MIT MANAGEMENT SLOAN SCHOOL

AI-Proofing The Board and The C-suite: Managing AI Business and Security Risk

Research Position Paper

Dr. Keri Pearlson • June 25, 2025

AI capabilities are advancing at breakneck speed, from preschool-level intelligence in 2020 to high school-level now and college level in the near future[1]. Industry leaders have compared AI's transformational potential to civilization's greatest technological leaps:

> *"[AI could be] as transformational as some of the major technological inventions of the past several hundred years: Think the printing press, the steam engine, electricity, computing and the Internet"* - said Jamie Dimon of JP Morgan[2]
>
> *"To the three great technological revolutions–the agricultural, the industrial, and the computational–we will add a fourth: the AI revolution"* - said Sam Altman of OpenAI[3]

However, leaders must view enterprise adoption with caution. The infrastructure to assess and manage AI risk has not kept pace with the evolution of the technology itself. Enterprise leaders must walk a tightrope between security and business risk and operational advantages. For example:

- **Not adopting AI opens the door for competitors to capture share.** Goldman Sachs boasts that AI can already do work in minutes that once took teams weeks - like drafting 95% of an IPO prospectus.[4]
- **Recklessly adopting AI creates financial and reputational risk**. Air Canada faced a lawsuit and PR disaster as their AI chatbot fabricated a refund policy and they refused to honor it.[5]
- **AI adoption can increase productivity but at the same time it can deskill the work force**. Junior lawyers used to learn about details of cases from review of documents during discovery in many litigation cases, but with eDiscovery platforms and advanced AI search capabilities, these skills are no longer learned.[6]

Enterprise adoption of AI is growing rapidly. Enterprise AI spending has increased 130% since 2023 and weekly business leader AI usage surged from 37% to 72% according to a Wharton survey[7]. However, the same study summarized key risks: customer data privacy and security breaches remain the top concerns preventing AI adoption. PwC's 2025 Global Digital Trust Insights Survey highlights the same trend: 67% of security leaders report that GenAI has expanded their attack surface[8]. This mirrors cybersecurity's early days when rapid technology adoption outpaced security controls. it took two decades of standardization, measurement, and accountability mechanisms to close this gap.

Existing approaches face shortcomings

Traditional risk frameworks weren't designed for AI's speed and complexity. This is both a people and technical problem - risk managers understand business impact but lack AI technical depth, while technical

[1] Situational Awareness (2025): https://situational-awareness.ai/from-gpt-4-to-agi/

[2] CNBC (2024): https://www.cnbc.com/2024/04/08/jamie-dimon-says-ai-may-be-as-impactful-on-humanity-as-printing-press-electricity-and-computers.html

[3] Sam Altman (2021): https://moores.samaltman.com/

[4] Fortune (2025): https://fortune.com/2025/01/17/goldman-sachs-ceo-david-solomon-ai-tasks-ipo-prospectus-s1-filing-sec/

[5] BBC (2024): https://www.bbc.com/travel/article/20240222-air-canada-chatbot-misinformation-what-travellers-should-know

[6] Association of Corporate Counsel (2025): https://www.acc.com/artificial-intelligence-new-junior-lawyer

[7] AI at Wharton (2024): https://ai.wharton.upenn.edu/wp-content/uploads/2024/11/AI-Report_Full-Report.pdf

[8] PwC (2025): https://www.pwc.com/gx/en/news-room/press-releases/2024/pwc-2025-global-digital-trust-insights.html

teams are moving fast to capture competitive advantage, often outpacing enterprise risk functions. Board oversight cycles are quarterly; AI development cycles are weekly.

Some enterprise approaches to managing the business and security risks associated with AI include:

- **Adapting existing frameworks to AI:** For example, some CISOs are attempting to extend SOC 2 to cover AI-related risks. However, this approach is ineffective because AI introduces fundamentally different types of risks, such as hallucinations, that SOC 2 was not designed to address.
- **Adopting new AI frameworks as they are developed**: For example, the NIST AI Risk Management Framework (RMF) introduces new tools to help businesses manage AI-related risks. However, it falls short in practice because it provides only high-level principles, making it difficult for organizations to assess or measure how well they are following the framework.
- **Waiting for new approaches to be developed**. But this isn't a viable strategy given the intense competitive pressure businesses face; they can't afford to stand still while others move ahead.

Questions to explore in our new multi-company research project

We need a new approach to enable enterprise adoption of AI at scale. To get there, we are launching a multi-company research project to bring together industry leaders and answer the key questions and identify emerging best practices. Our goal is to create a framework to guide executives to identify and mitigate risks. Some of the questions this research will ask include:

1. What makes AI a Board level risk and what are the implications for how to manage it?
2. How do executives address the risk associated with AI adoption today? What are best practices others can learn from?
3. What are the unique risks for AI projects that the C-suite/board must track? How can these risks be measured and managed?
4. What are the C-suite/board level considerations needed to make responsible decisions about AI?
5. What should be included in an AI framework for business and security risk management? How can we create an actionable scorecard for C-suite executives and the board of directors?

Our hypotheses for the research project

Hypothesis 1: A business and security framework must meet 5 design criteria:

1. **Adoption-focused**: action-oriented focus rather than theory-focused
2. **Transparent**: clear line of site from shared learning to actions and recommendations
3. **Auditable**: objective, testable criteria with clear thresholds for acceptance
4. **Adaptable**: continuously updated as new AI capabilities and threats emerge
5. **Comprehensive**: covers the full range of societal and commercial concerns in one place.

Hypothesis 2: There are 6 risk categories that are top of mind for enterprise leaders to include in the framework.[9]

1. **Data & Privacy:** AI leaks and unauthorized AI system access or usage of user and enterprise data.
2. **Safety:** AI generated harmful outputs to users and other safety concerns.

[9] Adapted from the AIUC-1 protocol shared with the author from interviews of founders of the Artificial Intelligence Underwriting Company. https://aiuc1.com/

3. **Security:** Adversarial attacks, data poisoning, unauthorized access and other potential security concerns.
4. **Reliability:** Hallucinations and other areas that cause the system to be unreliable.
5. **Accountability:** Unclear or missing oversite activities of AI systems.
6. **Society:** Societal harm or national security risks that arise from cyberattacks.

Our research project will test these hypotheses and build on them through surveys, interviews, case studies, research, and input from the participants in our multi-company research project. Phase 1 of our research will result in at least one white paper of findings and a proposed framework. Additional articles and academic papers are anticipated.

Solutions to managing business and security risk at scale will be a multi-faceted approach, likely including standards, insurance, new training approaches, technical oversight tools, and governance processes. This was the most successful approach for managing cybersecurity where the facets included SOC 2 compliance became table stakes for enterprises and for their vendors, MFA adoption reduced their breach risk by 99.9%[10], cybersecurity trainings and enterprise phishing tests became standard practice, and boards developed dashboards to enable them to oversee the risk and adherence.

Call to Action: Contribute Ideas and Funding to this Project

Funding: We are seeking corporate and individual funders to support this research initiative. Funders will have the opportunity to shape the research agenda, participate in the research process, gain early access to insights, and receive recognition in all project publications and presentations. Funders will also be invited to share published results on their own platforms and may adapt the findings for internal use, with appropriate attribution and authorship acknowledgment. We aim to build a cohort of 5–10 forward-thinking funders to support this work. Research will launch as soon as the minimum funding commitment is secured.

Participation: Once funding is secured, we invite board members and executives from other enterprises who are involved in assessing AI benefits and risk to participate. Participants will be invited to contribute to case studies, respond to surveys/interviews, participate in discussions, and provide feedback to reports and white papers.

This QR code will take you to a short Qualtrics survey where you can give us input and express your interest in this project. The direct link is:

https://bit.ly/AIBizSecSurvey1

For more information, please contact Dr. Keri Pearlson, kerip@mit.edu.

[10] HIPAA Journal: https://www.hipaajournal.com/multi-factor-authentication-blocks-99-9-of-automated-cyberattacks/