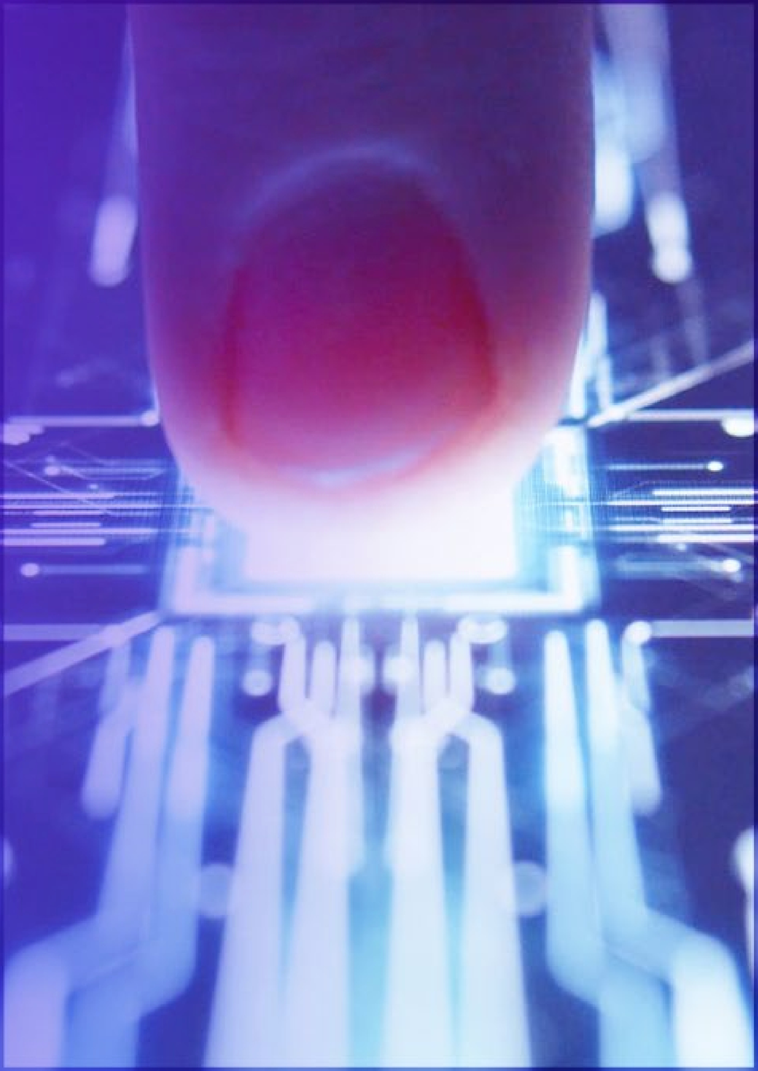
KPMG

The leadership guide to securing AI



kpmg.com

Dear Readers,

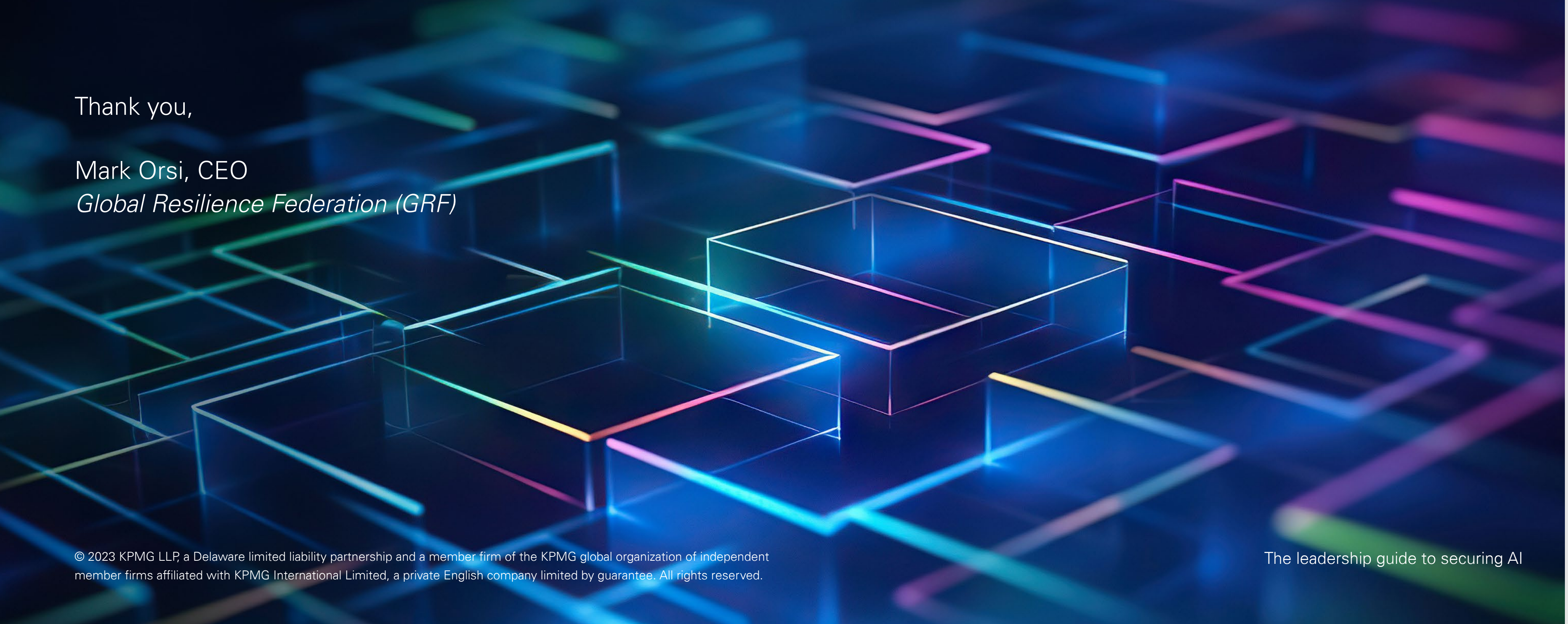
Artificial intelligence (AI) is critical to the future success and health of companies across industries. To empower emerging and current AI security leaders, the Global Resilience Federation (GRF) convened an exceptional working group and asked KPMG to facilitate in-depth ideation sessions and interviews with AI and security practitioners from more than 20 leading companies, think tanks, academic institutions, and industry organizations.

We believe that the output—*The Leadership Guide to Securing AI*—will be of tremendous value to the GRF community and to other organizations seeking to explore this groundbreaking technology.

Finally, we would like to express our sincerest appreciation for KPMG's leadership and our community partners' support. All invested significant time, thought, and energy to push forward an insightful and timely guide on how organizations can approach securing AI. I am confident that readers will derive insights that will guide them through the AI revolution, a trend that is poised to positively shape their futures.

Thank you,

Mark Orsi, CEO
Global Resilience Federation (GRF)



Introduction

This guide has been produced to support leaders who are responsible for adopting artificial intelligence (AI) while simultaneously navigating the uncharted waters of securing AI. Balancing a clear AI security strategy while keeping pace with innovation may very well distinguish those organizations that successfully adopt AI from those that don't.

Secure AI is central to innovation enablement:
AI technologies promise to improve organizations' ability to process and analyze large volumes of complex data, leading to strategic insights related to competitor activity, market fluctuations, the viability of new products and services, and more.

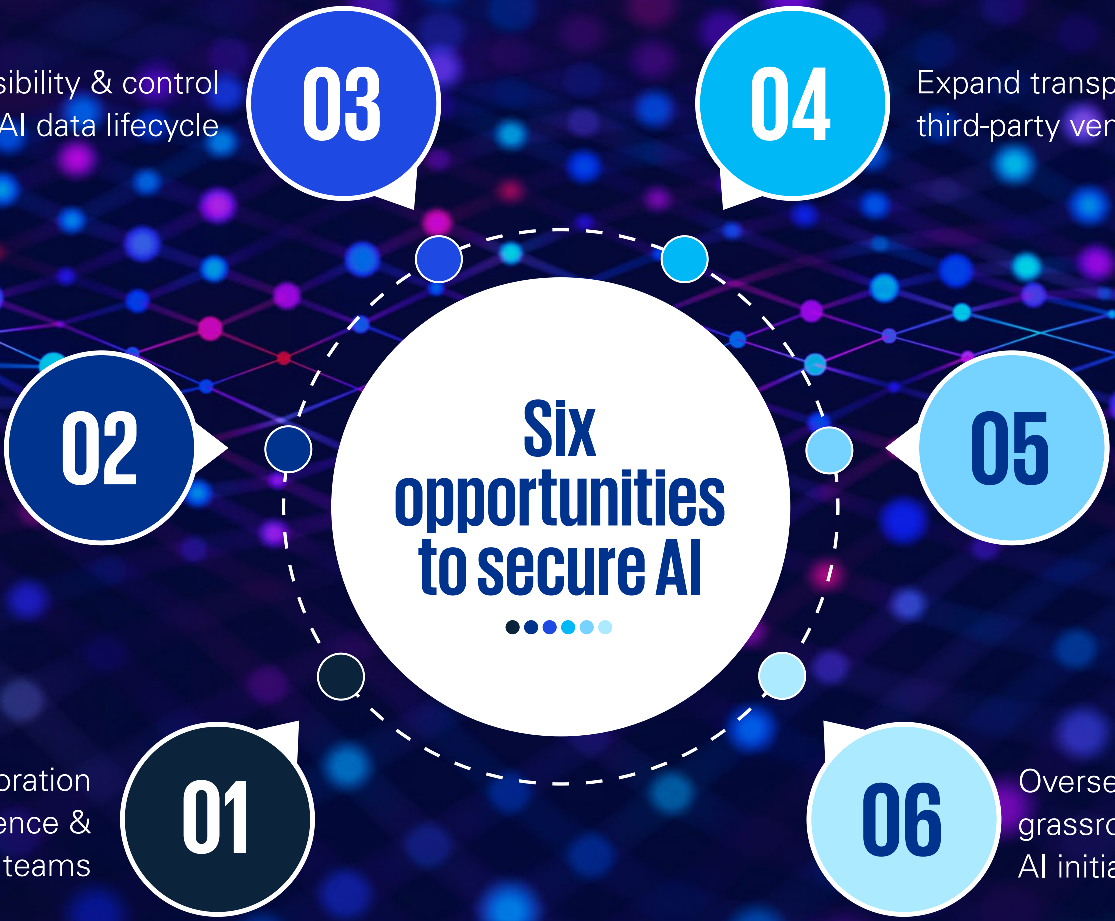
And yet, the demands for tangible AI security are real—from industry-specific Requests for Information to internal pressures to create Acceptable Usage Policies as employees test public AI tools. We see the industry shifting to respond, evidenced by the recent emergence of controls such as the MITRE ATLAS framework, NIST AI RMF 1.0, Google's Secure AI Framework, and the Microsoft Responsible AI Standard, all of which outline how to pursue Secure AI models.

> "Some companies are so excited about the possibilities of AI that the security and assurance of these evolving AI-enabled systems can be more of an afterthought."
>
> *—Dr. Christina Liaghati, AI Strategy Execution & Operations Manager, MITRE*

This paper is designed to expand upon those efforts by outlining six major opportunities to realize Secure AI, as well as six strategic imperatives corporate leaders should bear in mind as they make Secure AI central to their innovation efforts. This guidance is the result of months of KPMG- and Cranium AI-led interviews and working sessions with corporate AI leads and CISOs who are proponents of approaching AI innovation securely.

01
02
03
04
05

Six opportunities to secure AI



01 Promote collaboration between data science & cyber teams

02 Reduce & monitor a nuanced attack surface

03 Increase visibility & control over AI data lifecycle

04 Expand transparency into third-party vendors' AI use

05 Integrate evolving regulatory requirements

06 Oversee security of grassroots & enterprise AI initiatives

01
02
03
04
05

01 Promote collaboration between data science & cyber teams

Although some companies have been using AI for many years, cybersecurity standards around these efforts have been largely unregulated by the government. Further, security concerns haven't often been addressed in tandem with AI-centered solutions. Even today, the number of companies that can say their data science and security teams collaborate regularly are few and far between.

The approach to building AI models in Corporate America is a lot like the Wild West. At many companies, AI projects are pursued with a lack of definition around who bears responsibility for security, ethics and governance; a need for more rigor around budget and resourcing; and, perhaps most important, inconsistency when it comes to protecting intellectual property and organizations at large.

Research cited in the AI Index 2023 Annual Report from Stanford University validates the need to bring cybersecurity to the table early and often.[1] In the report, cybersecurity is identified as the most relevant risk when adopting AI technology (59%), followed by regulatory compliance (45%), and individual privacy (40%).

02 Reduce & monitor a nuanced attack surface

While AI opens a whole new world of innovation, it also introduces new attack vectors for cyber criminals. Attacks can range from adversaries focusing on the vulnerabilities of AI, to leveraging AI as an enabler of malicious schemes, such as those associated with WormGPT.[2]

In addition to traditional threats like ransomware and viruses, other threats specific to the AI environment are more insidious and include model evasion, data poisoning, inference, and functional extraction. [For more, please read our practitioners' guide to managing AI security, "Balancing ROI and Risk."] Further, nefarious spinoffs of Generative AI have the power of an LLM behind them, which could elicit much more convincing phishing emails.[3]

> **"The pace of AI adoption within enterprises demands that organizations ensure not only that their AI is secure, but also that training and awareness are provided to stakeholders and users to promote safe and responsible use of AI."**
>
> *—Katie Boswell, KPMG AI Security Leader*

[1] Nestor Maslej, Loredana Fattorini, Erik Brynjolfsson, John Etchemendy, Katrina Ligett, Terah Lyons, James Manyika, Helen Ngo, Juan Carlos Niebles, Vanessa Parli, Yoav Shoham, Russell Wald, Jack Clark, and Raymond Perrault, "The AI Index 2023 Annual Report," AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2023.

[2] THN, "WormGPT: New AI Tool Allows Cybercriminals to Launch Sophisticated Cyber Attacks," The Hacker News, July 15, 2023.

[3] Alex Stroxton, "Cyber-criminal AI tool WormGPT produces 'unsettling' results," Computer Weekly, July 19, 2023.

03 Increase visibility & control over AI data lifecycle

As organizations' willingness to leverage AI increases, a massive amount of data will be required to create valuable and accurate models. With a greater volume of data—both from individual organizations and, in an enhanced form, from data brokers—privacy risks will also increase. Unfortunately, when used for training, AI models that ingest intellectual property, trade secrets, PII, or PHI aren't always created with guardrails that protect against data loss.

And even when the same data is used for outcomes that will be made public or used in collaboration with another firm, it is not always anonymized.

Although AI's power to detect patterns and synthesize structured and unstructured data is clear, companies aren't always focused on the fact that getting full value out of AI tools requires robust data curation. Further, there are additional risks when multiple datasets are linked (intentionally or unintentionally). AI only knows what it is fed. If a dataset is poisoned—inherently or maliciously—it could take a long time for the human beings in the loop to detect whether the model is proffering inaccurate guidance that could inform critical business decisions. To protect against these missteps, companies need to thoughtfully consider how security fits into their AI data lifecycle.

> “Organizations can be less concerned about attacks like data poisoning if they are only harvesting data in their own environment. The situation is riskier when you blend data with external partners.”
>
> *—Ryan Boulais, Global CISO, AES Corporation*

04 Expand transparency into third-party vendors' AI use

In these early days of AI, some of the biggest risks companies face come from third parties with which they do business. Both from the perspective of merged datasets and of data that is less controlled, it is quite likely that software vendors and service providers will introduce AI into their interactions with client companies. And it is not uncommon for vendors to do so without disclosing exactly what AI models they are using and how they are being secured.

Risk introduced via an AI system used by a third party is more than a technological headache. It could open companies to a wide variety of liabilities, including:

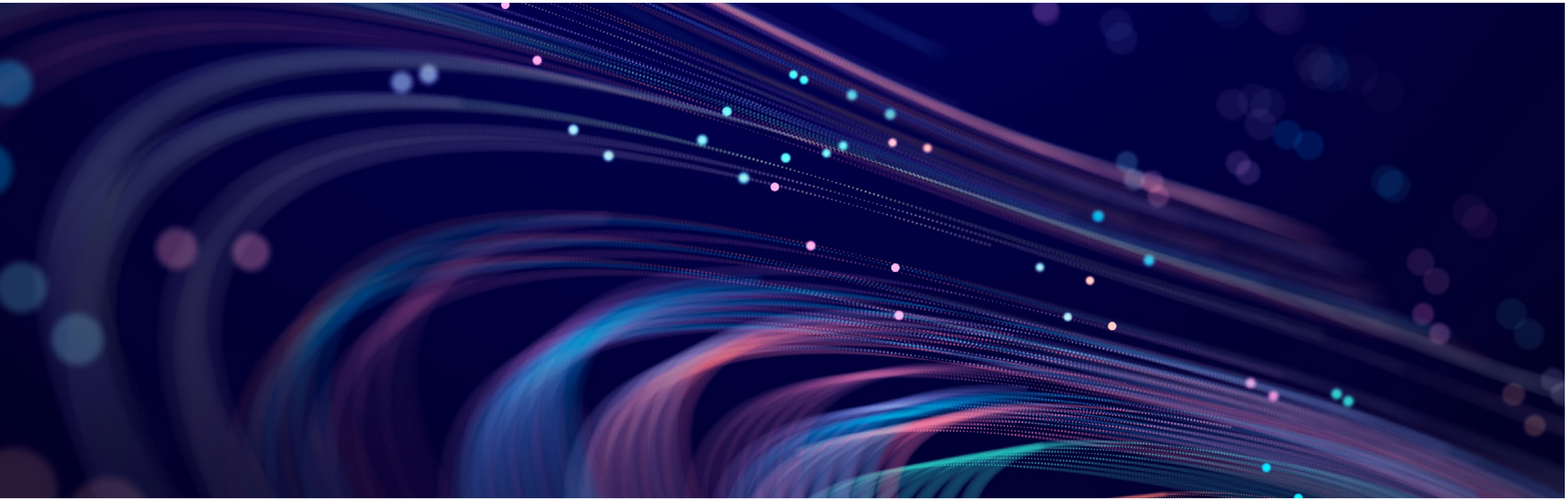
- Exposure of sensitive or confidential client data to the broader public if a vendor's AI solution pulls data back into its model
- Propagation of inaccurate or biased recommendations to clients who rely on data-driven decisions
- Running afoul of EU-based regulations, such as GDPR, if the vendor hasn't anonymized data from customers in the EU
- Potential legal consequences associated with fraud and abuse, even if a company is unaware that a vendor has vulnerabilities in their black box software

05 Integrate evolving regulatory requirements

At present, government agencies and Congress are trying not to be so restrictive with AI regulations that U.S. adoption of AI and pursuit of innovation solutions are stifled. Although the Biden-Harris administration has put forth a set of considerations to which it strongly suggests AI firms commit, none are set in stone, and they will evolve as public/private collaboration progresses.[4]

In the interim, the government is counting on private industry to provide learnings from their work that will likely inform future regulations.[5] Without this input, there are risks that regulations could be uninformed and fail to accomplish their original intent, or that they could be so stringent that innovation is stifled altogether.

"Collaboration between the private and public sectors will have a permanent impact on how AI is adopted for both every-day and consequential applications," said Dr. Christina Liaghati, AI Strategy Execution & Operations Manager, MITRE. "As industry leaders develop AI use cases, they have an opportunity to provide insights to the government and broader AI community on how best to protect and assure these systems. These insights can directly inform everything from standards to policy and regulations and help ensure AI is as effective as it can be for the entire community."



[4] FACT SHEET: "Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI," The White House, July 21, 2023

[5] Cecilia Kang, "OpenAI's Sam Altman urges AI regulation in Senate hearing," The New York Times, May 16, 2023

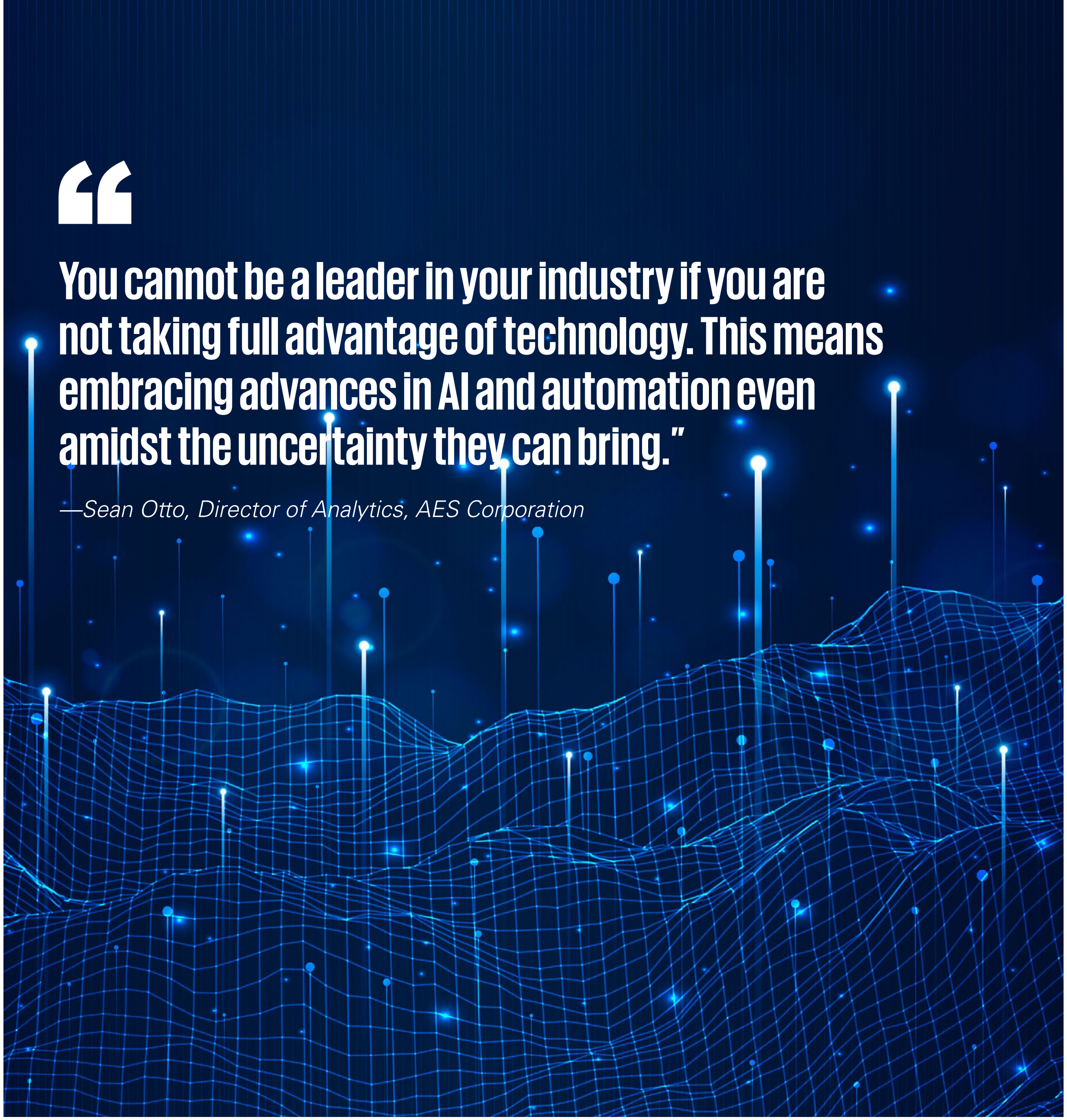
01
02
03
04
05

06 Oversee security of grassroots & enterprise AI initiatives

According to The AI Index 2023 Annual Report from Stanford University, the industry with the largest AI investments at present is healthcare ($6.1 billion); followed by data management, processing, and cloud ($5.9 billion); and Fintech ($5.5 billion).[6] By contrast, some industries—like energy, law, some professional services subsectors, and, to an extent, consumer packaged goods—are just starting to pilot use cases.

Nevertheless, across the entire maturity spectrum, AI algorithms are popping up via grassroots methods and, often, without engaging security. Before companies know it, these "sandbox tests" are going live and providing value to the business, even though the critical step of integrating security considerations has been skipped.

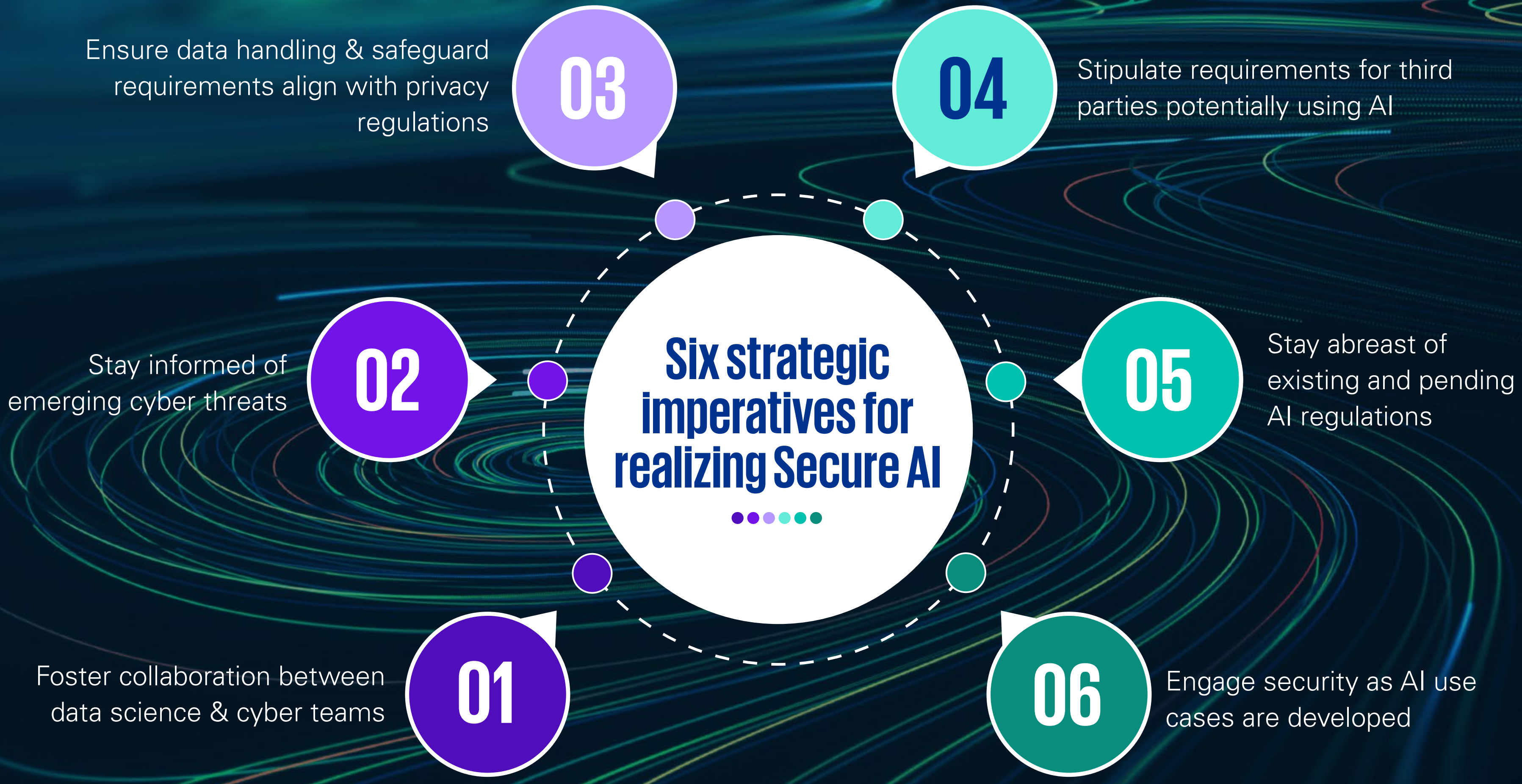
Over time, all industries can expect to utilize AI models to varying degrees. However, what must be consistent is a commitment to embedding security within the cycle at the same time the business is identifying use cases.



> "You cannot be a leader in your industry if you are not taking full advantage of technology. This means embracing advances in AI and automation even amidst the uncertainty they can bring."
>
> *—Sean Otto, Director of Analytics, AES Corporation*

[6] Nestor Maslej, Loredana Fattorini, Erik Brynjolfsson, John Etchemendy, Katrina Ligett, Terah Lyons, James Manyika, Helen Ngo, Juan Carlos Niebles, Vanessa Parli, Yoav Shoham, Russell Wald, Jack Clark, and Raymond Perrault, "The AI Index 2023 Annual Report," AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2023

Six strategic imperatives for realizing Secure AI



Six strategic imperatives for realizing Secure AI

01 Foster collaboration between data science & cyber teams

02 Stay informed of emerging cyber threats

03 Ensure data handling & safeguard requirements align with privacy regulations

04 Stipulate requirements for third parties potentially using AI

05 Stay abreast of existing and pending AI regulations

06 Engage security as AI use cases are developed

01 Foster collaboration between data science & cyber teams

To ensure that cybersecurity is considered as AI applications are developed, data science and security teams need to maintain direct lines of regular communication and combine their missions and cadences. Although many organizations are not collaborating this way yet, leading companies are taking guidance on this matter to heart.

According to Ryan Boulais, Global CISO, AES Corporation: “We put a ‘security by design’ program in place with the data science team for software development and analytics. As with other software, we ensure that our data science team is thinking about security while they’re innovating and building AI solutions.”

Ashish Shah, Technical Lead of Cybersecurity Engineering & Innovation, Chevron says, “IT and business teams need to speak the same language and continuously collaborate, bringing in cyber colleagues early so that risks can be evaluated and managed.”

Shelley Ivanko, Director of Analytics and Data Science, Campbell Soup Company, stresses that the company’s data science team has “been in lockstep with the security team. Now that we are on this journey together, we will apply this thinking to everything.”

Some important considerations for companies as they begin to align cybersecurity and data science teams are:

- **Create a framework** for collaboration on innovation and Secure AI.
- **Consider creating a Chief Trust Officer role**, so there is trust not only in data, but also in humans working on AI projects, as companies like SAP, Cisco, IBM, OneTrust, Microsoft, and Salesforce have done, many by promoting former CISOs.[7]
- **Be sure to loop in the risk management function** as new AI-based products are greenlit.
- **Stand up a cross-functional steering committee** comprising data science teams, security, legal, privacy, digital, etc.

[7] Carrie Pallardy, “The Chief Trust Officer Role Can Be the Next Career Step for CISOs,” InformationWeek, November 14, 2022.

02 Stay informed of emerging cyber threats

It is important for leadership to consider the fact that attacks on AI are likely to be either confidentiality-based or integrity based, in contrast to traditional cybersecurity attacks. Integrity-based attacks can include threat actors manipulating AI to generate content that could harm a company financially or reputationally. In these events, companies need to be prepared for the fact that remediation efforts could span beyond technical considerations and include repairing reputational damage.

Some methods for staying ahead of emerging threats include:

- Establish an AI risk assessment framework that includes formal processes for identifying and assessing risks associated with AI systems, including AI threat profiles and risks associated with each AI model use case.
- Enhance communication with relevant stakeholders, including senior leadership, IT teams, data scientists and engineers, software engineers, business unit leaders, procurement, legal counsel, and others through formal documentation of risks associated with AI systems.
- Regularly consult expert resources such as MITRE ATLAS, NIST AI RMF, and the OWASP AI Security and Privacy Guide to inform your perspective on the expanded attack surface associated with AI.
- Provide continuous security and threat awareness education and training to personnel who leverage and deploy AI algorithms.

"Bad actors and nation states are not going to slow down as they work to take advantage of the vulnerabilities of our AI-enabled systems," said Dr. Christina Liaghati, AI Strategy Execution & Operations Manager, MITRE. "Therefore, we need to stay engaged, at the cutting edge of AI and AI security, keeping the leaders who understand the vulnerabilities and risks of using AI-enabled systems at the forefront."

> "
>
> **If you only focus on preventing and detecting threats that are known to you, you are failing. You must be ahead of emerging threats, which AI can help accomplish.**
>
> Generative AI includes tools to help speed up investigations, remove a threat from the environment relatively quickly, traverse through petabytes of data, and understand relationships in the data faster than humans can."
>
> *—David LaFalce, Global Head of Operational Resilience, Wells Fargo*

01
02
03
04
05

03 Ensure data handling & safeguard requirements align with privacy regulations

When you are using data to train an AI model, remember that synthetic data will allow you to build your algorithms quicker and without the privacy restrictions required by real-world data. In medicine for example, synthetic data can be used to train an AI system to provide guidance on understanding population dynamics, selecting clinical trial sites, and tracking the typical progression of a disease. Conversely, clinical applications that require real patient data—such as drawing biomarkers to understand the genesis of a disease—will require much more stringent privacy and security protocols.

Across industries, leaders may want to rely on data governance teams to advise on what use cases are safe to pursue and align to a responsible AI framework. Ethical guidelines should encompass the transparency of AI algorithms, as well as privacy, data security, and assurance that AI models are free from any kind of algorithmic bias.

To ensure that the data employed in AI algorithms is useful, it may be prudent to pursue digital transformation in parallel. This can include creating data products, and undergoing data cleanup, classification, and tagging with the right level of confidentiality.

Finally, when bringing multiple datasets together, take the time to understand identification risk and create a privacy threshold for what can be combined. For example, solutions like Amazon Web Services' (AWS') Clean Rooms can help companies collaborate and combine datasets without exposing underlying data.[8]

> **"As we explore AI use cases, our top priority is making sure our data is protected and confidential."**
>
> *—Martin Bally, CISO,*
> *Campbell Soup Company*

[8] Press release: "AWS Announces AWS Clean Rooms," Amazon Web Services, November 29, 2022

04 Stipulate requirements for third parties potentially using AI

Organizations are concerned that third parties with which they do business may not always be transparent about the extent to which AI is present in their solutions. Such transparency should be a condition of doing business and, as such, should be included in new contracts and added to existing ones. Mandate that your vendors describe how they are validating their AI systems and ensure that their privacy standards are on par with yours.

Eventually, vendors that can demonstrate they are using responsible AI may be able to earn ISO certifications, as well as a compliance score similar to SOC2. Regarding the latter, such scoring will measure whether AI data is secure from confidentiality, availability, and integrity standpoints. Although this will likely require documentation that satisfies existing data protection and privacy regulations, such standards are apt to evolve to meet the realities of AI. Until such certifications or scoring models are in place, organizations should add AI-related questions to their third-party risk assessment questionnaires.

The AI third-party ecosystem will continue to expand and become increasingly integrated," said Jonathan Dambrot, CEO, Cranium. "New methods for ensuring transparency into risk and safety will continue to evolve. Not only do organizations need to monitor third-party use of AI technology, but many organizations themselves are third parties. Strategies such as AI card enable enterprises to publish the state of their AI environments, promoting trust in the use of AI throughout the supply chain and providing an AI Bill of Material to provide better visibility."

> “We are looking at three kinds of AI: models we create internally; AI we purchase, particularly for the supply chain; and emerging technologies, such as Generative AI. All models need to have guardrails to ensure that responsible AI is utilized.”
>
> *—Ashish Shah, Technical Lead of Cybersecurity Engineering & Innovation, Chevron*

05 Stay abreast of existing & pending AI regulations

Organizations should ensure they understand existing AI regulations and guidance and how they could impact their approach to AI security. For example, in the U.S.[9]:

- The number of proposed federal AI bills passed into law increased from two percent in 2021 to 10 percent in 2022.
- The White House's blueprint for the AI Bill of Rights was introduced in October 2022. The guidance establishes principles for managing AI risks, mitigating biases, and avoiding conflicts of interest.
- There are a number of privacy and AI-related bills on the state front, with 35 percent of all proposed AI bills passed into law.
- The Federal Trade Commission has warned that consumer protection laws now apply to AI, as evidenced by the recent investigation of OpenAI to determine whether the company has put personal reputations and data at risk.[10]
- The Consumer Financial Protection Bureau has warned that opaque AI systems used by credit agencies could violate anti-discrimination laws.

Importantly, the Biden-Harris Administration recently convened seven leading AI companies to secure their voluntary commitments to manage AI risks until more formal regulations are in place.[11] Stemming from this initial public/private collaboration, the government is encouraging the entire AI industry to demonstrate commitments to:

1. Conducting internal and external testing before AI systems are released
2. Sharing information across corporations, government agencies, and academia to reach consensus on best practices for minimizing risk
3. Investing in cybersecurity and insider-threat safeguards
4. Reporting vulnerabilities and flaws in AI systems as they arise
5. Exploring a watermarking system so the public knows when content is generated by AI
6. Committing to using AI systems for the greater good, including eliminating diseases, addressing climate change, and more.

[9] Nestor Maslej, Loredana Fattorini, Erik Brynjolfsson, John Etchemendy, Katrina Ligett, Terah Lyons, James Manyika, Helen Ngo, Juan Carlos Niebles, Vanessa Parli, Yoav Shoham, Russell Wald, Jack Clark, and Raymond Perrault, "The AI Index 2023 Annual Report," AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2023

[10] Cat Zakrzewski, "FTC investigates OpenAI over data leak and ChatGPT's inaccuracy," The Washington Post, July 13, 2023

[11] FACT SHEET: "Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI," The White House, July 21, 2023.

06 Engage security as AI use cases are developed

There is a wide variety of AI use cases across industries. While some companies are engaging security before they actively pursue these use cases, there is still a tendency to deploy exciting new models without the security team's input. This creates situations where the CISO's team must remediate issues after a model is deployed to production and informing critical business decisions. It is important to pay attention to what peers are doing so organizations can learn from each other's pitfalls and risks. The following are some examples of how Secure AI is being used across the life sciences and healthcare, financial services, and legal industries.

Life sciences and healthcare

The life sciences industry has been ahead of the pack, using AI in research & development and decision support long before the current excitement around the technology. As part of a regulated industry and one where life-or-death issues are at stake, life sciences companies are also relatively mature in terms of pursuing Secure AI.

While ML is being used to enhance devices and diagnostics, the risk of hacking into these devices requires a "security by design" approach through which security practitioners are at the table with R&D teams from the outset. In the drug research process, AI can allow pharmaceutical companies to enhance discovery and innovation and bring therapies to market faster.

As the use of AI increases in the life sciences industry, it is more important than ever to have safeguards in place to ensure that the data used in AI algorithms is accurate, indisputable, and free from tampering.

"When you use large language models in high-impact products and initiatives like medical devices, drug discovery, and surgical decision-making, you have to be absolutely sure that proper guardrails are in place," said Gary Harbison, Global CISO, Johnson & Johnson.

Regarding data privacy, life sciences organizations using PHI in AI products need to ensure that data usage aligns with the Health Insurance Portability and Accountability Act (HIPAA), as well as other state privacy and security laws and regulations. Among the ways to ensure that you are steering clear of potential HIPAA violations are de-identifying data before it is uploaded into an AI database; performing due diligence on vendors that may be handling PHI, including where it is being stored; and carefully monitoring external access points to ensure that networks connecting patient data and patient care are thoroughly secure.[12]

> **"With AI in pharma, there are three critical imperatives: The first is proper security related to whether AI systems represent a broader attack surface. The second is regulatory and compliance, such as whether a validated system making clinical recommendations or diagnoses should be treated as software as a medical device. The third is data privacy, which is, of course, a prime concern in healthcare."**
>
> *—Hal Stern, Head of Technology for Pharmaceutical R&D, Johnson & Johnson*

[12] Linda A. Malek, Pralika Jain, and Jason Johnson, "Data privacy and artificial intelligence in healthcare," Reuters, March 17, 2022.

Financial Services

While the financial services industry, and FinTech in particular, is considered a leader in AI adoption, organizations don't always bring cybersecurity into the mix when they are conducting model risk management. Instead, they tend to be more focused on creating applications that meet customer, employee, and stakeholder needs.

According to David LaFalce, Global Head of Operational Resilience, Wells Fargo, "We are narrowing down our potential use cases by really focusing on customer journeys and pain points and considering where it makes sense to plug in AI and ML to drive value."

However, it is critical to note that the company is doing so with a commitment to security. The company is building out a large data platform as a precursor to using ML and AI more strategically.

At the end of the day, financial services AI models shouldn't be created without governance and controls in place. Organizations need to embed security into the product team that owns end-to-end delivery, so a comprehensive team is designing together from the start and with the same vision. This can be a cultural change and mindset shift for companies that are accustomed to waterfall methodologies and project management, which follow a series of steps. Although some may fear that introducing security during the ideation process will extend the delivery deadline, the reality is that fixing defects retroactively is much riskier and takes much longer. Finally, it will be useful for teams to have a single place to access guidelines for securing AI and ML, as well as answers to questions about whether individuals can legally use certain data, how to avoid biases when it comes to model output, and more.

> "Given the possibility of influence that AI has, we don't want people experimenting with it. Security guidelines have to be defined as part of the model development lifecycle, not after it goes live."
>
> *—David LaFalce, Global Head of Operational Resilience, Wells Fargo*

01
02
03
04
05

Legal

Law firms handle a great deal of sensitive data, e.g., M&A documents, technical documents for patentability, medical records, etc., so they must be mindful about when and if this data can be used in AI solutions. For example, some attorneys are interested in using ChatGPT to help with research, but they still must double check the output for accuracy. Using AI to go through hundreds of pages of discovery data would, of course, be a tremendous time saver. However, until the technology is refined, firms must remember that the data won't always be 100 percent accurate or complete.

At the same time, the legal profession is highly motivated to find time- and resource-saving AI applications. According to Sherri Vollick, Senior Manager of Information Security and Governance, Hinshaw Law, "Our attorneys pull and analyze information from various resources, and AI can help us get to resolutions faster. If you're a law firm and you haven't yet started looking at AI, you're already behind."

Many law firms certainly view this as a competitive issue and understand that hesitating on AI could put them at a competitive disadvantage. On the other hand, data loss or data poisoning could have a catastrophic impact on law firms and their reputations. Therefore, it is critical that any AI use cases at law firms integrate robust security principles from the start.

Vollick continued: "The New York State CLE Program Rules now require attorneys to make themselves aware of new technologies, as well as cybersecurity, privacy, and data protection measures that their law firms and clients are using. I would recommend that all firms educate themselves on what AI is, what different models exist, and what is required from a risk assessment perspective so you can advise organizational leadership."

Top 10 recommendations for organizations new to AI Security

Take a collaborative approach to addressing AI risk

Remember that organizations must address AI risk as a community; even competing startups are communicating because they share a combined reputational risk.

Define governance roles and responsibilities

Be clear about who owns AI, who owns data, whether new roles are required, etc.

Invest in sufficient security protocols

Be sure security protocols align with financial and reputational risks.

Integrate AI Security processes

Pursue AI Security processes and development processes in tandem to ensure they are in sync.

Create an iterative process

Establish regular lines of communication between cybersecurity teams and AI teams to ensure that emerging cyber threats are known to all.

Ensure all businesses that use GenAI know when to contact security

As AI applications are piloted within business units, teams should know when and how to contact security for help.

Stay ahead of potential third-party risks

Put mechanisms in place to ensure you are aware of when AI is embedded in vendor products.

Utilize existing tools, technologies, and frameworks

Use both existing and cutting-edge resources to secure AI solutions.

Determine whether AI products should be developed in house or outsourced

Depending on internal capabilities and budgets for external vendors, AI solutions may be developed in house or via outsourcing.

Start with data provenance

Take an inventory of where data used in AI models originated and what you are allowed to do with it, as well as whether data hygiene and classification are robust.

Conclusion

There is no doubt that artificial intelligence is here to stay. From medicine to finance, and eventually across all industries, there is tremendous potential for growth, efficiency, and innovation. Therefore, companies seeking to institute AI in their operations need to ensure that their security, privacy, risk management, and governance efforts are in lockstep with their technological ambitions. This will require:

- A commitment to a close working relationship between data science and security teams to ensure visibility into how AI is embedded across business functions and engagement of security practitioners as soon as use cases are identified.
- A firm understanding of how AI expands the attack surface, as well as a concerted effort to stay ahead of emerging threats, including those introduced by third parties.
- Acknowledgment that data represents not only organizations' greatest opportunity, but also their biggest risk, requiring strict adherence to data protection and security guidelines.
- A commitment to self-governance until comprehensive regulatory guidance is in place, as well as thoughtful consideration of how to advise government entities on future regulations.

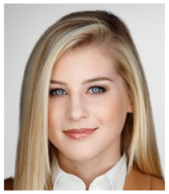
Finally, remember that, when it comes to AI, companies shouldn't be going it alone. It will take transparency, information sharing, and collaboration among corporations, government entities, and academia to ensure that any AI is Secure AI.

> **"Our adversaries are collaborating, so we need to do the same to stay ahead. We will always be stronger together."**
>
> *—Gary Harbison, Global CISO, Johnson & Johnson*

Contact us

Katie Boswell
Managing Director &
AI Security Lead, KPMG
katieboswell@kpmg.com

Matt Miller
Principal, KPMG
matthewpmiller@kpmg.com



Kristy Hornland
Director, KPMG
khornland@kpmg.com

Mark Orsi
CEO
Global Resilience Federation
morsi@grf.org

Jonathan Dambrot
CEO
Cranium AI
jdambrot@cranium.ai

We would like to thank the following individuals for their invaluable contributions:

KPMG
Donna Ceparano
John Hodson
Riley Richards
Kelsey Flynn
Kelly Combs

Cranium AI
Felix Knoll
Paul Spicer
Daniel Christman

AES
Ryan Boulais
Sean Otto

Amherst College
Dr. Scott Alfeld

Campbell Soup Company
Martin Bally
Shelly Ivanko
Mark Wehrle

Chevron
Ashish Shah
Margery Connor

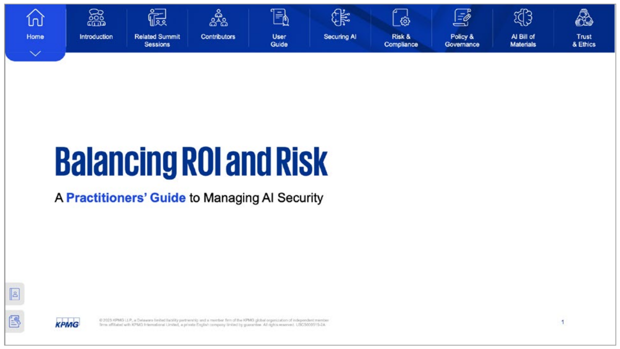
Hinshaw Law
Sherri Vollick

Johnson & Johnson
Gary Harbison
Hal Stern
Bill Janicki

Mitre
Dr. Christina Liaghati

Wells Fargo
David LaFalce

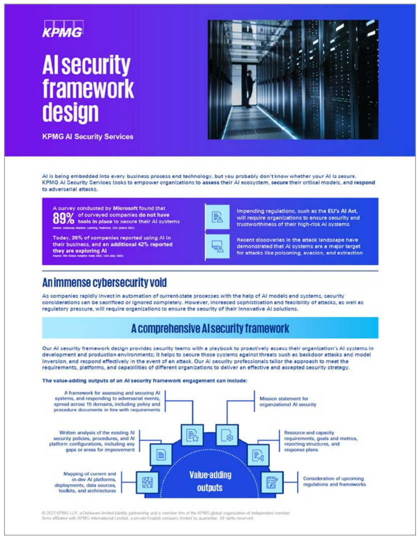
Related thought leadership:



Balancing ROI and Risk

A **Practitioners' Guide** to Managing AI Security

© 2023 KPMG LLP, a Delaware limited liability partnership and a member firm of the KPMG global organization of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved. USCS000515-2A



AI security framework design

KPMG AI Security Services

AI is being embedded into every business process and technology, but you probably don't know whether your AI is secure. KPMG AI Security Services looks to empower organizations to assess their AI ecosystem, **secure** their critical models, and **respond** to adversarial attacks.

A survey conducted by **Microsoft** found that **89%** of surveyed companies **do not have tools in place** to secure their AI systems

Source: Adversary Machine Learning, Redmond, USA (March 2021)

Today, **35%** of companies reported using AI in their business, and an **additional 42% reported they are exploring AI**

Source: IBM Global Adoption Index 2022, USA (May 2022)

Impending regulations, such as the **EU's AI Act**, will require organizations to ensure security and trustworthiness of their high-risk AI systems

Recent discoveries in the attack landscape have demonstrated that AI systems are a major target for attacks like poisoning, evasion, and extraction

An immense cybersecurity void

As companies rapidly invest in automation of current-state processes with the help of AI models and systems, security considerations can be sacrificed or ignored completely. However, increased sophistication and feasibility of attacks, as well as regulatory pressure, will require organizations to ensure the security of their innovative AI solutions.

A comprehensive AI security framework

Our AI security framework design provides security teams with a playbook to proactively assess their organization's AI systems in development and production environments; it helps to secure those systems against threats such as backdoor attacks and model inversion, and respond effectively in the event of an attack. Our AI security professionals tailor the approach to meet the requirements, platforms, and capabilities of different organizations to deliver an effective and accepted security strategy.

The value-adding outputs of an AI security framework engagement can include:

A framework for assessing and securing AI systems, and responding to adversarial events, spread across 15 domains, including policy and procedure documents in line with requirements

Mission statement for organizational AI security

Written analysis of the existing AI security policies, procedures, and AI platform configurations, including any gaps or areas for improvement

Resource and capacity requirements, goals and metrics, reporting structures, and response plans

Mapping of current and in-dev AI platforms, deployments, data sources, toolkits, and architectures

Consideration of upcoming regulations and frameworks

© 2023 KPMG LLP, a Delaware limited liability partnership and a member firm of the KPMG global organization of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved.

Some or all of the services described herein may not be permissible for KPMG audit clients and their affiliates or related entities.

The information contained herein is of a general nature and is not intended to address the circumstances of any particular individual or entity. Although we endeavor to provide accurate and timely information, there can be no guarantee that such information is accurate as of the date it is received or that it will continue to be accurate in the future. No one should act upon such information without appropriate professional advice after a thorough examination of the particular situation.

DASD-2023-13298

kpmg.com/socialmedia