Generative AI transformation

A rising risk awareness

Artificial Intelligence (AI) introduces a complex and multidisciplinary set of risk factors that demand new depth, expertise, and leadership from agency risk functions. Recognizing the urgency of these risks, the Office of Management and Budget (OMB) memorandum on Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence (M-24-10) emphasizes the importance of an integrated, agency-wide risk management function for most types of AI usage. In this paper, we explore the benefits of viewing this risk function as more than yet another cybersecurity exercise. Instead, we propose a strategy that empowers even those with limited AI expertise to evaluate their systems through a comprehensive, sociotechnical lens. This approach not only safeguards human rights and safety, but also equips agencies with the necessary tools to implement AI ethically and effectively. By viewing AI through the lens of the communities they serve, agencies can navigate the complexities of AI adoption with greater confidence and ensuring their actions align with both technological advancements and societal values. To that end, we introduce an intuitive, probabilistic methodology for quantifying risks despite the uncertainty inherent to AI, thereby enabling more strategic decision-making for AI risk management and portfolio governance at the enterprise level.

Classifying AI risks

One of the requirements in OMB's memo is to conduct periodic risk reviews of any safety-impacting or rights-impacting AI. Indeed, a clear understanding of the risks is the only way to know the true costs of an AI solution and whether its purported benefits are desirable considering those costs. Since the risks inherent to AI stem as much from their technological implementation as from their (mis)use, we've found it helpful to ground AI risk assurance in cross-functional, sociotechnical thinking about how an AI model fits into a business process. When you shift your attention from the bits and bytes of one particular AI model and instead conceptualize how that model integrates into a business process, four risk categories emerge:

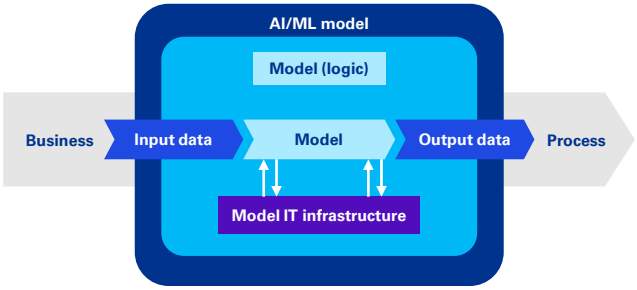

Exhibit 1. How an AI model fits into a business process

Business process risks

Focus on the purpose, scope, and ethical framework within which the AI model operates, including training data as well as risks that are already well-known in the line of business.

1

Input data risks

Concern the quality, relevance, and fairness of the input data, including the quality of instructions used to generate an output.

2

Output data risks

Address the importance of quality assurance for the model's diverse outputs, highlighting the potential for bias and legal issues, especially with GenAI's nondeterministic outputs.

3

Model logic and infrastructure risks

Emphasize understanding the model's operational logic, data sources, data processing, and output generation within its architecture / infrastructure.

4

When you think about a model in the context of a specific business process, it's easier to have a tactical conversation about the risks no matter their provenance, whether they be technical (e.g., cybersecurity), sociotechnical (e.g., misuse), systems-related (e.g., software supply chain), or organizational (e.g., reputational harm). For instance, the ethical issues inherent to using Large Language Model (LLM) for an HR function, like recruitment, will be different from those inherent to another function, like finance, even if both use cases leverage the same LLM.

Quantifying AI risks

While it's clear that AI risk assessments need to be tailored to each individual AI application, OMB—as well as the National Institute of Standards and Technology (NIST), Government Accountability Office (GAO), and other authorities on risk management—have not been prescriptive on how agencies ought to ***quantify*** AI risk exposure. Although useful as a heuristic, qualitative risk registers make it difficult to perform simulations or perform apples-to-apples comparisons between use cases, which is crucial for AI portfolio governance.

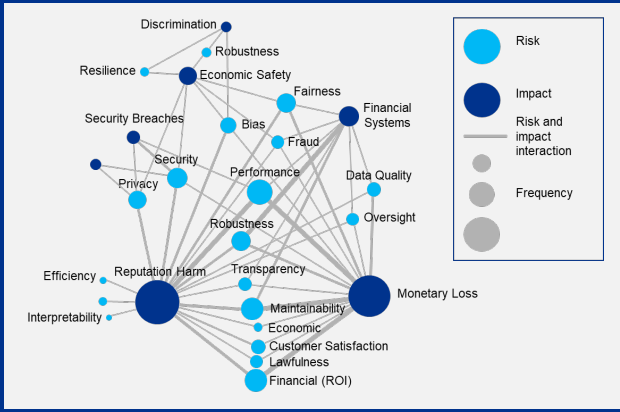

Exhibit 2: A Bayesian Network linking risks to impacts created by a KPMG data scientist

At KPMG, we leverage probabilistic risk assessment techniques coupled with leading AI methodologies to build atop Bayesian Networks. Our methodology goes well beyond traditional heat maps by combining probabilistic methods with graph data science to model complex, potentially interdependent risk factors in a fashion that can handle uncertainty, incorporate both quantitative and qualitative data, and provide visualizations of the complex interdependencies amongst risk factors and their potential impacts (Exhibit 2). For decision-makers in the AI risk management space, this unlocks the ability to better understand technical risks, such as your traditional cybersecurity vulnerabilities, alongside the more amorphous sociotechnical risks, such as (un)intended (mis)use. For instance, the methodology permits integration of expert knowledge from the business/mission function, which can be useful where empirical data is lacking, as well as updates when new information comes to light, which is crucial in such a fast-changing field as AI. For decision-makers, this flexibility enables what-if analyses to determine the impact of changes in information, assumptions, risk appetite, or all the above. The methodology also enables your most senior leaders to compare/contrast the risk postures of otherwise dissimilar use cases, which can facilitate those tough go or no-go decisions. By enhancing your understanding of the complex, multidisciplinary risk factors at play, the methodology ultimately helps all stakeholders uncover the right set of risk mitigation strategies and monitoring plans.

KPMG ranks #1 for quality AI advice and implementation in the US

KPMG RANKED 1ST FOR QUALITY in AI advice and implementation in the US 2024

Learn more >

visit.kpmg.us/SourceAIRanking

Source: *Perceptions of Consulting in the US in 2024*, https://www.sourceglobalresearch.com/

How KPMG can help

With our rich pedigree in assurance functions, KPMG has developed the Trusted AI Framework to help ensure fairness, transparency, explainability, accountability, data integrity, reliability, security, safety, privacy, and sustainability in AI adoption (Exhibit 3). Designed for use across all AI activities, including the definition of data quality standards, Trusted AI can help you strengthen AI governance, advance responsible AI innovation, and manage risks from the use of AI. Through this framework, our cross-functional professionals can help you quantify risk exposure for each AI use case and then leverage probabilistic simulations to model scenarios and their impacts. That transparency and flexibility will help you go above and beyond M-24-10's minimum practices for safety- or rights-impacting AI and make more confident, data-driven decisions when assessing and prioritizing AI use cases, ensuring not just the advancement of technology but also your mission and equitable outcomes.

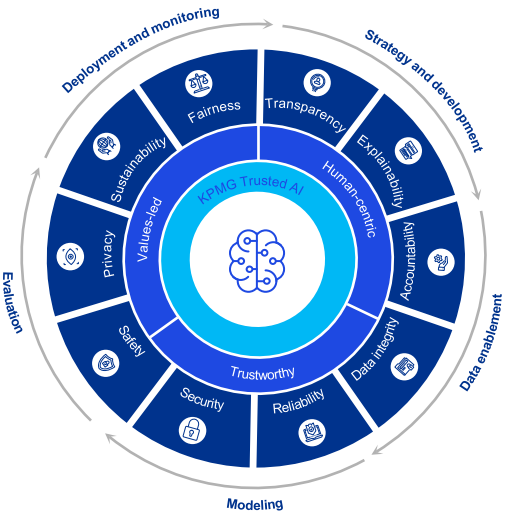

Exhibit 3: The KPMG Trusted AI Framework

Contact us

Viral Chawda
Principal, Advisory
Government Technology
KPMG US
E: vchawda@kpmg.com

Tim Comello
Partner, Advisory
Government Risk Services
KPMG US
E: tcomello@kpmg.com

Learn about us: kpmg.com

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