

This article was prepared for the Munk School—The Future of Money event held on April 4, 2024.



Introduction

Generative AI (Gen AI) continues to fascinate and excite users. In a recent Deloitte survey, more than 79% of respondents expect Gen AI to bring significant organizational change in less than three years.¹ But what's driving its popularity? Unlike many past technological advancements, Gen AI boasts a rare combination of user-friendliness and accessibility, making it approachable for a wide audience. This has led to the democratization of Gen AI in a way never imagined, where people of all ages with various backgrounds, and organizations across diverse sectors and disciplines, are experimenting with these tools and developing solutions at a scale and speed seldom seen for a new technology.

For organizations, unlocking the potential of Gen AI offers a goldmine of opportunities. Businesses of all kinds can benefit from capabilities such as hyper-personalized experiences for consumers, intelligent supply chain management to optimize manufacturing practices, and enablement of an AI-augmented workforce. Gen AI can also be a source of good, enhancing fields like health care through AI-powered diagnostics and drug discovery or helping solve complex problems like climate change through AI-generated scenario planning and analysis.

While many business leaders are excited by the possibilities that Gen AI can unleash for businesses, employees, and our communities at large, Gen AI also poses significant challenges and risks beyond the ethical dilemmas, social impacts, legal implications, and technical limitations already inherent in traditional AI. What marks the shift we are seeing now are emerging risks that are specific to Gen AI: explainability, hallucinations and inaccuracies, misinformation, deepfakes, algorithmic bias, blurring of the lines of authorship, and increasingly, ongoing legal debates regarding copyright protection measures.^{2,3} As AI's risk profile changes over time, our collective obligation towards understanding and proactively mitigating these risks will become paramount.

The path towards capitalizing on Gen AI's potential is not without obstacles. A one-size-fits-all AI governance approach won't work. Emerging global regulations will demand clear justifications for AI model function and data privacy compliance. A nuanced approach will be required that balances benefits with risks and fosters trust among stakeholders. Furthermore, as Gen AI solutions are being widely positioned as general-purpose technologies, the urgency for responsible AI governance and development has never been greater.

This paper will explore the urgent need to accelerate our path towards responsible AI. By doing so, we can unlock its tremendous potential while proactively mitigating risks and unintended consequences. We hope that this paper will stimulate a constructive and inclusive dialogue on the future of Gen AI, and its implications for organizations and society.

What is Responsible AI, and why is it so urgent to act now?

Though the sentiment toward Gen AI is at an all-time high, the trust gap with this technology remains equally prevalent. In a recent study conducted by the Boston Consulting Group, only 12% of non-management respondents who were not yet using Gen AI in the workplace, reported that they place trust in this new technology. This finding was consistent with senior managers and executives who were not yet actively using GenAI, with only 21%, and 22% respectively expressing a sentiment of trust.⁴

A similar trust gap can be extended to consumer sentiment. For instance, imagine a scenario where distrust in Gen AI persists among consumers, leading to reputational risks or eroding brand loyalty. Companies reluctant to invest in AI technology due to uncertain market demand may scale back their research and development efforts. Consequently, this reduced investment could slow down the pace of innovation, hindering the emergence of transformative AI applications that could benefit society.

To counteract these risks and negative outcomes, fostering trust in Gen AI necessitates prioritizing responsible AI development and enhancing education for all stakeholders. Responsible AI encompasses ethical considerations, such as fairness, transparency, and accountability, throughout the AI lifecycle.⁵ By adhering to responsible AI principles, stakeholders can mitigate risks associated with AI deployment and address concerns that contribute to low trust. For example, implementing transparent AI algorithms that allow users to understand decision-making processes can help alleviate fears of opaque and biased AI systems, thereby fostering trust among end-users and stakeholders.⁶ When responsible AI development practices are deployed at scale, the opportunities for societal good can be profound.

Exploring the potential of Gen AI for industry, government, and society

Beyond Gen AI's commercial applications, it holds immense potential for driving positive societal impact and addressing some of the world's most complex challenges. The growing acceptance of Gen AI stems from its demonstrable efficacy in addressing some of our most pressing societal challenges, such as incurable disease, global food waste, and climate change. Below are some sample case studies where imagination meets inspiration to ignite Gen AI as a source of good.

Health | Drug discovery and development

In a world where cures aren't delayed and treatment is truly personalized, Gen AI is poised to revolutionize drug discovery. By analyzing massive datasets and simulating drug interactions, researchers can pinpoint the most promising molecules to develop therapeutic solutions. This approach can cut through traditional roadblocks, slashing development times and costs. The expected results would see fewer failed trials, which would translate into faster breakthroughs and safer treatments for patients in desperate need.⁷

Supply Chain | Food Waste

Food waste is a global crisis, but what if we could end it? Intelligent supply chain management could be our untapped superpower. By analyzing data on inventory, demand, and expiration dates, we can optimize every step from farm to fork, ensuring that nothing goes to waste. We could plan for intelligent systems that manage food distribution more effectively, with the impact of reducing hunger and environmental impact.

Climate change | Geospatial modelling

Both traditional and Gen AI offer to be our shield against earthquakes, floods, and wildfires. Imagine using weather patterns and seismic activity to predict disasters before they strike, evacuating people and positioning aid for immediate response. Satellite images and sensor data become our eyes, mapping disaster zones and pinpointing damage for faster rescues. Agencies could then collaborate seamlessly, sharing data and resources in real-time. The result? Fewer lives lost, communities recovering quicker, and a future where nature's fury is met with human resilience.

Financial Crime: Fraud Detection

The ever-rising tide of data can cripple traditional fraud detection methods. Businesses and individuals are left exposed and vulnerable to attack. The change on the horizon is one where advanced analytics powered by Gen AI can become our defender. Imagine intelligent machines sifting through massive datasets, unearthing hidden patterns of fraud before they strike. Businesses and individuals become fortresses, protected by AI's watchful eye. The result? A safer financial landscape, where transparency reigns, individuals feel empowered, and the economy stabilizes.

These examples provide just a glimpse into the unique opportunity that Gen AI presents to reimagine the future of our economy and communities, where innovation serves as a catalyst for inclusive growth, sustainable development, and shared prosperity. This is where responsible AI development supports innovation at scale to enable AI-fueled organizations realize their full potential.

Tackling AI challenges

Traditional AI, which focuses on a rule-based approach, is designed to handle specific tasks that demand logical reasoning, pattern recognition, and rule-based decisioning.^{8,9} Beyond its pre-defined rules, traditional AI does not generalize information outside of its designated tasks and lacks the ability to learn from the data it receives to improve over time. As a result, the inherent risks that are associated with traditional AI can range from training-data biases leading to discriminatory outputs to a lack of explainability that makes it difficult to understand how the AI arrives at decisions, which can raise concerns about accountability.¹⁰

While Gen AI shares risks with traditional machine learning models, its capacity for original content creation and self-learning introduce an entirely new set of risks to understand and mitigate. The unique qualities of generative AI models make them inherently more challenging to understand and control, emphasizing the imperative for safety and trust.¹¹ Some unique risks related to Gen AI are listed below.



Explainability

- **Lack of explainability:** Generative AI models often function as black boxes—we know they produce outputs, but the internal decision-making process is opaque. This makes it difficult to understand how the model arrives at its results, or consider how biases can potentially weigh in. Traditional AI, particularly rule-based systems, can be relatively more transparent in its decision-making.
- **Limited human oversight:** The black box nature of these models can make it challenging to effectively oversee or audit a generative AI system. This raises concerns about accountability and potential misuse, as it becomes difficult to identify where issues might originate within the complex model. Traditional AI, especially for safety-critical applications, often has built-in mechanisms for human oversight and control.



Hallucinations / Inaccuracies

Generative AI models are based on probability estimates, which allows them to predict the likelihood of an outcome. When faced with information they don't know, these models tend to 'make it up.'¹² So why is this happening? There can be multiple factors at play:

- **Data limitations:** If the AI model is trained on incomplete or inaccurate data, it might reflect those flaws in its outputs.
- **Misinterpretation:** The AI might pick up on patterns in the data that aren't there, leading to nonsensical creations.
- **Model complexity:** The inner workings of complex AI models can be mysterious, and sometimes they can produce unexpected results.



Data

- **Fit for use:** the quality of the data used to train Gen AI models can significantly impact their performance and fairness. Additional due diligence is also required to assess the quality and relevance of the data used.¹³ Biased data can lead to discriminatory outputs in the Gen AI model and data loss, and corruption may also compromise the model performance.
- **Security and privacy:** Data security breaches can be catastrophic for an organization. At various stages of the model development lifecycle, protection measures will need to be in place to ensure robust security measures, alignment to relevant regulations, and protection of sensitive PII information. Establishing a data governance framework that outlines how data will be collected, stored, and used can help organizations plan for secure, private, and ethical use of the data.



Deepfakes

Deepfakes refer to manipulated or synthesized media, such as images, videos, or audio, that appear to be real but are created using AI. They have gained additional attention for their ability to deceive or impersonate individuals, posing serious ethical and security concerns. Threats that deepfakes pose can range from phishing attacks, tricking people into revealing sensitive information, to malicious actors creating fake videos or audio of public figures, all of which pose serious public risks.



Algorithm bias

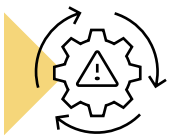
- **Amplification of existing biases:** Generative AI models trained on biased data will amplify those biases in their outputs. This can lead to discriminatory outcomes in areas like loan approvals, job applications, or even criminal justice. Unlike traditional AI focused on specific tasks, generative AI can unintentionally perpetuate these biases across newly created outputs.
- **Difficult to detect and mitigate:** The complex nature of generative AI algorithms makes it challenging to pinpoint the source of bias within the model. Unlike traditional rule-based AI, where bias might be easier to identify in the code, generative AI's bias can be deeply ingrained and require specialized techniques to detect and address.



Intellectual property (IP)

- **Ownership of generated content:** Who owns the copyright or rights to creative content generated by AI? This is an ongoing legal debate. If a generative AI uses copyrighted material in its training data, the ownership of the generated content becomes murky. Traditional AI tools typically wouldn't create original content, so IP concerns are less prominent.
- **Attribution and originality:** Generative AI can blur the lines of authorship. It's crucial to distinguish between human-created content and AI-generated content, especially for creative fields like music or writing. This is less of a concern with traditional AI used for analysis, where the source of the output (the algorithm) is clear.

Gen AI's impact on existential risk is likely to be indirect and rests heavily on responsible development and deployment.



Existential risk

- **Limited capabilities:** Gen AI, despite its unprecedented abilities, remains primarily focused on specific tasks, like content creation. To date, it lacks the level of general intelligence, reasoning, and judgement that would introduce an existential threat to society, like some form of super-intelligent AI.
- **Human control:** Gen AI systems are still under human control. We design, train, and deploy them. While potential for misuse and bad actors are certainly a concern, these risks are not unique to Gen AI.

However, Gen AI could indirectly contribute to existential risks by:

- **Fueling propaganda, bias, and disinformation:** As mentioned earlier, generative AI can create highly realistic deepfakes. This could erode trust in information and potentially destabilize societies. Bias in Gen AI may also be exacerbated, heightening social inequalities for marginalized or vulnerable populations. Focusing on known harms could result in ignoring long-term risks that are not well addressed by traditional AI policy objectives.¹⁴

The Ecosystem Model for Gen AI adoption

An ecosystem model approach to Gen AI emphasizes collaboration and interdependence across business, individuals, and technology platforms, and may provide a clearer path towards furthering the adoption and scaling of GenAI solutions. Unlike traditional market models that rely on businesses to operate independently and compete against each other, an ecosystem model could create greater value for organizations and customers, while solving for unique risks introduced by this technology.¹⁵

Organizations are already well down the path of making build, buy, or partner decisions regarding their development and adoption of Gen AI, each posing distinct considerations from a risk perspective (see Appendix 2).

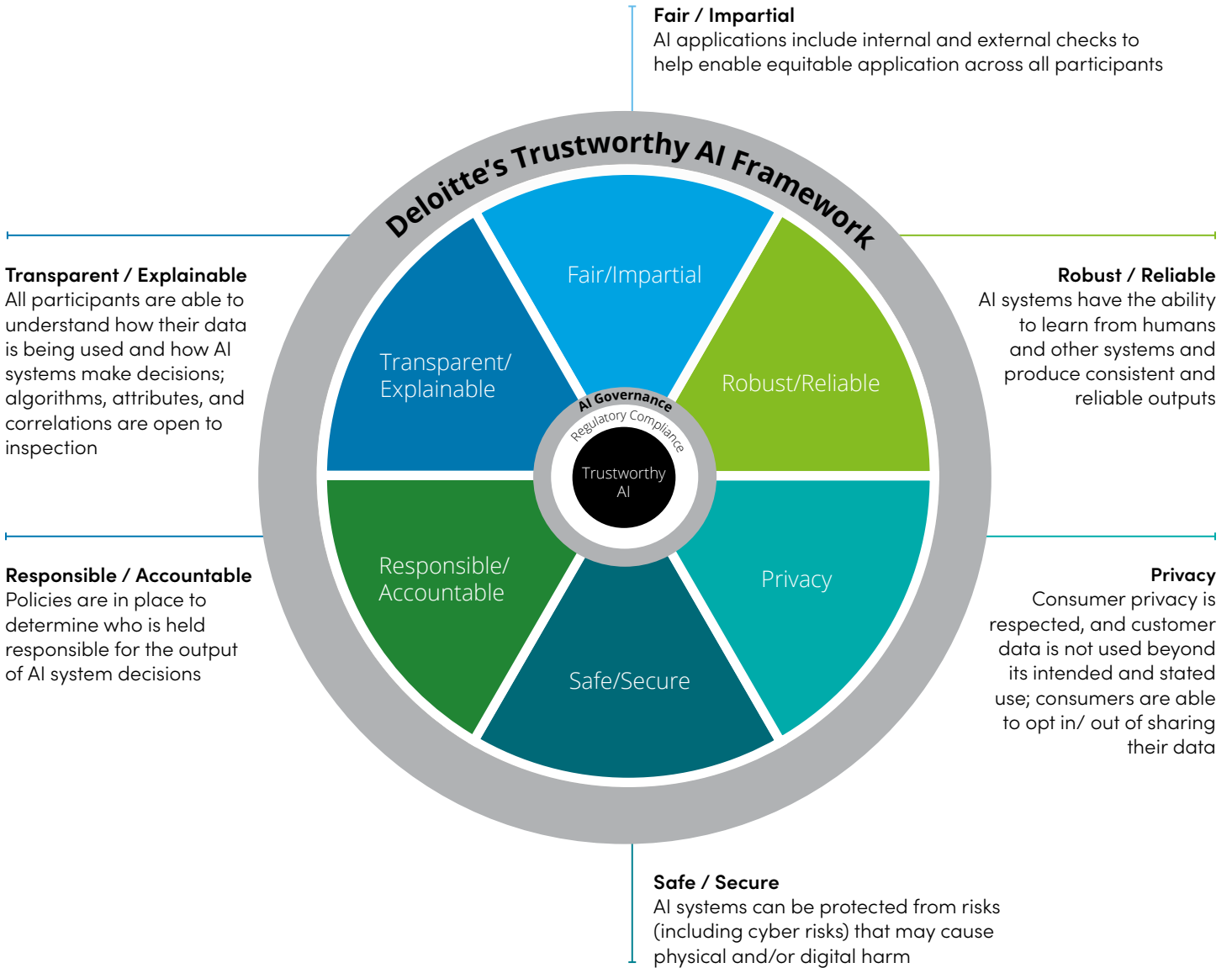
Navigating the Risk landscape

A well-defined risk framework can help organizations and policy leaders navigate complexities related to Gen AI risks. By considering all facets of risk, encompassing safety, privacy, transparency, fairness, reliability and responsibility, organizations can chart a clear path forward and harness the transformative potential of GenAI in a sustainable manner.

By adopting a holistic approach to risk management that integrates technical, ethical, and societal considerations, stakeholders can harness the transformative potential of generative AI while addressing the associated risks and challenges in a responsible and sustainable manner.



Figure 1: Trustworthy AI Framework addresses traditional AI risks





Constructing appropriate regulatory, policy, and governance guardrails

Navigating the delicate balance between innovation and regulation is essential to ensuring adaptability to technological advancements without compromising ethical standards.¹⁶ While the potential of Gen AI is immense, it is essential to temper enthusiasm with a pragmatic approach to AI governance and regulation.

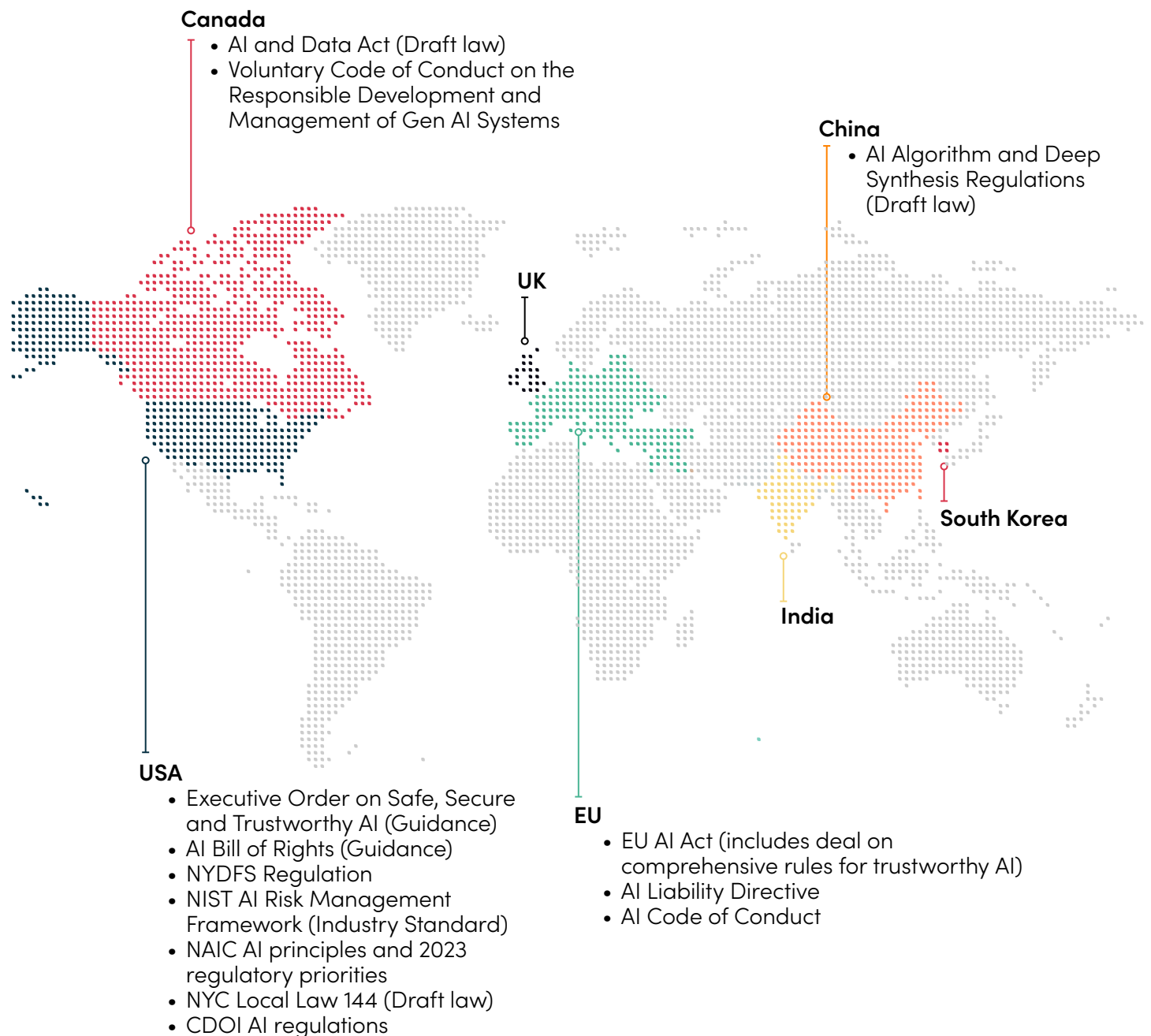
AI governance landscape

Since the public launch of ChatGPT, more than 1,600 policy initiatives—including regulations about other policies—have been cited as under development by the Organization for Economic Co-operation and Development (OECD) AI policy observatory.^{17,18} Based on a recent Deloitte survey, 47% of organizations reported that they were actively monitoring regulatory requirements as part of their risk management efforts.¹⁹

As technological innovation begins to outpace the development of regulatory frameworks, regulators worldwide are grappling with the challenge of ensuring that the deployment of generative AI across industries remains transparent, fair, and compliant with existing regulations.²⁰ Striking the right balance between fostering innovation and safeguarding citizen interests requires collaborative efforts between industry stakeholders, policymakers, and regulatory bodies.

Figure 2. Emerging global regulations and industry standards^{21, 22}

Regulatory frameworks play a pivotal role in shaping the rules of engagement in the ever-evolving landscape of AI. Regulators are taking a multidimensional view of AI risk, considering how AI is being used and what controls are in place to mitigate AI risks. When it comes to AI risk management frameworks, several structures are emerging, featuring key elements such as data governance, explainability, and driving accountability through processes and defined roles and responsibilities.





Striking a balance: Enabling protection measures to scale innovation

While global regulation and governance aims to address potential risks and ethical concerns, it's vital to implement them in a way that doesn't hinder progress. This delicate dance requires careful consideration of various factors to ensure that safeguards are implemented without stifling innovation.

One approach to achieving this balance involves a stepwise process that fosters collaboration among researchers, businesses leaders, and policy makers. This process can involve prioritizing use cases that hold the potential for significant societal and organizational benefits. By carefully assessing the potential impact and potential risks associated with each use case, leaders can prioritize those that align with ethical principles and responsible development practices. This measured approach can help ensure that innovation flourishes within a framework that safeguards against potential harms.

Spotlight | Crowdsourcing for Change

(Stanford University's AI Audit Challenge)

Challenge: The lack of necessary tools to independently analyze and audit deployed AI systems can pose a significant obstacle in ensuring that AI systems do not exhibit bias or carry the potential for discrimination.

Solution: Through crowdsourcing, a diverse group of stakeholders composed of researchers, industry leaders, policymakers, and technologists set out to design and create applied tools that can assess deployed AI systems and open-source models for bias and discrimination. The developed tool can enable humans to effectively audit AI systems, allowing for transparency and accountability in their deployment.

Impact:

- **Enhanced accountability:** Independent analysis and audit of AI systems can be used to enhance accountability by allowing external parties to evaluate fairness, ethics, and biases, aligning with societal values.
- **Mitigation of bias and discrimination:** the creation of tools to help identify and address biases and discriminatory patterns in AI systems, fostering inclusivity and fairness.
- **Improved trust and user acceptance:** Auditing tools build trust and user acceptance by assuring scrutiny for potential biases and discrimination, increasing confidence in AI technology.
- **Advancement of responsible AI:** Independent analysis and audit tools promote responsible AI practices, prioritizing fairness, transparency, and accountability for more ethical and inclusive AI systems.

The challenge isn't technology— it's people and culture

In a future where we look forward to navigating and mitigating the myriad Gen AI risks outlined in this paper, with a view to Gen AI's exciting potential, perhaps the most significant challenge ahead lies in cultivating trust at a human level.

There's no doubt that Gen AI will revolutionize the way we work across all fields and sectors. Studies by the International Monetary Fund (IMF) suggest that nearly 40% of global jobs have the potential to be affected by AI, with many routine tasks becoming prime targets for automation.²³ Some of us may see this as a risk, leading to the potential for unemployment, income equality and social disparity, particularly for those who lack the skills and resources to adapt to a new work reality. Others, in contrast, may see this as a tremendous opportunity to harmonize the interactions between humans and technology, where Gen AI serves as a powerful assistant and collaborator. By alleviating the more mundane elements of work in favor of problem solving, creativity, and innovation, Gen AI has the potential to elevate workers to achieve levels of satisfaction and productivity that we can only imagine.

How the talent landscape will evolve in the future depends on our commitment to education, reskilling, and above all, trust building. This is our choice and, like reducing other types of risk, maybe one of the most critical challenges we face as leaders now and into the future.

Questions for deployment across the dimensions of responsible AI

- 1. Transparency and explainability:** Can the decision-making processes of Gen AI solutions be understood by humans? How can we ensure transparency in these systems to build trust and identify potential biases?
- 2. Bias and fairness:** How can we mitigate potential biases present in the data used to train Gen AI models and ensure their outputs are fair and non-discriminatory across different demographics and social groups?
- 3. Safety and security:** How can we ensure the safety and security of Gen AI systems, preventing them from being hacked, manipulated, or misused for malicious purposes? What safeguards are needed to mitigate potential risks and unintended consequences?
- 4. Accountability and responsibility:** Who is accountable for the actions and decisions made by Gen AI systems? How can we establish clear lines of responsibility for their development, deployment, and potential harms?
- 5. Societal impact and human values:** How will Gen AI solutions impact society as a whole? How can we ensure their development and use align with human values such as privacy, equality, and justice? How can we manage the potential economic and social disruptions caused by widespread Gen AI adoption?
- 6. The future of work:** How can organizations design effective upskilling programs that address the evolving needs of workers in the era of AI? What role should governments play in promoting learning and providing accessible education and training opportunities to ensure individuals are equipped with the necessary tools for the future of work?

Conclusion

As we grapple with the implications involved with the safe, responsible, and ethical usage of AI technologies, how organizations respond now will echo across industry landscapes. The challenge before us is understanding how we can harness the potential of these technologies for the greater good, while mitigating risk and at the same time empowering and elevating a workforce of the future. It's a tall order, and thankfully, there are several initiatives underway that can help to guide our collective path forward.

Vector Institute Summit of Generative AI and Roundtables: In 2023, Vector convened 170 AI researchers, along with Canadian business leaders from startups to large organizations, to discuss how to harness the potential of Gen AI. Through collaboration and sharing of best practices, attendees could make connections across the ecosystem, share knowledge, and explore potential paths forward. These discussions highlighted the need for open collaboration and the healthy exchange of ideas to make AI more trustworthy. Proposed paths forward that were presented include increasing opportunities for open dialogue and creating a dedicated platform for companies working with Gen AI to share findings.^{24,25}

Global Risk Institute | OFSI Canadian Financial Industry Forum on Artificial Intelligence Responsible AI from a Canadian perspective: In 2023, the Canadian Financial Industry Forum on Artificial Intelligence (FIFAI) hosted a series of workshops that discussed how to use AI responsibly in the Canadian financial sector. The goal of this forum was to promote the dialogue on how to adopt Gen AI in the Canadian financial sector in a responsible way. The main outcomes of these discussions were that collaboration, a multidisciplinary approach, and trust in AI technology are essential for success. Trust can be fostered by ensuring adequate levels of explainability, disclosure, and a customer-centric approach that respects ethical values and protects privacy. Moreover, trust in the technology can help speed up adoption, and this can be done by applying the EDGE principles (Explainability, Data, Governance, and Ethics).²⁶

Deloitte Global AI Institute: As a global innovation hub, with regional offices around the world, the Institute services as a resource for business, researchers, and the community to navigate the AI landscape. With a focus on practical applications, knowledge sharing and real-world use cases, the Institute regularly publishes research and insights to foster discussions across industry sectors including risk and ethical considerations. The global perspectives shared provide awareness on enterprise adoption trends, and emerging areas of focus.^{27,28}

As taken from best practices like the initiatives above, we see a few common themes on how to actionably proceed:

- **Collaborating with industry stakeholders, policymakers, and regulatory bodies** to strike the right balance between fostering innovation and safeguarding citizen interests.
- **Recognizing the imperative to ensure that human skills are explicitly considered as part of our Gen AI revolution.** Ensuring that organizations invest in education, upskilling, and continuous learning will be key to enabling employees to work effectively alongside Gen AI tools and unlock their full potential.
- **Prioritizing use cases** that hold the potential for significant societal and organizational benefits, while aligning with ethical principles and responsible development practices.
- Adopting a **holistic approach to risk management** that integrates technical, ethical, and societal considerations, while **monitoring regulatory requirements** both globally and regionally as part of risk management efforts.

Unleashing the full potential of Gen AI will require us to work together, develop ethical guidelines, and apply a comprehensive risk management method to the solutions that emerge. This will assist us in our endeavors to ensure that this groundbreaking technology is developed in a responsible and sustainable manner.



Contributors

This paper was written by Deloitte on behalf of the following collaborators:

Scott McDonald

President, The Future of Money
Senior Fellow at the Munk School

Yiyi Shi

Master of Global Affairs, University of Toronto

We wish to thank the following Deloitte people for their contributions to this report.

Jas Jaaj

Managing partner, Data and AI (including the Office of Generative AI)
Global Innovation Leader

Shauna Emerson-O'Neill

Partner, National FSI Ecosystems and Alliance Leader

Mike Jancik

Director, Future of Canada Centre

Aisha Greene

Senior Manager, AI Institute Canada

Appendix 1

Aligning Gen AI risks with the LLM lifecycle helps identify potential mitigation strategies²⁹ for adoption and usage.

	Stage	Common Risks	Mitigation
1	Data Collection and Training	<ul style="list-style-type: none">• Biases in data• Data privacy concerns• Data Quality Issues	<ul style="list-style-type: none">• Implement bias audits• Data anonymization• Diversify datasets
2	Model Development	<ul style="list-style-type: none">• Model vulnerabilities• Algorithmic biases• Overfitting	<ul style="list-style-type: none">• Robust testing• Model interpretability techniques• Diversity in model architecture
3	Deployment and Integration	<ul style="list-style-type: none">• Unauthorized access• Privacy breaches• Intellectual property infringement	<ul style="list-style-type: none">• Access controls• Security vulnerabilities (encryption mechanisms)• Secure APIs• Copyright checks
4	Interaction and Use	<ul style="list-style-type: none">• Misuse• Spreading Misinformation and Fake Content• Unintended consequences	<ul style="list-style-type: none">• Education and awareness campaigns• Adherence to ethical guidelines• Content moderation
5	Monitoring and Maintenance	<ul style="list-style-type: none">• Drift in data distribution• Performance degradation• Model Staleness	<ul style="list-style-type: none">• Regular model updates• Continuous monitoring• Feedback loops• Model retraining

Appendix 2

Exploring the risk profile to build, buy or partner

For business leaders, the pressure is on to move quickly to adopt Gen AI solutions for their organization,³⁰ with 94% of 2,620 global leaders saying generative AI is critical to their strategy for success over the next five years.^{31,32} Many organizations have reacted by using ready-made Gen AI solutions in the short term. These include productivity applications with integrated generative AI (71%); enterprise platforms with integrated generative AI (61%); standard generative AI applications (68%); and publicly available large language models (LLMs) (56%), such as ChatGPT.³³

When deploying a GenAI solution for the enterprise, leaders must evaluate if the models used will be taken from open-source communities, bought from other third parties, or built in-house. Model choice will need to consider how well each model fits the use case, as well as time to market, necessary resource investments (e.g., capital and talent), licensing and acceptable use policies, risk exposure, and competitive advantage offered.³⁴

The organizational risk considerations related to buy, build, or partnership approaches are described in Figure 3.

Figure 3. Build vs Buy vs Partnership

Generative AI solutions & models have varying levels of customizability, which underscores the need to evaluate build vs. buy vs. partnership approaches to operationalize these tools. Explainability and transparency are critical across all approaches to ensure there is clarity regarding how these models reason and arrive at their output.

Custom solutions to evaluate whether to build or buy a foundation Gen AI model

GAI adoption via software applications	'As-Is" commercial or open source LLMs	Bespoke vendor-trained foundation model	Custom-build your own model
Out of the box	Generalizing		Level of effort and cost to maintain
GenAI on existing software (e.g., Salesforce, Workday, SAP) or stand-alone applications (Midjourney, GitHub copilot, RunwayML)	Vendor trained foundational model solutions, such as GPT4, trained on large, publicly available datasets	Vendor trained foundational models that mixes use case specific enterprise data with public data that is both use case specific and general purpose to train a foundation model	Developing in-house large language models for industry specific use cases where off-the-shelf solutions are lacking or don't yet exist (for non-language modalities)
Risk Considerations: <ul style="list-style-type: none"> • Time-to-market • Integration with technology platform • Clear IP ownership rights and usage policies 	Risk Considerations: <ul style="list-style-type: none"> • Applicability of realistic generalized models in business scenarios • Cyber threat actors 	Risk Considerations: <ul style="list-style-type: none"> • Additional fine-tuning requirements • Data privacy agreements • Compliance to Personal Identifiable Information (PII) 	Risk Considerations: <ul style="list-style-type: none"> • Training data availability • New business models • Cost and feasibility • Ability to track data lineage

Endnotes

1. Deloitte, "Now decides next: Insights from the leading edge of generative AI adoption", Jan 2024. <https://www2.deloitte.com/content/dam/Deloitte/at/Documents/presse/at-gen-ai-studie-2024.pdf>
2. See note 1.
3. Deloitte Germany, "Trustworthy AI." <https://www2.deloitte.com/de/de/pages/innovation/contents/trustworthy-ai.html>
4. International Monetary Fund, "Gen-AI : Artificial Intelligence and the future of work." <https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2024/01/14/Gen-AI-Artificial-Intelligence-and-the-Future-of-Work-542379>
5. Gartner, "Responsible AI." <https://www.gartner.com/en/information-technology/glossary/responsible-ai>
6. See note 3.
7. Deloitte US, Quartz Atlas AI™ for Drug Discovery. <https://www2.deloitte.com/us/en/pages/consulting/solutions/deloitte-nvidia-atlas-ai-for-drug-discovery.html>
8. Ibid.
9. Dan Hendrycks, Mantas Mazeika, Thomas Woodside, "An Overview of Catastrophic AI Risks." <https://arxiv.org/pdf/2306.12001.pdf>
10. See note 3.
11. See note 3.
12. See note 3.
13. See note 8.
14. See note 10.
15. Michael G. Jacobides, "In the Ecosystem Economy, What's Your Strategy?" Harvard Business Review. <https://hbr.org/2019/09/in-the-ecosystem-economy-whats-your-strategy>
16. Laura Weidinger, Jonathan Uesato et al. "Taxonomy of Risks posed by Language Models." <https://dl.acm.org/doi/pdf/10.1145/3531146.3533088>
17. See note 1.
18. Deloitte, "The AI regulations that aren't being talked about." <https://www2.deloitte.com/us/en/insights/industry/public-sector/ai-regulations-around-the-world.html>
19. See note 3.
20. The Lawfare Institute, "The EU's AI Act Is Barreling Toward AI Standards That Do Not Exist." <https://www.lawfaremedia.org/article/eus-ai-act-barreling-toward-ai-standards-do-not-exist>
21. See note 1.
22. Based on a recent Deloitte-facilitated panel with the U.S. Federal Reserve Board, Office of the Comptroller of the Currency, and Financial Industry Regulatory Authority.
23. See note 4
24. Vector Institute, "Vector Institute hosts first-of-its-kind Generative AI Leadership Summit." <https://vectorinstitute.ai/vector-institute-hosts-first-of-its-kind-generative-ai-leadership-summit/>
25. Vector Institute, "How to put Generative AI to good use?" <https://vectorinstitute.ai/generative-ai-discussions/>
26. Office of the Superintendent of Financial Institutions and Global Risk Institute, Financial Industry Forum on Artificial Intelligence: A Canadian Perspective on Responsible AI. https://www.osfi-bsif.gc.ca/sites/default/files/documents/ai-ia_en.pdf
27. Deloitte Global AI Institute, <https://www2.deloitte.com/us/en/pages/deloitte-analytics/articles/advancing-human-ai-collaboration.html>
28. Deloitte Canada AI Institute, <https://www2.deloitte.com/ca/en/pages/deloitte-analytics/articles/ai-institute.html>
29. Dan Hendrycks, Nicholas Carlini et al, "Unsolved Problems in ML Safety." <https://arxiv.org/pdf/2109.13916.pdf>
30. See note 1.
31. See note 1.
32. Deloitte, Fueling the AI transformation: Four key actions powering widespread value from AI, right now. <https://www2.deloitte.com/us/en/pages/consulting/articles/state-of-ai-2022.html>
33. See note 1.
34. World Economic Forum, AI Governance Alliance Briefing Paper Series, January 2024. https://www3.weforum.org/docs/WEF_AI_Governance_Alliance_Briefing_Paper_Series_2024.pdf



Disclaimer and Copyright

This publication contains general information only and Deloitte is not, by means of this publication, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your business. Before making any decision or taking any action that may affect your business, you should consult a qualified professional advisor. Deloitte shall not be responsible for any loss sustained by any person who relies on this publication.

About Deloitte: Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee ("DTTL"), its network of member firms, and their related entities. DTTL and each of its member firms are legally separate and independent entities. DTTL (also referred to as "Deloitte Global") does not provide services to clients. In Canada Deloitte refers to one or more of the Canadian member firms of DTTL, their related entities that operate using the "Deloitte" name in Canada and their respective affiliates. Certain services may not be available to attest clients under the rules and regulations of public accounting. Please see www.deloitte.com/about to learn more about our global network of member firms.