Avoid wrong turns on the road to responsible AI

EY Building a better working world

Microsoft

Author

Samta Kapoor
EY Americas Energy AI and Responsible AI Leader

Rani Bhuva
EY Americas Financial Services Responsible AI Leader

In a world where technology advances at a rapid pace, staying up to date with the latest leading practices can be a challenge. However, when it comes to generative AI (GenAI), staying on top of these practices is not just helpful — it is crucial to the success of your business. With GenAI making arguably the fastest technological leap in history, it has the power to transform entire industries. Along with this potential comes some significant challenges. Issues such as algorithm bias affecting hiring, targeting and even credit risk profiles; inaccuracies, and complex ethical considerations that can easily be mishandled, resulting in business complications; and reputational risk or even the spread of misinformation on a global scale.

67%

of CEOs cite potential errors as a top risk of AI and ML integration, reinforcing their concern about the trustworthiness of AI and ML.[1]

To help counter these issues, in October, President Biden signed an executive order to establish new standards for AI safety, security and privacy. In Europe, after a lengthy debate, European Union (EU) officials have established a groundbreaking AI safety and transparency framework, known as the AI Act. This legally binding framework imposes new transparency requirements for AI developers and bans controversial uses of AI, such as real-time facial recognition scanning, emotion recognition and predictive policing.

Now, as businesses around the world rush to embrace the possibilities and advantages of AI, there is an urgent call from governments, academics and forward-thinking business leaders to prioritize responsible AI usage and implement leading practices. Together with Microsoft, experienced Ernst & Young LLP (EY) teams, deeply involved in advanced machine learning (ML) and automation strategies, have taken proactive steps to develop a comprehensive framework. This framework serves as a guide for companies on their journey toward responsible AI.

In May of 2023, Microsoft proposed a five-point plan for the responsible governance of AI, aiming to foster accountability and trust in the emerging technology. The plan outlines the key principles and practices that organizations should follow to make sure that their AI systems are ethical, fair and transparent.

In collaboration with Microsoft, we have developed a comprehensive framework for responsible AI, drawing on their extensive experience and know-how in advanced machine learning and automation. The framework provides a practical and actionable guide for organizations to operationalize Microsoft's five-point plan and implement responsible AI in their business processes and solutions.

Here we share our insights and provide valuable perspectives on how together EY and Microsoft teams help organizations adopt a responsible AI approach.

1 Workday Newsroom, "*Workday Global Survey: 98% of CEOs Say Their Organizations Would Benefit from Implementing AI, But Trust Remains a Concern*," https://newsroom.workday.com/2023-09-14-Workday-Global-Survey-98-of-CEOs-Say-Their-Organizations-Would-Benefit-from-Implementing-AI,-But-Trust-Remains-a-Concern, 2 February 2024.

Mapping out responsible AI

Many companies are already experimenting with GenAI tools and integrating products into their operations at various stages. EY teams help organizations establish the appropriate GenAI governance framework with ongoing monitoring.

As rapid adoption takes place across more organizations, it is crucial to take time to make sure governance practices are in place. For example, GenAI is not 100% accurate or unbiased.

There are already instances of GenAI products introducing racial and gender bias into organizations' hiring processes. Therefore, it is important to ask, what guardrails do you have in place to protect against bias?

A lot of GenAI tools that are being integrated into daily business will soon become invisible and ubiquitous thanks to their easy-to-use interfaces and no-code/low-code architecture. With this widespread use comes significant risk that must be carefully assessed, managed and mitigated. Another important question to ask, are you assessing risk comprehensively across data inputs, user design, algorithms and usage?

Guidelines for responsible AI

1. **Follow AI legal and regulatory compliance** – Adhere to policies that regulate the use of AI systems and their output, such as explainability, data privacy and security, and bias. Consider approaching international organizations to influence greater AI global regulatory consistency.

2. **Maintain standards of cybersecurity** – Protect the data associated with the use of a GenAI tool, and deploy strategies and tactics to mitigate emerging risks, such as adversarial GenAI prompt stream intelligence attack vectors for individuals, companies and providers.

3. **Communicate AI usage or failure to use** – Create alerts for AI system usage and develop policies for obtaining consent to use data. Provide training programs for explaining how AI systems work, its purpose, what value it adds, and how the information is used. Outline the implications of using and not using an AI system.

4. **Check for bias** – Take action to prevent the creation or reinforcement of unfair bias through independent analyses of AI conclusions.

5. **Keep abreast of emerging technology capabilities** – Develop general technical knowledge of GenAI and GenAI prompt engineering skills. AI systems are continuously learning, making it critical to monitor how the performance of today's queries change over time.

6. **Maintain data privacy** – Actively make sure that inadvertent disclosure of confidential information does not occur. Prioritize transparency regarding your plan to use AI.

7. **Supervise technology applications** – Take responsibility for every aspect of the GenAI output by requiring a human verification of the output's legal and factual accuracy, competence and completeness. Maintain a log of decision-making considerations.

8. **Check facts** – Prioritize independent verification of output, as GenAI does not typically cite sources and sometimes delivers fictitious information. Be wary of overestimating the knowledge boundaries of GenAI by comparing a variety of GenAI prompts, particularly if the source data is unknown by the user.

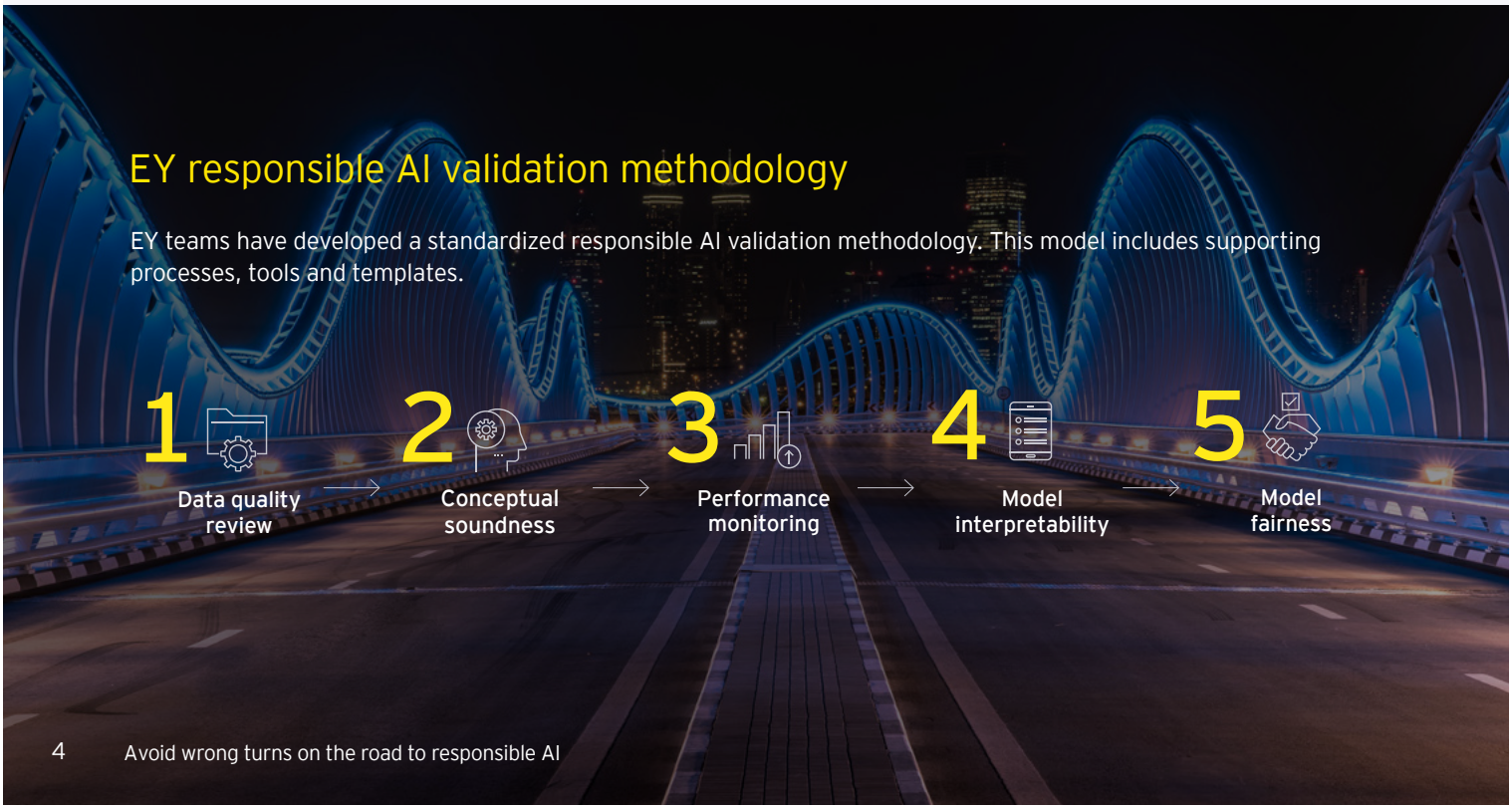
EY and Microsoft mitigate AI risks

EY teams help companies across industry categories embrace GenAI tools to responsibly propel their businesses forward in new ways. In retail, for instance, we're working with companies to harness AI capabilities to analyze competitive data, customer feedback, and thousands of features and design inputs to help accelerate product design cycle time and increased customer satisfaction. In banking we're helping companies develop real-time, AI-assisted knowledge bases that keep employees current on constantly evolving regulations.

We have developed an agile responsible AI framework building on top of Microsoft AI principles. EY principles include; **accountability, data protection, reliability, security, transparency, explainability, fairness, compliance, and sustainability**. This framework assists organizations in mitigating the risks associated with GenAI while providing compliance with rapidly emerging governmental regulations. The EY framework helps companies to innovate responsibly in alignment with their AI strategy and core principles. With implementation of the EY responsible AI framework, companies confidently, and responsibly navigate the complex landscape of GenAI while adhering to ethical practices, corporate compliance considerations and regulatory requirements.

Once your AI system, tool or platform meets the standards outlined in our framework, there are ongoing implementation considerations to address. It is imperative to instill trust and provide accuracy both in the input data and the outputs generated by the AI system. Making sure data is input accurately and the establishment of robust data governance practices within your organization is a priority. Ongoing monitoring of biases in the model is necessary to confirm fair outcomes across all business applications and to implement safeguards against any toxic AI responses. Additionally, it is important to assess whether the usage of GenAI will potentially expose private or sensitive company or client data, posing the risk of it becoming public and vulnerable.

In the evolving landscape of early generative pre-trained transformer (GPT) technologies, careful management of copyright issues, ownership rights and infringement is crucial. Without these safeguards, AI will be prone to risk.



Responsible AI in action with Microsoft

EY teams have capitalized on Microsoft cloud to create a Machine Learning Operations framework (MLOps) for companies scaling AI capabilities. By addressing the needs of all stakeholders in the ML life cycle, the framework provides principles of responsible AI while helping to accelerate value realization, streamlining use case model deployment and improving ML operations at scale. These stakeholders include subject-matter resources, data scientists, data engineers, ML engineers, data analysts and business leaders.

The framework consists of three main phases: **design**, **development** and **operations**.

- The **design phase** focuses on identifying high-value use cases and creating machine learning solutions that drive business value.
- In the **development phase**, specifically chosen data is used to iteratively experiment, develop and validate ML models resulting in swift deployment.
- The **operations phase** aids in the management and upkeep of ML models in the product, achieving ongoing business value realization while proactively monitoring risk as part of our responsible AI approach.

Key Microsoft Azure services utilized to provide responsible MLOps at scale include:

- **Azure ML**: AI/ML model management
- **Microsoft Fabric**: Integrated platform delivering AI and insights at scale
- **Azure Purview**: Data understanding and metadata management
- **Azure OpenAI**: Unlocking the power of GenAI

Operate with conviction and compliance

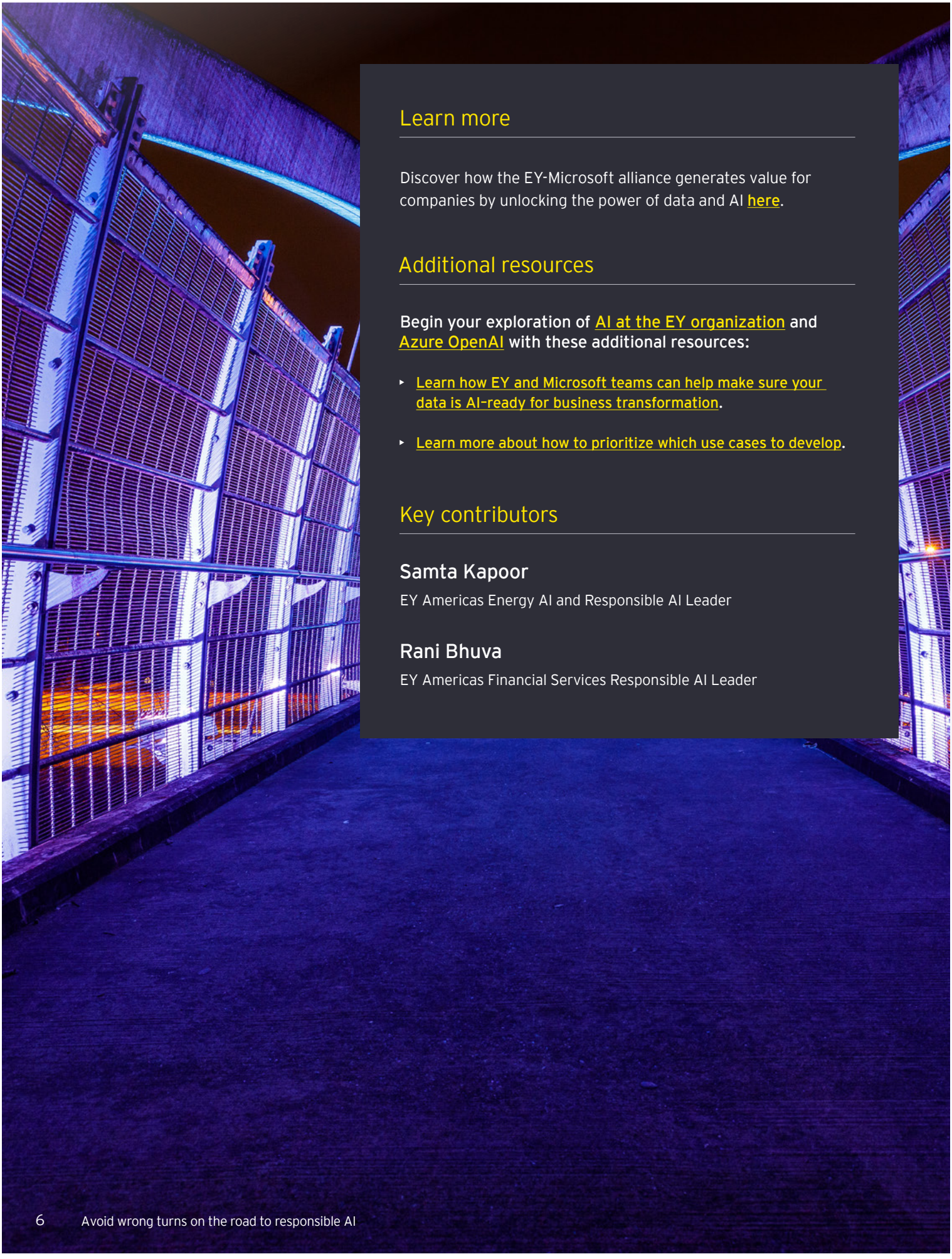
While it may seem tempting to follow competitors who are rapidly adopting GenAI tools and integrations, it is crucial for organizations to exercise caution and undertake due diligence. European countries have already begun regulating GenAI following the adoption of the AI Act from 2021 by the European Commission in April of 2023. Increased oversight in the United States is expected to follow President Biden's executive order on AI usage. Any organization involved in designing, building or integrating GenAI technologies into their operations must adhere strictly to responsible usage principles.

Key considerations, such as data quality management, bias prevention, conceptual soundness assessment, performance monitoring, guaranteed model interpretability and model fairness management, will soon become mandatory rather than voluntary.

By investing time, resources and attention into responsible GenAI practices now, organizations will be well positioned to operate within future compliance guidelines while also promoting the well-being of their businesses, their people and society at large.

CONCLUSION

Together with Microsoft, EY teams have developed a comprehensive framework to guide companies on their journey toward responsible AI. This framework emphasizes unbiased design, resilience, explainability, transparency and consistent performance. By adopting responsible AI practices, organizations can navigate the complexities of GenAI while providing ethical practices, corporate compliance and adherence to regulatory requirements. It is crucial to establish robust data governance practices, monitor biases, and protect private and sensitive data. As regulation increases and compliance becomes mandatory, organizations must prioritize responsible AI usage to operate with conviction and compliance. By investing in responsible GenAI practices, businesses can prepare for the future while promoting the well-being of their organizations, employees and society as a whole.



Learn more

Discover how the EY-Microsoft alliance generates value for companies by unlocking the power of data and AI here.

Additional resources

Begin your exploration of AI at the EY organization and Azure OpenAI with these additional resources:

- Learn how EY and Microsoft teams can help make sure your data is AI-ready for business transformation.
- Learn more about how to prioritize which use cases to develop.

Key contributors

Samta Kapoor

EY Americas Energy AI and Responsible AI Leader

Rani Bhuva

EY Americas Financial Services Responsible AI Leader

EY | Building a better working world

EY exists to build a better working world, helping to create long-term value for clients, people and society and build trust in the capital markets.

Enabled by data and technology, diverse EY teams in over 150 countries provide trust through assurance and help clients grow, transform and operate.

Working across assurance, consulting, law, strategy, tax and transactions, EY teams ask better questions to find new answers for the complex issues facing our world today.

About the EY-Microsoft alliance

Every day, throughout the world, businesses, governments, and capital markets rely on EY business ingenuity and the power of Microsoft technology to solve the most challenging global issues.

EY and Microsoft teams bring a compelling formula to spark the potential of the cloud and unlock the power of data. We solve our clients' most challenging issues by blending trusted industry expertise with innovative cloud technology. Our strategic relationship draws on decades of success developing visionary solutions that provide lasting value.

Together, we empower organizations to create exceptional experiences that help the world work better and achieve more.

EYG no. 001683-24Gbl
CSG no. 2312-4390176
ED None

ey.com