



AI-FIRST USING AI RESPONSIBLY

Artificial Intelligence (AI) technologies have immense potential to generate value for organizations. However, they also carry certain ethical risks, as decisions taken by systems, relying purely on data without any human involvement, could turn out to be inaccurate, biased, opaque, or hard to justify. Worryingly, 75 percent of organizations use third-party AI tools, but 20 percent of these organizations do not assess the risks stemming from them.¹ Enterprises need to make sure their AI solutions are built responsibly to mitigate various types of risk.

Enter Responsible AI

Responsible AI is an approach to developing and deploying AI technologies in an ethical, transparent, and fair manner in accordance with the law. The goal of responsible AI is to ensure that the decisions taken to create AI systems produce equitable and beneficial outcomes. In responsible AI, human well-being and values such as fairness, reliability, and accountability are at the heart of system design.^{2,3,4}

That being said, responsible AI doesn't only reduce the risk of flawed or biased outcomes, but also improves system performance to generate significant value. Since responsibly designed AI systems are less prone to error, they improve innovation and consumer trust. Besides system design issues, responsible AI also addresses various data-related concerns. For example, it seeks to "balance" training data so that the AI model (such as a loan approval system) bases its decisions only on relevant criteria to produce an outcome that is not biased against a certain gender, ethnicity, or race. Another example is data privacy protection. Responsible AI requires that organizations evaluate the relevance of Personally Identifiable Information to their use cases before deploying it, and make sure that once deployed, it is not exposed to other parties.

Responsible AI is particularly relevant to the highly regulated, risk-averse financial services industry, which is exploring use cases across banking, wealth management, and investment advisory. Some of the use cases that are seeing traction include, AI-based helpdesk/self-help systems or recommendation engines that use customer data to complete their tasks, while safeguarding it from misuse or exposure. Another use case is wealth management systems that allow only the designated relationship manager (and not other managers) to access a client's data.



Consider AI-First

With AI use cases expanding rapidly, financial institutions will need support to adopt AI solutions responsibly. An AI-First framework can guide them to try AI as the first option to solve any problem and prioritize the various use cases in order of value. Using an AI-First framework, financial institutions can adopt those use cases that are best-suited to their context and business objectives. A bank that wishes to adopt responsible AI should first strengthen the data foundation by ensuring that the data is clean, of high quality, and free from bias. For a bank whose goal is to use AI responsibly in core functions, use cases such as fair and ethical credit decisioning, fraud detection, or regulatory compliance would take priority. Organizations seeking to drive growth with AI can deploy responsibly designed AI solutions for cross-selling, portfolio management, and improving customer relationship.

At the same time, financial institutions going AI-First should expect a few challenges along the way. From an ethical and responsible perspective, a major challenge, as mentioned earlier, is protecting personal, confidential, or sensitive data from being exposed outside the AI system. A second challenge is controlling access to various data, based on role, need, permissions, function, and geography. For example, a wealth manager should not ordinarily be allowed to access clients' Personally Identifiable Information; building these "walls" where there is no human intervention can pose certain difficulties.

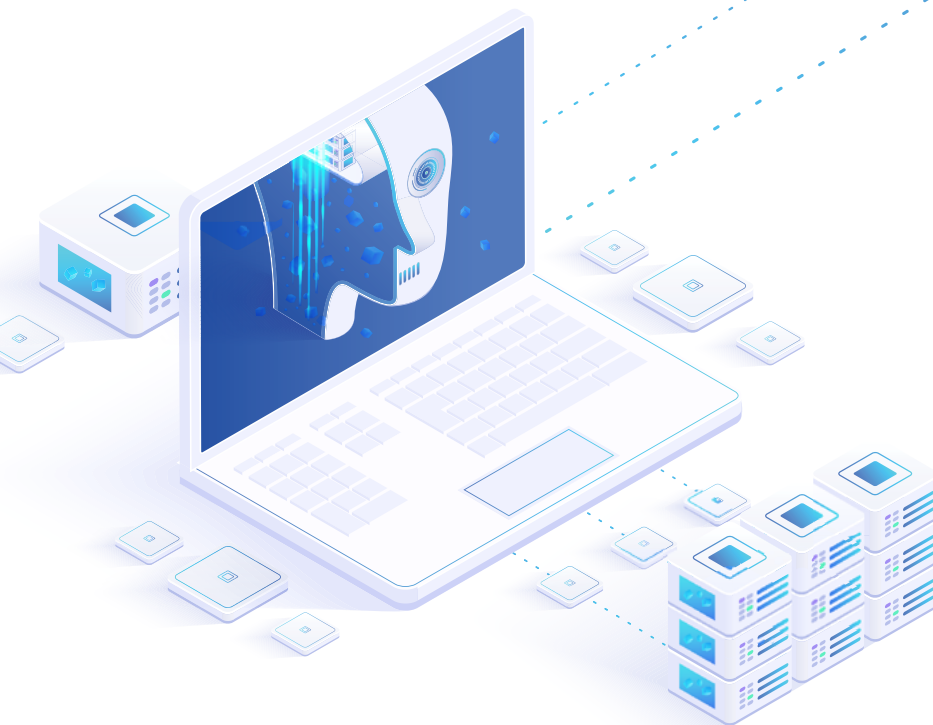
The organizations must also establish guardrails to prevent any toxic or profane data input at the user interface from reaching the core system at the back end; failure to do so can eventually corrupt the AI model (and training data) and skew its output with potentially serious consequences. Think of a rogue user telling a credit risk assessment system that people of a certain race are habitual defaulters; if this information is ingested by the system's AI model it may bias future decisions against creditworthy individuals belonging to that race. Negative stereotyping apart, jailbreaking and prompt injection attacks can bypass the AI system's safeguards to elicit unauthorized data or force the system to take a particular action. Finally, there is the problem of explainability – explaining why the system has taken a particular decision. Organizations must ensure their AI systems are transparent so they can justify a decision or outcome, or, finding that it is flawed, correct it immediately.



First Steps Towards Responsible AI

Organizations can take an AI-First approach to address every need. Starting out with high-quality data – clean, consistent, accurate, and complete – gets the best results from AI. For the same reason, it is important to build a robust AI-First foundation before moving to an AI-First core and AI-First growth applications.

However, factors, such as increasing regulations, rampant cyber threats, and the need to balance innovation and return on investment with ethics and compliance, can make implementing responsible AI very challenging for financial institutions. While the AI-First framework provides direction, the support of an experienced technology partner is required to navigate the technical, policy, and governance challenges of embedding responsible AI across the organization. Banks should work with a partner with strong AI credentials, including a suite of responsible AI offerings and services, who can guide them on the journey to successful adoption.



Next Steps Towards the Future

Few technologies have evoked as much excitement as AI. But the AI dream can shatter quickly unless the technology is designed and deployed responsibly, with respect for ethical values. In this backdrop, the findings of a recent survey of some of the largest banks in North America, Europe, and Asia are cause for concern. For example, 8 of the 23 largest banks in the US, Canada, and Europe are not publicly reporting their responsible AI development practices. The research also notes a lack of transparency around the banks' use of AI, as well as the absence of responsible AI reporting standards.⁵

Financial institutions need to increase the adoption of responsible AI without delay, not only out of respect for ethics, fairness, and other values, but also to maximize AI's transformative impact on their business.⁶

References

- 1 – <https://www.bcg.com/capabilities/artificial-intelligence/responsible-ai>
- 2 – <https://learn.microsoft.com/en-us/azure/machine-learning/concept-responsible-ai?view=azureml-api-2>
- 3 – <https://www.techtarget.com/searchenterpriseai/definition/responsible-AI>
- 4 – <https://www.pegacom/responsible-ai>
- 5 – <https://www.thebanker.com/Banks-need-to-do-more-to-ensure-responsible-AI-use-1681206011>
- 6 – <https://www.bcg.com/capabilities/artificial-intelligence/responsible-ai>

Author



Mandanna Appanderanda
Principal Consultant,
Responsible AI Office, Infosys
Mandanna_AN@infosys.com

Mandanna is an enthusiastic advocate for AI and its transformative potential. With a passion for ethical and responsible AI, he works closely with industry leaders to navigate the complex landscape of AI adoption. Drawing from his experience in product strategy and modernization initiatives, he has successfully shaped innovative technical solutions that are scalable and secure to serve the business needs.



Senthil Nathan
Responsible AI Domain Expert,
Infosys
senthilnathan.n@infosys.com

Senthil supports enterprises to embrace responsible AI throughout the supply chain through innovation, ethics, legal compliance, and maximizing return on investments.

For more information, contact askus@infosys.com



© 2024 Infosys Limited, Bengaluru, India. All Rights Reserved. Infosys believes the information in this document is accurate as of its publication date; such information is subject to change without notice. Infosys acknowledges the proprietary rights of other companies to the trademarks, product names and such other intellectual property rights mentioned in this document. Except as expressly permitted, neither this documentation nor any part of it may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, printing, photocopying, recording or otherwise, without the prior permission of Infosys Limited and/ or any named intellectual property rights holders under this document.