# HSBC - HKMA GenAI Sandbox Use Cases Summary

HSBC

# HSBC Experiments with Gen AI Powered Website Chat Agent

## Background

In December 2024, HSBC was selected to join the Generative Artificial Intelligence ("Gen AI") Sandbox ("Gen AI Sandbox"), launched by the Hong Kong Monetary Authority ("HKMA") in collaboration with the Hong Kong Cyberport Management Company Limited ("Cyberport"). Through the Gen AI Sandbox, HSBC has conducted experiments for potential Gen AI use cases in a risk-controlled environment. This document shares results of one of HSBC's use cases.

As a long-standing leader in Artificial Intelligence (AI), the Hongkong and Shanghai Banking Corporation Limited ("HSBC") has been investing heavily in AI technologies for many years, applying them across areas such as fraud detection, cyber security, transaction monitoring, customer services, and risk management. Building on this foundation, Gen AI offers vast potential to enhance how we work and serve our customers. Within the Gen AI Sandbox, HSBC has concluded Proof of Concept (PoC) experiments on Gen AI powered Website Chat Agent (also known as "the Chatbot"). These experiments demonstrated the effectiveness of the Chatbot in elevating customers' search experience on HSBC public website.

## Why it matters?

HSBC as Hong Kong's largest retail bank, it is essential to meet customer expectations[1]:

- **80%** expressed willingness to adopt Gen AI for financial management
- **68%** were interested in using a chatbot powered by Gen AI
- **70%** expected chatbot search experience to be fast and intuitive with relevant and concise responses

These insights highlight the importance for HSBC to explore the potential of integrating Gen AI into its services.

## Customer and business benefits

HSBC conducted PoC experiments with internal staff to explore the versatility of a Gen AI powered chatbot in interacting with customers autonomously. The PoC tested whether this Chatbot could:

- Enhance search experience with seamless and efficient navigation of information on HSBC's public website
- Provide 24/7 personalised, conversational, and contextual information search to foster customer engagement
- Gather data-driven insights from chatbot interactions to unveil emerging customer needs, driving new business growth opportunities

## Key capabilities tested in this Chatbot PoC1

- Intelligent contextual question answering: Addressing customer queries in natural languages with key functions:
    - Fact summarisation – Summarising public information on HSBC products and services as inquired by customers and providing digestible responses

[1] HSBC Community online polling for Gen-AI Chatbot conducted from July 17-21, 2025, with 1,053 respondents, and August 12-13, 2025, with 502 respondents

- Product comparison – Factual comparison of products by attributes with no suggestion on product selection
- Link retrieval – Provision of link to navigate customers to the relevant HSBC public website pages for understanding more information

- Scope management: Recognising customer queries that may lead to product recommendation or falling outside the Chatbot's designated knowledge domain, and redirecting to human agent politely

## PoC approach and results

- Approach: A two-pronged testing with HSBC colleagues and automated scripts was completed in June 2025 to:
  - Evaluate quality of chatbot responses in over 800 distinct scenarios across all HSBC public website pages
  - Validate robustness and effectiveness of risk control mechanisms
  - Identify improvement areas for future iterations
- Number of testers: Over 400 HSBC colleagues located in Hong Kong
- Results:
  - User experience[2]
    - **95%** testers agreed that the Chatbot is helpful
    - **85%** testers agreed that the overall experience with the Chatbot is satisfactory
    - **85%** testers would recommend the Chatbot to their friends and family
- Following the completion of the PoC, a press release was issued in August 2025 with extensive media coverage in more than 10 online and print publications.
- The PoC was also showcased to HSBC colleagues during HSBC's AI Day event in September 2025.


## What to expect next?

The PoC provided valuable insights and fresh perspectives on how HSBC can advance the role of Gen AI in the Bank's future digital banking roadmap. With proper risk assessment and appropriate controls, the Chatbot could become an integral part of HSBC's service offerings for delivering seamless and hyper-personalised customer experience.

[2] The user experience results are collected from testers who have completed the post-experiment satisfaction survey.

# Harnessing fine-tuned language models for financial market insights

## High-level summary

In December 2024, HSBC was selected to participate in the Generative Artificial Intelligence ("Gen AI") Sandbox initiative, launched by the Hong Kong Monetary Authority ("HKMA") in collaboration with the Hong Kong Cyberport Management Company Limited ("Cyberport"). Through the Gen AI Sandbox, HSBC has conducted experiments and test for potential Gen AI use cases in a risk-controlled environment. This document shares results of one of HSBC's use cases.

This initiative to fine-tune language models for financial applications is a key component of HSBC's efforts to improve financial decision-making. In this sandbox project, we sought to demonstrate that smaller, deployable language models can effectively handle specific financial queries posed by professional traders, such as identifying immediate investment opportunities from financial news.

Business analysts require rapid identification of actionable investment insights from news articles. However, current language models must be refined to efficiently and professionally interpret complex news contexts, enabling swift decision-making. Additionally, these models must accurately process intricate financial contexts, necessitating the adaptation of models to specific financial tasks.

- **Harnessing fine-tuned language models for enhanced financial market insights**

To address these challenges, HSBC sought to develop a small-size language model designed to identify investment opportunities within news contexts using specific reasoning logics. The model aimed to accurately parse and interpret news articles to extract actionable investment insights, such as identifying impacted industries and mapping relevant upstream and downstream sectors that may benefit or suffer from the reported events. HSBC's approach utilised knowledge distillation to empower compact models to execute complex logical processes, ensuring they are cost-effective and easily integrated into investment banking workflows.

Furthermore, our experiment aimed to validate a robust curation pipeline for creating high-quality reasoning trajectory datasets in the financial domain. The experiment concluded that the data curation pipeline—combining decontamination, difficulty filtering, and verified chain-of-thought distillation—can elevate a compact language model's proficiency in financial reasoning.

## Key findings and learnings

As part of this experiment, we employed an open-sourced third-party 7B model as the base model for a series of fine-tuning processes. We then used these small language models to identify investment opportunities and benchmarked each fine-tuned model's financial analysis using a full-power LLM as the evaluation model (LLM-as-a-judge).


- **Testing authentic financial logic with chain-of-thought finetuning:**

We found that the fine-tuned small language models **successfully adapted to** a professional analyst-like logic, significantly enhancing the length, logical rationality, and industry coverage in their outputs. Based on our testing, this structured logical framework, known as "chain-of-thought" (CoT), seemed to mirror expert financial analysis, indicating that the models were reflecting authentic trader-like reasoning patterns. Crucially, this accomplishment was achieved using only **approximately 1,000** synthetic training instances, highlighting the effectiveness of synthetic data in overcoming the challenges of acquiring real-world financial analyses directly from businesses.

- **Model iterations and improvements:**

The fine-tuning process in our experiment involved two more iterations after our initial fine-tuning attempt using Low-Rank Adaptation (LoRA). The iterative fine-tuning process led to **progressive improvements** in the model's performance. Model 2.0, trained with a more diversified augmented dataset, slightly expanded industry coverage. Subsequently, Model 3.0, which underwent full fine-tuning with the augmented dataset, achieved the most optimal performance in logical reasoning and industry coverage. These findings underscored the importance of diverse training data in enhancing model capabilities and demonstrated the necessity for continuous refinement to effectively address industry-specific challenges.

As a complementary test for financial domain adaptation, we also curated a new reasoning trajectory dataset as part of our experiment, comprising questions from public financial question-and-answer (Q&A) datasets. These trajectories, distilled from a full parameter LLM, were subsequently used to fine-tune the same 7B model. The curation pipeline was structured around three criteria to enhance quality: 1) Decontamination 2) Difficulty Filtering 3) Verification of Reasoning Trajectories.

- **Efficacy of domain-specific fine-tuning**: Upon evaluation against a test set of 645 examples, the fine-tuned model demonstrated **significant performance enhancements** over the base model. The improvement in correctness was consistently observed across various source datasets, with accuracy on certain subsets approaching a two-fold increase.
- **Correlation of answer correctness:** Furthermore, there was an increase in the average thinking trajectory length across datasets, suggesting longer reasoning depth. This indicated a **positive correlation** between increased thinking trajectory length and correctness.

## Next steps

To enhance the effectiveness and impact of this experiment, several areas for future research and development are recommended. These include:

- **RAG system integration**: The integration of Retrieval-Augmented Generation (RAG) systems can further enhance the investment opportunity associating applications by providing real-time access to diverse supply chain datasets, such as upstream and downstream sectors.
- **Development of multimodal reasoning capabilities**: Many complex financial tasks require the co-interpretation of textual data with visual representations. A future extension could expand the model's reasoning capabilities beyond the textual modality, involving the curation of specialised, paired datasets comprising text, numerical tables, and charts.

# HSBC experiments with fraud management innovation

## Background

In December 2024, HSBC was selected to participate in the Generative Artificial Intelligence ("Gen AI") Sandbox initiative, launched by the Hong Kong Monetary Authority ("HKMA") in collaboration with the Hong Kong Cyberport Management Company Limited ("Cyberport"). Through the Gen AI Sandbox, HSBC has conducted experiments and test for potential Gen AI use cases in a risk-controlled environment. This document shares results of one of HSBC's use cases.

### Safeguarding customers through innovation

At HSBC, protecting our customers against fraud remains one of our highest priorities. In today's fast-moving digital economy, fraudsters are adopting increasing sophisticated methods. To stay ahead, HSBC is investing in next-generation technologies – including Voice AI and Generative AI technologies – to strengthen fraud detection and improve operational efficiency. The tools discussed on this document are for internal staff use and there is human-in-the-loop.

Our vision is clear:

- **Enhanced customer protection** by identifying and responding to fraud risks earlier
- **Support our employees** with intelligent tools that improve productivity and decision-making
- **Enable safer financial journeys** through responsible innovation

## Experimenting with multistep use cases in fraud management

1. **Real-time speech-to-text transcription**

Real-time speech-to-text transcription can help capture conversations accurately and quickly when the customers report the fraud case to us. This enables:

- Instant record of call interactions
- More efficient handling of customer cases
- Better insights for investigators

Technology used: Voice AI

2. **Generative AI summarisation**

Building on transcribed conversation, **Generative AI summarisation** was applied to produce clear, concise summaries. This ensures:

- Complex case information is structured and easy to navigate
- Staff can focus on decision-making rather than manual note-taking
- Greater consistency across reports

Technology used: Generative AI (Large Language Models)

### 3. AI-supported fraud reports

When transcripts, documents, and case notes are combined, Generative AI can assist investigators in preparing structured fraud reports. While staff remain fully accountable for decisions, AI can act as a co-pilot – which can potentially help streamlining documentation, improving turnaround time, and ensuring higher quality outputs.

Technology used: Generative AI (Large Language Models)

What modifications have we made?

- To ensure superior quality of outputs while taking into consideration limitations of Large Language Models, we have conducted detailed testing and iterative prompt engineering to ensure the outputs meet the requirements of the business.

## Responsible AI and customer outcomes

| **Responsible innovation** | **Better outcomes for customers** |
|---|---|
| At HSBC, we recognise that adopting AI in fraud management must be done responsibly. All initiatives undergo a robust review and governance process, including but not limited to:<br><br>• **Risk assessments and compliance checks**<br>• **Bias, accuracy, and privacy** assessments before deployment<br>• If required, clear **human-in-the-loop design**, ensuring staff retain ultimate decision-making responsibility. | The ultimate goal is to reduce fraud losses and improve customer confidence. With these new tools, we expect:<br><br>• Customers can receive **faster resolution** when fraud occurs.<br>• Investigators can have **richer insights** at their fingertips.<br>• HSBC can strengthen its role as a **trusted partner** in safeguarding financial wellbeing. |

## Closing notes

HSBC continues to collaborate with regulators, partners, and industry peers to ensure technology is applied responsibly in fighting fraud. By combining human expertise with advanced AI, we are building a more secure and resilient banking experience for all.

# HSBC Experiments with GenAI-Powered Chatbot For Corporate Clients

## Project Overview

In December 2024, HSBC was selected to join the Generative Artificial Intelligence ("GenA.I.") Sandbox, launched by the Hong Kong Monetary Authority ("HKMA") in collaboration with the Hong Kong Cyberport Management Company Limited ("Cyberport"). Through the GenA.I Sandbox, HSBC has conducted experiments for potential GenA.I. use cases in a risk-controlled environment. This document shares results of one of HSBC's use cases.

As a long-standing leader in Artificial Intelligence (A.I.), HSBC has been investing heavily in A.I. technologies for many years, applying them across areas like fraud detection, cyber security, transaction monitoring, customer service, and risk management. Building on this foundation, GenA.I. offers vast potential to enhance how we work and serve our customers. Within the HKMA GenA.I. Sandbox, HSBC has concluded Proof of Concept (PoC) experiments of GenAI powered interactive chatbot, trying to address and resolve queries from commercial banking clients. No actual customers were involved in this project, and the primary objectives were to:

- Deliver responses to routine operational and product-related enquiries.
- Identify client-specific questions and craft scripts to access answers from internal databases, ensuring appropriate authority control that only destinated client information will be processed within the query session.

## Project Background

In this HKMA SANDBOX initiative, HSBC explores how GenAI could enhance client servicing in below aspects:

In the unique Commercial Banking landscape, customers tend to have different and more sophisticated questions towards the bank, **on both general operation and in-depth transaction or account status.** Currently the bank's customer service team invest large amount of time and efforts to check client status, related transaction details and remarks for accurated answers. This project targeted to address this issue, save time for customer service department and accelerate customer experience through GenAI technology.

## Project Solution Overview

### From data preparation to model optimization: how to release full potential of AI technology:

- Data preparation for unstructured data: knowledge content & conversation.
- Data privacy protection: apply masking on PII anonymized using token-based substitution.
- Embedding + Retrieval-Augmented Generation (RAG) to reduce model hallucination & increase answer accuracy. Average **43% improvement in Realiability & Accuracy** responses across various output quality metric.

## Our Sandbox Journey

### Internal & External collobaration

#### 1. Business alignment & data preparation

Based on the technical design, project team closely worked with HSBC commercial banking customer service and digital team to collect the training data, and transform unstructured materials into model friendly structure.

#### 2. Model finetuning and training

The project team first adopted a prompt engineering driven method with generic trained model. This design created high hallucination on the model output. Based on the result, project team pivioted to using RAG and golden source embedding method to retrain the generic model, achieved average 43% improvement in accuracy.

The HKMA SANDBOX remote development infrastructure was generally comprehensive, allow great flexibilities for the project team to adopt and pivot to different models and solutions. During the training process, we received invaluablt support from cyberport team on technical set up blockers.

#### 3. Output testing & evaluation by business

The chatbot output are tested by HSBC commercial banking digital team business users, evaluated based on their industry expertise. As this is an internal prove of concept project, no actual customer was participated in the testing.

### The Challenges we faced

- The main challenge stemmed from HSBC's NAT (Network Address Translation) setup which caused frequent connectivity issues that affect model finetuning & feedback efficiency.
- The PII masking weakened entity resolution, and the absence of tool-calling restricted the model's ability to answer client-specific queries.

## Conclusion

This experiment highlights GenAI's potential to enhance banking client services, emphasizing the crucial role of data governance, system integration, and security measures. GenAI technology scene is booming with various model to adopt, yet to release the full potential of the AI power, project team needs to select suitable model & method based on technology design, to better address the real-world business painpoint.

**Disclaimer:** This summary is produced by HSBC for the HKMA Gen AI Sandbox. It is important to recognise that HSBC's results and findings from the PoC presented in this summary should be interpreted within the context of the experiments which was conducted, and the limitations of the experimental scale and variety of data available. Despite these limitations, the results and findings are intended to support ongoing discussions and experimentation around Gen AI. All contents of this summary (including any views, insights and learning experience expressed in it) are provided "as is" intended for informational, educational and demonstration purposes only and do not reflect HSBC's policy stance on Gen AI implementation in Hong Kong. HSBC makes no representation or warranty, express or implied, that the contents in this summary are fit for purpose, timely, correct, accurate, complete or non-infringing. HSBC does not accept any responsibility or liability with respect to the use of or reliance on any contents or information contained in this summary, to the extent permitted by all applicable laws.