



Driving responsible innovation

Reflections on a year of AI governance

Introduction

Artificial Intelligence (AI) is transforming companies and industries at an unprecedented pace. From streamlining financial operations with automated fraud detection to boosting marketing strategies and customer experiences using data-driven insights to optimizing supply chain management with predictive analytics, AI's potential seems limitless.

And yet, at the same time, scaling from one-off AI use cases to full-scale enterprise-wide transformation presents what can seem like an ever-growing list of vulnerabilities for an organization to evaluate. Massive transformation without governance top of mind can open organizations to a whole host of risks from security violations to data breaches to irreparable harm to the brand.

Over the past two years, KPMG LLP has led interviews and working sessions with corporate AI leaders who believe that balancing a clear AI governance strategy while keeping pace with innovation will distinguish those organizations that successfully adopt AI from those that don't. While the past year has seen the concept of "responsible AI" gain currency and fewer organizations allowing "rogue" AI algorithms to be instituted without engaging security, there is still room for progress. This paper examines where we have been and where we are going by delving into evolving risks and priorities, as well as the increasingly multifaceted nature of AI governance as a driver of innovation.

Enormous potential comes with enormous responsibility

The global AI market is projected to reach \$267 billion by 2027,ⁱ highlighting the rapid adoption and investment in AI technologies. However, as organizations increasingly scale AI systems across their operations, the need for robust AI governance frameworks that ensure they align with regulations designed to ensure ethical and responsible use has become a strategic imperative.

The challenge is that the regulatory environment is a moving target: The Biden Administration released an executive order in October 2023 reflecting a commitment to the U.S. leading in AI development "while managing its risks and ensuring the technology benefits society as a whole." The provisions cover

security risks; ethical issues such as bias, discrimination, and disinformation; and national security/human rights. The White House has also released a Framework to Advance AI Governance, which aims to drive U.S. leadership in AI, while advancing governance and risk management practices on the department level so that AI becomes an effective component of the U.S. National Security System.ⁱⁱ

Although the SEC has not yet published a final rule, it has provided guidance to ensure that financial firms using AI put their clients' interests above their own, ensure transparency and fairness, and disclose how they are managing cyber risks. The Algorithmic Accountability Act requires companies to conduct impact assessments for AI systems that significantly affect consumers' rights or well-being.ⁱⁱⁱ These assessments must evaluate the potential for bias, discrimination, and other adverse effects, ensuring that AI systems are fair and transparent.

In the 2024 legislative session, at least 45 states, Puerto Rico, the Virgin Islands and Washington, D.C., introduced AI bills. Also, 31 states, Puerto Rico, and the Virgin Islands adopted resolutions or enacted legislation.^{iv} It is worth noting that the CA AI Safety Act was recently vetoed by the governor. Although many praised the act's focus on encouraging more robust attestation and providing a kill switch for AI systems, the act was deemed too burdensome for companies, which underscores the need for more nuanced and targeted approaches to AI governance.

The World Economic Forum's AI Governance Alliance emphasizes the importance of enhancing human capabilities, fostering inclusive growth, and promoting global prosperity through AI.^v Finally, the European Union's AI Act is a significant step towards comprehensive AI regulation. It categorizes AI applications based on their risk levels and imposes strict requirements for high-risk AI systems, including transparency, accountability, and human oversight.^{vi} This act is setting a global precedent and influencing AI governance strategies worldwide.

ⁱ Isaac Sacolick, What you need to know about AI governance, InfoWorld, September 23, 2024.

ⁱⁱ Framework to advance AI Governance and risk management in national security, The White House.

ⁱⁱⁱ Ben Chester Cheong, Transparency and accountability in AI systems: safeguarding wellbeing in the age of algorithmic decision-making, Frontiers, July 2, 2024.

^{iv} Artificial Intelligence 2024 Legislation, National Conference of State Legislatures, September 9, 2024.

^v AI Governance Alliance, World Economic Forum.

^{vi} Michael Borrelli, EU AI Act Published: A New Era of Trustworthy AI Begins, European AI Alliance, July 12, 2024.

Risks of inaction

These regulations and guidelines highlight the significant legal and security risks organizations can face if they don't address governance concurrently with AI implementations. As AI proliferates, cyber risks will certainly continue to grow, and organizations should anticipate and prepare for newer forms of cyber threats as potential attack surfaces expand.^{vii}

Legal risks: Since regulations are increasingly putting the onus on corporations to avoid risks associated with AI use, legal and compliance departments will be central to helping organizations avoid litigation. Legal departments also play a key role in ensuring that AI use cases align with the organization's brand values and do not pose reputational risks.

Further, Big Tech companies and industry consortiums are putting forth guidelines for avoiding legal risks. In 2023, an international AI Safety Summit was convened where leading AI companies, as well as international governments, civil society groups and researchers, met to discuss how AI risks can be mitigated through coordinated action.^{viii} NIST has established the U.S. Artificial Intelligence Safety Institute (USAISI), which is supported by the U.S. AI Safety Institute Consortium. The Consortium comprises more than 280 organizations that are working on guidelines and standards for AI safety across the world.^{ix} Steve Wilson, the project lead for OAS for LLMs, has developed a playbook for LLM security, providing valuable guidance for developers.

Involving legal and compliance teams early in the AI development process can help identify potential legal risks and ensure compliance with relevant regulations, thus minimizing the possibility of costly litigation and reputational damage. Further, collaboration with industry peers and government agencies, as well as participating in industry forums and working groups, can help organizations stay informed about emerging security threats and best practices for AI security.

Security risks: As mentioned above, although AI-specific cyber-threats currently represent only a small portion of overall threats, companies should still keep potential threat areas in mind as AI becomes more central to their operations. Threats such as targeted phishing attacks, deepfake videos, and botnets could pose threats to organizations' data security and, ultimately, their reputations.

Another set of potential concerns are adversarial attacks—where bad actors manipulate inputs to AI algorithms to produce incorrect or damaging results. This can have implications, especially in industries like healthcare and finance where incorrect output can have serious consequences for individuals. AI supply chain attacks, where attackers target the components and devices used in AI systems, can disrupt the production and delivery of critical goods.

Since AI systems process and collect both personal data and intellectual property, organizations need to pay close attention to how this data is being used and protected within AI models. Seventy-two percent of respondents to the KPMG AI Governance survey cited privacy concerns as the most significant AI risk faced by their organizations.^x Integrating AI risks into existing enterprise risk management frameworks can help ensure that AI risks are considered alongside other organizational risks and managed in a coordinated manner.

Organizations that fail to put effective governance frameworks in place to mitigate security risks may eventually run afoul of regulations such as the EU AI Act. The act includes provisions that require additional obligations for general purpose AI (GPAI) models that pose systemic risks; provisions include model evaluations, systemic risk assessments, adversarial testing, and reporting serious incidents.^{xi} Organizations need to help ensure that they comply with these regulations to avoid penalties and protect their reputations.

^{vii} Caleb Sima, The Real Story Behind AI Security Incidents, LinkedIn, October 29, 2024.

^{viii} AI Safety Summit

^{ix} Artificial Intelligence Safety Institute Consortium (AISIC), National Institute of Standards and Technology (NIST).

^x KPMG AI Governance Survey, October 2024.

^{xi} Martin Coulter, Exclusive-EU AI Act checker reveals Big Tech's compliance pitfalls, Reuters, October 16, 2024.

Governance as the foundation of innovation

One of the key lessons emerging as AI becomes ubiquitous is the necessity of taking a balanced approach to AI development and security. Currently, the allocation of most organizations' budgets reflects nearly equal investments. According to the KPMG AI Governance survey, organizations surveyed say 48 percent of their budgets are dedicated to developing innovative AI solutions while 43 percent is devoted to securing them.^{xii} Despite these investments, many organizations are still in the early stages of deciding on governance models, not to mention defined roles and responsibilities, codified policies, ethical considerations, and training/upskilling.

Governance Operating Models

Governance models that companies are exploring include centralized, decentralized, and outsourced. Each model has its own advantages and challenges.

A centralized governance model involves a clear AI center of excellence that sets standards for the entire organization; maintains universal guidance, policies, and procedures for AI usage; and drives which use cases are prioritized due to alignment with both the AI strategy and the AI governance objectives of the organization. Having clear lines of responsibility and oversight helps ensure alignment on strategic direction, drive consistent output, and avoid duplication of efforts across business units. Perhaps most important, having a central authority that is closely tied to legal and regulatory experts helps reduce the risk of regulatory noncompliance. Large enterprises, particularly in industries like Big Tech and finance, are more apt to adopt centralized governance models. According to a recent study, 77 percent of business leaders believe that centralized AI governance can help their organizations maintain a competitive advantage.^{xiii}

On the other hand, centralized governance models can sometimes deter real-time decision-making and stifle innovation due to the need for approvals from a central authority. Therefore, it is critical to ensure that the centralized body keeps a pulse on innovative use cases across the organization, providing direction—and funding when warranted—to ensure that AI serves as a true transformative force.

A decentralized governance model involves individual business units taking responsibility for AI governance. This approach is, perhaps, a bit more flexible than centralized models, as business units can tailor AI initiatives to their specific needs. Further, it may be an appropriate avenue for small- to mid-size firms that need to react with agility in response to market forces and customer demand. Being able to act without checking with a centralized authority will make it easier to actualize creative solutions and experimentation.

On the other hand, since decentralized governance can lead to inconsistencies, duplication of efforts, and challenges with regulatory compliance, many organizations that started with decentralized governance models are finding it necessary to shift to centralized models.

An outsourced governance model involves leveraging third-party expertise for AI governance. This approach can be beneficial for organizations that lack the resources or expertise to manage AI governance internally. According to a report by Atlan, outsourcing can provide access to specialized knowledge and skills, enhancing the overall quality of AI governance.^{xiv} Finally, for companies just beginning to ramp up their AI programs, outsourcing governance—and, perhaps, other aspects of AI management like cybersecurity—can allow scalability as needs change.

On the other hand, it is important to keep in mind that outsourcing requires careful management of third-party relationships to ensure that external partners adhere to the organization's governance standards and philosophies, while also maintaining expected levels of data security and confidentiality.

The takeaway: Choosing the right governance model depends on the specific needs and capabilities of an organization. In some cases, a hybrid approach, combining elements of centralized, decentralized, and outsourced models, may offer the best balance of consistency, flexibility, and expertise. However, as AI models and applications become more complex and widespread, it is likely that most organizations will gravitate toward a centralized governance model.

^{xii} KPMG AI Governance Survey, October 2024.

^{xiii} The enterprise guide to AI governance, IBM.

^{xiv} AI Governance: How to Mitigate Risks & Maximize Business Benefits, Atlan, June 24, 2024.

Roles and Responsibilities

Establishing clear and well-defined roles and responsibilities for AI governance is critical so that organizations can effectively handle the complexities of AI deployment. It is critical to have dedicated teams and subject matter experts who are responsible for overseeing AI initiatives, ensuring compliance with regulations, and addressing any ethical concerns that arise, thus minimizing legal and reputational risks. Transparent governance structures build trust among stakeholders, including customers, employees, and regulators. Finally, defined roles streamline decision-making processes, enabling organizations to respond quickly to emerging challenges and opportunities.

Some examples of critical roles follow:



Cross-Functional AI Governance Team: Cross-functional teams are essential for managing multiple aspects of AI governance concurrently. These teams typically include representatives from departments that include legal, ethics, IT, and regulatory. By bringing together diverse perspectives, cross-functional teams can ensure comprehensive oversight and decision-making.



AI Governance Leader: A dedicated AI governance leader oversees AI initiatives and ensures alignment across the organization. The role requires technical expertise and a deep understanding of the organization's operations and culture.



AI Compliance and Risk Management Team: Integrating AI risks into existing enterprise risk management frameworks helps ensure that AI risks are considered alongside other organizational risks and managed in a coordinated manner. According to Gartner, 27 percent of organizations have established an AI compliance and risk management team, and another 27 percent have an AI oversight committee.^{xv}



AI Ethics Officer or Committee: An AI ethics officer or committee develops and enforces ethical guidelines, conducts regular audits, and ensures that AI systems align with the organization's values and ethical standards.^{xvi} According to Gartner, the appointment of AI ethics officers is becoming more common, with 35 percent of large organizations expected to have such roles by 2025.^{xvii}



AI Validators and Auditors: AI validators and auditors are responsible for assessing the performance, fairness, and compliance of AI systems. They conduct regular audits and impact assessments to evaluate the potential consequences of AI decisions on individuals and society.

^{xv} Gartner AI Governance Frameworks

^{xvi} RTS Labs AI Governance Framework

^{xvii} Jasleen Kaur Sindhu, AI Ethics: Enable AI Innovation With Governance Platforms, Gartner, October 14, 2024.

Ethics in AI governance

A foundational tenet of AI development is a responsible AI team—which ensures AI systems are safe, ethical, inclusive, and transparent. Such a team will oversee an ethics-based framework that ensures alignment with core company values and principles and includes guidelines on sustainability, fairness, technical safety, accountability, explainability, and data stewardship. Responsible AI teams should conduct regular audits and monitoring to ensure that AI deployments align with ethical standards.

A key ethical issue centers around mitigating biases and striving for inclusivity. It is important to recognize that AI systems can inadvertently perpetuate existing societal biases if not properly managed. For example, a 2019 study by the national Institute of Standards and Technology (NIST) found that facial recognition systems had higher error rates for people of color.^{xviii} Lack of transparency in AI algorithms can also contribute to bias and discrimination, which can present a risk of legal or regulatory penalties.

Further, when using generative AI to create content that is similar to that produced by human writers, it is important to disclose how the content was generated, while also ensuring its authenticity, accuracy, and integrity. Failing to do so can cause companies irreparable reputational damage and lead to regulatory noncompliance as well. The EU AI Act specifically addresses this issue by including strict requirements for high-risk AI applications.^{xix}

Policies and documentation

Many organizations have found that without well-defined policies, it becomes challenging to manage AI risks effectively. These policies should cover various aspects of AI governance, including data privacy, cybersecurity, and ethical considerations, as well as vulnerability management processes, AI incident response protocols, and AI resiliency plans. Despite the risks of operating without formalized policies, a 2024 report from the Stanford Institute for Human-Centered Artificial Intelligence found that only 18 percent of companies have formal, comprehensive policies in place for managing AI technologies.^{xx} This highlights a significant gap in AI governance that needs to be addressed to mitigate risks, ensure responsible AI deployment, and stay ahead of the evolving regulatory landscape.

Organizations should ensure that their policies and procedures are regularly reviewed and updated as technology implementations expand and evolve. This includes incorporating feedback from audits and monitoring activities, as well as staying abreast of new regulations and industry best practices.

[See next page for more on policy development]



^{xviii} NIST Study Evaluates Effects of Race, Age, Sex on Face Recognition Software, NIST, December 19, 2019.

^{xix} Austin Chia, AI Governance in 2024: An Overview, Splunk, December 19, 2023.

^{xx} THE AI INDEX REPORT: Measuring trends in AI, Stanford Institute for Human-Centered Artificial Intelligence, 2024.

Effective AI policies should include the following components:

- 01 The scope of AI** for the organization and related definitions
- 02 Underlying principles for responsible governance** (e.g., accountability, reliability, security, safety, privacy, sustainability, explainability, data integrity, transparency, fairness)
- 03 Definitions of roles and responsibilities** in accordance with the governance program
- 04 Usage guidelines** for the workforce detailing permitted and prohibited uses of AI
- 05 Taxonomies for risk mitigation and controls** to guide tactical implementation of governance program objectives
- 06 Data privacy and security protection** measures, including data anonymization, encryption, and access controls to prevent unauthorized access and data breaches.^{xxi}
- 07 Supporting documentation**, such as AI resiliency plans, incident response playbooks, vulnerability management standard operating procedures, etc.

Workforce training and skill development

Organizations should provide mandatory training on AI governance, security, and ethical implications to ensure that employees understand that they have critical roles to play when it comes to AI usage and governance. This helps build a culture of ethical AI use and ensures that employees are equipped to handle the complexities of AI deployment. The World Economic Forum emphasizes the need for ongoing education and training to keep pace with the evolving ethical challenges posed by AI. Further, a report by ISACA highlights that only 10 percent of organizations have comprehensive policies for generative AI, underscoring the need for robust training programs.^{xxii}

To develop effective training programs, organizations should tailor training modules to different roles and responsibilities within the organization; use interactive learning methods such as workshops, simulations, and case studies to reinforce key concepts; underscore the importance of continuous learning to keep pace with the latest developments in AI governance and technology; and encourage cross-functional collaboration to bring together diverse perspectives and expertise.

[See next page for more on workforce training]



^{xxi} Mary Carmichael, Key Considerations for Developing Organizational Generative AI Policies, ISACA, November 1, 2023.

^{xxii} Euan Blair, The future of learning is working: How to boost skill development in the workplace, World Economic Forum, December 28, 2023.

Critical competences for employees include:

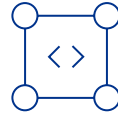
01 Understanding AI governance principles and the importance of following acceptable usage guidelines.

02 Considering ethical implications to protect from negative bias and discrimination and ensure fairness, transparency, and accountability.

03 Following best practices for data privacy and security, including data anonymization, encryption, and access controls to protect sensitive information.

04 Prioritizing use cases that fit the risk appetite of the company to ensure employee usage is in alignment with innovation and responsible governance strategies of the company.

One specialized area of training is **red team testing**. This involves conducting thorough testing of AI systems to identify and address any vulnerabilities or risks before they are deployed. This approach helps organizations build more resilient AI systems that can withstand adversarial attacks and other security threats, while building trust and public confidence in AI. According to a report by the National Institute of Standards and Technology (NIST), establishing guidelines and benchmarks for evaluating AI capabilities, including red team testing, are critical for identifying and mitigating risks.^{xxiv} Specific simulations include:



Adversarial testing through which teams simulate attacks on AI systems to identify weaknesses and improve defenses.



Scenario analyses to evaluate how AI systems perform under various scenarios, including worst-case situations.



Continuous monitoring to detect and mitigate emerging threats.

Advanced considerations

As AI governance matures, advanced considerations such as developing AI Bills of Materials (BOMs), and incorporating governance into a Responsible AI framework are crucial for ensuring the transparency, security, and reliability of AI systems. By adopting these practices, organizations can better manage risks, enhance accountability, and promote the ethical and responsible use of AI technologies. Further it has become increasingly important to engender trust in AI systems across stakeholders (including the workforce, customers, regulators, and the public at large) and a major step in this direction is offering explainability and transparency into AI components.

AI Bill of Materials: An AI Bill of Materials is a comprehensive inventory of the components that comprise an AI system, including hardware, software, data, tools, methodologies, and pipelines. The focus on AI-related hardware and software includes analyses of computational hardware used for training and inference, programming languages utilized, and machine learning library versions.

This approach fosters accountability as it helps ensure that AI systems are built using reliable components, that the origins of these components are clearly known, and that a chronological record of an AI model's training is included. Further, the AI BOM is also a critical factor in data governance, as it can be used for data source documentation that covers training and validation datasets, as well as classification and analysis of synthetic data to address any potential biases, limitations or ethical considerations.

Finally, organizations that use AI BOMs will find that the approach has many uses as AI models evolve. These include tracking performance metrics such as R2 scores and F1 scores and providing information that allows researchers and developers to reproduce models and results.

^{xxiv} Test, Evaluation & Red-Teaming, National Institute of Standards and Technology (NIST) Test, Evaluation & Red-Teaming initiative.

As we look forward

The rapid integration of artificial intelligence into organizations across the globe underscores the critical need for robust AI governance frameworks. As AI technologies continue to evolve, organizations must prioritize the development and implementation of comprehensive governance models that address legal, ethical, and security risks. By establishing clear roles and responsibilities, paying close attention to ethical considerations, codifying robust policies and documentation, and instituting continuous workforce training, organizations can ensure the transparency, accountability, and resiliency of their AI systems.

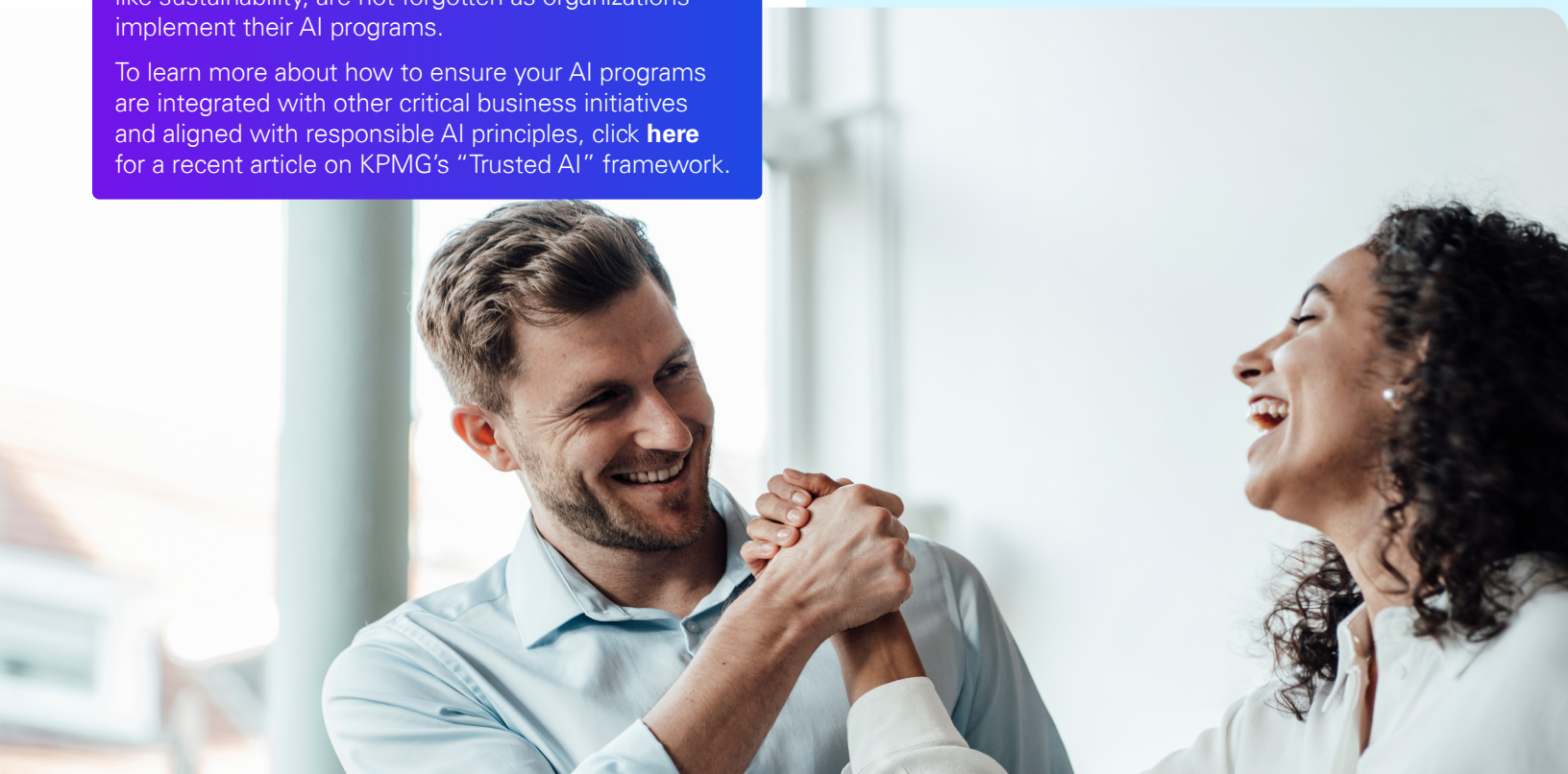
Ultimately, the future of AI and the innovation it fosters hinge on our ability to govern responsibly. The benefits of AI are immense, but so are the potential risks if not managed properly. By embracing a balanced approach to AI development and security, organizations will be empowered to use AI technologies to contribute positively to society and drive sustainable growth.

Responsible AI framework: It is encouraging that 56 percent of executives believe that AI governance is a critical issue that needs immediate attention, according to a report by the Stanford Institute for Human-Centered Artificial Intelligence.^{xv} However, not all companies are approaching governance comprehensively as part of a broader Responsible AI framework.

Such frameworks should be guided by three principles: values, human-centricity, and trust. A values-driven approach shapes a culture that is open and inclusive, operating to the highest ethical standards. A human-centric approach embraces AI to empower and augment human capabilities, while remembering that there will be opportunities to automate lower-risk use cases. A trustworthy approach ensures compliance with all relevant data privacy and protection regulations and confidentiality.

The takeaway: Many organizations today are facing the risk of noncompliance with expanding regulatory demands, which could force them to halt the use of AI in both immediate and long-term ventures. Additionally, the media has cautioned that organizations may struggle to secure sufficient data and/or computing power to support their AI ambitions, potentially leading to an “AI Winter”—a period of reduced funding and interest in AI. These concerns underscore the critical need to establish a robust and systematic AI governance strategy to ensure that business priorities, like sustainability, are not forgotten as organizations implement their AI programs.

To learn more about how to ensure your AI programs are integrated with other critical business initiatives and aligned with responsible AI principles, click [here](#) for a recent article on KPMG’s “Trusted AI” framework.

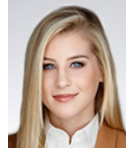


^{xv} James Fell, Stanford just released its annual AI Index report. Here’s what it reveals, World Economic Forum, April 26, 2024.

Contact us



Katie Boswell
KPMG LLP
Managing Director,
Cybersecurity & Tech Risk
katieboswell@kpmg.com



Kristy Hornland
KPMG LLP
Director, Cybersecurity &
Tech Risk
khornland@kpmg.com



Mark Orsi
CEO
GRF
morsi@grf.org

Contributors

Campbell's
Mark Wehrle

Cranium
Jonathan Dambrot
Felix Knoll
Daniel Christman
Josh Harguess
Chris Ward
Kelsey Flynn

GRF
Chris Denning
Pat McGlone

Hinshaw & Culbertson LLP
Sherri Vollick

**Independent Security
Consultant (formerly with Intel)**
Jean-Philippe Martin

Johnson & Johnson
Bill Janicki

Kenvue
David Merritt, PhD

Knostic/Cyber Defense Matrix
Sounil Yu

KPMG
Trevor LaNouette
Alexis Harrison-Wattley
Katherine Yang
Donna Ceparano
John Hodson

Manifest Cyber
Marc Frankel

Daniel Bardenstein

New York University
Joel Caminer

Recorded Future
Rupal Kharod

Shared Assessments
Andrew Moyad
Chris Johnson

**Sharpe Management
Consulting LLC**
Alex Sharpe

Stan Stoel Rives LLP
Jon Washburn

Wells Fargo
David LaFalce

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

Learn about us:



[kpmg.com](https://www.kpmg.com)

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.

© 2024 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. The KPMG name and logo are trademarks used under license by the independent member firms of the KPMG global organization. USCS022932-1A