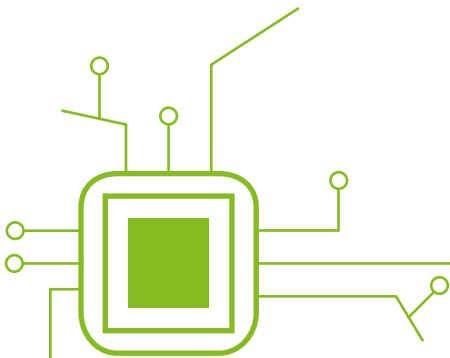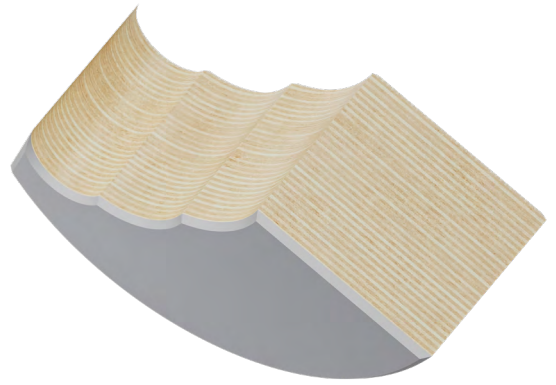# Google Cloud | Deloitte.

Ensuring the future of Generative AI is ethical

February 2024

Generative Artificial Intelligence (Gen AI) has emerged as a transformative force with the potential to reshape industries and enterprises, revolutionise the way people work and live, and address society's most pressing challenges and opportunities.

But, as the buzz around Gen AI grows, so do concerns about its potential risks. These include fears about AI models lacking accountability, breaches of people's right to privacy, irresponsible use of personal information, and potential unfair bias amplification that could reinforce harmful stereotypes.

Currently, Gen AI has a vast amount of knowledge, but it does not necessarily have the experience or ability to apply specific business acumen in decision-making or autonomously understand and respect social norms. It also tends to embellish the truth on occasion, or even make things up completely, to engage and keep your attention. This hallucination could potentially have serious consequences for enterprises using Gen AI systems to complement or augment human endeavours.

To help Gen AI reach its full potential as it matures, enterprises must put in place robust frameworks and guardrails to ensure the future of Gen AI is both ethical and beneficial to society.
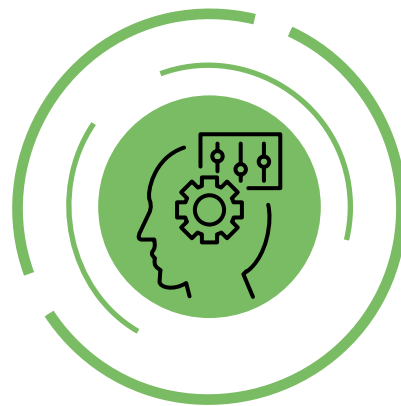
Understanding the challenges and opportunities of ethical AI

The ethical risks and challenges associated with Gen AI are complex and multifaceted. Ensuring Gen AI is responsible requires enterprises to see to it that systems are fair, safe, secure, socially beneficial, accountable, and designed with privacy in mind. It is also about ensuring AI systems are used in a way that respects human rights and values and does not cause harm to individuals.

This means actively mitigating biases in algorithms, being transparent about the data sources and processes used in training AI models, and regularly assessing and addressing potential ethical risks throughout the AI system's lifecycle. Enterprises must also seek individuals' informed consent to use their data and give them control over how their data is used.

It is also vital that organisations put in place safeguards to prevent the misuse of their AI systems, allow them to adhere to legal and regulatory frameworks, and ensure that oversight and accountability mechanisms are embedded in their AI practices.

Failing to address these issues can lead to severe consequences, including regulatory penalties, reputational damage, loss of customer trust and the erosion of share prices. However, the most compelling reason for enterprises to develop and deploy AI responsibly is that they have little choice. With Gen AI quickly becoming an essential capability

and key strategic differentiator for businesses, organisations that do not proactively operationalise how to use AI responsibly may find themselves continually running into roadblocks. On the other hand, taking a proactive approach to using AI responsibly can result in business benefits, including a better reputation, improved customer trust and greater market share.

As Reena Jana, Head of Content & Partnership Enablement, Responsible Innovation at Google, Deloitte's technology partner, pointed out recently, "Ethical AI isn't just a 'nice to have'. It is essential for any company that wants to use AI sustainably and build for long-term success.

"Organisations that are seen to be using AI ethically are more likely to be trusted by customers and other stakeholders. This can lead to increased sales, improved brand loyalty and a more positive public image.

"AI can also help organisations develop new products and services, improve their operations, and reduce costs, giving them a competitive advantage in the marketplace. Employees are also more likely to want to work for organisations that are seen to be using AI ethically."

Establishing a strong framework for responsible innovation

Despite the strategic imperative of addressing ethical issues around Gen AI, many organisations lack an effective strategy or processes in place for developing and using Gen AI solutions.

This is where having a robust framework for developing and using AI can help. For example, Google's AI Principles, first published in 2018, underscore the technology company's commitment to the responsible use of AI. By adhering to these principles, the company aims to develop AI systems that contribute positively to the world and uphold ethical values. This includes ensuring that its AI technology:

- is socially beneficial.

- avoids creating or reinforcing unfair bias.

- is built and tested for safety.

- is subject to human control.

- incorporates privacy principles.

- upholds high standards of scientific excellence.

- is made available for uses that accord with these principles.
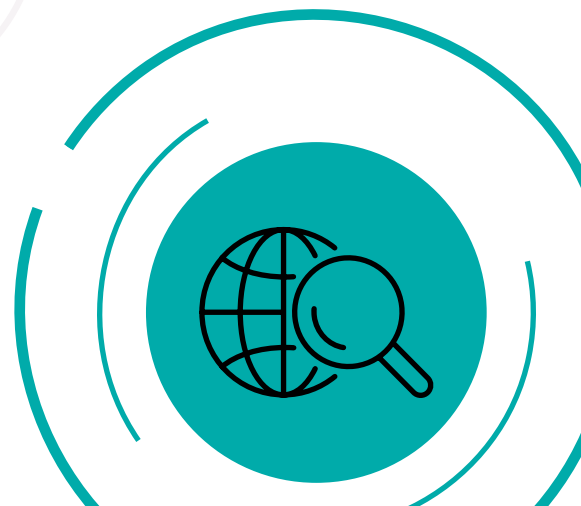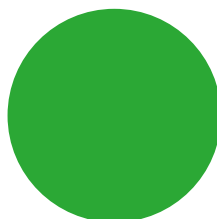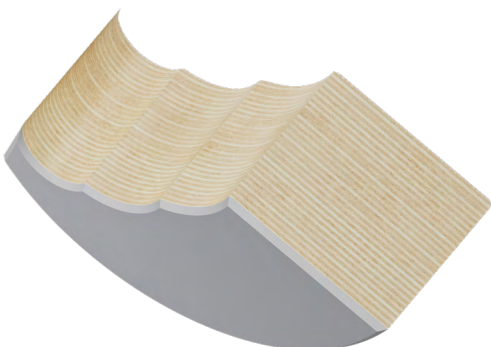
Google is also committed to limiting potentially harmful or abusive applications by evaluating all new AI applications through a critical lens. Furthermore, the company has stated that it would not develop AI applications that will cause, or are likely to cause, overall harm to society, such as weapons or surveillance systems or systems that contravene widely accepted principles of international law and human rights.

Google puts its AI Principles into practice in a variety of ways, including by developing tools and resources to help developers mitigate unfair biases in AI models, implementing internal processes to review new AI products and services, and funding ongoing research into AI's benefits and risks.

"Google's AI Principles are not a 'cure all' for solving all of the challenges associated with fairness, interpretability, privacy, and safety in AI systems," says Jana. "New challenges will evolve over time, just as technology will. However, the principles provide a framework for developing and using AI systems in a responsible way."

At Deloitte, our Trustworthy AI framework reflects our shared values and expectations about AI and guides decision making around its design, operation, and governance. The framework identifies seven key areas for organisations to consider when designing, developing, deploying and operating AI products.

Privacy
User privacy is respected, and data is not used or stored beyond its intended and stated use and duration. Furthermore, users can opt-in or opt-out of data sharing.

Accountability
Policies determine who is responsible for decisions that are made using technology.

Fairness and impartiality
The technology is designed and operated inclusively, with the aim of delivering an equitable application, access and outcome.

Robustness and reliability
The technology produces consistent and accurate outputs, withstands errors, and recovers quickly from unforeseen disruptions and misuse.

Transparency
Users understand how technology is being used, particularly in decision-making. These decisions are easy to understand, auditable, and open to inspection.

Responsibleness
The technology is created and operated in a socially responsible manner.

Safety and security
The technology is protected from risks that may cause individual or collective physical, emotional, environmental or digital harm.

Like Google, Deloitte believes that enterprises can guide the responsible use of AI by applying this type of ethical lens to every decision about its development and applications.

Taking the first steps towards a responsible future

While there is no doubt that Gen AI has the potential to deliver significant transformational benefits to enterprises, ensuring it is used responsibly can be enormously challenging. Hence, Deloitte's Generative AI practice has partnered with Google to help organisations take the first steps on their Gen AI journey.

Drawing on our combined expertise and extensive experience delivering tailored solutions for clients, we work with organisations to design, implement, and scale solutions that help them achieve their business objectives in an ethical and responsible way. This includes assessing clients' existing AI maturity, including their capabilities, data assets, and AI governance structures, and developing a custom AI strategy that aligns with their business goals.

We also work with clients to develop and implement an AI governance policy and processes to guide the development of all new AI products and services. Clients can also access Google's AI platforms, pre-trained AI models and consulting services, as well as any training and ongoing support they may need to optimise their AI projects' value.

By addressing the complex ethical issues surrounding Gen AI, enterprises are not only laying the foundation for their own future success, but are also contributing to a fairer, more equitable AI-driven future for all.

If you would like to find out how we can help you embark on your Gen AI journey, sign up for our Gen AI Workshop or contact the Deloitte Greenhouse team. We would love to hear from you.
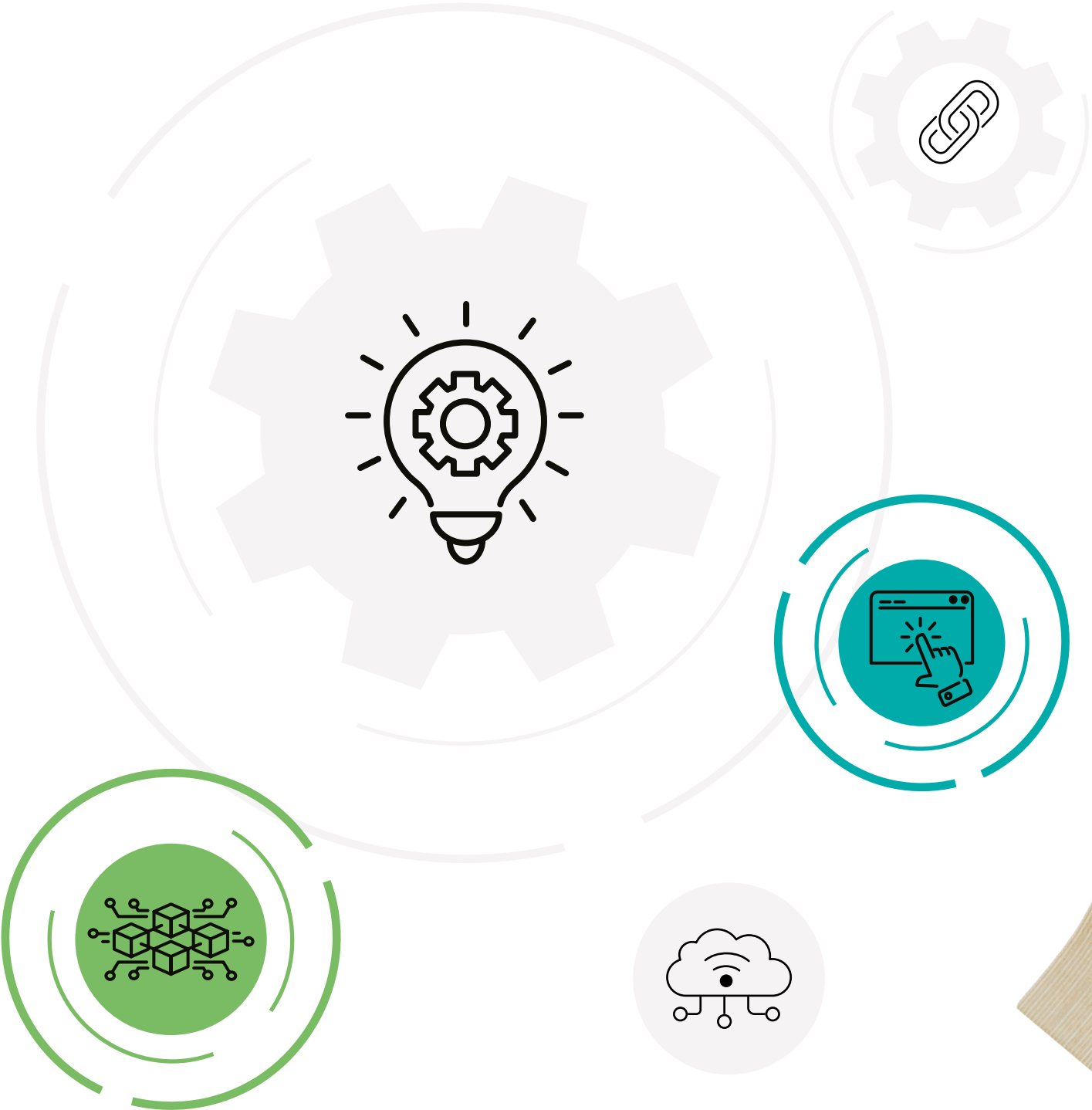
Karen Brown

Greenhouse Lead
Deloitte Southeast Asia
karenbrown@deloitte.com

Deloitte Greenhouse®

GENERATIVE AI LAB

Harness the potential of Generative AI with Deloitte demos adapted to unique industry nuances. Reimagine business futures and craft forward-thinking solutions.

Author

Leonard Jayamohan

Consulting Executive Director
Deloitte Southeast Asia
ljayamohan@deloitte.com

Deloitte.