intersections and Application Contexts

Framework	Core Focus	Mapped Domains
NIST AI RMF (2023)	Trustworthy AI governance and risk management	Enterprise AI, Cybersecurity Operations, Healthcare
EU AI Act (2024)	Regulatory risk classification and compliance	Healthcare, Defense, Public Administration
ISO/IEC 42001 (2023)	AI Management System requirements	Enterprise Governance, Cross-domain Adoption
ISO/IEC 23894 (2023)	AI Risk Management guidelines	Regulated Industries and Critical Infrastructure
ENISA Threat Landscape (2023)	Cybersecurity threats and resilience	Cybersecurity Operations, Industrial IoT
OECD AI Principles (2024)	Human-centric design and transparency	Public Policy and Cross-border Governance

Source: Aggregated results from coded dataset [1] - [51]

3.6 Summary of findings

The cross-analysis underscores that agentic AI readiness tends to increase with regulatory maturity and declines in sectors characterized by higher operational sensitivity. High-readiness, low-risk domains demonstrate rapid operationalization of autonomous workflows, while safety-critical sectors proceed with constrained pilot programs. The structured layering of governance ensures that technical autonomy remains bounded by transparent oversight mechanisms. These results validate the conceptual Agentic AI Security Adoption Matrix as a heuristic for assessing domain-specific adoption potential and guiding future research on secure autonomy scaling.

4. Discussion

The findings of this study reveal a clear divide between rapid adopters of agentic AI, largely within digitally bounded enterprise domains, and more cautious implementers in safety-critical sectors. This divide reflects both organizational readiness and governance maturity.

4.1 Governance-driven adoption

Domains characterized by established risk management frameworks demonstrate faster and safer adoption trajectories. The integration of NIST AI RMF [9], EU AI Act [12], and ISO/IEC [14], [13] standards underpins the formalization of governance practices that enable responsible innovation.

4.2 Enterprise implementation practices

Enterprise deployments such as Cato's MCPSaaS framework [20] exemplify how scalable, policy-controlled architectures can safely operationalize agentic AI. Its emphasis on network-native supervision and policy enforcement provides a model for embedding autonomy within compliance boundaries. Complementing this, Waizel et al. [23] described the Cato MCP Server as a unifying control layer for orchestration, policy distribution, and telemetry—ensuring transparent observability across AI-driven operations. This separation between framework and control layer reinforces secure autonomy while maintaining governance integrity.

4.3 Domain implications

In cybersecurity, agentic AI demonstrates immediate value through automated triage and adaptive response capabilities [30], [32, 33, 34, 35, 36, 37]. Conversely, in healthcare and defense, explainability, liability, and ethical governance remain barriers. The EU AI Act [12] and ISO/IEC frameworks [14], [13] explicitly classify these as high-risk domains, mandating human oversight and impact assessments.

4.4 Regulatory synchronization

Regulatory harmonization across NIST [9], ENISA [18], and OECD [19] guidance documents suggests an emerging consensus around layered governance and risk-based deployment. This global alignment supports interoperability across sectors, enabling innovation without compromising trust or accountability.

4.5 Theoretical implications

From an academic standpoint, the study extends the adoption models introduced by Alanen [3] and Sowmyanarayanan [5], integrating them with empirical insights from regulatory and enterprise ecosystems. The Agentic AI Security Adoption Matrix provides a conceptual tool that aligns with Creation–Disruption Theory, offering predictive capacity for assessing where agentic AI can responsibly scale.

4.6 Practical implications

For practitioners, these findings inform architecture design, risk mitigation, and adoption strategy. The alignment between layered governance (Fig. 2) and domain readiness (Table 2) demonstrates how regulatory depth and infrastructural maturity jointly determine safe adoption velocity. Cross-domain collaboration between regulators, academia, and industry is essential to maintain equilibrium between innovation and safety. Future research should focus on longitudinal assessment of agentic AI deployment outcomes, particularly as adoption expands into safety-critical and high-liability contexts.

4.7 Limitations

This research is situated at the early stage of the agentic AI era, during which empirical evidence and large-scale implementations remain limited. While the systematic review covered 51 sources spanning academic, regulatory, and industrial literature, many studies examined conceptual frameworks or pilot initiatives rather than longitudinal deployments. Consequently, the findings should be interpreted as indicative of current readiness rather than definitive measurements of adoption success.

Because the field is still maturing, cross-domain comparisons may reflect varying levels of documentation, regulatory evolution, and reporting transparency. Some sectors, particularly healthcare, defense, and public administration, possess fewer published evaluations of deployed systems, constraining the ability to assess post-deployment outcomes or long-term governance effectiveness.

Future research will benefit from longitudinal case studies, quantitative adoption metrics, and real-world validation across diverse domains. As agentic AI ecosystems stabilize, ongoing empirical investigation will be essential to refine the conceptual Agentic AI

Security Adoption Matrix and to evaluate its predictive validity across evolving technological and policy contexts.

References

[1] T. Li, "Computational Foundations of Multi-Agent Learning in Cyber-Physical-Human Networks Under Amorphous Information Attributes. Ph.D. Dissertation," New York University, 2025.

[2] S. N. N. Acha, "Cooperative Intelligent Control Through Reliable Trusted Communication: Enhancing Reliability and Data Pooling in Multi-Agent Systems (MAS). Ph.D. Dissertation," North Carolina A&T State University, 2025.

[3] J. Alanen, "Creation–Disruption Theory: How Innovators Create and Disrupt Markets in the Era of AI and Beyond. Doctoral Dissertation," Pepperdine University, 2025.

[4] S. Gupta, "Demonstration Selection and Task Formulation for Effective In-Context Learning. Ph.D. Dissertation," University of California, Irvine, 2025.

[5] S. Sowmyanarayanan, "Agentic AI Adoption in Enterprises: Maturity Model and Readiness Checklist," *Express Computer*, 2025.

[6] R. Sapkota and et al, "AI Agents vs. Agentic AI: A Conceptual Taxonomy, Applications and Challenges," *arXiv preprint arXiv:2505.10468*, 2025.

[7] J. Schneider, "Generative to Agentic AI: Survey, Conceptualization, and Challenges," *arXiv preprint arXiv:2504.18875*, 2025.

[8] A. K. Pati, "Agentic AI: A Comprehensive Survey of Technologies, Applications and Societal Implications," *IEEE Access*, 2024.

[9] National Institute of Standards and Technology (NIST), "Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST AI 100-1," 2023.

[10] National Institute of Standards and Technology (NIST), "The NIST Cybersecurity Framework (CSF) 2.0, NIST CSWP 29," 2024.

[11] National Institute of Standards and Technology (NIST), "NIST Cybersecurity Framework 2.0: Resource & Overview Guide, NIST SP 1299," 2024.

[12] European Union, "Regulation (EU) 2024/1689 Laying Down Harmonised Rules on Artificial Intelligence (EU AI Act)," Official Journal of the European Union, 2024.

[13] ISO/IEC, "ISO/IEC 42001:2023 — Artificial Intelligence Management System (AIMS) — Requirements," 2023.

[14] ISO/IEC, "ISO/IEC 23894:2023 — Artificial Intelligence — Guidance on Risk Management," 2023.

[15] S. Hosseini and H. Seilani, "The Role of Agentic AI in Shaping a Smart Future: A Systematic Review," *Array*, vol. 26, p. 100399, 2025.

[16] J. Tam, "United States: The Growing Importance of the NIST AI Risk Management Framework," *Baker McKenzie InsightPlus*, 2024.

[17] Cloud Security Alliance, "Implementing the NIST AI RMF: Everything You Need to Know," Cloud Security Alliance Insights, June 2025.

[18] European Union Agency for Cybersecurity, "ENISA Threat Landscape 2023," Oct. 2023.

[19] Organisation for Economic Co-operation and Development (OECD), "OECD AI Principles," 2024.

[20] G. Waizel, D. M. Attiya and S. Bamberger, "Cato CTRL™ Threat Research: PoC Attack Targeting Atlassian's Model Context Protocol (MCP) Introduces New 'Living Off AI' Risk," Cato Networks Blog, 2025.

[21] G. Waizel and A. Fershtman, "Introducing Cato's API Assistant: Your New Copilot for GraphQL," Cato Networks Blog, 2025.

[22] G. Waizel, A. Fershtman, D. Pienica and E. Plotnik, "Meet Cato's MCP Server: A Smarter Way to Integrate AI Into Your IT & Security Processes," Cato Networks Blog, 2025.

[23] G. Waizel and Z. Fried, "Secure Agentic AI Adoption in Enterprise Environments: Extending the MCPSaaS Framework with Policy-Controlled Governance," Cato Networks Blog, 2025.

[24] T. Lewandowski, "Artificial Intelligence in Organizations: Managing the Life/ycle of Conversational Agents. Doctoral Dissertation," Universität Hamburg, 2024.

[25] P. A. Olujimi, P. A. Owolawi, R. C. Mogase and E. V. Wyk, "Agentic AI Frameworks in SMMEs: A Systematic Literature Review of Ecosystemic Interconnected Agents," *AI*, vol. 6, no. 6, p. 123, 2025.

[26] J. Alanen, "Creation–Disruption Theory: How Innovators Create and Disrupt Markets in the Era of AI and Beyond. Doctoral Dissertation," Pepperdine University, 2025.

[27] S. T. Adapala and Y. R. Alugubelly, "The Aegis Protocol: A Foundational Security Framework for Autonomous AI Agents," *arXiv preprint arXiv:2508.19267*, 2025.

[28] Y. Liu and et al., "Secure Multi-LLM Agentic AI and Agentification for Edge General Intelligence by Zero-Trust: A Survey," *arXiv preprint arXiv:2508.19870*, 2025.

[29] P. Dal Cin and et al., "Three Essentials for Agentic AI Security," *MIT Sloan Management Review*, 2025.

[30] S. Hosseini and et al., "A Review of Agentic AI in Cybersecurity: Cognitive Autonomy, Adaptability, and Governance," *F1000Research*,, vol. 14, p. 843, 2025.

[31] L. Wang and et al., "A Survey on Large Language Model-Based Autonomous Agents," *arXiv preprint arXiv:2308.11432*, 2023.

[32] PwC, "The Rise of Autonomous AI in Cybersecurity," PwC Global, 2025.

[33] J. Marshall, "Agentic AI for SecOps Teams," *ReliaQuest Cyber Knowledge*, 2025.

[34] S. Manjrekar, "Implementing Agentic AI Architecture: A Technical Overview of Architecture and Frameworks," Fabrix AI Blog,, Sept. 2025.

[35] J. Alvarez and et al., "Systematic Review on Enterprise Agentic AI," ResearchGate, 2025.

[36] R. Kumar and et aL, "Agentic AI: Autonomous Intelligence for Complex Goals," ResearchGate, 2025.

[37] A. Rashid, "Agentic AI: The New Era of Autonomous SOC," DarkReading, May 2025.

[38] C. Schachtner, "Smart Government in Local Adoption – Authorities in Strategic Change through AI," *Smart Cities and Regional Development (SCRD) Journal*, vol. 5, no. 3, p. 53–62, 2021.

[39] E. Grabocka and E. Ndoka, "AI-Driven Innovation within the ICT Sector," *SCRD Journal*, vol. 9, no. 1, p. 77–97, 2025.

[40] R. M. Benshams, "Cultivating Organizational Culture for AI Integration," *SCRD Journal*, vol. 9, no. 1, 2025.

[41] D. M. Popa, "Smart States Deploy Collective Intelligence for Security and Surveillance," *SCRD Journal*, vol. 9, no. 2, 2025.

[42] J. A. Salas, M. Iqbal and L. Pallahidu, "Guarding Digital Health Data: Strengthening Healthcare Data Security in Indonesian Smart Cities," *SCRD Journal*, vol. 10, 2023.

[43] THE Journal, "Proposed NIST Cybersecurity Guidelines Aim to Safeguard AI Systems," 2025.

[44] y. Shi and et al., "Intelligent System for Automated Molecular Patent Infringement Assessment," *arXiv preprint arXiv:2412.07819v2,* 2025.

[45] P. Rasmussen and et al., "ZEP: A Temporal Knowledge Graph Architecture for Agent Memory," *arXiv preprint arXiv:2501.13956v1,* 2025.

[46] W. Xu and et al., "A-MEM: Agentic Memory for LLM Agents," *arXiv preprint arXiv:2502.12110v11,* 2025.

[47] N. Lee and et al., "RAG-Enhanced Collaborative LLM Agents for Drug Discovery," *arXiv preprint arXiv:2502.17506v2,* 2025.

[48] B. Gao and et al., "PharmAgents: Building a Virtual Pharma with Large Language Model Agents," *arXiv preprint arXiv:2503.22164v2,* 2025.

[49] Y. Xu and et al., "Scalable UAV Multi-Hop Networking via Multi-Agent Reinforcement Learning with Large Language Models," *arXiv preprint arXiv:2505.08448v1,* 2025.

[50] K. Ma, "AI Agents in Chemical Research: GVIM – An Intelligent Research Assistant System," *Digital Discovery*, vol. 4, p. 355–375, 2025.

[51] G. Waizel and Z. Fried, "Designing the Future of Agentic AI: Cato Engineering Details a New Practical, Secure, and Scalable MCPSaaS Framework," Cato Networks Blog, 2025.