Reimagining Banking Services:
An AI-Driven Strategy for Growth

Executive Summary

Leveraging on conversational generative AI technology to build autonomous AI agents, assisting in customer services and unlocking untapped revenue potential.

Benefits:

- Improve customer experience by reducing waiting time at bank branches, allowing capacity for market adoption
- Unlock sales potential from existing customer base with least friction to enquire
- Promote work-life balance for frontline receptionists by reducing work days and hours with the help of AI receptionists

The Growth Bottleneck - Time Consuming Customer Services

With the growing smart economy, banking industry in Singapore is expected to benefit heavily in the coming years. With more sales potential comes a greater need for customer service expansion, otherwise, growing dissatisfaction among unserved customers will lead to bad reviews and increasing termination of saving accounts, jeopardising sales expansion efforts.

The current capacity for customer services had been at saturation for years, with waiting time of one to three hours, and a long queue that requires additional manpower to manage. Most of these services on demand were administrative and non value-generating to the bank. With bank branches constantly at full capacity and making an appointment a hassle, it has dettered people who had intention of enquiring about financial products to cast it aside, resulting in opportunity costs for growth.

The AI Solution, Supervised by Humans

An AI-enabled two-pronged approach can be implemented to drastically reduce waiting time on-site.

1. Revamp online chatbot to make it a truely problem-solving tool, thereby taking off heavy loads from on-site branches
2. Extend opening days and hours of bank branches to dilute traffic volume per hour, by incorporating AI receptionists.

Online AI Chatbot

The current chatbot on the website was only able to offer limited assistance due to the inability to comprehend complicated cases, which were many.

Most customers required an understanding of their situations and to receive personalised suggestions from the staffs. Yet, the current chatbot was only able to provide generic online resources which customers had to discern by themselves, only to realise that those resources did not provide a solution to their specific needs, leading to frustrations. With problems un-solved, it forced customers to visit physical branches, adding pressure to on-site congestion, while putting themselves to another round of frustrations due to long waiting hours.

By using a GPT-powered AI chatbot, it can understand customer enquiries effectively and converse in a professional yet emphatetic manner, solving problems and reducing frustrstions at the same time.

The AI chatbot shall not be just a conversational tool that provides suggestions to customers, where customers still have to act on the problems themselves. The chatbot will be run by an autonomous AI agent, who is able to converse, provide feedback, and make adjustments on the customers' bank accounts just like a human receptionist. This will truly enable problem solving online instead of merely providing suggestions.

The AI chatbot agent will be trained on product knowledge, how to use the bank's platform, and the processes of permissiable problem solving.

Extend On-site Operating Time

The nature of customer enquiries at the bank are pre-dominantly administrstive. It requires providing advice and getting permissions to help act on their problems - not straightforward solutions that can be automated entirely. This means that reducing time spent per customer enquiry will remain a challenging task.

Instead of focusing on cutting down time spent with customers, a more viable alternative would be to spread customer traffic across longer opening hours. Referring to the table below, an extension of opening days and hours can lead to 39.5 hours or 88.8% more servicing time per week. If there were 30 awaiting customers at a branch, it could be reduced to 17 after time extension, greatly reducing waiting time for a better customer experience.

<u>Current Opening Hours</u>
Mon-Fri: 8.30am - 4.30pm (8 hrs)
Sat: 8.30am - 1pm (4.5hrs)
Total per week: 44.5hrs

Proposed Opening Hours
Mon-Sun: 8.30am - 8.30pm (12 hrs)
Total per week: 84hrs

The same AI chatbot agent technology used on the website can be replicated on-site as AI receptionists, cutting down the need for human receptionists. However, a human receptionist manager will still be needed per branch to monitor the AI receptionists and intervene when there are technical errors or when the AI receptionists could not handle certain requests from customers.

Current Human Receptionist Allocation
Per Branch: 4
No. of Branches: 25
Total Human Receptionists: 100

Proposed Human Receptionist Allocation
Per Branch per Shift: 1 (with 4 AI recepts)
Shifts per Day (8.30-2.30, 2.30-8.30): 2
Teams per Week (Wed-Sat, Sat-Thur) : 2
No. of Branches: 25
Total Human Receptionists: 100

The extension of opening hours will not only reduce on-site traffic, it opens up sales opportunities from working adults who were unable to visit the banks during office hours, therefore casting less important enquiries aside - which were often sales enquiries; focusing on solving key functional problems instead which were administrative and non-revenue generating. The opened-up night hours and weekend windows will allow working adults to make enquires with least friction.

3-year Blueprint: Propelling into a Human-AI Productivity Model

Phase 1: The Chatbot on Beta

The key driver of this strategy relies on the functionality of the AI agent running the chatbot on the website and on-site. Thus, it is paramount that the chatbot agent is given sufficient training before rolling out to the public.

The AI chatbot agent will be trained on product knowledge, how to use the bank's platform, and the processes of permissiable problem solving. All these processes, including the way of speaking and numerous case studies, will be provided by the Head Receptionist to the AI developer.

The AI developer will insert the processes and speaking styles into a GPT model by prompt engineering, and connects the GPT into the bank's staging website and dummy bank accounts via an

API. Both the Head Receptionist and AI developer will work closely together and test out all kinds of customer enquiries.

Phase 1 may take half to one year.

Phase 2: The Live Online Chatbot

When the beta chatbot has passed all scenerios, it is ready to live. The developer will connect the GPT API to the live website and real bank accounts.

Close attention will be given upon the roll out to check for technical errors, incompetent responses, hallucinations and unauthorised actions. Fine tuning will be done in parallel to live actions. However, the live chatbot will be disabled immediately should there be a critical breach of security or privacy. The chatbot will be fine-tuned and put on stress test, re-living it again after it hashas been proven resilient to prompt manipulations.

Phase 2 will last for one year to ensure the chatbot has been battle-tested by the online customers.

Phase 3: Human-AI Receptionist Collaboration

By now, the AI chatbot agent has proven its functionality and is ready to be deployed on-site.

The smallest bank branch with the least traffic will be chosen for the pilot programme. All four counters will each be installed with a digital screen that displays AI-generared human-looking avatar. Human-looking instead of cartoon avatar will be used to create a more realistic impression of human interactions, forging relationship with customers.

The AI receptionists will be capable of speaking in four languages, namely English, Chinese, Malay and Tamil using the GPT advance voice mode, catering to the needs of elderly customers. Prior to phase 3, language experts will work with AI developer to ensure the GPT model understands colloquial and speaks in proper syntaxes; professional voice actors will aid in providing the recorded voices.

Each counter will have both human and AI receptionists. While AI receptionists will perform most of the tasks including signaling the next customer in the queue (API connects to queue system), the human receptionists will be in charge of monitoring the AI receptionists and feedback to the developer for any glitches and incompetencies. At the same time, human receptionists get to familiarise themselves with AI companions and be equipped with managerial skills.

After 3 months with no issues, the programme will be rolled out islandwide on all other branches.

Phase 3 will last for one year to ensure all human receptionists are familiar with the new work system.

Phase 4: Human Manager with AI Receptionists

In the last of phases, AI receptionists will take independence with no human sitting beside. All counters across all branches will handle customer enquiries with full automation. The branches will commence the extension of opening hours. Human staffs will be deployed in shifts with reduced work days and shorter working hours, assuming the role of branch manager with one manager supervising four AI receptionists.

Risks and Safeguards When Using Generative AI

Generative AI has been a new but transformative technology, the use of GPT-powered autonomous AI agents are even newer in the world. While it enables a boost in productivity through automation, the associated risks should not be ignored.

Security

While the bank can enhance its cybersecurity effort, it cannot guarantees the same security level from the third-party GPT vendor, but to accept that it will provide highest level of security on a "at-best" basis.

Privacy

With the chatbot AI agents exposed to huge customers' banking data and still be put on the frontline, some criminals may try to extract sensitive information from the chatbot using jailbreaking prompts.

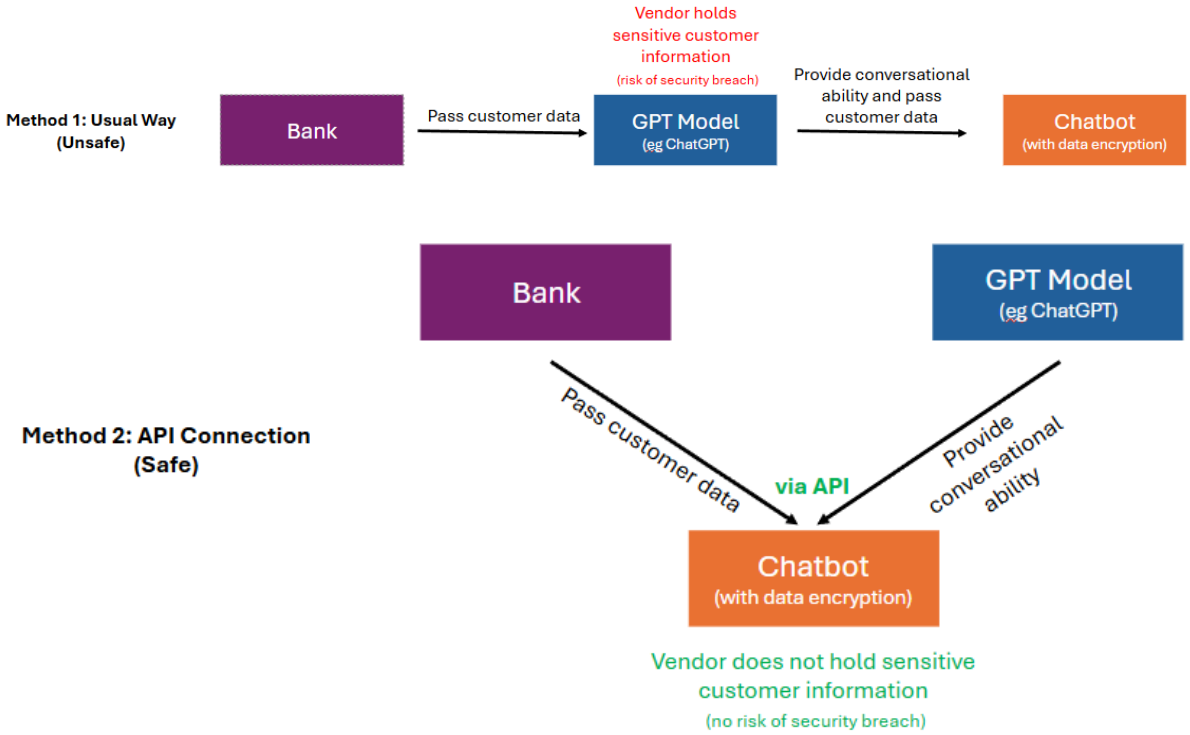
Hallucination

All LLM models are unable to get rid of this side effect whereby it may make up an answer when it does not know the answer, leading to misinformation.

A systemic structure can be put in place to overcome the known risks.

1. API Connection

Instead of pouring all banking information to the GPT based model, connect the GPT to the bank's data via an API. This ensures that vendor has no access to any bank information, hence no risk of data leakage should the security of third-party vendor be breached.



2. REST API + Timer and Sensor

REST stands for Representational State Transfer. It is a “stateless” architecture that has no memory. When a customer uses the chatbot online, with permission given by customer, the AI chatbot agent may extract customer’s information from the system to help solve the problem. When the session ends, the information in the chat are all erased and no longer retrivable. Another customer using the chatbot from another computer will not be able to extract information since different chatrooms do not share the same memory, and the memory will be erased after every session.

To strengthen the safeguard, each online chatbot will be logged out after a short period of inactivity, and each on-site AI receptionist will restart when the customer in front leaves his seat (using sensor technology). This will ensure all memories erased before anyone could infiltrate.

3. Constant Monitoring of LLM

AI hallucinates when it does not have an answer. The Head of Receptionist and AI developer should constantly observe and feed it with proper information whenever hallucinations occur. The more information it has, the less likely it hallucinates. After multiple testing, it should be equipped with sufficient knowledge for proper functioning.

Tackling the Seven Dimensions of Risks – Challenges and Solutions

The use of generative AI in this strategy paper adopts MAS' Fairness, Ethics, Accountability and Transparency (FEAT) Principles and is planned while considering the expanded seven dimensions of risks highlighted in Veritas initiative.

Fairness and Bias

The AI receptionists can speak in four languages, catering to the key ethic groups in Singapore, ensuring no under-representation of minor groups.

Ethics and Impact

The chatbot will be trained with financial literacy and risk management knowledge, to prevent misleading customers into taking on unjustified risks or leverages. As a safety precaution, it will only be allowed to provide basic financial knowledge and will direct customers to consult the professional sales team for financial advices.

Accountability and Governance

The AI developer team and receptionist team will constantly be providing oversight on AI performance to ensure accountability on top of AI's autonomous work.

Transparency and Explainability

The training data and guardrails used on the third-party LLMs would likely be uncontrollable and not transparent. To further filter potential malicious output, the AI developers will craft and input secondary guardrails through custom instructions. A new SOP will be created to mandate these instructions be listed out to the management and compliance team for transparency.

Legal and Regulatory

RESTful API will be used to ensure all conversations be erased after each chat session, eliminating the risk of PDPA and privacy data infringement. All customers' data shall remain secure within the bank's internal system, with no way of leaking through conversations with chatbot.

Monitoring and Stability

The AI developers will constantly monitor to prevent nonsensical outputs and minimised hallucinations. When it does due to a lack of knowledge, the developers will provide the necessary information to patch the gap. A routine audit will also be carried out to ensure no malicious add-on or negligence on the custom instructions.

Cyber and Data Security

The bank will continue to ensure highest level of cyber security across its systems for maximum data security. API connection will be used to prevent banking data from flowing into third party vendor.

A Role Model for Generative AI Usage: Correcting the AI-Replacing-Human Stigma

Social sentiments can turn negative when retrenchments happen due to AI replacing human work, which the branding can take a hit.

By changing the goal of using AI, it can effectively turn the table around. Instead of replacing human receptionists, promote them to be managers that monitor AI receptionists. With the surplus of human receptionists, fit them into new vacancies by extending opening hours. This initiative leads to better customer experience with shorter waiting time, which in turn helps in driving sales, and at the same time, empowering human receptionists into managerial level with no need for retrenchment. It also creates a four-day work week with only six working hours per day at no additional costs, setting the benchmark for a new work-life balance job model.