KPMG

Generative AI will help banks accelerate digital transformation

Generative AI is both a spur for banks to accelerate digitization and an opportunity for banks to leapfrog to more efficient operations and better customer experience. The good news: bank leaders are moving quickly to make the investments to get their first generative AI applications off the ground. Read about how they are doing it.

kpmg.com

Introduction

Digital-native fintechs have disrupted the banking sector by rapidly delivering scalable, customer-friendly services, in part by using cloud platforms. Banks, many of them still wrestling with inflexible and complex legacy systems, have raced to adapt, moving more and more work to the cloud to improve agility and responsiveness, enable innovation, and accelerate their digital transformation.

Now, just as banks are narrowing the technology gap with fintechs, the goal posts are moving. Advances in artificial intelligence—particularly new, easily deployed generative AI tools—are redefining what digital transformation means in banking operations.

To compete, many banks will need to accelerate their digital transformations by using AI technology more widely. "AI is part of the''how'—the digital business is the outcome," says Andy Brown, CEO of Sand Hill East, a strategic venture advisory group.

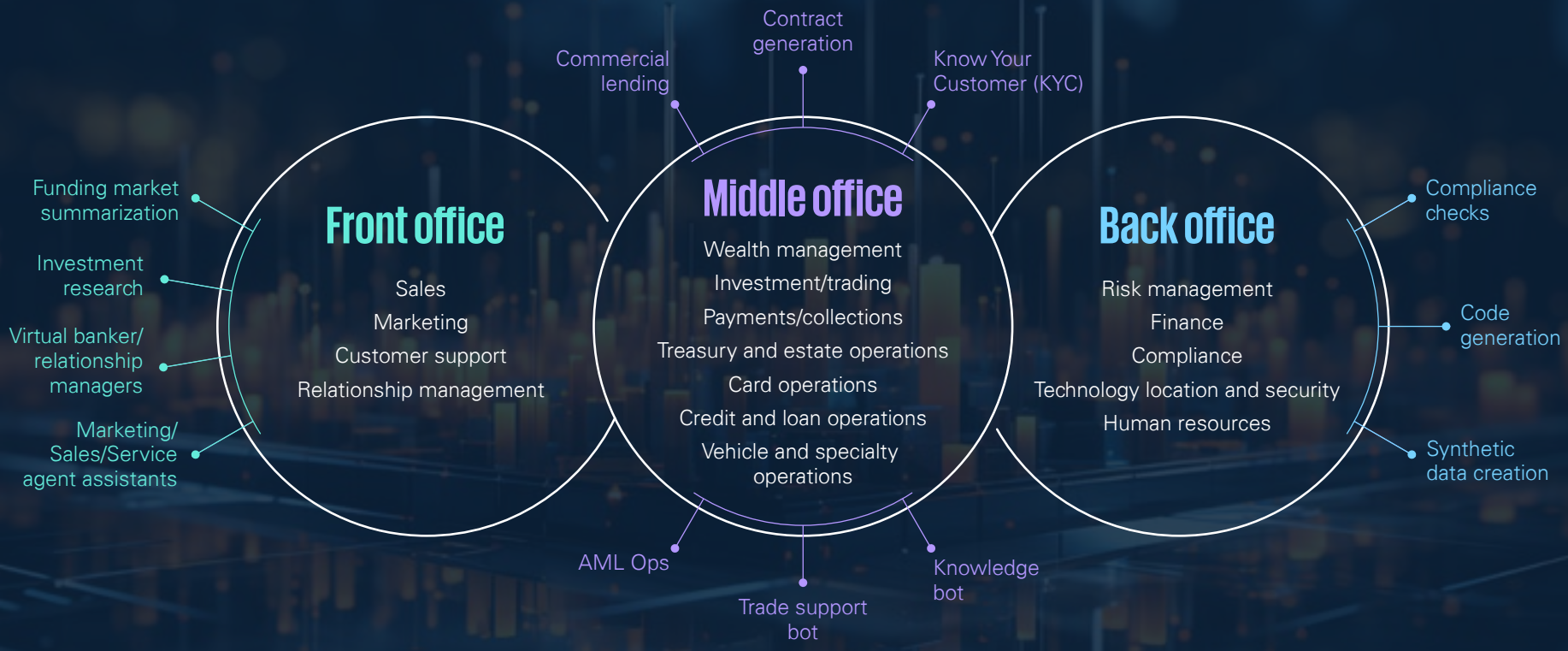
We believe that generative AI can help established financial institutions match the agility and innovation of nimbler fintech players in the next phase of AI-driven innovation. The good news is that bank technology leaders—and other C-suite executives—understand the growing potential of AI and are making the investments and forging the partnerships that will help them avoid being left behind.

In this paper, we explain how banks can build a repeatable process for qualifying and delivering new generative AI projects, factoring in the operational realities of legacy banking and the extra risks that a highly regulated industry must consider to protect its stakeholders. We'll share some examples of what banks are already working on and, through use-case descriptions, explore the art of the possible within the industry. With this, banks can begin to move more confidently toward an AI-integrated future and build a regular cadence for delivering business value to customers using generative AI.

Real-world progress in GenAI applications

Ever since companies like OpenAI, Microsoft, and Google have brought out GenAI products that make the power of AI comprehensible and palpable to everyone, banks have been aggressively investing in the technology, confident that fast implementation of high-impact generative AI-supported solutions can deliver improvement in scale, productivity, and innovation, and level the digital banking playing field. Some of the GenAI adoption opportunities across the front, middle and back office are illustrated below (Exhibit 1).

Exhibit 1. GenAI at work across front, middle, and back office



Banks have already made massive investments in most of these categories and have multiple pilots underway. Some examples of current pilots include:

A large financial institution is using document summarization/text generation to improve call center efficiency and to help agents in its consumer bank resolve customer inquiries quickly and accurately.

Many banks are using generative AI to improve software quality, write documentation, and ensure compliance with internal design guidelines. Banks are also using Microsoft Copilot to run code-change experiments.

Investment banks are experimenting with Gen AI in key front office processes such as helping prepare pitch books and summarizing earnings calls.

Several institutions are using generative AI for knowledge management activities. One bank is piloting a GenAI system to help new hires with policy interpretation during loan underwriting. Another is using GenAI to help new software engineers learn coding policies and standards for development on key internal platforms.

We also see potential for generative AI assistance in Know Your Customer (KYC), a particularly difficult and time-consuming process for banks. Generative AI can help with summarizing applicable KYC policies, news screening, data extraction and summarization, and flagging risks using all available data points. Capital markets functions, which often still rely on time-consuming processes, are also fertile ground for GenAI experimentation.

Building the foundations

Selecting the right large-language model (LLM) is an important early decision for generative AI programs. The LLM should be chosen to support your initial use cases but with flexibility for your future needs. While larger institutions might consider building their own LLM as an option, we expect most financial services companies to license or buy their first LLM.

KPMG has evaluated various foundational models to determine the best for each type of capability. Here, we summarize our findings and estimate the hardware needs for running different models:

Exhibit 2. Evaluating LLMs

Model type & examples	Capabilities
Representational Language Model • BERT	• Semantic search & clustering • Classification • Basic question answering and information extraction
Generative models	**Smaller size** • Summarization • Medium complexity question answering
	Medium size • Summarization • Complex question answering • General purpose text generation
	Larger size • Summarization • Complex question answering • General purpose text generation

Take aways from our testing

1. Flan-T5 is good for summarizing documents
2. Dolly V2 is good for question and answer
3. Llama2 is superior in document summarization and questions and answer
4. Finetune is required for domain-specific tasks. The vanilla public- trained base model will not perform well on domain specific tasks
5. OpenSource LLM can outperform general models after domain-specific training

Missing the mark with the first LLM in no way dooms the GenAI program. In the same way that most big banks' cloud journey leads them toward multi-cloud as an architectural standard, banks will eventually adopt multi-GAI as the norm. Organizations aren't limited to one platform or innovation path. After all, new LLMs are constantly in development, specialty function and industry models are emerging, and new functionality is changing the decision criteria of organizations' selection.

The steps in LLM development

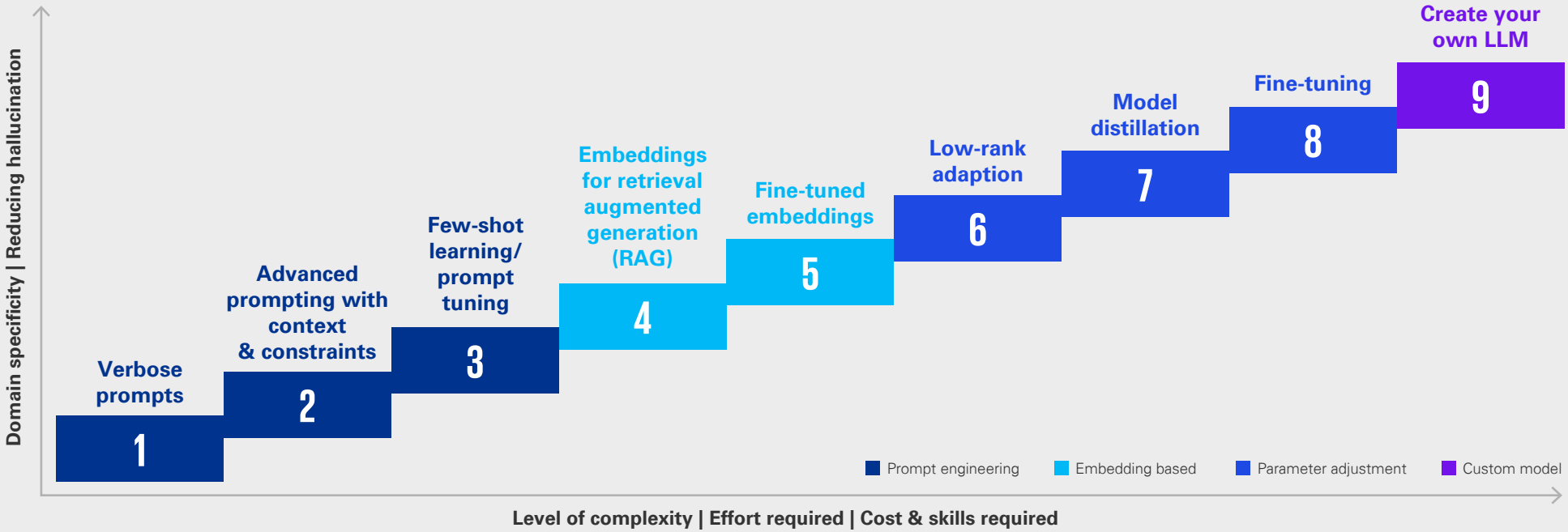
Alongside picking the right LLM, foundational steps include building knowledge about generative AI across the organization to spawn creativity in applying it. In our practice, we see top use cases from the front office, which include creating pitchbooks, credit underwriting on commercial loans, and cross-selling. Eventually, the goal everywhere is to adapt processes, establish best practices, and put the necessary controls in place to ensure the enterprise is capturing the right opportunities and is prepared for the peculiarities, risks, and management load that come with generative AI.

Generative AI promises to put powerful new tools in the hands of many employees. However, banks must first address issues of transparency, accountability, fairness, reliability, privacy, and security before generative AI is embedded in jobs and functions, particularly where partners or the public are involved.

Putting your data house in order is also an important prerequisite to a successful GenAI program. Data is the bottleneck to creating great content using GenAI. It's about context, quality, and utility. If there's data you haven't used and isn't of value, then don't bother training it into your model. Dump it, and use the right data for the right job, protecting it, of course, from unintended disclosure.

Once your data is in shape, there are various ways to integrate it into an LLM:

Exhibit 3. From general to specific LLM development



Build a Trusted AI framework

Organizations that are planning to use generative AI in their business must understand the importance and strategic imperative of ensuring that all AI applications, including those using generative AI, are trustworthy and responsible. Reputations are at stake, and without governance in place to ensure the technology is operating ethically and reliably, businesses risk damaging their relationships with customers, employees, partners, and the market.

The AI risk landscape calls for an agile, rapid, granular, and focused response grounded in a collective thought process from stakeholders across the organization. This is where an organization's risk functions can step in. Risk functions should be empowered to help ensure that AI development and deployment aligns with ethical and legal principles while being accountable and transparent to stakeholders. This is where creating and operationalizing your Trusted AI framework comes in. KPMG has developed a Trusted AI framework that stresses fairness, transparency, explainability, accountability, data integrity, reliability, security, safety, privacy, and sustainability.

Building institutional momentum for a generative AI program isn't easy. Getting from where you are today to embedding GenAI in your institutional culture is complicated, expensive, and time-consuming. But companies can ill-afford to sit on their hands as their competitors—and the technology itself—evolve around them. The best first step is to get started.

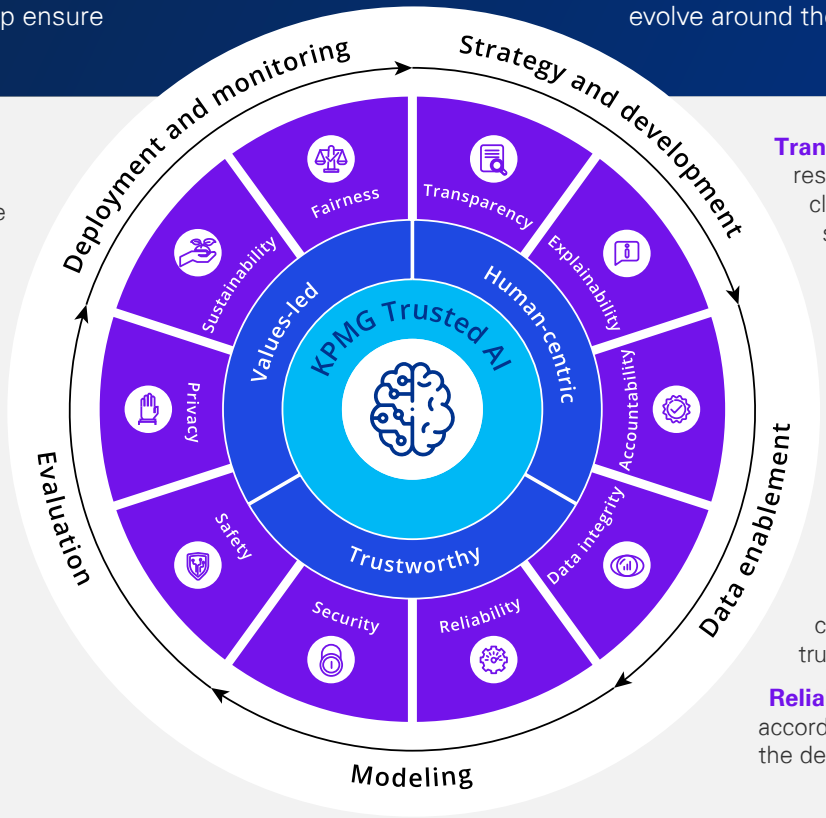


Exhibit 4. The Trusted AI framework

Fairness. AI solutions should be designed to reduce or eliminate bias against individuals, communities, or groups.

Sustainability. AI solutions should be designed to be energy efficient, reduce carbon emissions, and support a cleaner environment.

Privacy. AI solutions should be designed to comply with applicable privacy and data protection laws and regulations.

Safety. AI solutions should be designed and implemented to safeguard against harm to humans and/or property.

Security. Robust and resilient practices should be implemented to safeguard AI solutions against bad actors, misinformation, or adverse events.

Transparency. AI solutions should include responsible disclosure to provide stakeholders a clear understanding as to what is happening in the solution across the AI lifecycle.

Explainability. AI solutions should be developed and delivered in a way that answers the questions of how and why a conclusion was drawn from the solution.

Accountability. Human oversight and responsibility should be embedded across the AI lifecycle to manage risk and comply with applicable laws and regulations.

Data integrity. Data used in AI solutions should be acquired in compliance with applicable laws and regulations and assessed for accuracy, completeness, appropriateness, and quality to drive trusted decisions.

Reliability. AI solutions should consistently operate in accordance with their intended purpose and scope and at the desired level of precision.

Forging a path to production

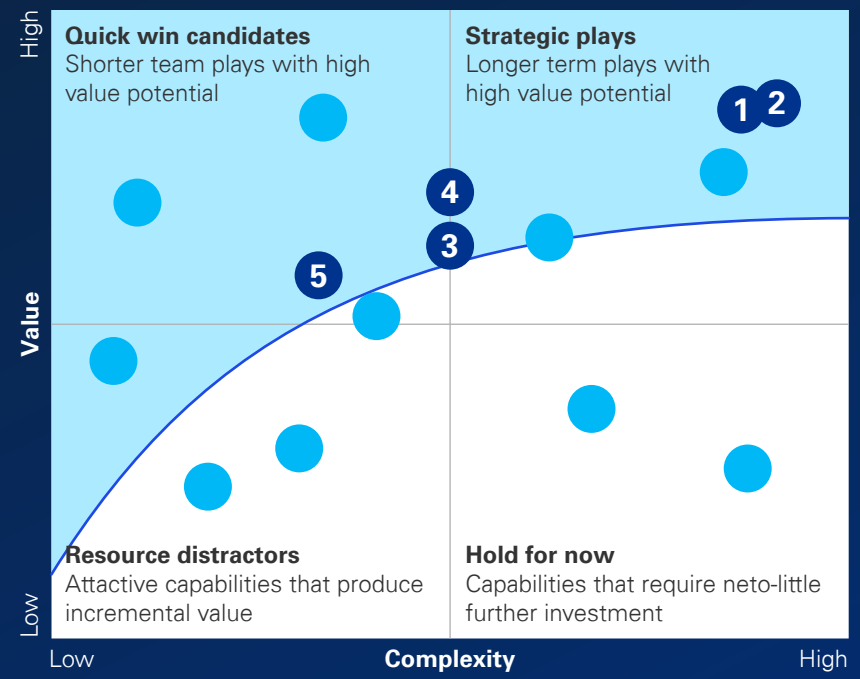
With the foundations in place, establishing a reliable, efficient path to production for use cases is where the rubber meets the road in delivering organizational value. The use case selection and prioritization considerations should include business value and value to customers and consider cost optimization and efficiency levers, revenue enhancement opportunities, improve customer satisfaction through better customer experience and consider speed to market. At the same time, companies need to factor in risks involved and controls needed around regulatory compliance, cyber security, privacy, intellectual property issues and other similar dimensions. Ideally, the re-usability and adjacency of use cases and the tech infrastructure and data needed across these use cases will result in further optimization of the organization's efforts to deliver increased value.

Prioritize the use cases.

Leadership should appoint a multi-disciplinary team, including heads of business, risk, engineering, financing, HR and operations, to function as a brain trust to review GenAI initiatives. The team should prioritize ideas that address a business need, solve a business problem, and have a clear payoff. These should include descriptions of the benefits, any required data, stakeholders, and key KPIs—the more information the better. Technical teams need to weigh in on questions of LLM model suitability, deployment, and data access, among other things. Then, it's up to an Innovation Team, including representation from legal, risk, and compliance, to assess the value and determine the project's feasibility. Addressing critical risk factors, such as privacy and intellectual property issues, is key.

With a clear, repeatable process in place, organizations can establish a reliable cadence for prioritizing gen AI projects, assessing their merit and feasibility, and ensuring all the necessary guardrails are in place for a smooth, successful project.

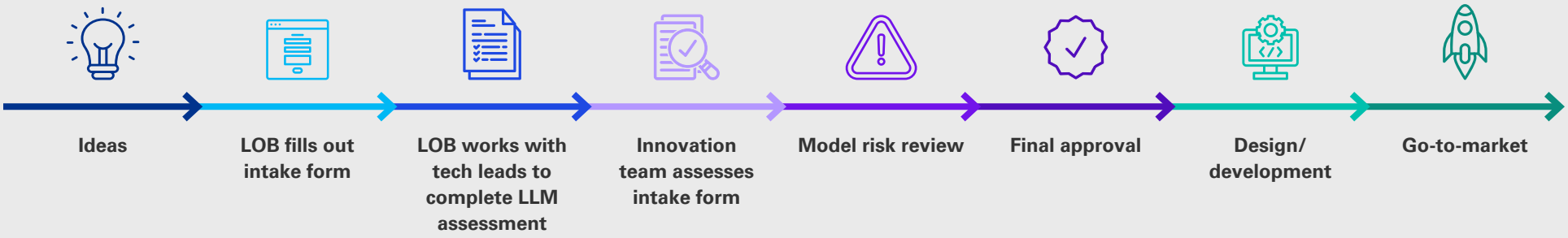
Exhibit 4. Use case prioritization dimensions



Dimensions

Complexity	
Data privacy	How sensitive is your data landscape?
Data quality	What is the quality of your data sources?
Data access	How easily accessible is the data you need?
Model usability	To what extent is there a model ready for use?
Value	
Economic value	What benefits are realized by this use case?
Business impact	How scalable is the impact to the business?
Risk	
Generative accuracy	How accurate do the results need to be?
Generative security	How secure does the information need to be?

Exhibit 5. A structured approach to approving generative AI applications



Managing the risks and accelerating the digital transformation.

The highly regulated banking industry is full of risks, some common to all companies and some particular to banks. Issues such as poor model performance, sub-par content quality, and results bias can expose banks to regulatory and legal risks (e.g., a loan underwriting program using AI that has a hidden gender bias). When implementing AI solutions, banks should have clear processes for ensuring data integrity and security, fine-tuning models, evaluating content, monitoring results, and vetting AI-generated product for consumption. Nothing should go to customers or regulators without human review.

With the right controls in place, organizations can thoughtfully manage these, and other risks, while pursuing an action plan for accelerating organizational and business innovation using this transformational new technology. A cross-disciplinary team, including Risk, Finance, Technology, and HR, can help accelerate the pace of adoption by establishing an innovation framework that prioritizes the top revenue and operational use cases while being mindful of the human, reputational, and business sensitivities that generative AI introduces. A clear ethical framework is vital to any AI program.

AI in all its forms can be a game changer for banks. It offers the opportunity to match or beat nimble upstarts in the race to deliver better customer experience, tap new sources of growth, and operate more efficiently. But there is no time to lose. The race has begun.

How KPMG can help

KPMG is at the forefront of digital transformation in Banking & Capital Markets, with the integration of AI and GenAI being a crucial aspect of an organization's digital journey. Our comprehensive and reliable approach incorporates both top-down and bottom-up strategies, using predictive AI and GenAI to support our clients in various areas, including:

Enterprise value stream redesign, enabling them to rethink and reshape end-to-end processes and customer journeys

Operating model design, encompassing business models and processes, technology and infrastructure, talent strategy, and digital ecosystems

Cloud transformation, technology infrastructure assessments, and execution

Enterprise data strategy and application modernization

Trusted AI strategy and implementation

Establishment or transformation of an AI Center of Excellence (CoE) or Control Tower, including use case selection and execution

For every AI and GenAI project, KPMG brings a wealth of industry knowledge, cutting-edge technical expertise, innovative solutions, and a strong partner ecosystem to help clients harness the power of AI and GenAI in a trusted manner.

Authors

Binoy Palakkal

Principal, Advisory
Customer and Operations, Financial Services

Binoy helps clients achieve operational efficiency and cost savings by focusing on automation opportunities. He has over 20 years of experience advising financial services firms, with deep knowledge about how to transform largely manual and inefficient processes to a more automated and customer centric operation.

Mark Shank

Principal, Advisory

Mark leads Banking Technology Modernization at KPMG, responsible for architecting KPMG's response to technology change in the industry, including Generative AI. Mark is leading alliance relationships around Gen AI as well as asset development and client PoCs.

He has 20 years of experience in modern architectures, applications and processes to ensure that the technology stack defined and integrations outlined will enable the success of the target operating model for the business, drive value and cost take out and lessen risk.

Kalpana Ramakrishnan

Principal, Advisory,
Financial Services

Kalpana leads KPMG's US Financial Services Advisory line of business and serves as the Global Lead Partner on one of the firm's largest asset management clients. She has over 25 years of broad based advisory experience in aligning information technology strategy with business strategy, and has assisted several large clients in the areas of large digital, technology and business transformations. Kalpana served on the KPMG US and KPMG Americas Board from 2017 to 2021.

Contact us

Binoy Palakkal
Principal, Advisory, Customer and Operations, Financial Services
857-225-0975
bpalakkal@kpmg.com

Mark Shank
Principal, Advisory
301-514-7018
mshank@kpmg.com

Kalpana Ramakrishnan
Principal, Advisory, Financial Services
949-885-5590
kramakrishnan@kpmg.com

Related thought leadership:

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

Please visit us: **kpmg.com**

DASD-2024-14504