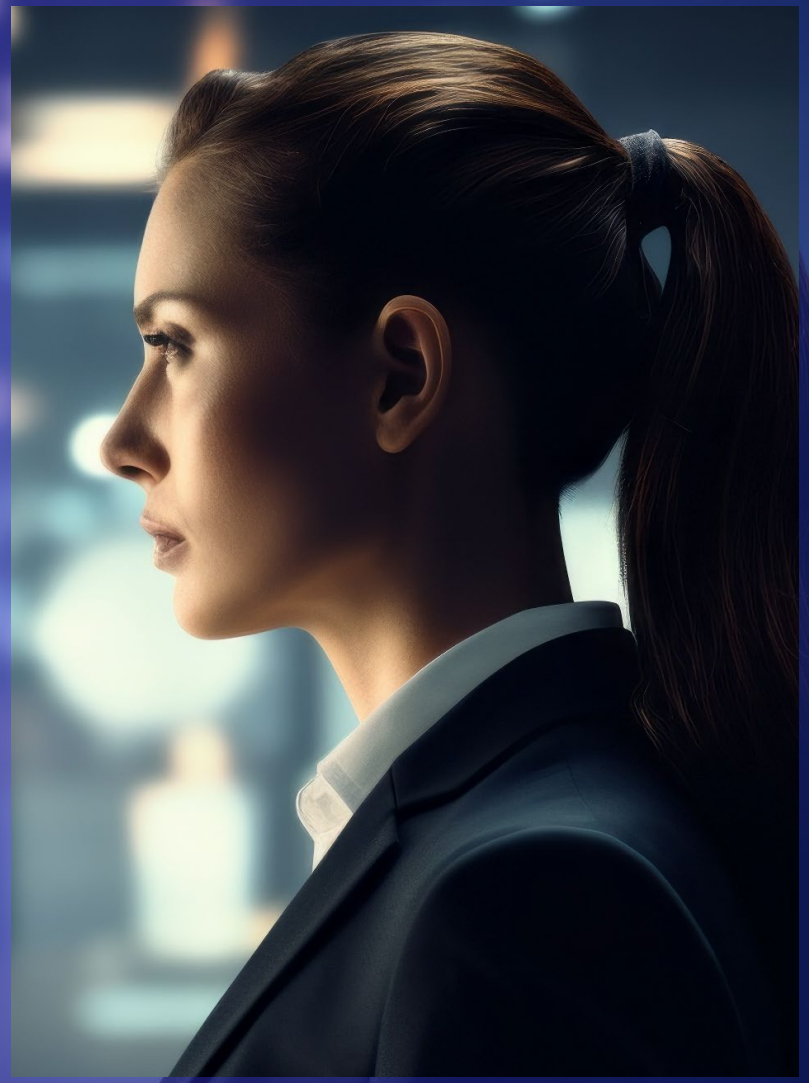
You can with AI – But what should your priorities be?



24 April 2025

KPMG in Thailand

With you Today

Itthipat Limmaneerak
Tech- Data and AI
KPMG in Thailand

Silvester Liu
Digital Transformation Leader and Head of Lighthouse, ASPAC Connected Enterprise

Sheldon Goh
Industry Director
Worldwide Financial Services

Thanayut Sriariyawat
Business Transformation – AI Solutions Architect
KPMG in Thailand

Agenda

What is the value of generative AI adoption and where can I find it?

Making generative AI happen in practice; experience from China market

CoachIQ: Telesales assistant agent and its risk assessment

What's next the shift to Agentic AI by Microsoft

01

What is the value of GenAI adoption and where can I find it?

Itthipat Limmaneerak

The revolutionary potential of AI for banks

High expectations

80% believe that banks that **embrace AI** will develop a **competitive edge** over those who do not.

62% expect a moderate to very high ROI from AI investments.

But the pressure is on to prove ROI

70% of shareholders expect to see immediate ROI on AI investment.

AI spending will likely increase significantly

70% plan to increase the percentage of global budget spent on AI,

of which **62%** say up to 20%

and **38%** say 20%+

AI goals are clear

68% are seeking to reduce costs.

42% aim to enhance the customer experience.

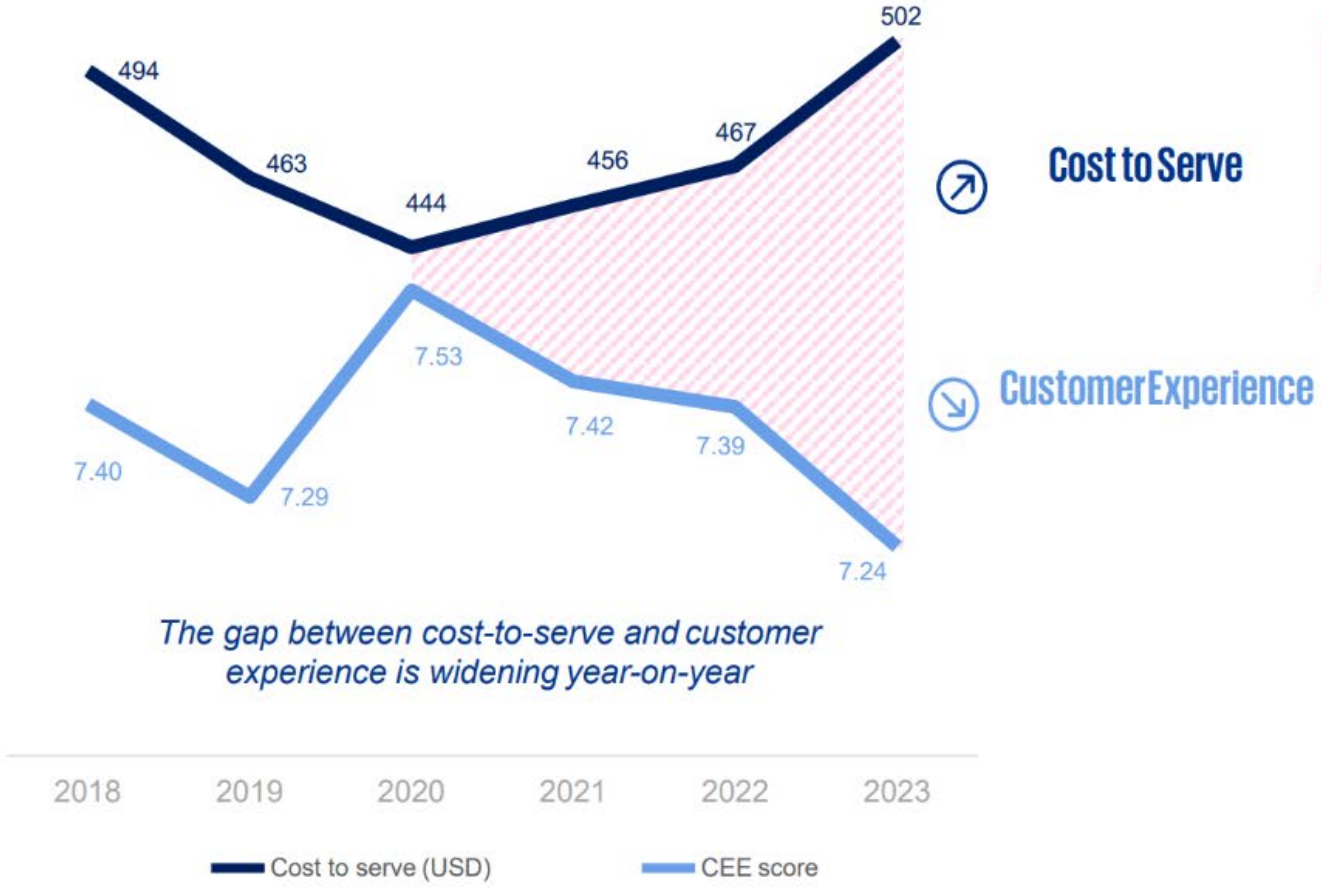
The initial benefits

66% have achieved cost savings.

Only **26%** have experienced revenue growth,

and only **13%** have experience a high revenue contribution from AI.

Are your investments creating the value you envision?



Sources:
KPMG Cost & Value Report: 2023, KPMG Customer Experience Reports (2018-2023)

Golden Rules to Customer Experience

	Net Promoter Score
Empathize – Tune in to my particular situation	10
Personalize – Recognize me as an individual	8-9
Simplify, automate and look ahead – Value my time	5-7
Set expectations clearly – Let me know in advance what you need from me	4-5
Own the resolution – Give me a warm and sincere apology when things go wrong	3-4
Act with integrity – Standing for something beyond profit	0-3

What is the value?

17 million companies globally were assessed.

After looking in depth at **7,074** companies

employing **72 million** people

and pressure-testing results with **500 clients**

GenAI opportunity across all sectors equates to…

19-23% of salary cost and

4-18% of EBITDA

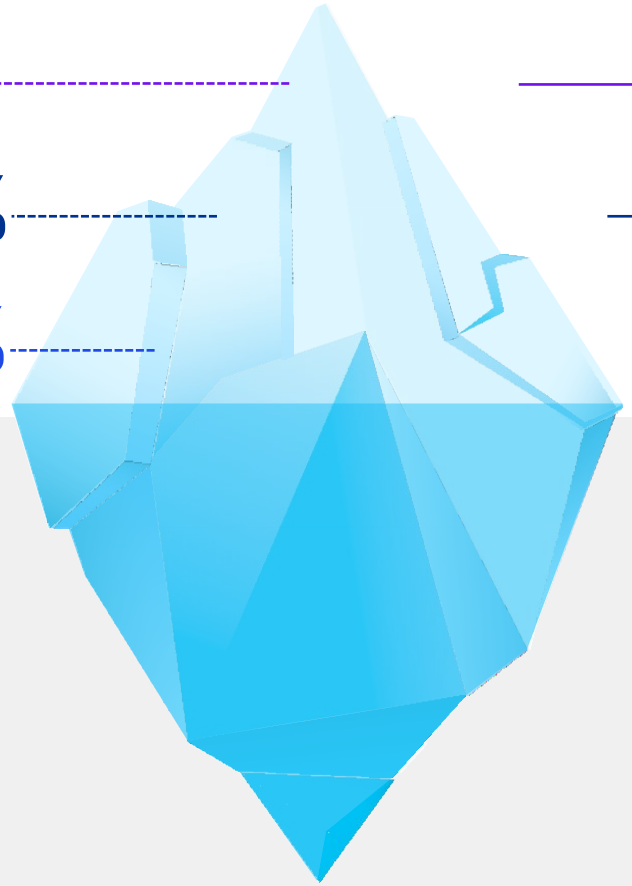
For Financial Services…

22.5% of salary cost and

10.8% of EBITDA

A conservative approach to opportunity sizing

Financial services



GenAI augmentation opportunity – cross sector

9%	Low complexity	Simple tasks that can be effectively augmented using out-of-the-box GenAI tool, e.g. Copilot, SO LLMs
39%	Medium complexity	Tasks that require more integration and customization, e.g. requiring more data piping and enrichment
52%	High complexity	More integrated, sophisticated and tailored solutions, e.g. agents, comprehensive process re-design

Other AI/digital automation opportunity not captured in this assessment:

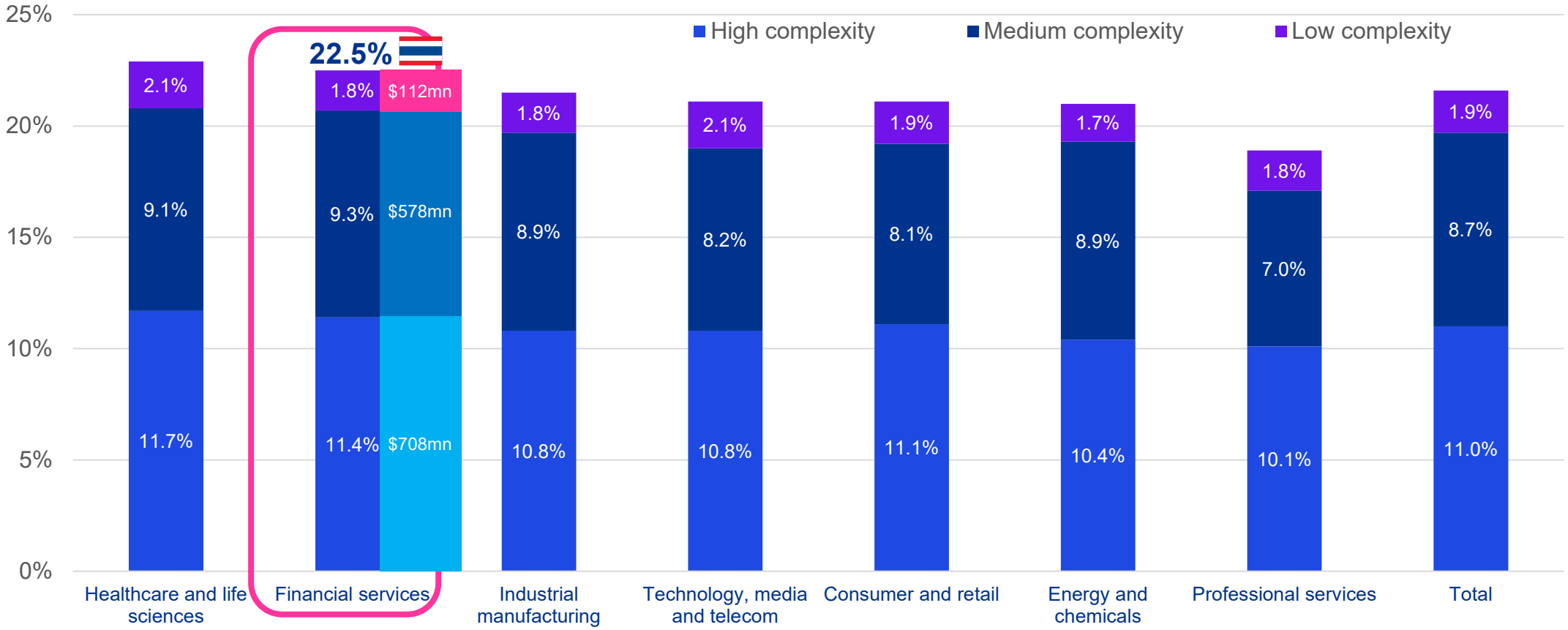
- tasks best automated by AI/digital technologies other than GenAI, e.g. RPA
- value that can be captured without GenAI but is more likely to be accelerated as part of a GenAI-triggered transformation

Estimated GenAI impact by role: Computer system analyst

Computer system analyst

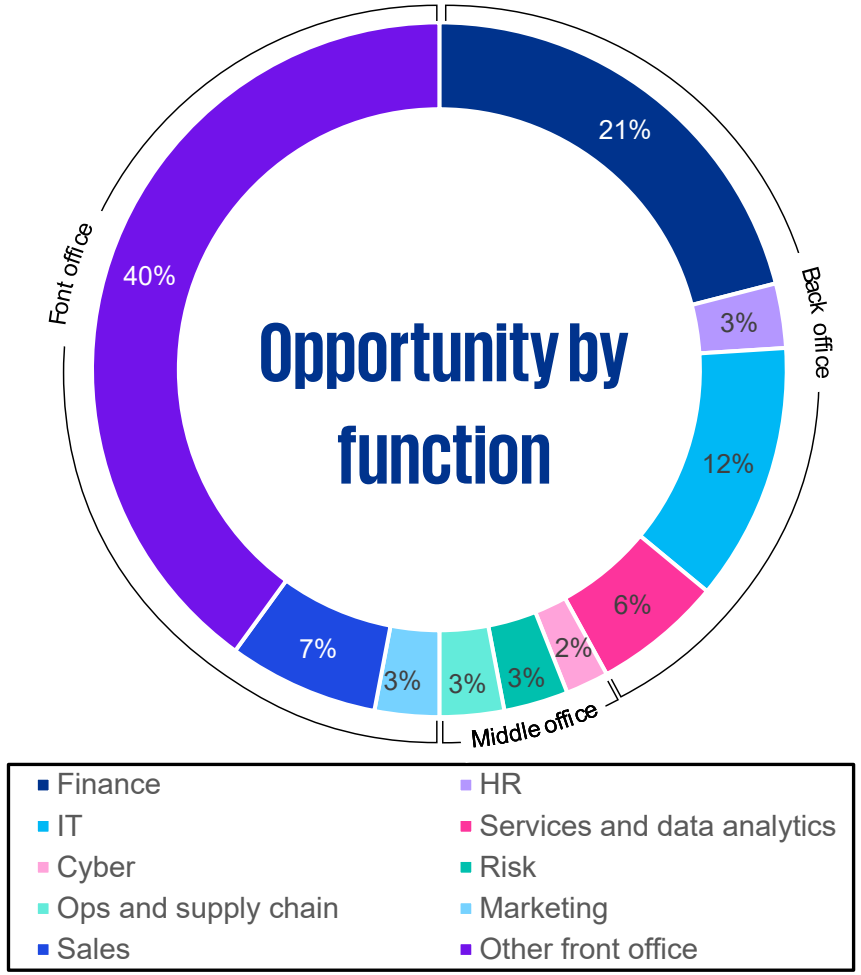
Work tasks	Activities	Time spent prior	GenAI saving	= Time saved
Code optimization	• Review and analyze computer printouts and performance indicators to locate/correct code	4%	25%	1%
Financial analysis	• Prepare cost-benefit and ROI analyses to aid system implementation decisions	2%	50%	1%
Systems design	• Analyze information processing or computation needs and plan and design computer systems	7%	43%	3%
Test and monitor	• Test, maintain and monitor computer programs and systems, including installation	7%	29%	2%
Team coordination	• Provide staff and users with assistance solving computer-related problems • Train staff and users to work with computer systems and programs • Supervise computer programmers, other systems analysts or lead particular systems projects • Coordinate and link computer systems to increase compatibility for information sharing	24%	33%	8%
Documentation	• Specify inputs accessed by the system and plan the distribution and use of the results • Interview/survey workers, observe job performance or determine information processing • Consult with management to ensure agreement on system principles • Confirm information processing or computation needs with clients	13%	31%	4%
Systems improvements	• Analyze and solve business problems (e.g. integrated production and inventory control) • Expand or modify system to serve new purposes or improve workflow • Recommend new equipment or software packages	14%	36%	5%
Other		29%	0%	0%
		100%		24%

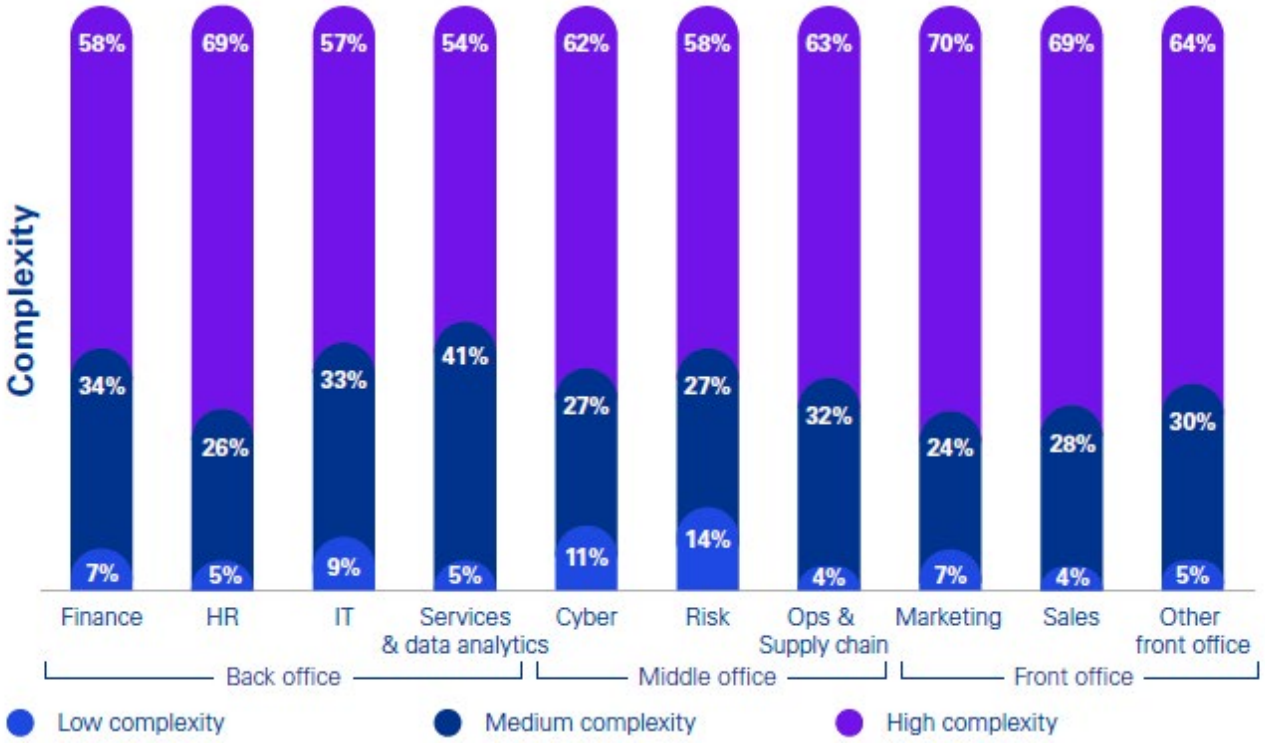
The GenAI opportunity equates to 19-23% of salary cost



Source: Intelligent banking: A blueprint for creating value through AI-driven transformation, KPMG International, 2025
Estimate based on salary cost impact applied to 2024 published operating profit of 10 Thai listed banks

A closer look at the GenAI opportunity for the banking sector





Source: Intelligent banking: A blueprint for creating value through AI-driven transformation, KPMG International, 2025

Where are we today?

Current state

The foundational infrastructure is still in development

Only **25%** have
wide cloud or hy
platforms strateg
supporting data
services.

Leaders are concerned about ROI and costs

It is myth to conclude that AI solutions will always deliver a positive ROI – selecting the right approach is critical.

Leaders and employees still coming to grips with AI's potential

de ethics and
training, only
more advanced
provide deeper
ding of potential.

Leaders are concerned about control

58% express nervousness about the influence of AI tech providers on their operations.

AI is transforming business functions in pockets

Integrated value stream driven approaches are not as dominant as they need to be.

- **Result-oriented technologies – select the right model for the job**
- **Data – hygiene, storage and transfer**
- **Cost models – free vs PPT vs subs**
- **Future proofing – expandability**
- **Resource allocation – containerization, serverless architecture**

Source: Intelligent banking: A blueprint for creating value through AI-driven transformation, KPMG International, 2025

Is the biggest challenge really a human one?

The fundamental steps towards success...

(1) Embrace the power of AI

Emerging technologies are the most powerful change agents shaping the world of work. We need to embrace them.

66% expect an increase in their productivity in the next three years.

But **31%** believe emerging technologies have impacted their work-life balance.

(2) Shape the workforce

Traditional workforce planning is no longer sufficient given the need for faster and more dynamic decision-making.

36% think their organization doesn't know their future workforce needs.

60% say their employer is not using emerging technology to match skills to job opportunities.

(3) Learn in the flow

The pace and dynamism of change affecting organizations is rapidly evolving the skills and capabilities needed, driving a requirement to learn in the flow.

78% are optimistic that their skills will be suitable for future roles.

72% agree that continuous upskilling will be crucial to stay relevant in their field.

(4) Lead from the middle

Middle managers are central in translating strategy into action – they need to be equipped with the skills and attitudes to lead in the future of work.

77% say support from their manager is more important than ever.

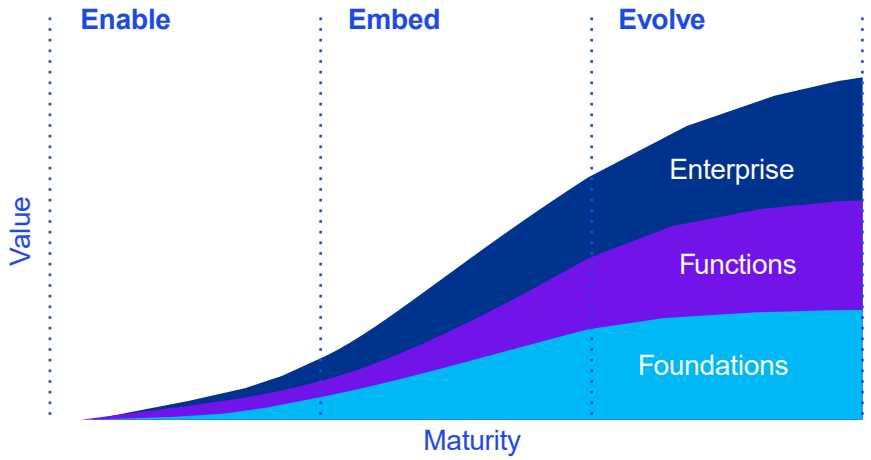
80% of managers are accidental, with no proper leadership training.

Key elements of the human experience at work

A non-human-centered AI transformation		A human-centered AI transformation
• **Job theft** – "AI will replace me" • **AI distrust** – "I don't trust the outputs of AI" • **Agency loss** – "I'm being surveilled by AI"	**Safety & Security** *"I feel secure in my work, certain of what's coming next & my place in it*	• **Job security** – "I understand how AI will complement, not replace me." • **AI trust** – "I understand how AI works and can trust its outputs" • **Agency** – "AI is there to help, not monitor me"
• **Task theft** – "I am left with tasks that don't interest or challenge me" • **Autonomy loss** – "AI makes decisions without me"	**Job Design** *"I feel accomplished, challenged & like I have autonomy"*	• **Task augmentation** – "I bring uniquely human skills alongside AI"; "I can now focus on more meaningful work" • **Autonomy** – "I am in control of the AI I work with"
• **Job status erosion** – "I don't matter anymore"; "my efforts aren't valued" • **Professional identify erosion** – "I don't know what value I bring"	**Reward & Recognition** *"I feel valued, invested in & have a clear sense of professional identify"*	• **Value recognition** – "My organization cares about me and invests in my development alongside AI" • **Evolved professional identify** – "I know how I contribute and what unique value I bring alongside AI"
• **Purpose disconnection** – "I no longer connect to how my organization is shaping / pursuing its purpose" • **Culture alienation** – "I feel like I don't 'fit' here anymore" • **Distrust in leadership** – "I don't trust my leadership to make the 'right' decisions with AI"	**Culture & Purpose** *"I feel connected to our purpose, understood & that I belong"*	• **Purpose connection** – "I connect to my organization's purpose and AI's role in this" • **Culture evolution** – "We are a team who embraces innovation" • **Trust in leadership** – "I know my leaders can and will ensure we use AI responsibly"

The journey to becoming an intelligent financial institution

Focus on 3 layers of the organization



Enable people →	Embed AI in work →	Evolve the enterprise
• Define highest value use cases • Model value opportunities • Deploy in operating model • Initiate early AI guardrails • Invest in AI literacy • Jumpstart initial program	• Align strategy and OKRs with AI • Define value and investments • Redesign operating model • Strengthen trust in AI • Reshape the workforce • Orchestrate enterprise change	• Define an ecosystem strategy • Model value of the ecosystem • Redesign business model • Institute always-on AI trust platforms • Extend to partner workforce • Orchestrate ecosystem change
• Implement functional use cases • Test and learn, and refine • Augment people with AI skills • Treat AI as "copilot"/"assistant" • Focus on rapid learning • Build and deploy in sprints	• Embed AI in value streams • Embed AI in process workflows • Embed AI agents as they mature • Use AI agents as they mature • Use AI to transform products and experiences • Undertake agile change	• Use AI to power ecosystems • Fuel inter-organization workflows • Deploy agents across ecosystems • Evolve new experience possibilities • Focus on end-to-end value • Continuous, agile changes
• Select AI strategic alliances • Implement AI applications • Configure and tailor • Introduce simple models first • Access AI through the cloud	• Build AI development "factory" • Select and train domain models • Curate enterprise-wide data • Invest in AI infrastructure • Invest in increased cybersecurity	• Deploy AI across ecosystem • Compete using domain models • Compete using ecosystem data • Combine cloud with AI-optimized chips • Consider AI with quantum

Then maybe you can be like one of the most AI-driven organizations in the world… a financial services conglomerate…

AI-enabled, customer servicing experience

93%
policies UW in seconds

10 seconds
fastest claim settled

CNY9.1 billion
fraudulent claims blocked

CNY28 billion
AI-driven boosted sales

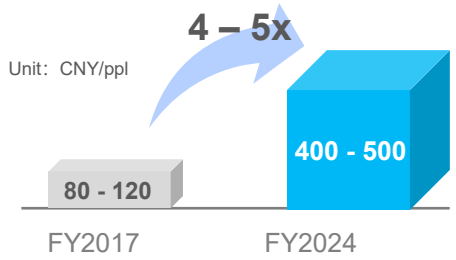
2.2 billion
AI customer service volume

CNY10 billion
lowered costs - CTI 27.9%

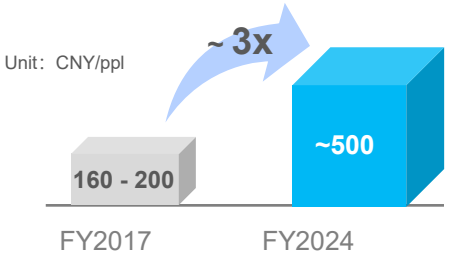
Source: Ping An Annual Report 2023, pingan.com investor relations

The saturated market drives Chinese banks to strengthen its digital capabilities in sales and customer-centric experience and product offerings

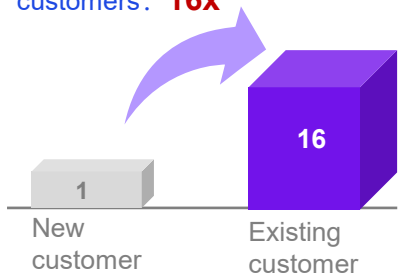
Average cost of customer acquisition for credit card



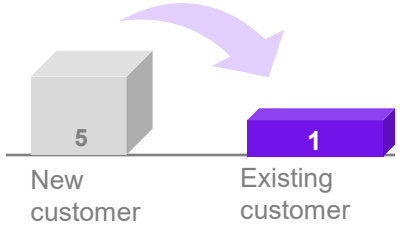
Cost of customer acquisition of a leading wealth management platform



Profit contribution of existing customers outperforms new customers: **16x**



5x cost saving of customer retention compared to client acquisition



Customer-centric experience optimization

- With deepened customers' awareness of financial services and an increase of their exclusive experience demands, clients obtain a richer and diverse range of services. **Retail banking keep investing in building financial and non-financial eco-system to improve customer satisfaction.**

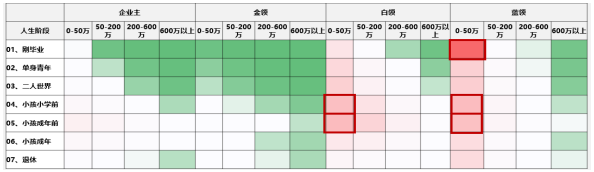
Integrated financial product and service offering

- Banks transformed from traditional perspective of financial product selling to integrated financial service offering through enrichment of exclusive service experiences to meet the comprehensive and differentiated needs of high-end customers.

Data-driven customer segmentation and precision marketing

- Most banks formed a comprehensive customer segmentation framework based on the NSM (North Star Matrices) of AUM and MAU and configured differentiated marketing strategies.

PING AN
Expertise Creates Value



Digital intelligent retail banking 3.0

Formatted 112 customer segmentations based on 3 key dimensions :

- Occupation: blue collar, white collar, business owner, etc.;
- Wallet: client's market investable AUM (model prediction);
- Life stage: new graduate, young professional, young parents, mature couple, retired, silver-haired

Customer segmentation evolved from strategic to tactic among leading retail banks, in search of precision marketing and personalized services

	Strategic segmentation	Tactic segmentation
1 Sources	➢ Client's LTV (RAROC, EVA, AUM) , demographic labels	➢ Diversified customer label: client lifecycle, age, gender, risk sensitivity, product preference, other non-financial attributes (i.e. consumption behaviors, societal attributes, channel preferences), etc.
2 Key customer segmentations	➢ Ultra high net-worth clients, family office, trust fund beneficiary, private banking, priority private	➢ On-boarding new customer, price-sensitive customer, wealth-management product preference customer, high risk resilience customer, market volatility averse customer
3 Marketing target	➢ Used for medium to long term, for precisely identification of the client base composition and targeting, steering the operation mode and resources allocation.	➢ Used for short-term, aiming to provide more detailed and precise segmentation for specific business target, i.e. product cross-sale and up-sale, service bundle design, MAU retention, etc.
4 Marketing strategy	➢ Key segmentation + asset portfolio design + customer benefits + channel strategy	➢ Data driven precision marketing and customer marketing list generation ➢ CRM or CMM supporting ➢ Marketing automation
5 Key actions & Digital requirements	➢ Based on the internal data, continuously optimize the data-driven identification of key customer segmentations and provide comprehensive financial and non-financial services accordingly.	➢ Formation of data-driven marketing automation system, composed by various types of marketing scenarios, including event-based, target name list based, auto triggered, etc.

Disruptive change on technology diffusion led by open-source ecosystem such as DeepSeek-R1

In recent years, AI researchers have developed pre-trained language models (PLMs) based on Transformer architectures trained on massive corpora. These models have demonstrated exceptional performance in natural language processing (NLP) tasks. Studies reveal that **as model parameters scale up, novel capabilities such as in-context learning emerge.** To distinguish language models of varying parameter scales, the concept of large language models (LLMs) was introduced.

Trends in large language model development

Currently, LLM development is characterized by three key trends:

1. **The tension between open-source sharing and commercial monopoly**
   - Open-source models like DeepSeek-R1 are driving technological democratization.
2. **The symbiosis between computing efficiency and model capabilities**
   - Breakthroughs in multimodal learning and reinforcement learning (RL) are reshaping human-AI collaboration.
3. **The balance between data quality and privacy-security concerns**
   - Structural leaps in foundational computing resources
   - Continuous refinement of data corpus ecosystems
   - Expanding and deepening real-world applications

Looking ahead, Sino-US competition will center on AI governance influence, **with the open-source ecosystem (exemplified by DeepSeek-R1) accelerating technology diffusion, redefining national innovation systems and industrial dynamics.**

- Model architecture evolution
- Model capability evolution
- Basic resource development
- Application scenario development

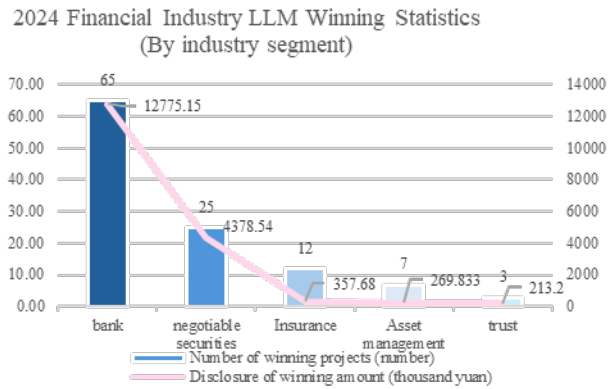
DeepSeek's strategic positioning aligns with OpenAI's established framework, that is, base models still follow the GPT-style pretraining paradigm, while inference-optimized models adhere to the RL-centric post-training route.

DeepSeek's innovation lies in engineering breakthroughs along these two tracks and cost-efficiency advancements in scaling and deployment.

Overview of LLM/ GenAI deployment in China's financial services industry

With the continuous optimization and upgrading of computing resources and the vigorous development of AI technology, AI technologies that centered on LLMs are emerging, evolving, and iterating in the financial industry at an unprecedented speed. **Strategic-driven and value-oriented principles have become the dual engines for the financial sector to deploy emerging AI scenarios, demonstrating a new trend of deep integration between finance and AI.**

The application of large models in the financial industry has formed a **tiered development pattern** of "banking sector dominance, securities and insurance sectors catching-up, and trust and asset management sectors exploring." Both the number and value of public procurement projects exhibit a **concentration trend toward leading institutions.**

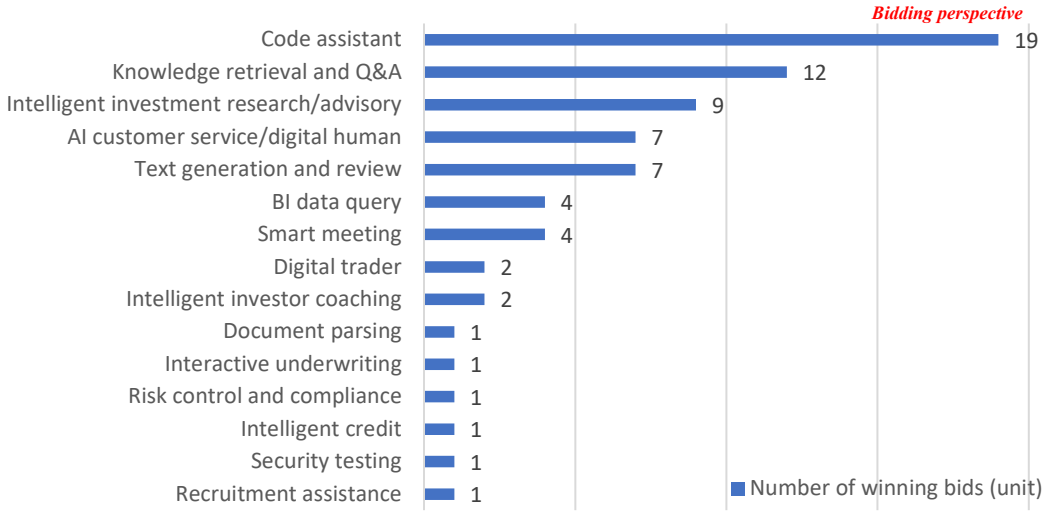


In 2024, procurement projects related to large models in the financial industry were predominantly driven by banking, securities, and insurance sectors, accounting for approximately 91% of total demand.
Banking sector, in particular, holds an absolute dominant position with 58% of awarded projects and a total disclosed contract value of CNY127.75 million, *demonstrating significant head effects.*

From financial sector perspective:

- **Banking sector:** Being the financial sector with the highest adoption rate of LLMs, the penetration of LLMs has rapidly expanded from state-owned banks and joint-stock banks to leading regional banks. State-owned banks, leveraging strong capital and technological resources, aims for full-stack control of the technology. Joint-stock banks exhibit more flexible and diversified implementation approaches. Although regional banks started later, they have taken solid steps in small-scale pilot applications for specific use cases.
- **Securities and insurance sectors:** Both sectors display similar trends in LLM adoption, with leading institutions taking the initiative and employing diverse implementation models.
- **Asset management and trust sectors:** Their approach resembles that of regional banks, primarily focusing on introducing tool-side capabilities for specific business scenarios rather than establishing systematic large model frameworks.

Distribution of LLM application scenarios in the financial industry (2024)



From application scenario perspective:

- AI applications pursue self-controlled infrastructure and rapid business value realization, while mid-layer capabilities rely on ecosystem partnerships. This further reflects the **financial industry LLM construction/procurement strategy: "short-term scenario effectiveness + long-term computing power reserves".**
- High-investment scenarios include code assistants, knowledge Q&A, intelligent investment research, AI customer service and text generation/review, indicating a **technological penetration path evolving "from internal to external, from efficiency tools to decision support".**
- Recently, **middle-office management (e.g. risk management) is also a key focus area**, while mostly developed in-house or through vendor partnerships, thus not directly reflected in bidding data.

Panoramic view of the LLM applications in China's banking sector

General tools

Office management

- Document/picture parsing
- Search and Q&A
- Intelligent generation of knowledge base
- Audio/video text image generation
- Intelligent writing
- Content review
- Mail composition
- Real time translation
- Meeting assistant

Data analysis

- Interactive reports
- Spectrum generation reasoning
- Data modeling decision

Front-office business engagement

Retail banking

Channel operation

- Intelligent outbound call/customer service
- Channel tagging/tracking reporting
- Client sentiment/behavior analysis
- Intelligent application processing
- Personalized client engagement guidance
- Intelligent recommendation
- Intelligent consumer protection processing
- Intelligent claim settlement

Client retention

- Client segmentation
- Client portrait
- Client relationship mapping
- Client rating
- Equity management
- Loss management

Marketing management

- Personalized marketing content generation
- Marketing strategy design
- Marketing opportunity delegation

Product service

- Emerging product trend insight
- Product customization
- Product component matching
- Differentiated pricing strategy
- Product analysis and evaluation

Corporate banking

Channel operation

- Intelligent outbound call/customer service
- Intelligent recommendation
- Credit management
- Intelligent due-diligence
- Client opportunity capture
- Scenario-based credit service
- Intelligent claim settlement

Client retention

- Corporate client portrait
- Corporate entity relationship mapping
- Industry research analysis
- Client rating
- Segmented management
- Loss management

Marketing management

- Customized marketing content generation
- Marketing strategy design
- Marketing opportunity delegation

Product service

- Emerging product trend insight
- Product customization
- Product component matching
- Differentiated pricing strategy
- Product analysis and evaluation

Treasury and money markets

Investment banking and asset management

- Intelligent investment advisor
- Intelligent investor coaching
- Fund portraiting
- Intelligent investment research
- Intelligent due-diligence
- Asset appraisal/allocation
- Self-service transaction Q&A
- Business opportunity/success rate prediction
- Aggregated information capture
- Issuance pricing forecast

Middle-office management

Risk control compliance

- Risk identification
- Risk feature generation
- Risk assessment
- Risk alert and resolve
- Public opinion monitoring
- Anti-money laundering
- Anti-fraud
- Intelligent audit
- Risk report generation
- Risk & compliance knowledge base

Operation management

- Self prepared process
- Operational priority assessment
- Operation index generation and verification
- Operation efficiency evaluation
- Theme report generation

Back-office operations

Financial management

- Financial data processing and Report generation
- Prediction and analysis of key indicators
- Financial monitoring and review

Development, operation and maintenance management

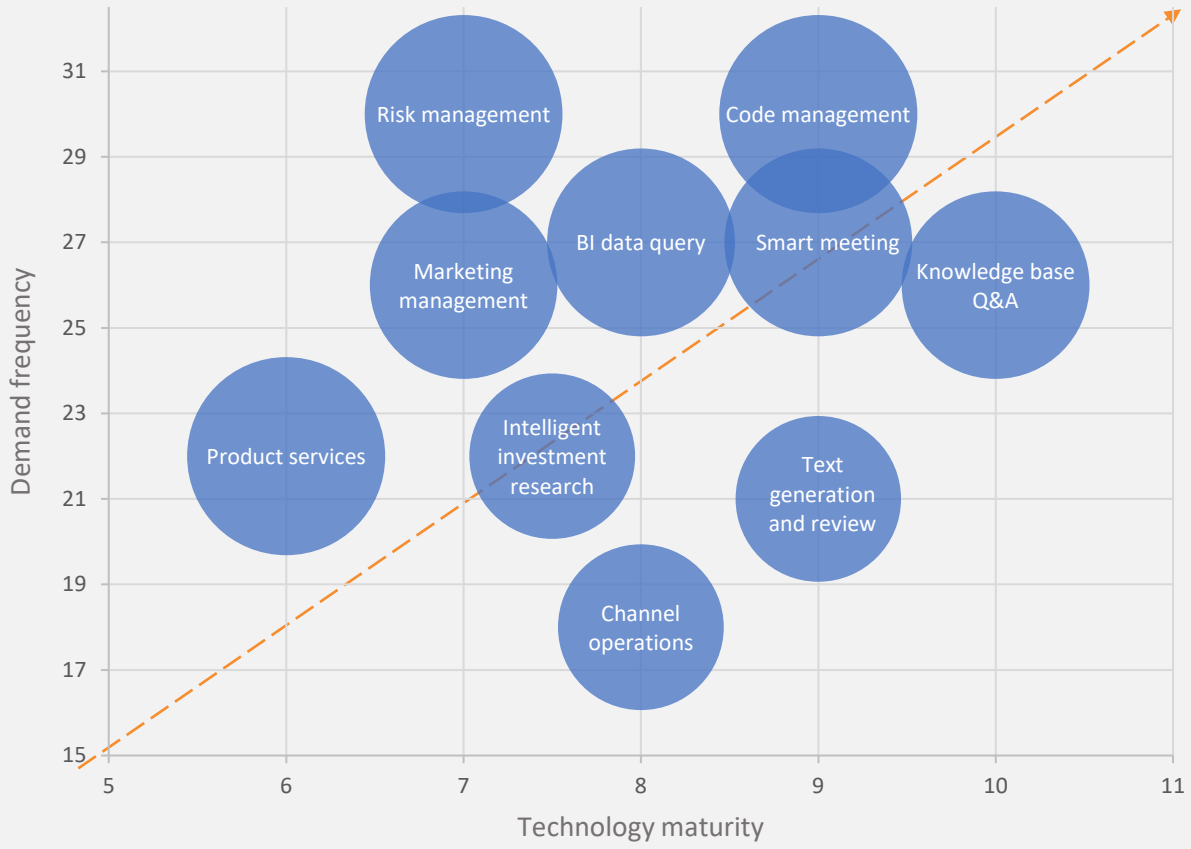
- Code generation and supplement
- Test and function level bug repair
- Operation and maintenance monitoring, analysis and alert
- Equipment control and report generation
- Code comment/translation
- Code review/refactoring
- Test case generation
- Network security analysis
- Intelligent work ticket

Personnel management
Investment supervision
Discipline inspection and supervision
……

General tools | Front-office business engagement | Middle-office management | Back-office operations

Application roadmap of banking LLM

Application roadmap of LLM in China FS market



- *Scene value, proportional to the radius*
- *Application route, linear representation of horizontal and vertical coordinates*
- *Demand frequency: 30 demanders in total (sample size)*
- *Technology maturity: Commercially available as the standard of high technology maturity*

KPMG insights

Develop focus

Three key focus based on nearly two years of scenario testing and technological development:

- **"Controllability"** remains the primary constraint on the large-scale deployment of large model applications.
- **"Explainability"** has become a core concern when expanding pilot scenarios to broader implementations.
- **"ROI"** has emerged as a key factor in banks' selection of large model use cases, particularly in decision-making scenarios where comparisons with traditional AI solutions (e.g. reasoning efficiency, computing resource investment and scenario value) are critical.

Product offerings

Currently validated or near-mature application models:

Cost reduction and efficiency improvement:

- AI provides localized efficiency gains but struggles to integrate into core business processes.
- Example: copilot-style office automation (e.g. meeting summaries, email drafting).

Productivity revolution:

- AI-assisted or AI-led content generation.
- Example: marketing copywriting, analytical reports, data dashboards, promotional visuals.

Still undefined or unvalidated, immature application models:

Process reengineering:

- New business workflows involving AI agents (e.g. risk management agents, marketing agents, asset management agents). Currently in exploratory stages in finance.

Traffic innovation:

- Significant improvements in customer acquisition or engagement within existing digital platforms.
- Example: AI-driven retail business growth strategies, content recommendation systems

Disruptive transformation:

- Entirely new mass-market products built on AI-native logic. Not yet observed in the financial sector.

Key challenges and countermeasures for banking industry under the LLM opportunity (Challenge I: Data assets and capitalization)

Imbalance between fragmented high-value data resources and insufficient large-scale, high-quality training corpus supply

Data fragmentation

- **Lack of full lifecycle data governance mechanisms,** where poor data quality hinders customer service, insight analysis, and business innovation
- **Inadequate standardization of unstructured data,** preventing high-value assets (e.g. contracts, call logs) from being converted into training corpus
- **Missing Chain-of-Thought (CoT) reasoning datasets** for LLMs, with no established manual annotation or feedback mechanisms for scenario-specific data assets

Privacy and sharing dilemmas

- Institutional hesitancy due to **data sovereignty and compliance risks**, limiting cross-institution collaboration to experimental phases
- **Disputes between data owners (institutions) and AI providers (tech firms)** over data usage rights and profit-sharing from model outputs
- **Cross-border data flow restrictions**, forcing global institutions to navigate conflicting data sovereignty regulations

Impact analysis

- **Model performance constraints**
- **Innovation lag**
- **Compliance risk accumulation**

Countermeasures

Building a full-lifecycle data governance system

Integrated governance framework:

- Embed data governance into business innovation, R&D, project management and IT development.
- Establish an enterprise-wide data asset framework for systematic data accumulation and refinement.

Breaking privacy sharing barriers:

- Adopt privacy-preserving technologies (e.g. federated learning, homomorphic encryption).
- Leverage regulatory sandboxes to pilot cross-institution data collaborations.
- Deploy distilled open-source base models for private, compliant fine-tuning.
- Proactively manage cross-border data compliance, ensuring adherence to security and legal obligations.

Key challenges and countermeasures for banking industry under the LLM opportunity (Challenge II: Organization and talent structure)

The cutting-edge nature of large model technology and its rapid iteration cycles impose new demands on organizations and talent

Composite talent bottleneck

- Technical teams lack deep understanding of financial business logic, while business units struggle to assess technical feasibility, **resulting in disjointed demand identification and product design.**
- High-caliber LLM engineers and algorithm researchers command premium costs, with **talent allocation often mismatched to implementation models and pathways**.

Organizational inefficiencies

- Strategic-level gaps: absence of strong lead departments or stakeholders for centralized governance, with **inadequate cost allocation and validation/rollout strategies**
- Execution-level gaps: **lack of agile, tech-savvy cross-functional units to maximize ROI per scenario**, causing promising use cases to stall in lab phases

Impact analysis

- **Structural talent imbalances**
- **Lack of agility**
- **Rising probability of project failure**

Countermeasures

Two-way embedment mechanism:
- Upskill tech teams in business literacy (e.g. credit risk modeling)
- Train business teams on technical fundamentals (e.g. prompt engineering)

Industry-academia-research fusion:
- Develop interdisciplinary talent with quant finance + LLM expertise through targeted programs

Agile organization restructuring:
- Strengthen top-down strategic planning and oversight
- Deploy frontline agile squads for scenario-driven implementation

"Test-learn-evolve" flywheel:
- Build MVPs with 6-week iterative cycles
- Quantify cross-department synergy metrics (e.g. decision latency reduction)
- Eliminate redundant approval layers

03

Coach IQ: Telesales assistant agent and its risk assessment

Thanayut S.

Introducing Coach IQ – a real-time sales companion

Challenge

Sales agents often spend too much time on non-selling activities – searching for client information, crafting pitches or navigating complex systems. They lack timely insights on the best products, clients to focus on and the next best action to drive results. This leads to missed opportunities, low productivity and inconsistent performance across the field force.

Target solution overview

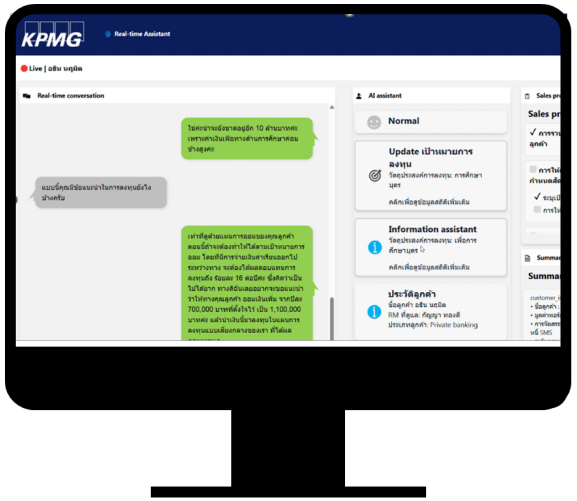
The **Coach IQ** is an AI-powered assistant that helps agent:

- **Serve next best product or action** based on customer data and interaction
- **Generate personalized sales scripts** and follow-up content
- **Provide real-time tips** in responding to customer objections
- **Automate CRM inputs and admin tasks** to free up time for selling
- **Auto-detect** non-compliant activities

This use case requires the following input:

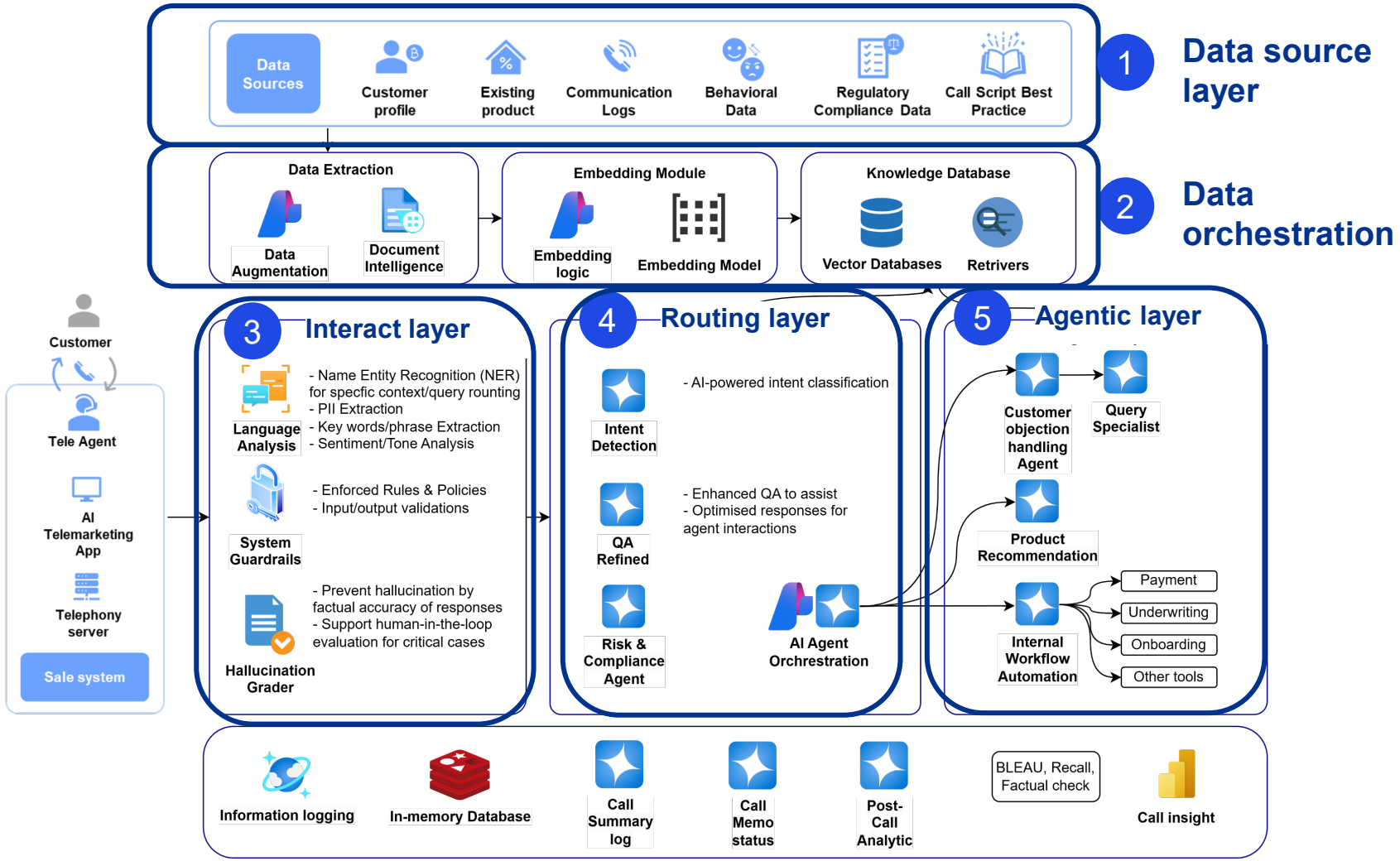
- **Audio streaming from telephony systems**
- **Sales script and compliance requirement**
- **QC checklist**
- **Product specification**

Coach IQ: virtual assistant agent



- Real-time transcription and call summary
- Real-time sentiment identification
- Customer objection assistance
- Real-time sale quality control
- Integration with Tele-sale workflow

Coach IQ – 5-layer solution architecture



KPMG tele-solution consist of 5 layers:

1. **Data sources layer: API ready to connect with your data sources**
2. **Azure data orchestration layer**
3. **Integration layer: taking input from your telephony system and control hallucination**
4. **Intent classification (routing layer)**
5. **Agentic layer: connect with your workflow to perform next actions**

Coach IQ: solution demonstration situation

Agent situations

- Bank relationship manager (agent) is trying to recapture the abandoned wealth client due to previous agent resignation.
- The agent observes that client has withdraw 80% of his AUM from the bank.
- The agent is incentivized to offer either endowment product or unit link product.

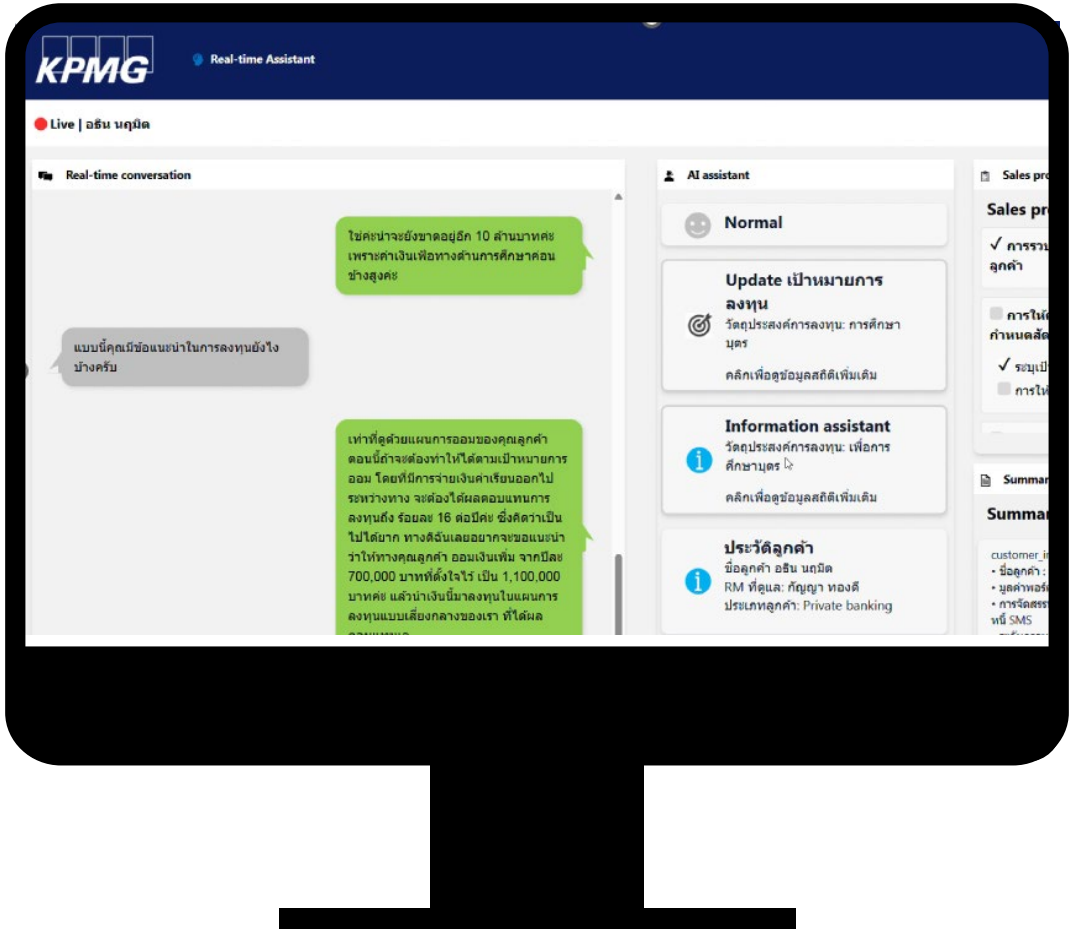
Customer situation

- Client has not been contacted for more than one year.
- Client just has a child and consider his saving plan for his child education.

Expected outcome

- Update customer investment objective
- Offer investment product that suitable to customer investment objective
- Regain trust from customer and convince customer to invest with the bank

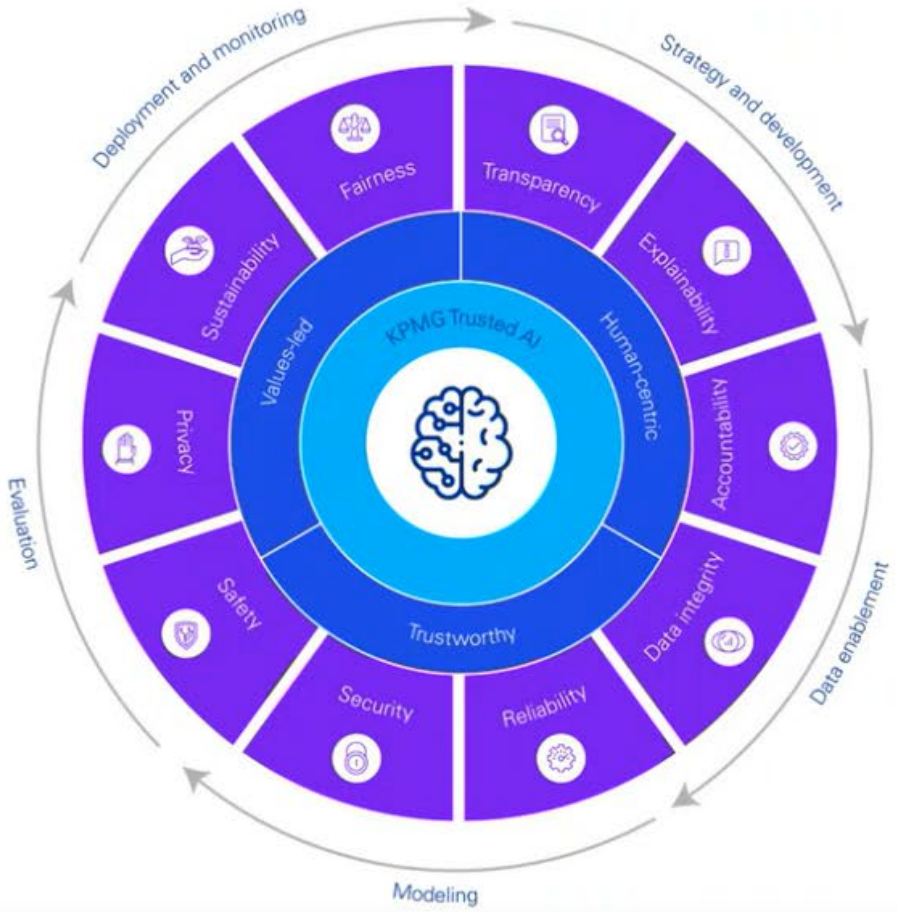
Virtual assistance AI: wealth advisory demonstration



- Real-time transcription and call summary
- Real-time sentiment identification
- Customer objection assistance
- Real-time sale quality control
- Integration with tele-sale workflow

KPMG Trusted AI Framework

Establishing a risk and control framework that covers the key risks across the end-to-end lifecycle. We have developed Trusted AI Framework consisting of 75+ sub risks and 200+ controls.



213+

Unique controls within our Trusted AI framework

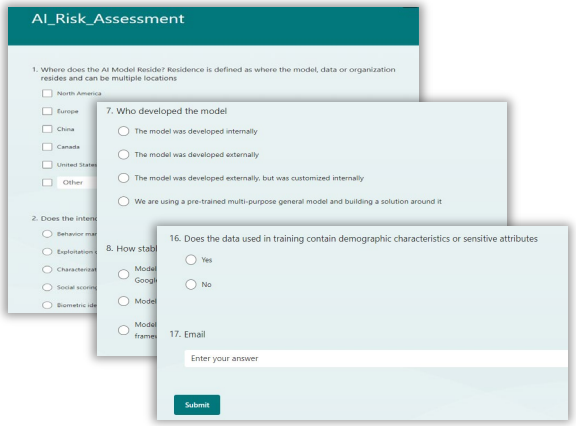
92+

Unique risks mapped to the controls

Sample of Frameworks Covered

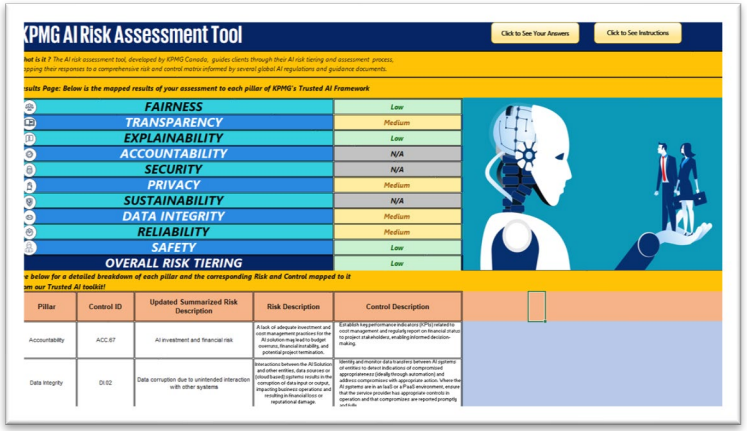
- NIST AI Risk Framework
- EU Artificial Intelligence Act
- 37 COBIT5 processes
- ISO
- Microsoft responsible AI guidelines
- Institute of Internal Auditor standards
- World Economic Forum
- IEEE, FED, MAS

Introducing KPMG AI risk-assessment tools



Step 1: Complete the questionnaire

- Users begin by accessing the AI Model Risk Assessment tool, which guides them through a series of questions designed to evaluate the risk level of their AI models.
- Each question is tailored to assess specific aspects of the AI model, such as data quality, model complexity and potential impact on business operations.
- Users are required to assess each question in relation to a single AI model, providing responses that reflect the model's characteristics and use case.
- By completing the questionnaire, users provide valuable insights into the risk profile of their AI models, serving as a basis for generating customized controls to mitigate identified risks.



Step 2: Generate controls based on risk assessment

- Once users have completed the questionnaire for a specific AI model, the tool generates a risk score based on the responses.
- The risk score categorizes the AI model as high, medium or low risk against each of KPMG's 10 Trusted AI pillars, providing a clear indication of the level of risk associated with the model's deployment and operation.
- Based on the identified risk score, the tool automatically maps a set of controls from KPMG's Trusted AI Risk Control Matrix (RCM) which are tailored to the unique risk profile and criteria of the organization and the specific AI model under assessment.
- Users can review the generated controls, which may include recommendations for enhancing data governance, model validation processes, transparency, privacy and security.
- By implementing the recommended controls, organizations can effectively manage and mitigate the risks associated with their AI models, contributing to compliance with regulatory requirements while maintaining trust in AI-driven decision-making processes.

Questions 1: Does the intended use case contain any of the following characteristics?

- ❑ Behavior manipulation without end user's awareness
- ❑ Exploitation of vulnerabilities of persons resulting in harmful behavior
- ❑ Characterization of persons based on sensitive characteristics
- ❑ Social scoring of individuals based on sensitive characteristics
- ❑ Biometric identification in public spaces
- ☑ Assessing the emotional state of persons
- ❑ Predictive policing
- ❑ Untargeted scraping of facial images on the internet or from video surveillance footage
- ❑ None of the above

According to EU AI Act, it is prohibited for company to use or provision of AI system incorporating **behavior manipulation, social scoring and prediction of criminal offense, biometric and emotion analysis of individuals, and biometric identification based on untargeted scraping of facial images.**

Assessing the emotional state of persons without using biometric information is deemed to be "low risk" as long as it is purely categorizing non-sensitive information.

Question 2: Does the use case involve any of the following?

- ❑ **Managing critical infrastructure**
- ❑ **Educational or vocational training, that may determine the access to education and professional course of someone's life**
- ❑ **Safety component of products**
- ❑ **Employment, management of works and access to self-employment**
- ❑ **Essential private and public services (including insurance and banking)**
- ❑ **Law enforcement that may interfere with people fundamental rights**
- ❑ **Migration, asylum and border control management**
- ❑ **Administration of justice and democratic process**
- ☑ **None of the above**

The scope of AI use in public will be deemed to be in "high-risk categories" and will require comprehensive risk mitigation.

Examples of essential private and public services include using AI in assessing creditworthiness and risk assessment for pricing of health and life insurance policies.

Question 3: Who is the intended users of the use case and who does it impact?

- ☑ Used by internal users, the model's outcomes only impact internal user
- ☐ Used by internal users, the model's outcomes impact individuals external to the organization
- ☐ Used by users external to the organization, the model's outcomes impact external users

Question 4: Does the use case specifically target vulnerable populations?

Vulnerable populations are groups of people who might be disproportionately affected by the outcome of AI. These groups are at higher risk of experiencing negative consequences due to factors such as socioeconomic status, age, disability, race, gender or lack of digital literacy

- ❑ **Yes**
- ☑ **No**
- ❑ **Unknown**

According to the EU AI Act, if the target group is **vulnerable,** the company is required "***to design and develop the system in such a way that natural persons are informed that they are communicating with an AI system, Art. 50 (1).***"

Question 5: Is there a potential harmful bias in AI systems that can perpetuate social inequalities or discriminatory outcomes which may lead to public trust and cause of legal, reputational risk or financial loss?

- ❑ Yes
- ☑ No
- ❑ Unknown

Results that are socially discriminatory (e.g. making decisions based on factors such as gender and age) will be treated as non-compliance with equal rights laws.

Potential scenarios include:

- The AI model developer, tasked with training the system, introduces biased or discriminatory data, perpetuating potential bias in the system's output.
- The provider implements algorithms and third-party AI models that tend to make discriminatory decisions based on gender, race or age-related characteristics.
- The user instructs the (generative) AI system to generate content with discriminatory elements and harmful information.
- The user, lacking a clear understanding of the accessed data or the (generative) AI system's biases, unintentionally spreads discriminatory statements.

Question 6: In this use case, how will users interact with AI?

- ❑ Users interact with the model through a built-in interface like a chatbot.
- ☑ Users interact with the model outputs that are generated automatically.
- ❑ Users interact with the model outputs that are prompted by a technical team.
- ❑ Users do not interact with the model, the model interacts directly with another system/model.

There is a risk of **inaccurate outputs** which can lead to operational damage, reputational damage or event financial loss. The model's outputs should be used in a careful manner with human supervision to mitigate the risk of misused.

Additionally, use cases with high complexity would require performance monitoring on a regular basis, according to the firm's internal model governance policies.

Question 7: What stage in the AI lifecycle are you currently in with the solution?

- ❑ Initial ideation
- ❑ Model development
- ❑ Evaluation and testing
- ☑ Deployment and monitoring

Question 8: Are you building or training an AI model?

Building an AI Model involves designing the architecture and structure of the model, defining how it will process input data to generate output Training an AI Model involves teaching the model to make accurate predictions or decisions based on input data.

- ☑ **The organization will use a pretrained model in their solution (for example, leveraging ChatGPT to deploy a chat bot).**
- ❑ **The organization will train a model to perform specific tasks (for example, using an object detection model and training it to identify cats by training it on images of cats).**
- ❑ **The organization will build and train an AI model.**

Breach of confidentiality

Scenario that could also apply:

- User utilizes an (generative) AI system from an unauthorized third-party and shares confidential client information in the prompt.
- The user utilizes approved (generative) AI system and shares client confidential information in the prompt, despite the customer not having consented.
- The AI model developer uses or used protected (third-party) data or content (e.g. PII, IP) without consent to train the model.
- The AI model developer, responsible for training the system, uses confidential client/third-party information for the purpose of creating a suitable data foundation.

Question 9: What is the level of complexity of the model?

- ❑ Rule-based, linear regression model, basic decision tree
- ❑ Random forest, support vector machines (SVMs)
- ☑ Deep learning, transformer models, reinforcement learning models, complex architecture

Question 10: Where is the AI solution stored?

- ❑ On a local machine like a developer's laptop
- ❑ On a local server
- ❑ On a cloud server owned by the enterprise
- ☑ On a cloud server owned by a third-party vendor

Question 11: Are there backups of the solution?

A backup of the solution includes components of the solution beyond data such as the knowledge database, system configurations

- ❑ Yes, a backup of the solution is taken periodically in an ad hoc fashion.
- ☑ Yes, a backup of the solution is taken regularly through an automate backup tool.
- ❑ No, there are no backups. However, there is a formalized runbook to guide the organization through a restoration of the solution.
- ❑ No, there are no backups.

Question 12: Which data sources can the AI system access?

- ❑ None - static training data only
- ☑ Access limited to curated company data
- ❑ Broad access to company data files
- ❑ Internet-search access
- ❑ Third-party database access

Question 13: Are there encryption standards implemented for the solution for data in transit and at rest?

- ❑ For data at rest only
- ❑ For data in transit only
- ☑ Both at rest and transit
- ❑ No encryption standard has been implemented.

Question 14: What are the potential consequences of a malfunction in the AI system, and how severe could they be?

- ❑ Severe consequences for life or health of individuals
- ❑ Damage to property or to the environment
- ☑ Disruption of service, production or logistics
- ❑ Severe economic consequences
- ❑ None of the above consequences are relevant.

Question 15: Who is or will be responsible for monitoring the AI model's performance?

- ❑ Dedicated data scientist team
- ❑ Team from IT function
- ☑ External contractors
- ❑ There is no formal responsibility for monitoring the solution.
- ❑ Unsure

Question 16: Is there a documented process for responding to performance issues detected in the solution?

- ☑ Yes, with clear steps and assigned responsibilities.
- ❑ No, responses are handled adhoc.
- ❑ Unsure

Question 17: Who is responsible for approving changes to the AI model?

- ❑ No formal approval process
- ❑ The data science team approves all changes.
- ☑ A cross-functional committee reviews and approves changes.
- ❑ The IT department approves changes.

Coach IQ: AI risk assessment result

Below is your mapped risk assessment result from the KPMG Trusted AI framework

	Pillars	Descriptions	Inherent risk	Number of control
	Fairness	Models have reduced or no bias.	Low	3
	Transparency	Provide stakeholders with a clear understanding	Medium	3
	Explainability	Can explain how recommendations or conclusions are made	Medium	3
	Accountability	Human oversight and responsibility embedded across the lifecycle	Medium	22
	Security	Safeguarded against unauthorized access	Medium	8
	Privacy	Compliance with data privacy is maintained	High	5
	Sustainability	AI solutions are improved to limit negative environmental impact	N/A	N/A
	Data integrity	Standards for data quality and governance are upheld	Medium	2
	Reliability	AI systems perform at the desired level of precision and consistency	High	6
	Safety	AI systems perform at the desired level of precision and consistency	Low	9
	Overall risk tier		**High**	**61**

Detail breakdown of each risk pillar – Fairness

Pillar	Summarized risk description	Control ID	Control description
Fairness	Potential bias and lack of inclusivity in solution development can arise from failing to identify and assess group sensitivities, impacting the fairness of outcomes.	FAIR.12	• Evaluate all datasets for inclusivity, identifying and addressing gaps with a remediation plan, including public databases, to eliminate existing biases. All steps and findings are documented.
	Lack of attention to bias and inclusivity in AI systems, along with failure to identify and assess group sensitivities during system development, may result in discriminatory outcomes, reduced fairness, and exclusion of certain user groups, impacting the fairness of outcomes and consumer trust.	FAIR.05	• Conduct periodically fairness assessments, documenting outcomes and comparing them against pre-defined risk tolerance levels to ensure ongoing adherence to fairness objectives. Remediation strategies are deployed and documented as necessary.
		FAIR.04	• Evaluate and record the AI system's capability to process diverse sub-population data accurately, both before and after deployment, using bias assessments. Mitigation strategies are implemented for any identified biases to prevent algorithmic discrimination. All findings, actions, and rationales are thoroughly documented, alongside any counterbalancing measures.

Detail breakdown of each risk pillar – Transparency

Pillar	Summarized risk description	Control ID	Control description
Transparency	Insufficient transparency in the development and use of AI systems may result in a lack of accountability, making it difficult to understand the rationale behind the system's behavior, raise ethical concerns, and erode consumer trust.	TRA.02	• Information regarding the intended use, limitations, permissibility, and data sources of the AI system are published by the appropriate authority in language that is understandable to relevant stakeholders, promoting transparency and facilitating informed engagement with the AI system.
		TRA.03	• Users or those impacted by emotion recognition or biometric categorization AI systems are notified of the system's operation prior to their use.
		TRA.05	• Informed consent is obtained for by clearly stating data collection purposes, including AI model training, and disclose AI use in decision-making processes for transparency.

Detail breakdown of each risk pillar – Explainability

Pillar	Summarized risk description	Control ID	Control description
Explainability	Lack of explainable AI solution environment	EXP.07	Develop approved Policy and Procedures as part of the Quality Management Framework that includes maintaining comprehensive records of pertinent documentation and information and the regular review and approval by a designated official to ensure ongoing compliance and accuracy. Ensure training and awareness to the relevant stakeholders to enforce compliance. The policies and procedures are reviewed and updated, as needed, periodically.
	Insufficient review of AI outputs	EXP.08	Document and evaluate the integration of significant human oversight in AI-driven decision processes, detailing the nature of human input, the reviewer's details, supplementary data influencing the final verdict, and specific scenarios prompting a system pause or manual override.
		EXP.09	Develop and conduct role-based training for human oversight, focusing on the AI system's optimal applications, effective result interpretation, troubleshooting techniques, combating automation and other detrimental biases, and complying with Automated Decision-Making rights and their related documentation needs.

Detail breakdown of each risk pillar – Security

Pillar	Summarized risk description	Control ID	Control description
Security	Lack of adherence to security principles in AI design, development, and deployment, in line with the organization's existing policies and procedures, may result in security vulnerabilities, malicious attacks, data breaches, and development of unsecure or unreliable AI system.	SEC.08	Implement data security measures throughout the AI system lifecycle to protect deployment and training code, training data, and input data sets from unauthorized modifications and potential attacks, preventing sensitive or personal data leaks in system output.
	Lack of audit and effective monitoring capabilities in AI system operations may impact the ability to monitor system performance and respond to incidents timely.	SEC.09	Alert mechanisms are implemented to continuously identify, track, and alert any security breach and/ or malfunction that may impact the operation, performance and safety of the AI system. The AI system is superseded, disengaged, deactivated, or decommissioned, as needed. When required by international regulatory bodies, alerts are reported to the appropriate governing body.
	Lack of adherence to security principles in AI design, development, and deployment, in line with the organization's existing policies and procedures, may result in security vulnerabilities, malicious attacks, data breaches, and development of unsecure or unreliable AI system.	SEC.11	Conduct periodic resiliency and security assessments of the AI system, adhering to organizational best practices and encompassing a range of tests to ensure comprehensive security and sustainability.
	Lack of effective security vulnerability management over AI systems/environment may lead to exploitation of weakness, resulting in unauthorized access, increased attacks such as malware, data breaches, and operational disruptions	SEC.13b	Vulnerability management processes include the identification, reporting, and monitoring of relevant vendor and national cyber authority alerts, and other open-source information channels to support the timely response to new vulnerabilities.
		SEC.15	Perform detailed risk assessments for AI vulnerabilities to determine the root cause and considering impact and the likelihood of reoccurrence.
	Adversarial attacks exploiting models, data sets, or algorithms may result in unauthorized access to confidential data, model tampering, data corruption or loss, misuse, inappropriate access, or non-compliance with underlying regulations.	SEC.20	Prior to launch and periodically thereafter, perform penetration tests and/or "Red Team" exercises for the AI system and its environment to identify potential vulnerabilities. Any identified exposures are promptly reviewed and addressed to ensure the system operates as expected.
		SEC.21	Implement training dataset expansion techniques as part of data cleaning process to ensure the performance and robustness of algorithms/systems and their resilience to adversarial and poisoning attacks.

Detail breakdown of each risk pillar – Privacy

Pillar	Summarized risk description	Control ID	Control description
Privacy	Lack of operational infrastructure to enable individuals to exercise their Data Subject Access Rights timely may result in a loss of consumer trust, regulatory non-compliance, or cause financial harm.	PRI.06	Launch awareness programs aimed at educating data subjects about their rights in relation to AI technologies, explaining how to exercise these rights and the implications of AI decision-making on their personal data.
		PRI.07	Personal data is indexed in the AI system to expedite responses to Data Subject Access Requests.
	Potential data breaches may result in the unauthorized access or disclosure of personal, official use, confidential, and strictly confidential data which could compromise user or organization privacy, violate data protection laws, lead to reputational damage, or cause financial harm.	PRI.10	Document rationale and explicit approval when obtaining data for training. Special precautions are implemented for AI use cases that may directly or indirectly affect vulnerable individuals or have safety or rights implications.
		PRI.11	To a degree appropriate for the model and use case, a controlled amount of randomness (i.e. differential privacy) is added to training and prompt data to protect data privacy.
	Lack of compliance and alignment with organization directives and/ or regulations on processing data subjects may lead to financial penalties, market losses, and reputational damage.	PRI.15	Establish criteria for processing personal or sensitive data in high-risk AI systems, integrating strong privacy protections, including data use limitations and encryption.

Detail breakdown of each risk pillar – Data integrity

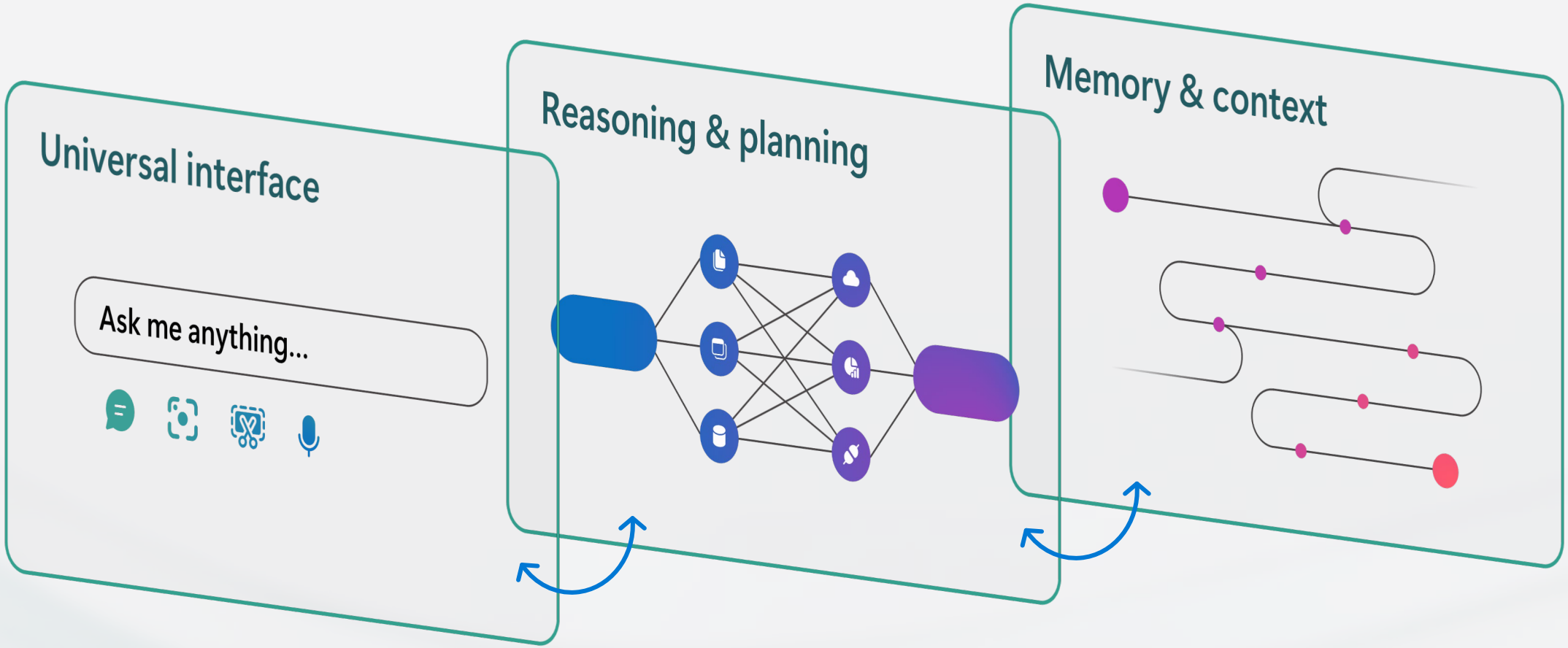
Pillar	Summarized risk description	Control ID	Control description
Data integrity	Lack of appropriate methods to facilitate and control data interactions (e.g. transfers) between the AI systems and data sources or other entities (e.g. applications, APIs) may result in data corruption or loss, system misuse, or inappropriate access.	DI.07	Continuously monitor and document any changes to real-time data sources - internally or externally - that interact with the AI system during operation. Change alerts are investigated and addressed timely.
		DI.08	During the change management process for an AI system, the training and testing data used is evaluated for relevancy and accuracy with the change. As needed, additional data is introduced to train and test new system capabilities or features.

Detail breakdown of each risk pillar – Reliability

Pillar	Summarized risk description	Control ID	Control description
Reliability	Lack of a comprehensive and systematically documented quality management system for high-risk AI systems may lead to non-compliance with regulatory requirements, resulting in the deployment of AI systems that are unsafe, ineffective, or violate ethical standards and a loss in consumer trust.	REL.01	Develop and approve a Quality Management System to ensure continuous operational support and maintenance on the AI system and includes aspects of resource management and supply security measures for high risk AI systems to ensure compliance to regulatory requirements. Ensure training and awareness to the relevant stakeholders to enforce compliance.
		REL.02	Develop and approve a Quality Management System to include the design control and verification of AI systems to ensure integration of the AI system within the wider IT landscape in which it operates. Ensure training and awareness to the relevant stakeholders to enforce compliance.
		REL.03	Prior to system development, the AI deployment team develops and enacts a strategy that includes detailed procedures for maintaining regulatory compliance throughout the AI system's lifecycle, including during the assessment of conformity and management of any modifications, to ensure ongoing compliance with regulatory requirements and effective change management for high-risk AI systems.
	Lack of audit and effective monitoring capabilities in AI system operations may impact the ability to monitor system performance and respond to incidents timely.	REL.11	Automated correction, fallback, or stop/loss mechanisms are implemented in the AI system's design to ensure the AI system corrects, or when necessary, halts unintended behavior. Humans are alerted and the issue(s) are remediated timely.
	Lack of resiliency in AI systems and services, including inadequate backup and restore capabilities and insufficient availability in case of a disaster, may result to extended downtimes and failure to provide critical functions and/or services in a safe, accurate, and timely manner.	REL.16	Include advanced support and warranty arrangements in contracts with AI vendors, ensuring system availability and effectiveness via clear service levels and monitoring.
		REL.18	Implement failover mechanisms such as automatic backup system switching and frequent system backups, including component snapshots and rollback capabilities, as a fail-safe against unexpected failures to ensure the AI system has the ability to manage unforeseen circumstances without compromising its overall performance or reliability.

04

What's next? The shift to agentic AI


Universal interface
Ask me anything...
Reasoning & planning
Memory & context

Generative AI Adoption Patterns

Wave 1 and Wave 2 – 2023 to 2024, moving to scale in 2025

Wave 1

Use Cases Inside the Organization

Internal Copilot deployments

Wave 2

Infusing Generative AI to your products

+

Customer Experience

Wave 3

Agentic AI

Autonomous and Semi-Autonomous agents

Wave 4

Artificial General Intelligence

New Product Offerings

Human 'in the loop' review of generative content

AI augmenting workgroups

Human Supervision of multiple agents

Building an agentic world

Personal agents

Organizational agents

Business process agents

Cross-org agents

Agentic multi-agents in action for Insurance Policy Renewal

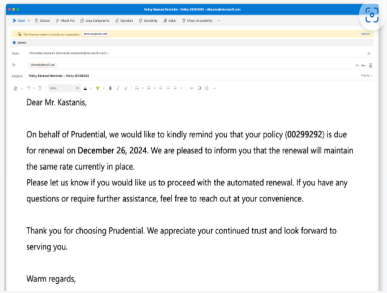
1. ORCHESTRATOR AGENT
- **Operating in the background 24/7**

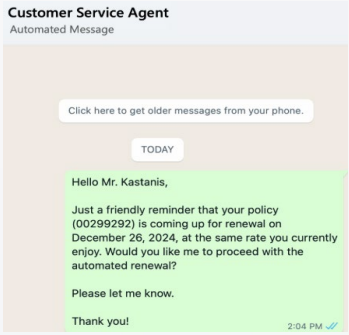
ORCHESTRATOR AGENT invokes POLICY AGENT

2. POLICY AGENT
- **Identify** policies expiring in **2 weeks time**
- **Get** policy details, such as insured name, policy type/number, expiry date, premium amount, contact information etc.

ORCHESTRATOR AGENT invokes CUSTOMER SERVICE AGENT

3. CUSTOMER SERVICE AGENT
- **Generate SMS and email** to inform customer that policy is expiring
- **Human in the loop to send** SMS and email to customer. CUSTOMER SERVICE AGENT can send directly to customer too if permissible.





4. CUSTOMER SERVICE AGENT
- **Track** customer's response for renewal
- **Receives customer's confirmation** to renew and updates ORCHESTRATOR AGENT

ORCHESTRATOR AGENT invokes PAYMENT AGENT

5. PAYMENT AGENT
- **Track and reconcile** payment from customer
- **Confirm** payment is successful

ORCHESTRATOR AGENT invokes POLICY AGENT

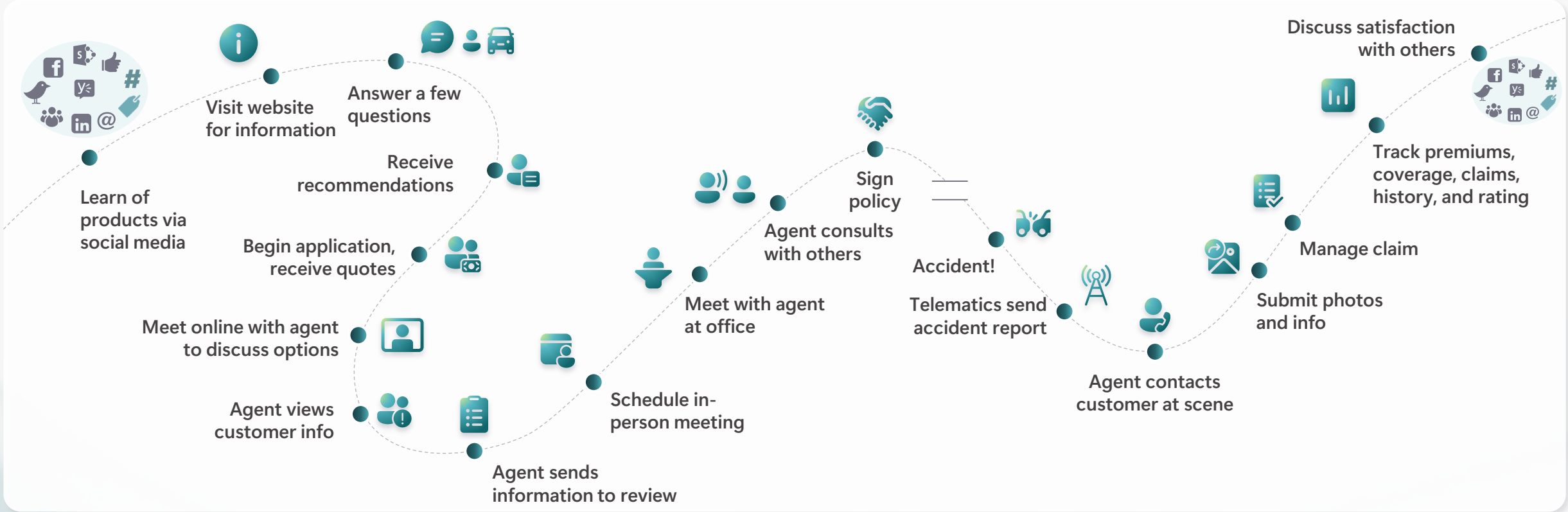
6. POLICY AGENT
- **Updates policy documents** with renewed status

ORCHESTRATOR AGENT invokes CUSTOMER SERVICE AGENT

7. CUSTOMER SERVICE AGENT
- **Generate SMS and email** to inform customer that policy is renewed
- **Human in the loop to send** SMS and email to customer. CUSTOMER SERVICE AGENT can send directly to customer too if permissible.

8. ORCHESTRATOR AGENT
- **Checks that everything is in place and closes this case.**

Insurance Customer Journey – Policy lifecycle



Digital Hotspots

Pre-purchase

- Learning of company on social media
- Exploring website
- Receiving personalized offers
- Chatting with agent online
- Applying for insurance

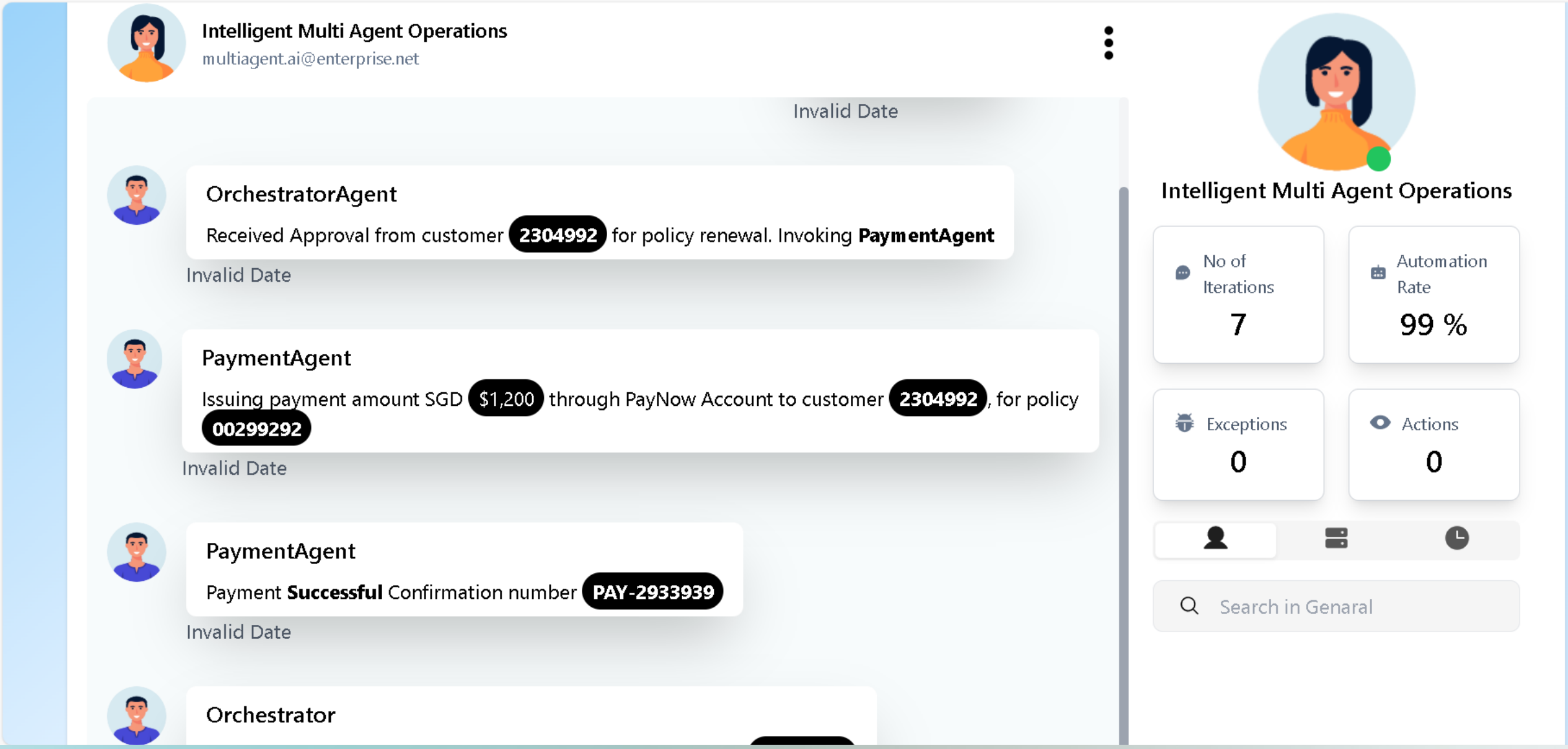
Account

- Tracking premiums, history, rating
- Reviewing coverage with agent
- Adding new products
- Setting up payments
- Updating profile

Policy services

- Contacting agent for help
- Reporting a problem
- Submitting a request to Amend policy
- Managing a policy life cycle
- Customer service

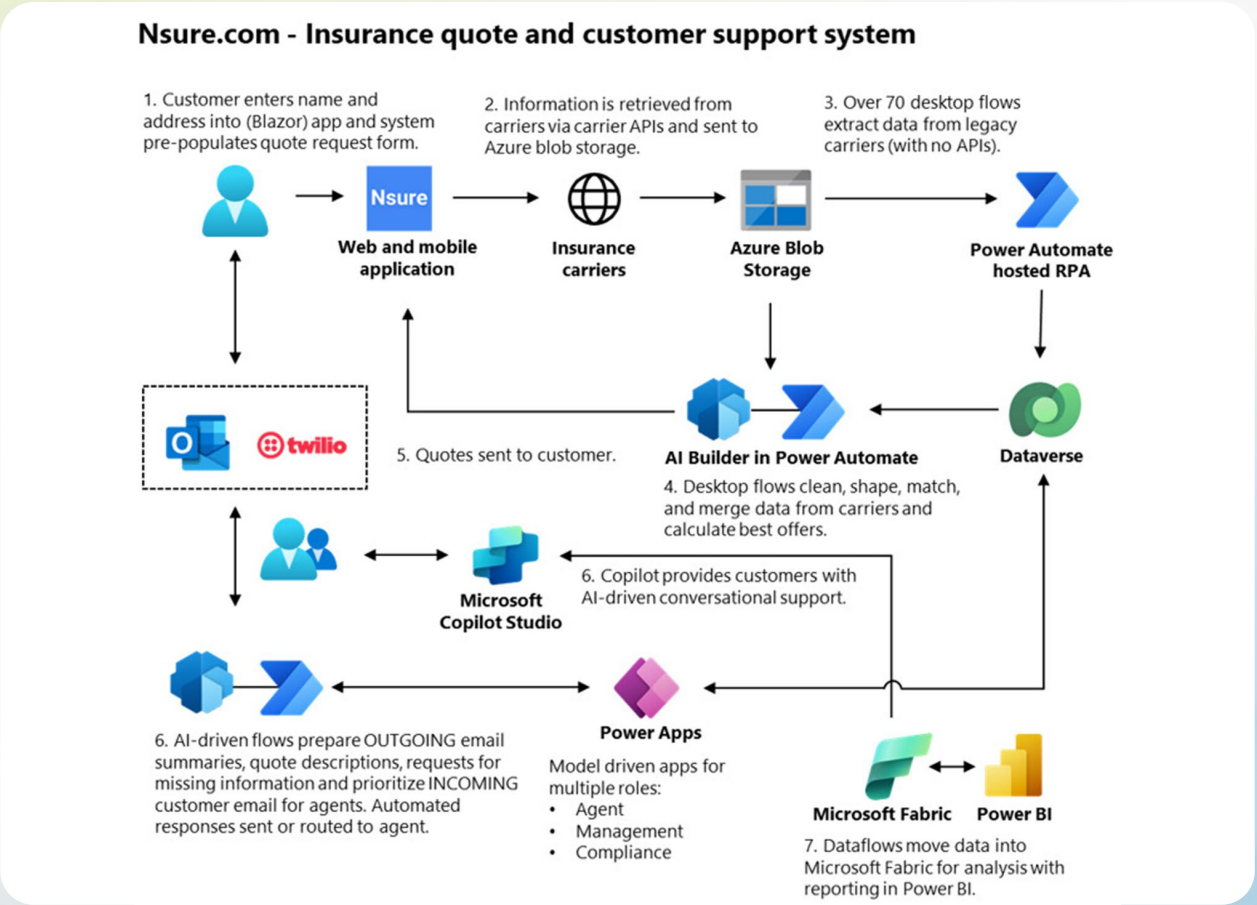
Multi-Agents at Work



Nsure.com Digital Insurer – Automates Quotes + Support

- **Automation and Efficiency:** Nsure.com used Microsoft Power Platform to automate business processes, reducing manual processing time by 60% and associated costs by 50%.
- **AI Integration:** The company leveraged generative AI capabilities in Power Automate to handle complex tasks, improving customer satisfaction and operational efficiency.
- **Rapid Growth:** Since implementing these technologies, Nsure.com has experienced significant growth, with a Compound Annual Growth Rate (CAGR) of over 100% from 2020–2023.

Link



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

KPMG in Thailand
48th-50th Floor, Empire Tower
1 South Sathorn Road
Bangkok 10120
T: +66 2677 2000

KPMG in Thailand

kpmg.com/th

Document Classification: KPMG Public