California Review Management

TECHNOLOGY

Cyber Risk Governance (CRG) in the Age of AI-Driven Open Innovation

by Chon Abraham

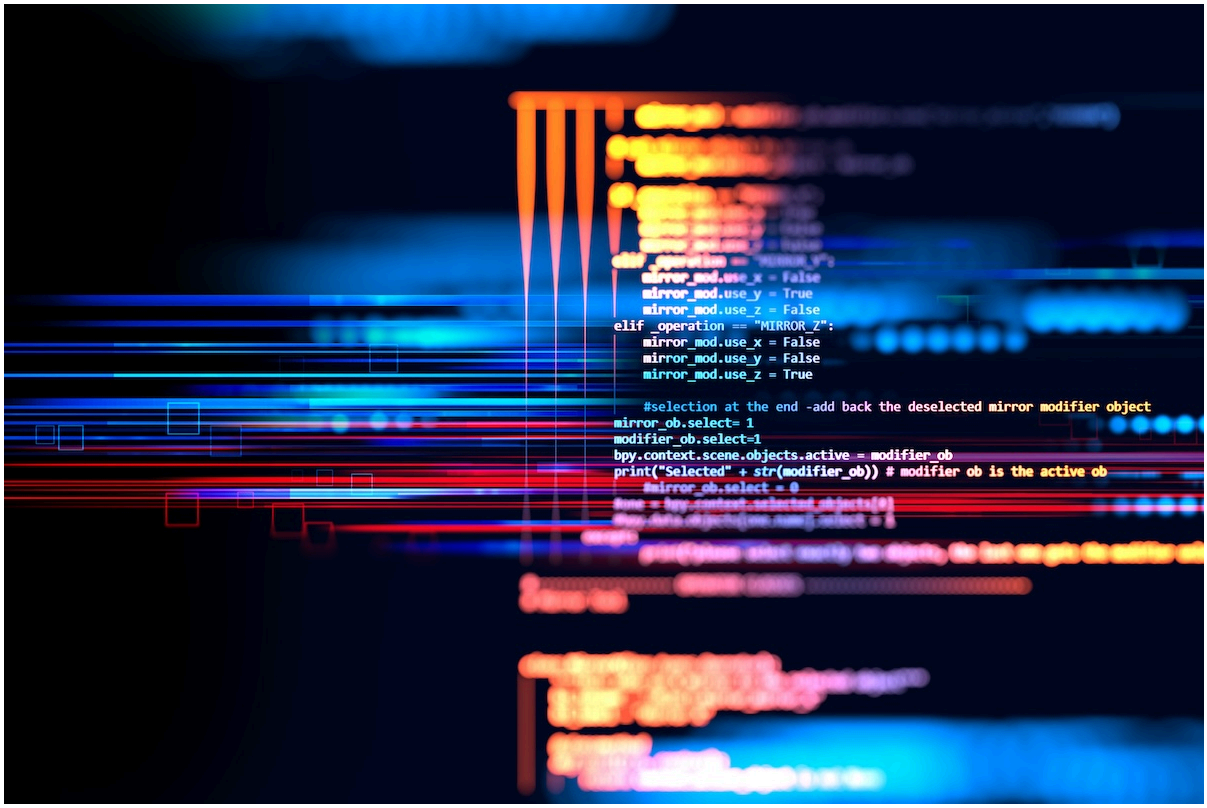

Image Credit | monsitj

AI supercharges OI with unparalleled benefits but unleashes cyber risks—demanding bold and proactive governance.

INSIGHT | OPINION 01 Apr 2025

Introduction

The rapid integration of artificial intelligence (AI) into open innovation (OI) is transforming industries with its efficiency and scalability (Holgerson et al., 2024). However, it also introduces unique cyber risks that require robust governance. CRG encompasses strategic oversight and accountability to mitigate cybersecurity risks while aligning innovation with growth objectives. As AI reshapes OI, the increased reliance on external knowledge flows, advanced models, and collaborative ecosystems creates new vulnerabilities. Proactively addressing these risks is essential to maintaining trust, efficiency, and competitive advantage. This article provides recommendations to guide C-suites and boards in managing CRG within AI-enabled OI.

RELATED CMR ARTICLES

M. Holgersson, L. Dahlander, H. Chesbrough, & M. L. Bogers, "**Open Innovation in the Age of AI**," California Management Review, 67/1 (2024): 5-20.

Dimensions of Cyber Risk

AI's integration into OI introduces opportunities for cyber exploitation across three equally important key dimensions: data security and privacy, algorithmic integrity, and collaborative ecosystems and markets.

1. **Data Security and Privacy**: AI's role in OI heightens concerns about safeguarding data and protecting sensitive information. Federated learning—collaborative training of AI models without sharing raw data—preserves privacy but is vulnerable to "data poisoning," where attackers introduce corrupt data to disrupt performance. Synthetic data—artificial datasets mimicking real ones—offers some protection but can expose sensitive patterns or inaccuracies. For instance, synthetic patient data in healthcare

projects may reveal vulnerabilities if manipulated. Generative AI systems, which create content like text or images, also face risks if their training data is insecure, leading to breaches of intellectual property or innovation strategies.

2. **Algorithmic Integrity**: AI models in OI are susceptible to bias and tampering, undermining trust. Attackers may insert a backdoor—malicious code embedded during training—to manipulate outputs or degrade performance. For example, in financial OI, tampered models could enable fraud by altering risk assessments. Robust auditing and validation processes are crucial to ensure algorithmic transparency and resilience.
3. **Collaborative Ecosystems and Markets**: Platforms supporting AI-driven marketplaces rely on trusted partnerships. Federated learning can spread harmful inputs, reducing reliability. Synthetic data, though protective, can be exploited for counterfeit datasets. AI systems using web-scraped data risk including sensitive or proprietary information. In AI-enabled markets, malicious actors may distribute counterfeit content or embed malware, leading to legal disputes or reputational damage if compromised datasets result in flawed products or IP theft.

Recommendations for CRG in AI-Driven Open Innovation

The recommendations are tailored to address the unique risks of AI-enabled OI while fostering alignment between cybersecurity and innovation goals.

1. **Understand and Prioritize AI-Specific Risks**

C-suites must adopt a proactive approach to identify and prioritize AI-related risks within OI ecosystems. This involves mapping critical assets, dependencies, and threats, ensuring that cybersecurity measures align with the organization's innovation strategy and entails the following:

- Adopting AI-Centric Risk Frameworks: Leverage established frameworks like the NIST AI Risk Management Framework (NIST, 2023) or ISO/IEC 42001 (ISO, 2023), adapted for AI-related challenges such as data poisoning and adversarial attacks.

- Mapping Critical Innovation Assets: Conduct a comprehensive inventory of AI-enabled assets, including datasets, algorithms, and collaborative tools, categorizing their significance to the innovation process. Define the crown jewels (e.g., critical assets such as data stores, AI model training environments, proprietary synthetic data generators, federated learning frameworks, and API integrations) of the business that are involved in AI-driven innovation to ensure targeted protection.
- Assessing Ecosystem Dependencies: Analyze interdependencies within the OI ecosystem to identify potential vulnerabilities in shared platforms, partner networks, and data pipelines. Establish mechanisms to evaluate and mitigate third-party risks.

1. **Quantifying and Communicating AI Cyber Risks**

Effective CRG in AI-driven OI requires quantifying cyber risks in financial and operational terms to foster clarity and actionable decision-making that require the following:

- Integrating Cyber Risk Quantification (CRQ) Tools: Employ CRQ tools to model potential financial impacts of AI-related risks. For instance, Factor Analysis of Information Risk for Artificial Intelligence Risks (FAIR -AIR) (FAIR, 2024) offers a methodology to quantify and prioritize AI-specific threats (e.g., algorithmic poisoning, model theft, etc.) that incorporate scenario modeling to simulate AI-specific breaches and evaluate mitigation strategies.
- Translate Risks into Business Contexts: Communicate AI-related risks using tangible metrics using a method like FAIR AIR, such as potential financial losses or reputational damage. For example, explain that a compromised AI model could lead to $2 million in regulatory fines, justifying a $200,000 mitigation investment.
- Enhance Risk Reporting for Stakeholders: Develop concise, scenario-driven reports that bridge the technical and strategic aspects of AI-related cyber risks, enabling C-suites and Boards to make informed decisions.

1. **Building and Sustaining Governance for AI-Driven Open Innovation**

A mature CRG framework ensures resilience against AI-related threats while promoting sustainable OI involves:

- Defining Governance Structures: Assign accountability for AI-related cyber risks across board, executive, and operational levels. Empower C-Suite data and process owners, including CISOs, CIOs, CDAOs, CPOs, and OI business leaders, to integrate cybersecurity into AI innovation strategies. Establish AI risk committees to define risk tolerance and align cybersecurity with innovation goals. Embed ethical principles into AI development using practices like federated learning and explainable AI to enhance transparency and trust.
- Implement Adaptive Policies and Metrics: Develop dynamic policies that evolve with emerging threats and AI technologies. Incorporate lessons from incidents and audits to refine governance practices. Use key performance indicators (KPIs) to monitor governance effectiveness, focusing on metrics like model reliability, compliance, and risk mitigation ROI.
- Promote Collaborative AI Security in Innovation Ecosystems: Develop multi-stakeholder governance models tailored to the unique risks of AI in open innovation. Foster secure collaboration by implementing AI-specific cybersecurity measures, such as secure federated learning frameworks, encrypted model-sharing platforms, and AI-based threat detection systems to strengthen collective resilience while enabling innovation.

Conclusion

AI-driven OI is transforming industries, but it also magnifies cyber risks that can jeopardize trust and operational stability. By adopting this three-component framework for CRG—focused on understanding risks, quantifying them, and maturing governance—C-suites and boards can ensure that innovation thrives securely.

References

Holgersson, M., Dahlander, L., Chesbrough, H., & Bogers, M. L. (2024). Open Innovation in the Age of AI. California Management Review, 67(1), 5-20.

International Standards Organization (ISO) (2023). ISO/IEC 42001 Information technology -Artificial Intelligence Management System. **https://www.iso.org/standard/81230.html**

National Institute of Standards and Technology (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0). **https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf**

Factor Analysis of Information Risk (FAIR) (2024). FAIR Artificial Intelligence Risk (AIR) Approach Playbook. **https://www.fairinstitute.org/blog/fair-artificial-intelligence-ai-cyber-risk-playbook**

Chon Abraham Follow

Chon Abraham, PhD, Mansfield Term Professor at William & Mary, focuses on AI, cybersecurity, and health IT. She is a retired cyber officer (last serving in the Air Force Chief Data & AI Office), cybersecurity Fulbright Scholar to Japan, and published author in leading journals.