



The flip side of generative AI: Challenges and risks around responsible use

This advanced machine learning technology offers quick and low-cost content creation. It can also expose organizations to IP theft, fraud, and reputational damage.



Introduction

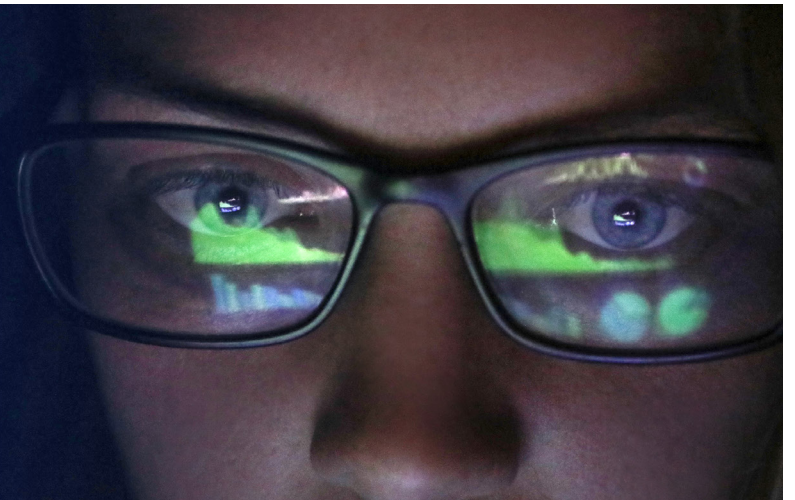
Businesses are taking a close look at generative artificial intelligence (AI), since it allows for the quick production of high-quality content—text, images, video, code—with minimal human effort. With less need for human resources, companies expect to produce content at faster speed and lower cost, enabling the creation of new kinds of content that were too costly to develop before. For example, generative AI can handle high-volume, low-value activities, such as summarizing articles, drafting emails, creating first drafts of code, or high-level activities, like creating images and video or debugging code, that would leverage intellectual property, but expose the company to the risk of fraud or theft of proprietary or private information.

Creating a generative AI tool takes a lot of time, money, and effort, so for now, most organizations will need to rely on third-party generative AI solutions, like OpenAI and Stability AI. Weighing potential risks, including error,

potential fraud, and loss of intellectual property, will likely remain a major consideration when deciding whether to use a third-party generative AI or relying on another type of AI in house.

With generative AI gaining rapid momentum and entering the mainstream, its growing popularity is yet another reason responsible AI—developing and deploying AI in an ethical manner—should be a top concern for organizations looking to use it or protect themselves against its misuse. In this paper, we'll look at the risk management challenges raised by generative AI and underscore the importance for organizations to identify these risks early, form governance constructs to help adopt generative AI responsibly, and understand how risks may impact building trust in the use of generative AI. We'll also provide our perspectives on steps companies can take now before adopting generative AI.

Responsible AI should be a top concern for organizations looking to use it or protect themselves against its misuse.



Risks from within

Intellectual property

Among the top risks around the use of generative AI are those to intellectual property. Generative AI technology uses neural networks that can be trained on large existing data sets to create new data or objects like text, images, audio, or video based on patterns it recognizes in the data it has been fed.¹ That includes the data that is inputted from its various users, which the tool retains to continually learn and build its knowledge. That data, in turn, could be used to answer a prompt inputted by someone else, possibly exposing private or proprietary information to the public. The more businesses use this technology, the more likely their information could be accessed by others.

Amazon has already sounded the alarm with its employees, warning them not to share code with ChatGPT.² A company lawyer specifically stated that their inputs could be used as training data for the bot and its future output could include or resemble Amazon's confidential information.³

Also, generative AI content created according to an organization's prompts could contain another company's IP. That could cause ambiguities over the authorship and ownership of the generated content, raising possible allegations of plagiarism or the risk of copyright lawsuits.

Organizations will need to figure out how to protect their intellectual property, while still being able to gain the benefits of generative AI. Given organizations' desire to use AI for competitive advantage and to harvest their existing data, specific considerations around how their data is used for training and public consumption should be evaluated. One solution is data anonymization and de-identification, but that would have to be a service offered by the vendor, or applied prior to sending data to the vendor, and based on contract agreements. The downside would be that businesses could be limited in reaping the benefits from other organization's data.

Employee misuse

Using generative AI offers business great efficiencies but also powerful temptations for misuse by employees. Educators have voiced concern that students could use generative AI to write their essays and other assignments; there have already been cases where a teacher has alleged a student has used ChatGPT to write an essay for a class.⁴ Employees, too, might be tempted to use generative AI and pass off the result as their own.

A related misuse would be for contract workers to pass off generative AI work as their own and billing the company for hours of work they didn't in fact perform.

A more serious example of employee misuse is using generative AI to automate legal confirmations or reviews that may skirt appropriate ethics and compliance, independence, or other programs, which may affect regulatory culpability.

¹ Source: Unlock the Potential of Generative AI: A Guide for Tech Leaders. Forbes.com. January 26, 2023

² ChatGPT is an online tool created by San Francisco start up OpenAI. GPT stands for Generative Pre-trained Transformer.

³ Source: Amazon Warns Employees to Beware of ChatGPT. Gizmodo.com. January 23, 2023

⁴ Source: Professor catches student cheating with ChatGPT: 'I feel abject terror.' The New York Post. December 26, 2022.

Risks from within *continued*

Inaccurate results


Even the legitimate use of generative AI carries risks.

Consider the inaccuracy of outcomes. Employees using generative AI will need to be vigilant in applying professional skepticism and an extra emphasis on quality assurance to the results. Generative AI has limitations on learning “new” outcomes, meaning additional training and research is needed, along with monitoring of end results to verify that it is delivering in line with expectations.

Should the generative AI content contain inaccuracies, it could cause any number of failures that could impact business outcomes or create liability issues for the business. Meta’s generative AI bot Galactica, for instance, was created to condense scientific information to help academics and researchers quickly find papers and studies. Instead, it produced vast amounts of misinformation that incorrectly cited reputable scientists.⁵

Lack of transparency in the use of generative AI content can also create reputational issues for organizations. Tech publisher CNET recently came under criticism for quietly using the technology to write 73 articles since November 2022, some of which contained errors, even though the publisher said on its website that a team of editors is involved in the content “from ideation to publication.”⁶

Other risks around generative AI include perpetuating or even amplifying societal biases that may be present in the data used to train the tool or the possibility that the technology could generate sensitive information, such as personal data, that could be used for identity theft or invade privacy. Even a disgruntled employee or angry customer could create fictitious material that could malign a company’s reputation or that of one of its employees or executives.



Should the generative AI content contain inaccuracies, it could cause any number of failures that could impact business outcomes or create liability issues for the business.

⁵ Source: Meta AI Bot Contributed to Fake Research and Nonsense Before Being Pulled Offline. Gizmodo.com. November 22, 2022.

⁶ Source: CNET Cops to Error Prone AI Writer, Doubles Down on Using It. Gizmodo.com. January 25, 2023, and CNET Has Been Quietly Publishing AI-Written Stories for Months. Gizmodo.com. January 11, 2023.


External risks

Risks from generative AI can come from the outside as well, with unscrupulous users having the potential to create a lot of mischief and headaches for companies. Many of these malicious actions can already be perpetrated without generative AI. But generative AI can make the deed that much easier and quicker to pull off—and much harder to detect.

Generative AI also can be used to create so-called deepfake images or videos with uncanny realism and without the forensic traces left behind in edited digital

media, making them extremely difficult for humans or even machines to detect.⁷ A deepfake image could be created depicting a company executive in a scandalous situation. Or an individual could use generative AI to create fictitious images or video and use them to file fraudulent insurance claims.

Finally, generative AI also raises cybersecurity risks. Cybercriminals can use the technology to create more realistic and sophisticated phishing scams or credentials to hack into systems.



Cybercriminals can use the technology to create more realistic and sophisticated phishing scams or credentials to hack into systems.

⁷ Source: Deepfakes: An insurance industry threat. Propertycasualty360.com. September 14, 2021.

What to do

The business use of generative AI is still in the early days. But there are steps companies can take now to begin building responsible AI governance and set accountability for strategic decisions around the use of data and generative AI.

To start, the development of new governance constructs to address risk and ethical implications of using AI should involve the evaluation of critical questions across multiple functions from across the organization.

- **Risk Management/Compliance/Internal Audit**—What policies and procedures do we need to put in place? What risk and controls does the business need to consider for the use of generative AI?
- **Legal**—What can or should we be able to do with generative AI? What data or IP is acceptable to be used in prompts, etc.? How do we protect the IP we make using generative AI? What contractual considerations should be in place to safeguard data?
- **Public Affairs**—What is our plan for potential external misuse that impacts our company?
- **Regulatory affairs**—What are regulators saying for our industry/company about how to consider generative AI?
- **Representatives from the business**—How might the organization be looking to use generative AI and what should we be on the lookout for? What types of levers can be used to track content generation for internal and contingent workers? What should our workforce know about generative AI, both risks and benefits?

From these inquiries, organizations can move forward to building a governance construct and set of principles to guide them in making decisions about the ethical use of data and AI, improving the digital literacy within the organization to build confidence in using advanced analytics techniques (such as generative AI), and creating automated workflows and validations to enforce AI standards throughout the development through production lifecycle.

With a Responsible AI program in place, organizations can begin moving forward with developing processes and procedures around the use of generative AI. These efforts should include:

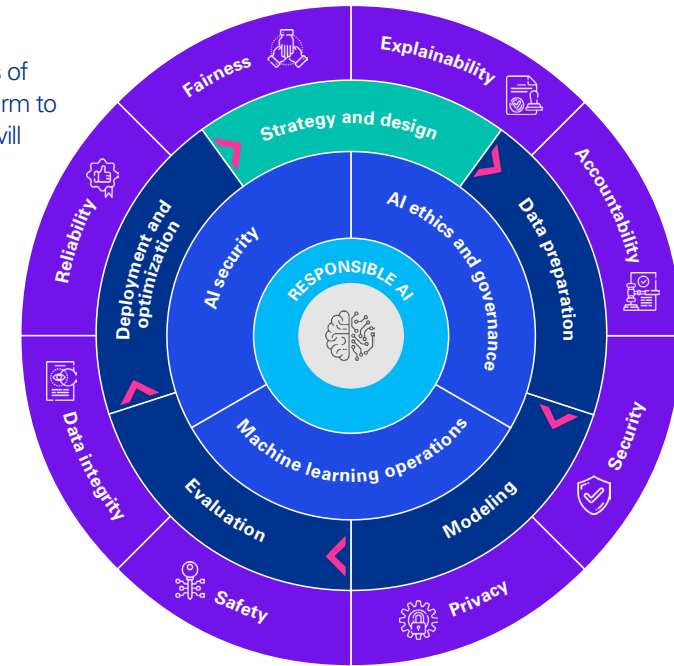
- Identifying appropriate stakeholders from the organization and driving initial education on the technology, its potential benefits, and risks
- Developing an initial internal POV for the organization to get employees thinking about generative AI
- Prioritizing risks and assigning ownership to stakeholder groups to address across the development through production lifecycle
- Taking into consideration the organization's AI governance principles (aligned to regulatory landscape) that will encourage the responsible use of generative AI

Generative AI is becoming another powerful technique in an organization's toolbox of advanced analytics and AI. However, with this growing set of capabilities comes the need to develop literacy and governance to ensure the use of generative AI builds trust in the outcomes it creates and is done in line with the organization's standards while considering business and customer impacts.

How KPMG can help

We all have a responsibility to learn about the risks of generative AI and control these risks to prevent harm to customers, businesses, and society. These risks will grow and evolve as AI technology advances and becomes more pervasive, and as public pressure from regulators increases.

KPMG's responsible AI offering is a set of frameworks, controls, processes and tools to ensure AI systems are being designed and deployed in a trustworthy and ethical manner so that companies can accelerate value. We understand responsible AI is a complex business, regulatory, and technical challenge, and we are committed to helping clients put it into practice.



Our eight core principles guide our approach to responsible AI across the AI/ML lifecycle



1

Fairness

Ensure models are free from bias and equitable.



2

Explainability

Ensure AI can be understood, documented, and open for review.



3

Accountability

Ensure mechanisms are in place to drive responsibility across the lifecycle.



4

Security

Safeguard against unauthorized access, corruption, or attacks.



5

Privacy

Ensure compliance with data privacy regulations and consumer data usage.



6

Safety

Ensure AI does not negatively impact humans, property, and environment.



7

Data integrity

Ensure data quality, governance, and enrichment steps embed trust.



8

Reliability

Ensure AI systems perform at the desired level of precision and consistency.

Wherever you are in your responsible AI journey, our 15,000+ technologists can tailor our vast experiences, field-tested approach, and cutting-edge solutions to your unique needs and challenges, helping you to accelerate the value of generative AI with confidence.

Contact us

To learn more about how KPMG can help your organization adopt a responsible AI program, please contact:



Matteo Colombo

Principal, Advisory,
Lighthouse
KPMG in the U.S.
matteocolombo@kpmg.com



Kelly Combs

Director, Leader of Responsible AI
Lighthouse
KPMG in the U.S.
kcombs@kpmg.com



Jordan Seiferas

Managing Director,
Data & Analytics, Lighthouse
KPMG in the U.S.
jseiferas@kpmg.com



Aisha Tahirkheli

Managing Director,
Advisory, Lighthouse
KPMG in the U.S.
atahirkheli@kpmg.com



Bart Van Rompaye

Head of Advanced Analytics and
Machine Learning
Lighthouse | Advisory
KPMG in Belgium
bvanrompaye@kpmg.com

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

© 2023 KPMG Advisory, a Belgian BV/SRL 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.

kpmg.com/be/socialmedia

