A comprehensive AI security framework for generative AI systems on smart cities cloud platforms

Federik Premti
Mediterranean University of Albania, Tirana, Albania
federik.premti@umsh.edu.al

Alma Hyra
Mediterranean University of Albania, Tirana, Albania
alma.hyra@umsh.edu.al

Abstract
This article, details a comprehensive AI security framework for generative AI systems deployed on smart cities cloud platforms. By integrating principles from Google Secure AI Framework, the OWASP AI Security and Privacy Guide, and the IBM Framework for Securing Generative AI, the framework proposes a layered, risk-based approach that covers identity and access management, data security, network isolation, application security, and rigorous monitoring and audit processes. This framework provides a method for mapping risk, threat, and control throughout the AI life cycle; secure cloud environment configuration; application of architectural best practices including zero trust; and continuing micro-adjustments of artificial intelligence to mitigate adversarial threats. In addition, the framework also deals with governance, required compliance with regulation, ethical guidelines and the operational issues related to the implementation of advanced AI capabilities in enterprises. Organizations can confidently implement generative AI systems in the cloud using this comprehensive AI security framework, which provides them with a strong layer of technical controls, good governance practices and flexible risk management practices. The comprehensive AI security framework creates a safe place for an organization to run its critical assets, by providing protection against both outside and inside threats, and also creates an environment for innovation through cutting-edge AI technology while maintaining the highest level of privacy, responsibility and ethical accountability. This comprehensive security model will protect organization against known attack vectors and adapt to changing attack vectors through ongoing learning and collaboration between humans and AI, thereby enabling businesses to develop resilient, trustworthy, and ethically aligned enterprise-wide AI systems. In this article also we talk about the next stage of cloud computing, smart cloud platforms that combine elastic infrastructure, artificial intelligence, automation, edge integration, security analytics, and cloud-native architectures. This comprehensive AI security framework is also used in this smart cities environment.

Keywords: “AI Security Framework“, “Generative AI systems“, “Risk management“, “Smart cities“, “ Smart cloud“.

1. Introduction
The rapid adoption of generative AI systems across numerous industries has driven a parallel need for robust security frameworks that protect these advanced systems from evolving threats. Cloud-based AI services, given their complexity and scale, require frameworks that not only offer protection at various security domains but also ensure regulatory compliance and ethical governance. In response to this challenge, we propose a comprehensive AI security framework that merge best practices from the Google Secure AI Framework, the OWASP AI Security and Privacy Guide, and the IBM Framework for Securing Generative AI [1], [2], [3]. This article outlines the architectural components, governance models, technical controls, and operational processes essential for securing generative AI pipelines within cloud environments [2], [4]. We will also introduce you with a smart cities cloud platform, that is a cloud computing environment integrated with AI, ML, and advanced automation into the core architecture. Smart cloud platforms provide

data-based intelligence to enhance the performance of any given system using a large amount of data, to detect anomalies, to automate processes, to assist in making decisions in real time, and to increase resiliency against cyber attacks. Smart cloud platforms will therefore be critical for the development of tomorrow's smart cities, healthcare systems, education systems, financial systems, manufacturing systems, and government systems where systems have to be able to scale, be secure, be reliable, and adapt. This dynamic context necessitates a comprehensive and adaptable security strategy that acknowledges the interconnected nature of smart city infrastructure and prepares for emerging threats [29], so the proposed AI security framework in cloud platforms is used for smart cloud platforms too.

2. Methodology

A qualitative, conceptual, and framework development approach was adopted for this article to examine smart cloud platforms used in smart cities and recommend an AI-based security framework for cloud computing environments.

There are three main approaches within the methodology:

1. Literature-based analysis covering smart cities, cloud computing, IoT, edge computing, and AI security.
2. The standards-based approach using recognised Cyber Security and AI risk management frameworks.
3. The framework design will concentrate on securing AI-enabled smart city cloud platforms.

The approach is appropriate because smart city cloud platforms are complex socio-technical systems comprising technological infrastructure and governance systems; public service delivery systems; citizen data; critical infrastructures; and cyber security risks.

3.Problem Statement

The smart city landscape is increasingly utilizing various technologies such as cloud computing, IoT (internet of things), AI (artificial intelligence), edge computing, big-data analytics for delivering efficient, intelligent and citizen-focused urban services. As such, there are many opportunities for utilizing these technologies to deliver real-time monitoring capabilities, predictive decision making, automated public service delivery, intelligent mobility systems, energy management, environmental monitoring, digital governance, smart economy that represents a transformative evolution of economic systems driven by AI [30], etc.

However, creating smart cloud platforms that can be integrated with existing urban infrastructure introduces several complex challenges including technology integration security, privacy, and governance.

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 [28].

The main issue for smart cities arises from the significant amount of heterogeneous and sensitive data created and processed by their systems which come from multiple different input sources including IoT sensors, traffic systems, smart meters, surveillance cameras, public-service applications, healthcare systems, and citizen-based platforms. This type of data is stored in a cloud computing environment where there are many risks including: unauthorized access, data leakage, cloud service mis-configurations, unsecured APIs, poor identity management and cyberattacks that threaten the confidentiality, integrity and availability of smart city services.

Moreover, there is also an increasing risk of insecure, unregulated and therefore dangerous urban systems due to the implementation of artificial intelligence (AI) in the cloud computing of smart cities. AI is used to manage many urban services, such as traffic management systems, emergency services, public safety, energy distribution, and environmental monitoring. Respective to the significant risks associated with using insecure AI, if any aspect of a smart city's operation is compromised, it has broad societal and ethical ramifications. The security vulnerabilities of AI include adversarial attacks, adversarial data poisoning, model manipulation through training data, bias in decision making due to bias in training data, loss of privacy through access to personal information, and usage of unreliable automated decision making through machine-based systems.

Lastly, a lack of a comprehensive and consistent security framework for AI that relates to smart city cloud computing is a contributing factor to the increasing risks associated with many smart cities worldwide. Therefore, the majority of current city-wide smart city cloud computing platforms exist as separate entities which make them more susceptible to potential risks associated with cloud infrastructure, IoT devices, edge computing, citizen privacy, AI model development, and data governance.

The integration of multiple frameworks into AI development is instrumental in addressing diverse array of issues and concerns inherent in the field, thereby providing nuanced insights into the management of AI advancement [32].

The purpose of the article is to present a new AI Security Framework specifically designed to protect cloud-computing environments within smart cities. The proposed framework in other research provides both diagnostic and developmental tools for practitioners seeking to improve transformation outcomes [31].

4. Proposed AI security framework

4.1 Overview and design principles

A multi-layered, depth-in-depth defense system that combines conventional cybersecurity practices and controls with AI-related threat protection controls will be the basis for the proposed framework. In particular, risk identification, threat modelling, continuous monitoring, and automated remediation are the main principles of the framework.

Emphasizing the importance of pre-emptive controls, as well as rapidly adapting to new forms of attacks, such as prompt injection, adversarial data poisoning, and model inversion attacks [1]; the framework creates a reference architecture that connects together the five different security domains related to AI: identity and access management (IAM), network security, data security, application security and, finally, audit, covering the entire AI life cycle from training and deployment through ongoing operations [1], [2].

4.2. Five security domains and controls

4.2.1. Identity and Access Management [IAM]

The purpose of effective IAM is to protect AI systems from unauthorized users. Since generative AI uses large amounts of sensitive data, effective IAM is a key component of protecting those systems. Therefore, using a centralized provisioning system and fine-grained Role-Based Access Control [RBAC] across all cloud services will create an IAM framework for all Generative AI applications. In addition, by deploying Multi-Factor Authentication [MFA] and Just-In-Time [JIT] access mechanisms, you can reduce the risk of unauthorized access to AI applications and privilege escalation [1]. To provide security for API endpoints and interfaces to AI systems, developers will use OAuth2 with JSON Web Tokens [JWT] to provide stateless authentication and provide tightly scoped access controls to only allow legitimate users and services to access sensitive AI functions [5].

4.2.2. Data security

Data security involves implementing robust encryption protocols for data at rest and in transit. The framework advocates for the use of cloud-native key management services to enable AES-256 level encryption and supports advanced techniques such as homomorphic encryption for sensitive computations, digital signatures, and cryptographic hashes to ensure data integrity [2], [6]. Data classification and labeling by sensitivity, combined with dynamic data masking and redaction techniques, help protect information during both training and operational phases. In addition, the architecture mandates regular data backups, disaster recovery processes, and secure data disposal methods to guarantee data availability and integrity while preventing unauthorized data exfiltration [1].

4.2.3. Network security

Establishing a strong network security posture in a cloud setting is important to ensure control over communication channels and isolation of AI workloads. The framework requires all deployments to use dedicated Virtual Networks (VNET) or Virtual Private Clouds (VPC), while following specific firewall rules, default deny security postures, and segmenting the network in zones to limit lateral movement after a breach [3], [1]. In addition to these steps to reinforce a secure network, advanced network intrusion detection systems and distributed denial-of-service (DDoS) protection, which help protect the infrastructure from external threats [7], [8], should be utilized. In order to keep all network communications between microservices confidential and maintain their integrity, a combination of secure tunnel protocols and mutual TLS (mTLS) should be used when communicating [9], [10].

4.2.4. Application security

For application security, the framework mandates a multi-layered approach that includes secure coding practices, comprehensive API security, vulnerability scanning, and runtime application self-protection. By leveraging Web Application Firewalls [WAFs] and automated tools that perform static and dynamic code analysis, vulnerabilities are identified and mitigated during the development cycle [3], [11]. Adversarial training is employed to harden models against prompt injection, evasion attacks, and data poisoning, ensuring that the AI system maintains robustness under adversarial conditions [1], [11]. Additionally, human-in-the-loop mechanisms are integrated to validate model outputs before full automation takes control to prevent erroneous or harmful responses [1].

4.2.5. Audit and monitoring

To detect and investigate security incidents, comprehensive audit logging and continuous monitoring are critical. The framework employs cloud-native services [i.e., Google Cloud Logging, AWS CloudWatch, Azure Sentinel] to aggregate telemetry data from all levels of the AI application [1], [2]. Log aggregation systems, combined with anomaly detection technologies and behavioral analytics, provide both real-time detection of attacks and subsequent forensic analysis of the incident [12]. Periodic audits and compliance assessments have been automated to minimize manual involvement and to ensure the effectiveness of security controls remains consistent with industry best practices [6].

4.3 Cloud specific controls and multi cloud considerations

Modern cloud architectures are distributed by nature; therefore, the framework integrates cloud-specific security features from providers such as AWS, IBM Cloud, Google Cloud, and Azure (where the framework uses cloud provider IAMs, Key Management Service (KMS), and managed firewalls) to ensure that a cloud provider's services implement the security "guardrails" that are required [1], [3], [13]. The architecture encourages the creation of dedicated cloud accounts and regions for the purpose of creating isolated environments to reduce the risk of lateral movement across service boundaries [1], [5]. Additionally, centralized dashboards and automated compliance tooling maintain cross-cloud visibility and consistency throughout multi-cloud deployments by providing visibility across all three cloud providers [2], [6].

5. Risk management and threat analysis

Identification a fundamental element of the framework is the systematic identification of risks and threats throughout the generative AI lifecycle. The first step of this process is enumerating all possible risks through 5 critical stages of the acquisition pipeline: data collection, preprocessing, training, inference, and output handling [1]. Threat modeling involves mapping each identified risk to technical controls and assigning distinct responsibilities to the various enterprise teams that comprise the organization via RACI Matrices in order to improve accountability across the entire organization [1], [11]. Some examples of the potential risks mentioned above are model poisoning, adversarial manipulation [e.g., prompt injections and data poisoning], excessive model privileges, and supply chain vulnerabilities, all of which will be mitigated through the use of multi-layered defenses [11], [12].

The framework uses a broad range of threat models, which take into account both internal and external risks. For example, external threat vectors are automated attacks, such as black box adversarial inputs that could alter the behavior of the model; phishing and malware-based attacks that have been amplified by the advent of generative AI [1]. Conversely, internal threats such as insider attacks, giving excessive privileges to AI components and inadvertent data leakage, are mitigated with strict adherence to the principle of least privilege and zero trust architectures [1]. Incorporating adversarial training and red teaming also enhances the overall capability of the framework to detect and respond to new adversarial tactics [7], [11]. In addition, the use of ensemble models which utilize a variety of AI architectures, adds to overall resilience by eliminating single points of failure and reducing the false negative rate during threat detection [8], [11].

6. Security governance, compliance, and ethical considerations

Governing effectively ensures a safe AI ecosystem. A framework aids with defined boundaries of responsibility between various types of participants involved with the development of an AI solution; such as AI developers, information technology security professionals, data governing officers, etc., thus creating a continuing ability to oversee every phase of AI program deployment [1],[4]. A comprehensive governance program involves regular review of established policies, risk assessments and coordination between departments to manage technical as well as regulatory issues associated with an AI program [2],[4]. In addition, governance models suggest that a human be in the "loop" when making decisions based on AI output that would create consequences associated with a high degree of risk, thereby assuring that all actions taken by the participants in an AI program will have accountability associated with those actions [1],[11].

It should be underlined that AI systems should always be regularly assessed for biases, vulnerabilities, and compliance with ethical and legal standards [18].

Aligning with international regulations like GDPR, HIPAA and the proposed EU AI Act, the framework includes controls for compliance with data protection or privacy [1], [6]. Ethical principles are part of Privacy by Design [PbD] principles, which embed systems with these principles from the beginning to ensure data minimization, transparency and fairness [14], [15]. AI systems produce responses that can be explained and audited by humans, enhance user trust and provide a means for human intervention in case of an unexpected event. This compliance with ethics and regulations will improve user trust in AI and reduce potential financial liability of data breaches or non-compliance with laws [2], [11].

7. Implementation considerations and operational best practices

In order to decrease the administrative load associated with responding to security incidents, this framework implements an automated security orchestration system that utilizes continuous threat intelligence feeds, dynamic policy enforcement and AI-driven incident response methods [1], [12]. Reinforcement learning algorithms can adaptively change the rules for firewalls, access settings, and thresholds for alerts based on real-time analysis of the behavior of users, devices and systems on the network [7], [12]. Automated

compliance monitoring solutions can always be used to determine whether an organization's cloud services have been configured in accordance with established security standards and would allow an organization to more quickly address any issues of non-compliance that may arise [3], [6].

The effectiveness of the security controls implemented throughout the AI lifecycle must be thoroughly tested and validated. The framework requires both static and dynamic analysis tools, continuous vulnerability scanning, and adversarial testing [1], [3], as part of a holistic approach to security. In addition, it is common practice to conduct periodic red team exercises that simulate potential attacks and allow for ongoing refinement to detection and response strategies [7], [11]. Finally, the use of formal methods, such as White Box Fuzzing and probabilistic inference methods, will provide some degree of heuristic confidence (in addition to empirical testing) and will result in the continued reduction of the attack surface over time [7], [11].

There are three overall types of blueprints for developing applications utilizing generative artificial intelligence: using an external AI API, self-hosting a model that is developed externally, and training and deploying your own model entirely in-house. Each type of blueprint has its own unique risks and responsibilities, which will affect the security controls, contractual agreements, and operational procedures that you choose [2], [12].

When using an external AI API, you will need to have a contract in place that requires the API provider to operate under the same security standards as you do. For self-hosted applications, you will have more control over the application, but this will come at an additional cost in terms of the resources you need to invest in setting up and maintaining the application [2], [11]. The framework outlines specific recommendations for addressing the inherent risk trade-offs of each blueprint as well as providing guidance on securely configuring, continuously monitoring, and incident reporting on your deployments [2], [12].

8. Challenges, future directions, and integration with emerging technologies

Securing generative AI systems while ensuring scalability is among the most significant challenges that face organizations today. While the continued evolution of AI models will lead to an increasing amount of data, the architecture must also be able to provide large-scale processing for these models while providing tight control of sensitive operations [6], [13].

Distributed cloud-based architectures provide a scalable solution for organizations, but they rely on advanced orchestration to ensure a consistent policy across multiple environments [3], [6]. Interoperability among different cloud service providers can be accomplished through the use of standardized security APIs and communication protocols that allow for seamless integration and a unified view across heterogeneous platforms [5], [16].

Security mechanisms driven by artificial intelligence are becoming more integrated into the security architecture to provide adaptive, proactive responses to dynamic threats. Federated learning and differential privacy allow collaborative threat model development

while supporting data sovereignty, and ensemble techniques improve resiliency against adversarial attacks through the combination of multiple detection models [11], [12]. Additionally, continuous retraining and self-learning capabilities within the security system allow it to adapt to new attack methods and reduce the number of false positives, thus improving the resiliency of the overall security posture over time [7], [12].

Several different deployment model systems exist today. They are all beginning to focus on providing structured access to AI functionality so that all the AI functionalities can be delivered without exposing end-users with all of the model's parameter values or internal data. The model proposed within this framework employs a controlled-access cloud-based deployment model where all of the interactions with AI functionality occur through secured application programming interfaces (APIs) thus providing a significant reduction in possible unauthorized copies and or reused algorithm outputs [6], [15].

This model provides a very strong foundation since the deployment model is based on the zero-trust model. Zero trust requires that all access requests must continuously be validated and that segmentation and isolation are enforced even if a user is connected inside of the trusted network perimeter [1]. The benefits of implementing structured access methodologies into the deployment model as demonstrated in many recent studies indicate that they not only provide operational flexibility but also provide a very high level of security by limiting the amount of exposure to only essential functionalities [1], [15].

9. Smart cloud platforms

9.1. Smart cloud computing

There is a broad definition of cloud computing as the delivery of conveniently available and scalable shared resources, such as networks, servers, storage, applications, and services. NIST has defined 5 characteristics of cloud computing as part of its definition: on-demand self-service, broad access across networks, resource pooling, dynamic scaling, and measuring your services [19]. By adding intelligence to infrastructure, software, security, and service operation smart Cloud Platforms extend this definition further.

Thus, smart Cloud Platforms aren't just a way of hosting applications. They have their own digital environment that may monitor their current line of operations, predict an appliance failure, optimize how much resource is being used, enforce security standards, and integrate data from a variety of outside sources (such as IoT devices, edge nodes, mobile applications and enterprise systems). To accomplish these goals, modern Cloud architectures are built according to several absolute principles: operational excellence, security, reliability, performance, cost-effectiveness, and sustainability [20], [21].

9.2. Benefits of smart cities cloud platforms

Organizations can benefit from using smart cloud platforms in a variety of ways. Some of these include:

1. Scalability - Organizations can adjust the size of their services based on demand, which will reduce large capital investments in infrastructure.

2. Operational Efficiency - With AIOps and automation, companies will have less manual intervention, quicker incident response times, and more reliable service uptime. There are numerous studies proving how AIOps can automate the detection, prediction, diagnosis and execution of actions in complex cloud operations [22].
3. Cost Optimization - Intelligent allocation of resources helps organizations avoid wasting infrastructure that is not being utilized effectively. Intelligent resource allocation allows companies to pay only for what they use, which is an essential part of any cloud architecture framework [20], [21].
4. Cyber Resilience - Smart cloud platforms can monitor their data centers for abnormal activity, correlate security incidents, and support automated responses. The NIST Cybersecurity Framework 2.0 provides a means for measuring cybersecurity outcomes by organizing them around six areas: Govern, Identify, Protect, Detect, Respond, and Recover, all of which closely align with cloud security governance [23].
5. Innovation Acceleration - Developers can take advantage of managed services, serverless computing (where the developer has less involvement with infrastructure), ML platforms, and DevOps pipelines to bring new applications to market more efficiently. Serverless computing is particularly useful for developers because they can concentrate on developing the "business logic" of an application while leaving the underlying infrastructure management to the cloud provider [24].

9.3. Security challenges of smart cities cloud platforms

Despite their benefits, smart cloud platforms introduce serious security and governance challenges. These risks are especially serious because smart cities often manage critical infrastructure and sensitive citizen data.

One of the greatest challenges that cloud customers face are issues related to identity and access management. Misconfigured identities, excessive permissions, weak authentication, and inadequate privilege management could all leave sensitive systems vulnerable.

The second challenge relates to data protection. Modern smart cloud platforms typically process large amounts of personal, institutional, and operational data; thus, encryption, key management, data classification and privacy-by-design should be foundational components of secure data processing.

Thirdly is a challenge in relation to shared responsibility. While cloud providers are responsible for ensuring the security of their underlying infrastructure, customers would still remain responsible for securing workloads, identities, configurations, and data. Energy, transport, water, and emergency systems may become targets of cyberattacks. As highlighted in the Cloud Security Alliance's Security Guidance v5, modern practices in securing cloud solutions; risk management; compliance and the shared responsibility model are required [25].

Fourthly, cloud providers using increasingly sophisticated AI to automate and make decisions creates potential risk to organizations. Examples of AI-associated risk include, biasing in models, unreliable model outputs, data poisoning, adversarial attacks, and lack of transparency. The NIST AI Risk Management Framework helps organizations that use

AI govern, map, measure and manage risks associated with using AI appropriately and being able to create trustworthy AI solutions [26].

Finally, the complexity of the environment in which cloud providers operate has changed significantly over the past few years. The use of multi-cloud environments, hybrid cloud environments, edge computing, microservices, and serverless systems has introduced a range of distributed potential attack vectors. Without effective governance and observability organizations can lose visibility of the assets, dependencies, vulnerabilities, and compliance status of the systems that they utilize.

According to the Cybersecurity Framework 2.0 from NIST, their cybersecurity functions are organized as follows: Governance, Identification, Protection, Detection, Response, and Recovery [23]. Being able to use these functions provides a structured way of managing risk (including for smart city cloud platforms), being resilient to disruptions, and holding yourself accountable for the security of your technology.

10.Future directions

The next few years are going to be exciting times for smart cities using intelligent cloud platforms. Here are some trends that we expect to see in years to come:

1. We will see a rise in the use of AI-native cloud platforms as these platforms utilize artificial intelligence for tasks such as predictive maintenance, anomaly detection, automated incident response, and urban planning.
2. There will be a significant increase in cloud-edge integration with cloud computing providing the analytics and storage necessary for large-scale data retention, while edge computing helps provide the low-latency, localized processing necessary for real-time applications.
3. The age of digital twins is here and continuously updates its internal models based on incoming environmental signals [27], but expect these digital representations of things to develop much further as cities begin to use these down-the-road simulations during policy testing.
4. The widespread adoption of Zero Trust security protocols will mean the implementation of continuous verification, least privilege access, microsegmentation, and strong identity management of smart city systems.
5. Open standards and interoperable platforms will become essential for cities to minimize dependency on closed vendor ecosystems by ensuring their data can be reused through multiple departments and services.

Layered design of the framework and mapping of risks − threats − controls offers sufficient means to detect, mitigate, and adapt to emerging security threats on an ongoing basis; thus, allowing AI systems to stay innovative yet secure as rapidly changing cyber threats develop [11], [12]. The framework has clearly defined governance structures, automated policy enforcement, and robust regulatory and ethical foundations, and it provides a scalable solution that meets today's security challenges and lays the groundwork for future advancements in securing AI [4], [17].

This framework continues to be revised on an ongoing basis based on continuous research, adversarial test evaluations and testing during deployment in real life; thus, it will develop with new and developing threats and advancements in technology [6], [7], [11].

11. Conclusion

This article presents an integrated AI security framework that provides a interconnected, multi-domain defense-in-depth strategy concerning generative AI systems used in cloud-based environments. The integrated AI security framework combines best practices from Google's Secure AI Framework, OWASP's AI Security and Privacy Guide, and IBM's Security Framework for Generative AI; it covers the entire spectrum of risk, including identity management, data security, network security, application hardening, and continuous monitoring.

To summarize, a comprehensive security model employing multiple levels of architecture, proactive risk management, and a strong governance system will provide the means to deploy generative AI on cloud platforms in a secure manner from start to finish. This comprehensive security model will protect against known attack vectors and adapt to changing attack vectors through ongoing learning and collaboration between humans and AI, thereby enabling businesses to develop resilient, trustworthy, and ethically aligned enterprise-wide AI systems [1], [2], [6], [7].

To ensure the entire development and implementation cycle of artificial intelligence systems is considered along with the initial threat assessment, and throughout the final audit and ongoing, variable-monitoring-automatic feedback loops for risk reduction [7], [12]; different hybrid security methods will be used so that the systems can prevent outside security breaches in addition to being able to find any internal abnormal situations that can be caused by either an insider threat or sophisticated adversarial manipulation [2], [7].

With the continued evolution of organizations toward edge AI and distributed computing architectures, the framework's scalability, interoperability and compliance capabilities become even more critical [6], [13]. The ability to maintain a unified security posture across heterogeneous systems and multiple geographic areas will be accomplished through a centralized security dashboard, with an automated compliance verification protocol as a means for measuring compliance [3], [5].

Additionally, the framework can continue to develop by enhancing the approach to adversarial training methods, employing new structures such as structured access controls, and utilizing new alternatives like zero trust architectures to further limit the accessibility of sensitive AI features [6], [15]. To improve the overall performance and effectiveness of this framework, researchers will need to work together to continue refining the techniques used in this framework and build them into enterprise-level AI deployments [8], [12].

To operationalize this framework, significant organizational transformation is required including: upskilling IT security personnel; adapting the development lifecycle to incorporate robust cybersecurity test cycles; creating cross-functional teams with a focus

on ongoing cybersecurity oversight [1], [4]. As the threat landscape changes, the hybrid model of automated defenses complemented with expert judgment from humans will continue to be an essential component of a strong operational security posture [1].

It is essential to have explicit incident response processes and continual improvement cycles associated with AI security measures, which incorporate technical standards and ethical standards [2], [7]. The dynamic relationship of technology, compliance with regulations, and ethics ensures that the overall security framework can rapidly adapt to mitigate emerging risks in addition to creating a secure, inclusive and transparent AI ecosystem [14], [17].

Combining Google's Secure AI Framework, OWASP's AI Security and Privacy Guide, and IBM's Framework for Securing Generative AI results in a holistic approach to securing the expanding AI ecosystem that is also adaptable for the foreseeable future. The security framework provides a robust defense against security threats, including through rigorous identity management, data protection, network segregation, application security, and advanced auditing measures, as well as by establishing governance and ethical structures, all designed with the capability to evolve along with the fast-paced rate of innovation found in the growing AI industry [1], [2], [7]. With the security framework in place, organizations would have the confidence, security, and regulatory compliance necessary to realize the full potential of using generative AI, thereby contributing to a more sustainable and responsible approach to AI-driven innovation as we enter the cloud era [4], [6], [12].

Finally, Smart Cities Cloud Platforms have taken traditional Cloud Computing (which was historically a "provisioning model") and evolved it into intelligent, adaptable, and secure Digital Infrastructure. The true value of Smart Cities Cloud Platforms is from the combination of AI, automation, observability, cloud-native design, cyber security, and edge computing. To successfully adopt Smart Cloud Platforms, organizations need to implement effective governance processes, develop robust security architecture, responsibly govern their use of AI, and continuously monitor the use of these technologies.

Modern smart cities rely heavily on smart cloud platforms to facilitate the development of smart city services because smart cloud platforms provide the necessary infrastructure to collect, process, analyze, secure and govern urban data. These smart cloud platforms enable the provision of smart city services by combining cloud computing, IoT, AI, edge computing, cyber security and open standards.

However, in order to be successful, the implementation of smart cloud platforms must be considered with regard to: interoperability; privacy; cyber security; data governance; and ethical use of technology. A smart city should be more than just a technologically advanced city. A smart city should be secure, transparent, inclusive, sustainable and accountable also.

This proposed AI security framework is used to strengthen security elements in smart cities cloud platforms as well.

References

[1] H. Hayagreevan and S. Khamaru, “Security of and by generative AI platforms,” arXiv, Oct. 2024.

[2] D. Bogdanov, P. Etti, L. Kamm, and F. Stomakhin, “Artificial intelligence system risk management methodology based on generalized blueprints,” in Proc. 16th Int. Conf. Cyber Conflict [CyCon]: Over the Horizon, May 2024.

[3] H. Booth, “Secure software development practices for generative AI and dual-use foundation models,” 2024.

[4] N. Lekota, “Governance considerations of adversarial attacks on AI systems,” in Proc. Int. Conf. on AI Research, Dec. 2024.

[5] P. Paidy and K. Chaganti, “Securing AI-driven APIs: Authentication and abuse prevention,” Int. J. of Emerging Research in Engineering and Technology, Jan. 2024.

[6] H. Tibebu, “Framework for data protection, security, and privacy in AI applications,” 2024.

[7] D. Dalrymple, J. Skalse, Y. Bengio, S. Russell, M. Tegmark, S. Seshia, et al., “Towards guaranteed safe AI: A framework for ensuring robust and reliable AI systems,” arXiv, May 2024.

[8] S. Joshi, “Gen AI in financial cybersecurity: A comprehensive review of architectures, algorithms, and regulatory challenges,” Int. J. of Innovations in Science Engineering and Management, Jul. 2025.

[9] H. Padmanaban, “Privacy-preserving architectures for AI/ML applications: Methods, balances, and illustrations,” Journal of Artificial Intelligence General Science [JAIGS], Apr. 2024.

[10] D. Patel, G. Raut, S. N. Cheetirala, G. N. Nadkarni, R. Freeman, B. S. Glicksberg, et al., “Cloud platforms for developing generative AI solutions: A scoping review of tools and services,” arXiv, Dec. 2024.

[11] A. Gurram, “Generative AI for enhanced cybersecurity: Building a zero-trust architecture with agentic AI,” World Journal of Advanced Engineering Technology and Sciences, Apr. 2025.

[12] D. A. Ademilua and E. Areghan, “AI-driven cloud security frameworks: Techniques, challenges, and lessons from case studies,” 2022.

[13] N. Mungoli, “Scalable, distributed AI frameworks: Leveraging cloud computing for enhanced deep learning performance and efficiency,” arXiv, Apr. 2023.

[14] E. Hashmi, M. M. Yamin, and S. Y. Yayilgan, “Securing tomorrow: A comprehensive survey on the synergy of artificial intelligence and information security,” AI and Ethics, Jul. 2024.

[15] T. Shevlane, “Structured access: An emerging paradigm for safe AI deployment,” Preprint, Jan. 2022.

[16] T. Rintamäki, “Comparing OWASP mappings,” 2025.

[17] P. Balasubramanian, S. Liyana, H. Sankaran, S. Sivaramakrishnan, S. Pusuluri, S. Pirttikangas, and E. Peltonen, “Generative AI for cyber threat intelligence: Applications, challenges, and analysis of real-world case studies,” Artificial Intelligence Review, Aug. 2025

[18] A. Hyra, and F. Premti. “The Double-Edged Sword of AI in Cybersecurity: Boosting Security While Addressing Privacy Risks“. Smart Cities and Regional Development (SCRD) Preprints. 1, 1, Vol. 1, Dec. 2024.

[19] P. Mell and T. Grance, “The NIST Definition of Cloud Computing,” National Institute of Standards and Technology, Special Publication 800-145, 2011.

[20] Amazon Web Services, “The Pillars of the AWS Well-Architected Framework,” AWS Documentation.

[21] Google Cloud, “Google Cloud Well-Architected Framework,” Google Cloud Architecture Center, 2026

[22] Q. Cheng, D. Sahoo, A. Saha, W. Yang, C. Liu, G. Woo, M. Singh, S. Savarese, and S. C. H. Hoi, “AI for IT Operations — AIOps — on Cloud Platforms: Reviews, Opportunities and Challenges,” 2023

[23] National Institute of Standards and Technology, “The NIST Cybersecurity Framework 2.0,” 2024.

[24] Y. Li et al., “Serverless Computing: State-of-the-Art, Challenges and Opportunities,” IEEE Transactions on Services Computing, 2023.

[25] Cloud Security Alliance, “Security Guidance for Critical Areas of Focus in Cloud Computing v5,” 2024.

[26] E. Tabassi, “Artificial Intelligence Risk Management Framework — AI RMF 1.0,” National Institute of Standards and Technology, 2023

[27] F.ZISKO and B.MEMA. “Digital twins and MIS integration for urban resilience: A strategic approach to smart city evolution “. Smart Cities and Regional Development (SCRD) Journal. 10, 2 (Apr. 2026), 35–50.

[28] G. WAIZEL. “The agentic AI security adoption matrix: Understanding readiness and resistance across domains“. Smart Cities and Regional Development (SCRD) Journal. 10, 2 (Apr. 2026), 75–84.
[29] P.Teimoory. “Towards robust security in smart payment systems: challenges and solutions “. Smart Cities and Regional Development (SCRD) Journal. 9, 3 (May 2025), 29–38.
[30] C. LI and I.VIRTOSU. “Competition law in the age of AI: Confronting algorithmic collusion in the smart economy“. Smart Cities and Regional Development (SCRD) Journal. 9, 2 (Mar. 2025), 41–54.
[31] R.M. BENSHAMS. “Cultivating organizational culture for AI integration: A framework for Smart Cities and Regional Development“. Smart Cities and Regional Development (SCRD) Journal. 9, 4 (Nov. 2025), 125–137.
[32] E. GRABOCKA and E. NDOKA. “AI-driven innovation within the ICT sector“. Smart Cities and Regional Development (SCRD) Journal. 9, 1 (Jan. 2025), 77–97.