GTI

AI Security Governance Technology White Paper

- Security Rooted in Networks, Governing AI with Intelligence, Shift Left Co-Governance

http://gtigroup.org/

GTI AI Security Governance Technology White Paper

Version:	V1.0
Deliverable Type	☐ **Procedural Document** ☑ **Working Document**
Confidential Level	☐ **Open to GTI Operator Members** ☐ **Open to GTI Partner** ☑ **Open to Public**
Program	5G-A×AI Development Program
Working Group	
Project	Research on the Technical Framework for AI Security Governance
Task	
Source members	China Mobile, HKT, YTL, Beijing Baidu Netcom Science Technology Co., Ltd., Beijing Qihoo Technology Co., Ltd., Huawei Technologies Co., Ltd., New H3C Technologies Co., Ltd., Novare, iFLYTEK Co., Ltd., Anhui SparkShield Intelligent Technology Co., Ltd., ZTE Corporation Co., Ltd., Beijing University of Posts and Telecommunications, Fudan University, Zhejiang University, AI EdTech Association, Hong Kong Web3.0 Standardization Association
Support members	MediaTek, Tsinghua University, TÜV Rheinland, Young AI Leaders Community - San Francisco Hub
Editor	
Last Edit Date	2026-08-21
Approval Date	2026-09-07

Document History

Date	Meeting #	Version #	Revision Contents
2026-09-01		1.0	Version 1.0

Foreword

Artificial intelligence is rapidly integrating into economic and social operations and powering digital transformation across industries. As AI systems scale, modern AI applications depend on data pipelines, computing infrastructure, software ecosystems, external tools, digital identities, and interconnected business processes. Security risks therefore extend far beyond model outputs. Vulnerabilities in infrastructure, misuse of permissions, malicious data manipulation, autonomous agent behavior, and supply-chain dependencies can all affect the reliability and trustworthiness of AI systems. Traditional governance mechanisms, which were largely designed for conventional information systems, are often insufficient to address these emerging challenges.

Around the world, governments, standards organizations, enterprises, and research institutions are accelerating the development of governance frameworks for AI. Safety, accountability, transparency, robustness, privacy protection, and value alignment are increasingly recognized as essential requirements for responsible innovation. Effective governance must balance innovation with risk management while ensuring that AI remains controllable, trustworthy, and beneficial to society. To this end, a “1-3-9” AI Security Governance Framework is proposed that integrates security requirements throughout the entire life-cycle of AI systems by focusing on 'Security rooted in Networks', 'Governing AI with Intelligence', and 'Shift Left Co-Governance' to support large-scale AI deployment. This proposal summarizes the major development trends, governance challenges, strategic approaches, implementation pathways, and collaborative initiatives necessary to establish a secure, trustworthy, and sustainable AI ecosystem.

Table of Contents

1 Trends in Artificial Intelligence Development and the Global Governance Landscape

Artificial intelligence is becoming a core driver of economic and social transformation. Technological progress, industrial adoption, and governance development are occurring simultaneously, creating both opportunities and challenges for stakeholders across industries.

1.1 Development Trends: Accelerating Technological Evolution Toward Systemic and Agentic AI

AI development is shifting from single-model performance to systemic capabilities that combine algorithms, computing power, data, tool chains, agents, and engineering platforms. LLM, multi-modal interaction, long context, and tool use move AI toward planning, coordination, and complex execution. Agents are evolving into systems that understand goals, decompose tasks, invoke tools, interact with environments, and learn from feedback. As AI enters R&D, finance, public services, and network operations, deployment requires process adaptation, data governance, integration, access control, and secure operations.

1.2 Global Perspective: Governance Rules Are Taking Shape

From the perspective of global governance, major countries, international organizations, standards bodies, and industry stakeholders are accelerating the development of AI governance rules and practical approaches. Current efforts focus on areas such as risk classification, trustworthiness assessment, data governance, model transparency, accountability boundaries, content provenance and labeling, technical standards, and international coordination.

AI security governance is also developing along two closely related fronts. The first is **AI for Security**. AI is becoming an important enabler for strengthening security capabilities, supporting vulnerability discovery, threat analysis, security operations, and risk assessment through intelligent analysis, automated detection, and decision support. The second is **Security of AI**. As the capabilities and deployment boundaries of AI systems continue to expand, the security of AI itself is receiving growing attention. Risks related to model security, data security, agent behavior, inter-agent communication leakage, and supply-chain dependencies are becoming key

governance priorities. Addressing these risks requires risk assessment, life-cycle management, and runtime control mechanisms that improve the security and trustworthiness of AI systems.

AI governance is moving from high-level principles and voluntary initiatives toward a more mature stage that combines institutional arrangements, standards development, technical evaluation, and regulatory implementation. Although governance approaches differ across countries and regions, a broad consensus is emerging around safety and trustworthiness, risk control, clear accountability, human-centered development, and open collaboration.

1.3 Industry Perspective: Collaborative Exploration by Diverse Stakeholders

AI governance requires collaboration among governments, infrastructure providers, technology companies, security vendors, standards organizations, research institutions, and end users. No single stakeholder possesses sufficient visibility or authority to manage all aspects of AI risk.

Digital infrastructure providers play a particularly important role because AI systems rely heavily on connectivity, computing power, identity management, and operational monitoring. Trusted infrastructure provides the foundation for secure AI deployment. At the same time, collaboration across industries promotes knowledge sharing, standardization, interoperability, and coordinated risk management.

The long-term success of AI governance depends on establishing mechanisms that encourage cooperation while maintaining flexibility for innovation.

2 Risks and Challenges: AI Security Governance Faces Five Major Shifts

The large-scale deployment of AI is fundamentally changing the nature of security risks. Governance challenges increasingly emerge from interactions among models, data, infrastructure, permissions, and operational processes. As a result, governance approaches must evolve beyond static controls and compliance-based supervision.

Five governance shifts at a glance

01 The Unknown Outweighs the Known

Emergent behavior and complex interactions make risks difficult to predict before deployment.

02 Attacks Outpace Defense

AI lowers the cost of vulnerability discovery, social engineering, and coordinated attack campaigns.

03 Evolution Outpaces Governance

New architectures and scenarios evolve faster than standards, rules, and governance mechanisms.

04 Divergence Outweighs Consensus

Shared principles exist, but regulatory tools, risk boundaries, and standards remain fragmented.

05 Behavior Over Content

Risk moves from compliant outputs to controllable actions, permissions, and execution chains.

2.1 The Unknown Outweighs the Known

Modern AI systems exhibit emergent behaviors that are often difficult to predict. Capabilities such as long-context reasoning, memory management, external tool invocation, and multi-agent collaboration introduce uncertainty into operational environments. Risks may emerge through complex interactions rather than isolated failures.

Examples include hallucinations, prompt injection attacks, objective misalignment, and unexpected behavioral outcomes. Because these risks cannot always be identified during development, governance frameworks must support continuous monitoring, runtime assessment, and adaptive mitigation strategies.

Governance must therefore evolve beyond static rule matching toward pre-deployment risk assessment, continuous runtime monitoring, dynamic risk evaluation, and runtime calibration. Reassessment should be triggered when key

elements such as models, data, tool permissions, or application scenarios change, ensuring that governance measures remain aligned with the evolution of AI systems.

2.2 Attacks Outpace Defenses

AI is becoming deeply embedded in software development, security testing, and cyber operations. Vulnerability discovery is shifting from expert-driven analysis toward intelligent workflows that support large-scale screening, automated validation, and continuous operation. Malicious actors may use AI to accelerate vulnerability discovery, exploit validation, and attack-chain orchestration, lowering preparation costs and execution barriers. Defenders, by contrast, must balance business continuity, system compatibility, and operational stability, which lengthens response chains and widens the tempo gap between attack and defense.

As AI is further integrated into cloud computing, mobile networks, and edge computing, defense models that rely on fixed perimeters and local detection are increasingly challenged. Edge intelligence, distributed inference, and autonomous agent interaction are weakening traditional network boundaries. Security governance is therefore evolving from perimeter-based protection toward dynamic trust enabled by distributed telemetry. In the future, security capabilities will be further embedded into communication networks and computing infrastructure, forming adaptive runtime guardrails for AI systems through operational-state sensing, entity verification, and dynamic policy adjustment.

2.3 Evolution Outpaces Governance

Technological innovation often advances faster than standards, regulations, and governance mechanisms. AI is evolving from single-model invocation into an integrated technical system that combines multiple capabilities, connects multiple systems, and becomes embedded across multiple operational stages. Its capabilities are expanding from text generation to multi-modal understanding, code generation, tool use, and task planning, while applications are extending from assistant-style Q&A to office collaboration, software development, customer service, operations management, and production control.

As technical forms, application boundaries, and operating modes continue to change, governance rules, evaluation methods, accountability boundaries, and audit mechanisms face continuous pressure to adapt. Governance therefore needs to move beyond fixed rule matching toward a model that combines risk-based management, principle-guided governance, and dynamic adaptation. Governance requirements

should also expand from content compliance to data flows, permission use, external connections, behavioral processes, and accountability chains. Full life-cycle access control, dynamic behavior calibration, real-time risk correction, and continuous evaluation and auditing are needed to narrow governance gaps.

2.4 Divergence Outweighs Consensus

Global actors broadly endorse principles such as common good, people-centered development, safety, reliability, transparency, accountability, fairness, inclusiveness, human oversight, and risk governance. However, differences in development stages, industrial foundations, and regulatory systems have led to different governance pathways in terms of regulatory intensity, policy tools, risk boundaries, responsibility allocation, and standards systems. The EU places greater emphasis on risk classification, compliance obligations, and mandatory regulation; the US focuses more on risk management frameworks, technical guidance, industry practice, and the balance between innovation and governance; Singapore and the UK are exploring more flexible approaches.

Representative governance frameworks and initiatives have also taken shape. The EU Artificial Intelligence Act focuses on risk classification and compliance requirements, NIST AI RMF emphasizes risk management processes, and ISO/IEC 42001 addresses the establishment of AI management systems. The Bletchley Declaration, the UNESCO Recommendation on the Ethics of Artificial Intelligence, and ITU AI for Good further contribute to global consensus on safety, ethics, and sustainable development.

Looking ahead, greater efforts are needed to promote interoperable standards frameworks, evaluation methodologies, interface rules, and coordination mechanisms. This will help different stakeholders build baseline consensus on risk identification, information sharing, incident response, and accountability tracing, thereby reducing regulatory friction and coordination costs.

2.5 Behavior Over Content

As AI enters operations, manufacturing, autonomous driving, and inspections, model judgments and instructions increasingly become system configurations, equipment actions, and process decisions. Risk therefore shifts from compliant content to controllable behavior. Error codes, configuration mistakes, perception anomalies, corrupted data, or adversarial inputs can propagate through business or

perception-decision-execution chains, causing leaks, disruptions, production failures, or real-world safety issues.

Consequently, organizations must ensure that AI actions remain within authorized boundaries. Governance mechanisms should include access controls, approval workflows, anomaly detection, rollback capabilities, and comprehensive auditing. The goal is to ensure that AI behavior remains predictable, controllable, and accountable.

3 Overall Approach to AI Security Governance

To address the five major shifts in AI security governance, we propose a "1-3-9" AI Security Governance Framework. Traditional point-based approaches centered on model security, content safety, and application protection are no longer sufficient to address cross-model, cross-data, cross-tool, cross-computing, and cross-stakeholder risks arising from systemic and agentic AI. AI security governance is therefore moving toward a more integrated approach that covers infrastructure support, intelligent governance, application operations, and ecosystem collaboration.

Building on the evolving global AI governance landscape and the security needs of large-scale AI deployment, the framework is guided by Security Rooted in Networks, Governing AI with Intelligence, and Shift-Left Collaborative Governance. It consists of one shared vision, three strategic approaches, and nine development pathways. By embedding security requirements into infrastructure support, intelligent capability governance, and industry application ecosystems, the framework works in coordination with international frameworks such as NIST AI RMF and ISO/IEC 42001, providing a practical reference for translating AI security governance principles into implementation.

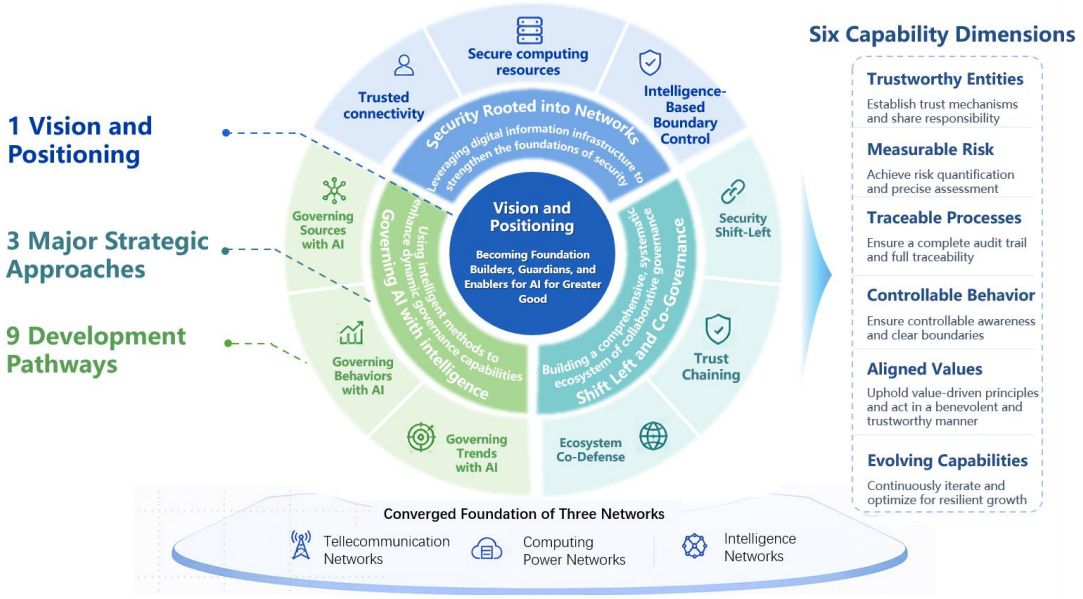

Figure 1: "1-3-9" AI Security Governance Framework

3.1 Vision: Becoming Foundation Builders, Guardians, and Enablers for AI for Greater Good

AI is rapidly integrating into production, daily life, and social governance, becoming a vital force driving the development of new productive forces, the construction of digital societies, and the modernization of governance systems. AI security governance serves as a critical pillar for ensuring that AI remains safe, trustworthy, inclusive, beneficial, and high-quality in its development. Across key scenarios such as digital and intelligent production, digital and intelligent living, and digital and intelligent governance, all industry stakeholders need to collectively fulfill their roles in providing foundational support, security assurance, and scenario empowerment. Together, they should act as foundation builders, guardians, and enablers for the beneficial development of AI.

For Digital and Intelligent Production- Acting as Foundation Builders for Trustworthy AI Development

As AI becomes deeply embedded in key areas such as R&D and design, production and manufacturing, supply chain coordination, and operations management, security capabilities have become an essential foundation for industrial intelligent transformation. It is necessary to leverage network, computing, data, and platform capabilities to promote the coordinated development of security capabilities alongside AI infrastructure, thereby providing a stable, reliable, safe, and trustworthy foundation for industrial intelligent upgrading.

For Digital and Intelligent Living — Acting as Guardians for Secure AI Applications

As applications such as smart homes, intelligent terminals, and virtual digital avatars spread, AI services require stronger safeguards for individuals and families. There is an urgent need to strengthen comprehensive public-facing security safeguards. Security governance must be embedded throughout the entire process of intelligent product design, service operation, and user interaction, reinforcing protection for user data, service content, identity authentication, and terminal access. This ensures that security boundaries are maintained for AI applications serving the public, making intelligent applications more trustworthy and reassuring.

For Digital and Intelligent Governance — Acting as Enablers for Integrated AI Innovation

As government services, urban operations, public safety, and emergency management adopt AI, cross-industry and cross-regional collaboration becomes vital. It is necessary to promote deep integration of AI capabilities and security capabilities with industry scenarios, providing robust support for governance modernization and intelligent social operations.

By establishing secure foundations for connectivity, computing, and intelligent services, the framework seeks to support the long-term development of a resilient and trustworthy AI ecosystem.

3.2 Strategic Approach

The framework is built upon three strategic principles: Security rooted into Networks, Governing AI with intelligence, and Shift Left Co-Governance. Together, these principles provide a holistic approach to managing AI risks throughout the life cycle of systems and services.

01 Security Rooted in Networks	**02 Governing AI with intelligence**	**03 Shift Left Co-Governance**

3.2.1 Security rooted in Networks

Security Rooted in Networks emphasizes digital infrastructure as the foundation for AI security governance. From the perspective of large-scale AI deployment, AI depends on the coordinated support of connectivity, computing power, trust, data, models, and applications. Among these, connectivity, computing, and trust are essential to secure and reliable AI operations, and are key areas where network operators can provide foundational support.

Security requirements should therefore be embedded into connectivity links, computing environments, interaction entry points, and business workflows. This enables security capabilities to be deployed, monitored, and enforced in step with AI operations, forming infrastructure-level protection across real operational chains.

3.2.2 Governing AI with intelligence

Governing AI with Intelligence emphasizes the use of AI capabilities to strengthen AI security governance. As model capabilities continue to emerge, agents become more autonomous, attacks become increasingly automated, and risks propagate across chains, governance capabilities must match the speed and complexity of evolving risks.

Security-oriented AI models, intelligent sensing, dynamic assessment, behavioral analyses, situational awareness, and policy orchestration can be used to continuously identify, assess, and intervene in the origins, behaviors, and trends of AI-related risks. Combined with continuous and automated security validation, including red teaming, attack simulation, and risk verification, these capabilities help uncover potential risks, refine protection strategies, and form a closed loop of risk discovery, validation, assessment, and security improvement.

This enables governance to move from experience-based manual oversight toward human-AI collaboration, intelligent judgment, and proactive defense.

3.2.3 Shift Left Co-Governance

Shift-Left Collaborative Governance emphasizes embedding security governance early into the full process of AI development and application, reducing risk accumulation through source prevention and process control. Given the systemic nature of AI security governance, it requires coordinated management of core elements, life cycle stages, and diverse stakeholders, moving governance toward full-scope coverage, early-stage prevention, and multi-party collaboration.

Responsibilities should be clearly defined among developers, deployers, users, and other relevant actors. In high-risk scenarios, human oversight, intervention, and accountability mechanisms should be strengthened to ensure that AI applications remain controllable and responsibilities remain traceable. This approach helps establish a long-term governance model characterized by clear responsibilities, coordinated action, and effective human-AI collaboration.

3.3 Development Path: Three Directions and Nine Pathways

The framework defines Three Directions and Nine Pathways that collectively support infrastructure protection, intelligent governance, and ecosystem collaboration. These pathways translate strategic principles into practical implementation mechanisms.

Nine development pathways

01 Trusted Connectivity	02 Secure Computing Resources	03 Intelligence-Based Boundary Control
04 Governing Sources with AI	05 Governing Behaviors with AI	06 Governing Trends with AI
07 Security Shift-Left	08 Trust Chaining	09 Ecosystem Co-Defense

3.3.1 Infrastructure Security Direction

This direction focuses on trusted connectivity, secure computing resources, and intelligent boundary control.

Trusted Connectivity serves as the first layer of defense, using trusted connections as the entry point for secure AI operations. It establishes trusted identity and access mechanisms for users, endpoints, devices, agents, and industry applications, and strengthens access control through continuous identity verification, dynamic authorization, and behavior verification, thereby blocking unauthorized access and identity spoofing.

Secure computing resources strengthens the middle layer by making secure computing power the foundation for training, inference, and execution, supporting resource isolation, dynamic scheduling, resilience, and stable AI operations.

Intelligence-Based Boundary Control embeds inner defenses through value and ethical constraints, defining behavioral boundaries and interaction rules, so AI systems remain aligned, controllable, and inherently safer.

Together, these measures create a multi-layer security architecture capable of supporting large-scale AI operations.

3.3.2 Intelligence-Driven Governance Direction

This strategic direction encompasses three pillars: governing sources with AI, governing behaviors with AI, and governing trends with AI.

Governing Sources with AI -Addressing Risks at the Origin

This pillar focuses on the origins of AI risk. It treats key elements such as training data, model weights, knowledge graphs, objective functions, and external component

dependencies as endogenous sources of risk. Through trusted data sourcing, model validation, risk detection, and provenance tracing, it helps identify potential value biases and security hazards at an early stage, enabling intelligent correction before risks materialize and strengthening intrinsic security.

Governing Behaviors with AI - Dynamically Regulating Behavioral Trajectories

This pillar focuses on the continuous action chains of AI systems, particularly AI agents, including intent decomposition, task planning, cross-domain permission use, tool invocation, and result propagation. These chains are treated as behavioral trajectories for security control. By continuously monitoring and dynamically constraining task planning, permission usage, tool invocation, and result execution, and by applying policy evaluation and authorization checks at critical execution points, AI behaviors can remain within authorized boundaries and security rules.

Governing Trends with AI — Orchestrating Holistic Risk Situational Awareness

This pillar focuses on the evolving trends of AI risks. By using intelligent capabilities for situational awareness, risk assessment, and policy orchestration, and by strengthening risk information aggregation, threat intelligence analysis, and cross-stakeholder coordinated response, it improves the ability to anticipate and actively intervene against emerging attacks, chained risks, and systemic threats.

3.3.3 Full-Scope Collaborative Governance Direction

This strategic direction focuses on three aspects: security shift-left, trust chaining, and ecosystem co-defense.

Security Shift-Left - Embedding Security Upstream Across the Full Life-cycle

From a temporal perspective, this pillar drives security requirements upstream into the entire AI system life-cycle including design, development, training, evaluation, deployment, operation, and iteration. The goal is to enable governance capabilities to intervene before risks materialize and to continuously calibrate security measures as the system evolves.

Trust Chaining - Establishing Continuous Trust Across the Full Chain

From a spatial perspective, this pillar extends the scope of governance across key layers including infrastructure, data, models, applications, and AI agents. It ensures that AI systems maintain a continuous chain of trust spanning source, identity, behavior, and accountability, preventing security capabilities from being confined to isolated nodes.

Ecosystem Co-Defense - Enabling Multi-Stakeholder Collaborative Governance

From a stakeholder perspective, this pillar promotes coordinated participation from government agencies, infrastructure operators, model providers, application service providers, security enterprises, research institutions, and industry customers. The objective is to form a governance framework characterized by jointly established rules, mutually recognized capabilities, shared risk prevention, and collective accountability.

3.4 Governance Assessment: Six Capability Dimensions

Addressing the security governance needs of large-scale AI applications, governance expectations are defined across six dimensions - entities, risks, processes, behaviors, value, and capabilities, which leads to the proposal of the Six Capability Dimensions: trustworthy, measurable, traceable, controllable, alignable, and evolvable. These objectives aim to drive AI systems toward achieving trustworthy entities, measurable risks, verifiable processes, controllable behaviors, aligned value with the public good, and continuously evolving capabilities, thus providing a clear road map for the safe and stable operation of AI across a wide range of business scenarios.

Trustworthy Entities: Ensuring that all participating entities are trustworthy and verifiable, and that their behaviors are traceable. Relying on a unified digital identity system, trusted access authentication, credential verification, and an end-to-end traceability framework, this ensures that the identities of participating entities including users, terminal devices, computing nodes, model services, intelligent agent applications, and industry systems are authentic and verifiable, their behaviors are traceable to their sources, and security responsibilities are clearly defined, thereby supporting the trustworthy and collaborative operation of AI across entities, networks, and systems.

Measurable Risks: Establish quantifiable, visible, and evidence-based security assessment. A multidimensional and measurable evaluation mechanism should be built around risk identification, risk assessment, control effectiveness, and governance outcomes. Through security testing, risk classification, adversarial testing, privacy compliance verification, robustness testing, and operational situation analysis, AI system security can be continuously assessed, ensuring that risks are identifiable, risk levels are determinable, and governance effectiveness is verifiable. This supports the transition from experience-based judgment to standardized and data-driven security governance.

Traceable Processes: Ensures that the entire AI operational chain is logged and accountability is traceable. Through mechanisms such as log retention, call chain tracing, data flow recording, model version management, tool call credentials, and audit evidence preservation, the system records key stages of AI operations from input, inference, invocation, and execution to output, ensuring that the source of anomalous behavior can be pinpointed, the process reconstructed, and responsibility determined, thereby meeting the needs of risk investigations, compliance audits, and liability determination.

Controllable Behaviors: Enforces closed-loop constraints on all operations performed by intelligent agents throughout the entire process. Through identity authentication, access control, call auditing, behavior monitoring, dynamic circuit breakers, and cross-domain coordinated response mechanisms, all activities including task planning, data access, system calls, and command execution are ensured to remain within predefined security boundaries at all times.

Aligned Values : Ensures continuous alignment between the objectives of intelligent systems and human values. Through mechanisms such as value alignment testing, ethical preference detection, human-machine collaboration calibration, and high-risk scenario reviews, this approach ensures that the design objectives, decision-making logic, and behavioral outcomes of AI systems remain consistent with laws and regulations, ethical guidelines, industry standards, and the public interest. This guarantees that AI applications do not deviate from a development path that prioritizes safety, compliance, and the common good, thereby reducing risks associated with goal drift, value deviation, and inappropriate behavior.

Evolving Capabilities: Achieve dynamic adaptation and long-term iteration of the governance system. By integrating standard updates, routine platform operations, comprehensive monitoring, policy orchestration, rule calibration, and industry ecosystem collaboration, the governance system can adapt to model capability upgrades, changes in AI application forms, the evolution of attack methods, and the expansion of business scenarios. This ensures that governance capabilities are updated in tandem with technological advancements, shifting risks, and business needs, continuously reducing governance lag and security gaps.

Going forward, industry practice can be used to further explore maturity assessment methods for AI security governance based on the “Six Can” capability dimensions. Different maturity levels may be defined according to governance system development, technical capability deployment, operational effectiveness validation,

and ecosystem collaboration.

4 AI Security Governance Initiative

AI security governance is a systematic, long-term endeavor requiring alignment among development orientation, security foundations, endogenous capabilities, and ecosystem collaboration. We hereby put forward this initiative on AI security governance, aiming to build industry consensus, pool collective efforts, and promote the establishment of an open, inclusive, safe, trustworthy, and collaboratively governed AI ecosystem.

4.1 Balance High-Quality Development with High-Level Security

Large-scale AI requires a dynamic balance between innovation and security. Industry stakeholders should treat security and trustworthiness as prerequisites for technological innovation, product services, and scenario applications, incorporating risk prevention, compliance requirements, and accountability constraints throughout AI development. Strong governance builds industry trust, while regulated development releases innovation vitality, enabling AI to serve high-quality economic and social development in a controllable, trustworthy, and reliable environment.

4.2 Build Robust and Reliable Operational Foundations

AI deployment depends on reliable infrastructure and a trustworthy operating environment. We call upon all stakeholders to strengthen coordination across key layers including networks, computing power, data, models, and applications, thereby enhancing the level of trusted access, reliable support, and stable operation of AI systems. This will provide a solid foundation for the deployment of innovative AI applications.

4.3 Embed Endogenous Governance Throughout Lifecycles

AI security governance cannot rely only on external constraints or post-incident remedies. We call upon all stakeholders to embed security requirements into the entire process of development, deployment, operation, and iteration, promoting the simultaneous construction and operation of security capabilities alongside business functions. This will enhance the trustworthiness and resilience of AI systems over the course of long-term operation.

4.4 Foster Open and Collaborative Security Ecosystems

AI security governance requires the participation and coordinated efforts of multiple stakeholders. We call upon stakeholders to cooperate on rule-making, standardization, evaluation, risk notification, capability recognition, and application demonstrations. Openness promotes innovation, collaboration strengthens security, and shared governance guides AI toward industrial upgrading, social governance, and public welfare.

AI is rapidly evolving into a highly interconnected and increasingly autonomous agentic ecosystem, driving a new phase of transformation in global digital and communications infrastructure. As modern AI deployment becomes increasingly dependent on continuous data flows, distributed edge computing, and real-time execution across connected environments, network operators play a foundational role in supporting large-scale AI applications. As key builders and operators of global connectivity infrastructure, network operators and ecosystem partners are well positioned to embed security capabilities into communications infrastructure and operational systems, making security and trustworthiness an endogenous foundation for AI development.

The “1-3-9” AI Security Governance Framework proposed in this white paper provides a systematic reference for industry stakeholders to advance AI security governance collaboratively. Going forward, GTI will continue to bring together global operators, technology partners, and industry stakeholders to promote standards coordination, capability sharing, and practical innovation. Together, we aim to build an open, inclusive, secure, trustworthy, and collaboratively governed AI ecosystem, and support the continued development of AI for the greater good.

Acronyms and Abbreviations

Abbreviation	Full Name
AI	Artificial Intelligence
LLM	Large Language Model(s)

References

[1] OWASP Foundation. OWASP Top 10 for Agentic Applications 2026. Available at: https://genai.owasp.org/

[2] State Council of the People's Republic of China. Opinions on Deepening the Implementation of the "AI+" Initiative. August 2025. Available at: https://www.gov.cn/zhengce/content/202508/content_7037861.htm

[3] European Parliament and Council of the European Union. Regulation (EU) 2024/1689 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act). 2024. Available at: https://eur-lex.europa.eu/eli/reg/2024/1689/oj

[4] The White House. National Policy Framework for Artificial Intelligence: Legislative Recommendations. 2026. Available at: https://www.whitehouse.gov/wp-content/uploads/2026/03/03.20.26-National-Policy-Framework-for-Artificial-Intelligence-Legislative-Recommendations.pdf

[5] California Legislature. SB-53 Frontier Artificial Intelligence Transparency Act. Available at: https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=202520260SB53

[6] Infocomm Media Development Authority (IMDA). Model AI Governance Framework for Agentic AI. Available at: https://www.imda.gov.sg/-/media/imda/files/about/emerging-tech-and-research/artificial-intelligence/mgf-for-agentic-ai.pdf

[7] Department for Science, Innovation and Technology (DSIT), UK. Code of Practice for the Cyber Security of AI. Available at: https://www.gov.uk/government/publications/ai-cyber-security-code-of-practice

[8] National Institute of Standards and Technology (NIST). Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1, 2023. Available at: https://doi.org/10.6028/NIST.AI.100-1

[9] International Organization for Standardization (ISO). ISO/IEC 42001:2023 Information Technology — Artificial Intelligence — Management System. 2023. Available at: https://www.iso.org/standard/42001

[10] 3GPP. TS 33.501: Security Architecture and Procedures for 5G System. Available at: https://www.3gpp.org/ftp/Specs/archive/33_series/33.501/

[11] National Institute of Standards and Technology (NIST). Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile. NIST AI 600-1, 2024. Available at: https://doi.org/10.6028/NIST.AI.600-1

[12] MITRE. Adversarial Threat Landscape for Artificial-Intelligence Systems (ATLAS). Available at: https://atlas.mitre.org/

[13] International Organization for Standardization (ISO). ISO/IEC 23894:2023 Information Technology — Artificial Intelligence — Guidance on Risk Management. 2023. Available at: https://www.iso.org/standard/77304.html

[14] National Institute of Standards and Technology (NIST). Zero Trust Architecture. NIST Special Publication 800-207, 2020. Available at: https://doi.org/10.6028/NIST.SP.800-207

[15] UNESCO. Recommendation on the Ethics of Artificial Intelligence. 2021. Available at: https://www.unesco.org/en/artificial-intelligence/recommendation-ethics

[16] Department for Science, Innovation and Technology (DSIT), UK. The Bletchley Declaration by Countries Attending the AI Safety Summit. 2023. Available at: https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration

[17] International Telecommunication Union (ITU). AI for Good Global Summit. Available at: https://aiforgood.itu.int/