

The background is a dark blue digital space filled with glowing hexagons, some containing icons like gears, a shield, a cloud, and a globe. A large, glowing blue shield shape is formed by two hands, one robotic on the left and one human on the right, holding bright blue light beams. In the center of the shield is a white outline of a person's head and shoulders. A yellow rectangular frame is positioned in the upper left, containing the main title and date.

Thriving in the AI era: Ensuring security and compliance for sustainable success

6 March 2025

■ ■ ■
The better the question. The better the answer.
The better the world works.

The EY logo consists of a yellow chevron pointing right, followed by the letters 'EY' in a bold, white, sans-serif font.

EY安永

Shape the future
with confidence
聚信心 塑未来



Thriving in the AI era: Ensuring security and compliance for sustainable success

6 March 2025

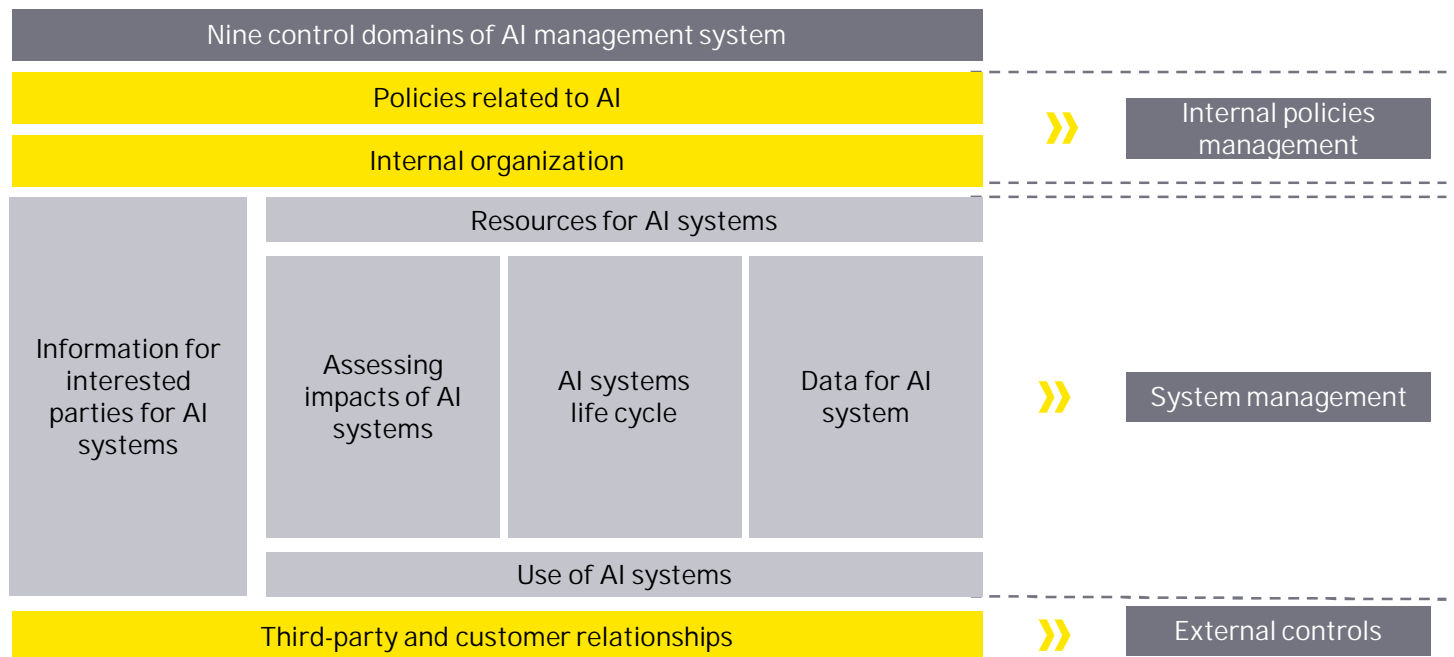


During the process of digital transformation, artificial intelligence (AI) has emerged the core driving force for innovative development across various industries. From intelligent risk management in finance and AI-assisted diagnostics in healthcare to personalized recommendations in retail and smart production scheduling in manufacturing, AI has swiftly and profoundly infiltrated the fundamental operations of organizations. However, this technological advancement reveals emerging crises. Security and compliance issues, including data breaches, algorithmic biases and model vulnerabilities, persistently threaten sustainable business growth.

Consequently, establishing a comprehensive AI security governance system throughout the entire lifecycle has become an inevitable strategy for organizations to thrive in the AI era.

Establishing an AI security and compliance management system: Setting the groundwork for security

With the rapid advancement of AI technology, constructing an AI security and compliance management system has emerged as the key to the steady growth of organizations. Leading the way as the world's first international standard for AI management systems, ISO/IEC 42001:2023 provides a definitive framework for organizations to guarantee the security, dependability and adherence of AI systems to standards throughout their lifecycle – from design and deployment to maintenance.

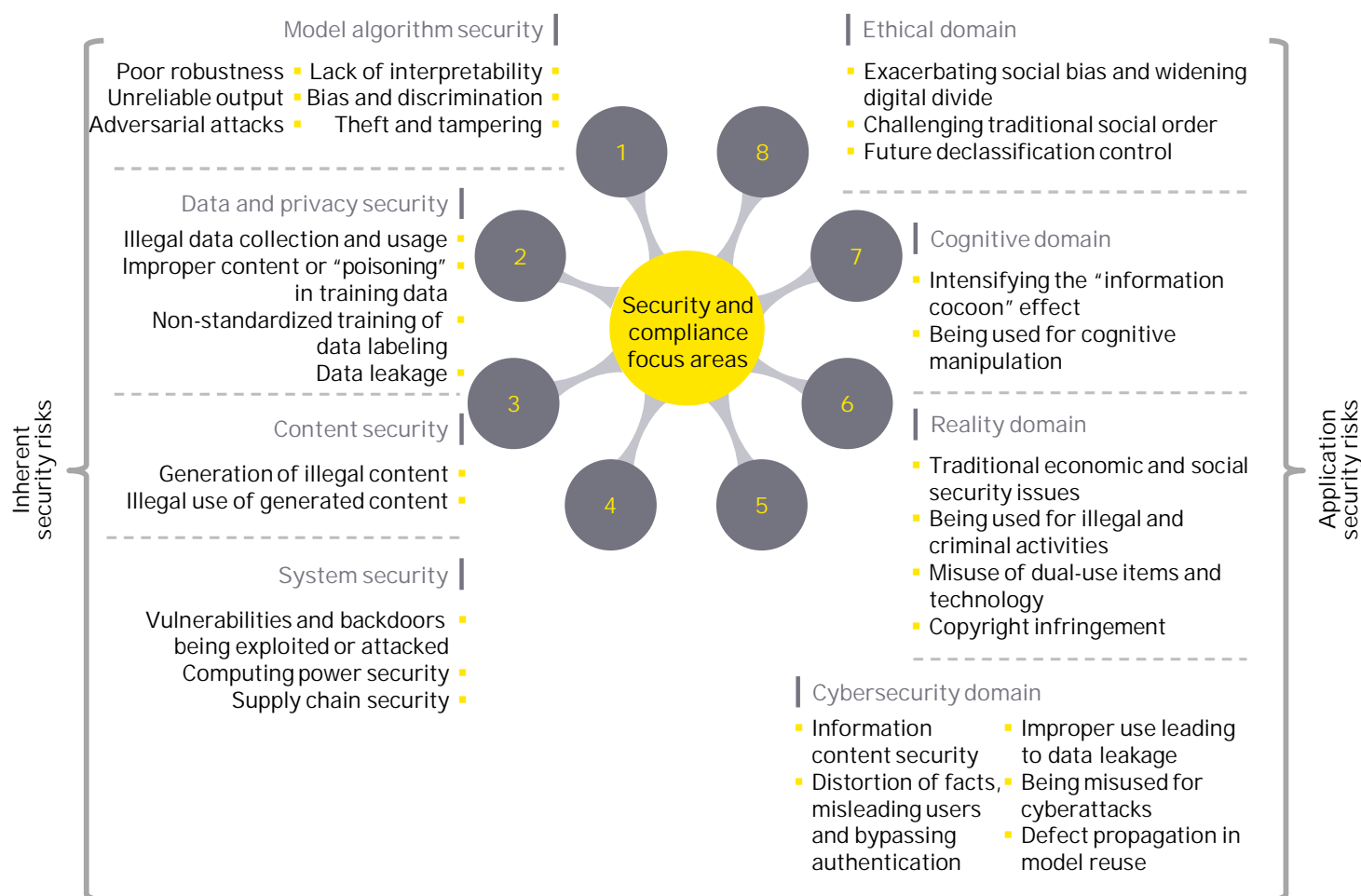


- Prepared with reference to ISO/IEC 42001:2023 appendices

Organizations should establish clear AI strategies and objectives while systematically identifying AI-related risks in alignment with relevant standards such as ISO/IEC 42001:2023. By implementing a management framework that addresses critical elements such as data, algorithms and models, organizations can strengthen process oversight while ensuring the transparency and explainability of AI systems. Furthermore, continuous personnel training programs and the cultivation of a security-conscious culture that prioritizes compliance awareness across all levels are vital for fostering robust AI governance. Through these coordinated efforts, organizations can develop a resilient AI security and compliance management system, enhance the effectiveness of AI applications, secure first-mover advantages in competitive markets and establish a sustainable foundation for long-term growth.

AI security evaluation: Identifying potential risks

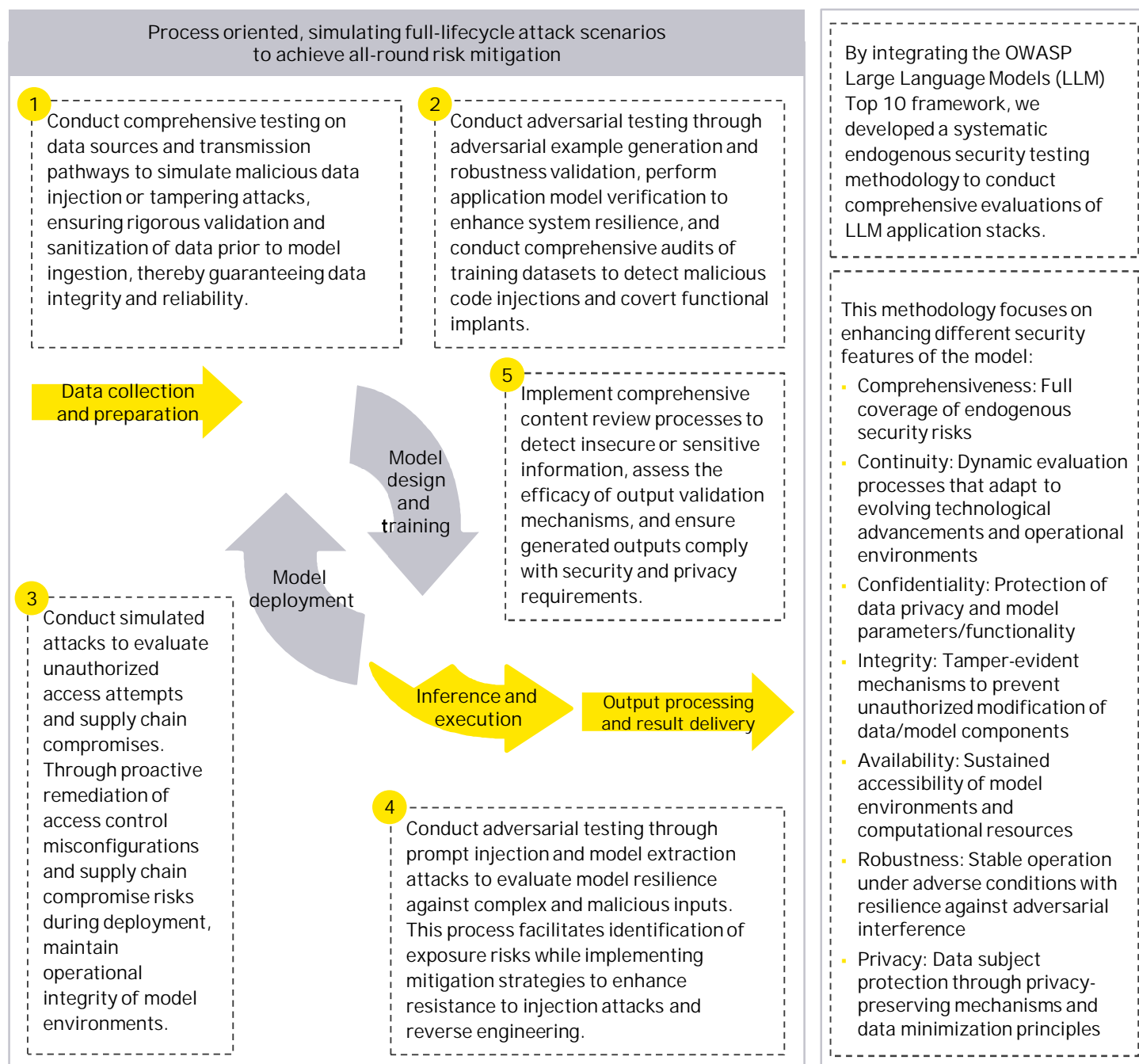
AI technology presents transformative opportunities while introducing multidimensional security challenges that can be systematically categorized into inherent security risks and application security risks. Inherent security risks primarily originate from the architecture and operational mechanisms of AI systems, encompassing hazards related to model algorithms, data privacy, content and system security. On the other hand, application security risks involve external threats spanning domains such as ethics, cognition, reality and network security.



To effectively mitigate these risks, organizations should comply with applicable laws and regulations while incorporating their unique business characteristics to establish a comprehensive and targeted security and compliance baseline. This baseline should address key areas including data security, algorithm transparency and model robustness, providing clear guidance for the development and deployment of AI systems. Organizations must regularly conduct thorough security and compliance assessments based on this baseline, identify potential risks, evaluate the efficacy of existing control measures and systematically refine risk management strategies in response to the assessment outcomes. This approach enables timely remediation actions and supports the implementation of a closed-loop risk management system for addressing risks at early stages.

Evaluating AI model technology: Fortifying defense capabilities






AI model security constitutes the cornerstone of organizational AI security practices. Organizations should implement a process-oriented approach to simulate the attack scenarios of whole data lifecycle, covering the full spectrum from data input to model output. This methodology enables comprehensive analysis of specific threat vectors including data poisoning, adversarial example attacks and model theft, thereby ensuring end-to-end risk mitigation across all phases of the AI model pipeline.



Concurrently, organizations can develop a customized framework for AI technology testing protocols aligned with their specific requirements, referencing industry standards such as the OWASP Top 10 for large language model (LLM). This methodology should incorporate rigorous evaluations across multiple risk dimensions including injection attacks, training data exposure and model misappropriation, aiming to identify vulnerabilities and define targeted protective measures. Through this systematic evaluation process, organizations can substantially enhance the security and resilience of their AI models, establish robust foundations for reliable deployment and drive innovative advancements within a secure framework.

AI algorithm and model documentation: Ensuring compliance operations

China's ongoing regulatory reinforcement in the AI sector is increasing mandatory requirements for algorithm and model filing as part of organization compliance obligations. The algorithm filing mandates comprehensive documentation for synthetic technologies such as image, video and audio generation. This ensures technical transparency and accountability to prevent misinformation dissemination and technology misuse. On the other hand, model filing focuses on overseeing technologies such as text and image generation to verify compliance with ethical standards and legal regulations, thereby mitigating the generation of harmful or deceptive content.

	★ Deep synthesis service algorithm filing (Algorithm filing)	Generative AI service filing (Model filing)
Coverage	 All commercial AIGC enterprises, including technology providers (offering APIs or SDKs) and service providers (via websites, apps, plugins, or mini-programs)	All commercial AIGC enterprises, organizations engaged in or providing generative AI services, technologies
Audit department	 Cyberspace Administration of China (CAC)	Branches of CAC at provincial level, CAC
Audit method	 Submitting materials online and likely of being selected for inspection, particularly in key or sensitive sectors (such as healthcare, manufacturing, and public big data industries). One-time review process.	Online filling, practical testing, offline audit
Filling cycle	 Taking two to six months, with an average processing time of approximately three months. Due to the increasing number of enterprises applying for filing, the required timeline may be extended.	Taking around three months, depending on the requirements of different provincial CAC
Filing stage	 Within 10 working days after launch	Prior to official launch, after completing internal testing and with a product or demo ready for external release.
	★ Mandatory filing Non-mandatory requirements, relevant departments are preparing implementation rules	With reference to the Administrative Provisions on Deep Synthesis in Internet-based Information Services and the Internet Information Service Algorithm Filing System (https://beian.cac.gov.cn)

Organizations should actively engage in the filing process, recognizing it as an opportunity to improve their management capabilities and overall competitiveness. This process enables organizations to conduct a comprehensive review of their algorithm logic, data sources and use cases, identify potential risks and devise mitigation strategies. Through algorithm and model filing, organizations can not only fulfill obligations to prevent legal penalties and financial liabilities from non-compliance, but also strengthen technical transparency and cultivate public confidence. This proactive engagement contributes to developing secure and trustworthy AI ecosystem, fostering the healthy growth of the industry. Moreover, it establishes standardized compliance frameworks for organization development pathways, ensuring AI technology maintain a balance between innovation and security, thereby advancing sustainable technological progress.

AI empowering security and compliance management: Innovative governance strategies

Amid accelerating AI technological advancements, organizations must maintain vigilance against emerging risks while strategically leveraging AI to augment security and compliance capabilities. For example, in cybersecurity and data protection, organizations can deploy AI-driven systems for real-time cyberthreat detection. By utilizing machine learning algorithms to analyze extensive network data, AI can effectively detect abnormal traffic patterns and potential cyberattacks. Furthermore, AI technologies can facilitate intelligent data encryption and anonymization protocols by automatically identifying and processing sensitive data, thereby strengthening data protection measures. Beyond threat mitigation, AI empowers organizations to address evolving security threats while optimizing compliance workflows, thereby improving operational efficiency and regulatory adherence.

Conclusion

EY teams play a pivotal role in aiding organizations in navigating the complex landscape of AI security and compliance management. By delivering profound insights and guidance, EY team can assist organizations in leveraging AI to strengthen their security and compliance frameworks. Additionally, EY team is adept at formulating pragmatic strategies tailored to the unique needs of each organization.

Organizations must embrace a forward-thinking perspective, replacing the conventional “after-fact remediation” approach with integrated “prevention-detection-response” governance model. By proactively addressing challenges and swiftly seizing opportunities, organizations can establish a solid security infrastructure on the path of technological innovation, achieving competitive differentiation and reaping the full benefits of AI technology in a fiercely competitive market.

For more information, please contact us:



Kelvin Gao

Partner
Cybersecurity, Technology Consulting
Ernst & Young (China) Advisory Limited
kelvin.gao@cn.ey.com



Alice Bao

Partner
Cybersecurity, Technology Consulting
Ernst & Young (China) Advisory Limited
alice.hx.bao@cn.ey.com



Chao Zuo

Director
Cybersecurity, Technology Consulting
Ernst & Young (China) Advisory Limited
chao.zuo@cn.ey.com



Serena Zhang

Manager
Cybersecurity, Technology Consulting
Ernst & Young (China) Advisory Limited
serena.s.zhang@cn.ey.com

EY | Building a better working world

EY is building a better working world by creating new value for clients, people, society and the planet, while building trust in capital markets.

Enabled by data, AI and advanced technology, EY teams help clients shape the future with confidence and develop answers for the most pressing issues of today and tomorrow.

EY teams work across a full spectrum of services in assurance, consulting, tax, strategy and transactions. Fueled by sector insights, a globally connected, multi-disciplinary network and diverse ecosystem partners, EY teams can provide services in more than 150 countries and territories.

All in to shape the future with confidence.

EY refers to the global organization, and may refer to one or more, of the member firms of Ernst & Young Global Limited, each of which is a separate legal entity. Ernst & Young Global Limited, a UK company limited by guarantee, does not provide services to clients, nor does it own or control any member firm or act as the headquarters of any member firm. Information about how EY collects and uses personal data and a description of the rights individuals have under data protection legislation are available via ey.com/privacy. EY member firms do not practice law where prohibited by local laws. For more information about our organization, please visit ey.com.

© 2025 Ernst & Young, China.
All Rights Reserved.

APAC no. 03022722
ED None

This material has been prepared for general informational purposes only and is not intended to be relied upon as accounting, tax, legal or other professional advice. Please refer to your advisors for specific advice.

ey.com/china

Follow us on WeChat
Scan the QR code and stay up-to-date with
the latest EY news.

