DSI-42-SG Capstone

AI Voice Assistant for Bank customers

Conrad Aw

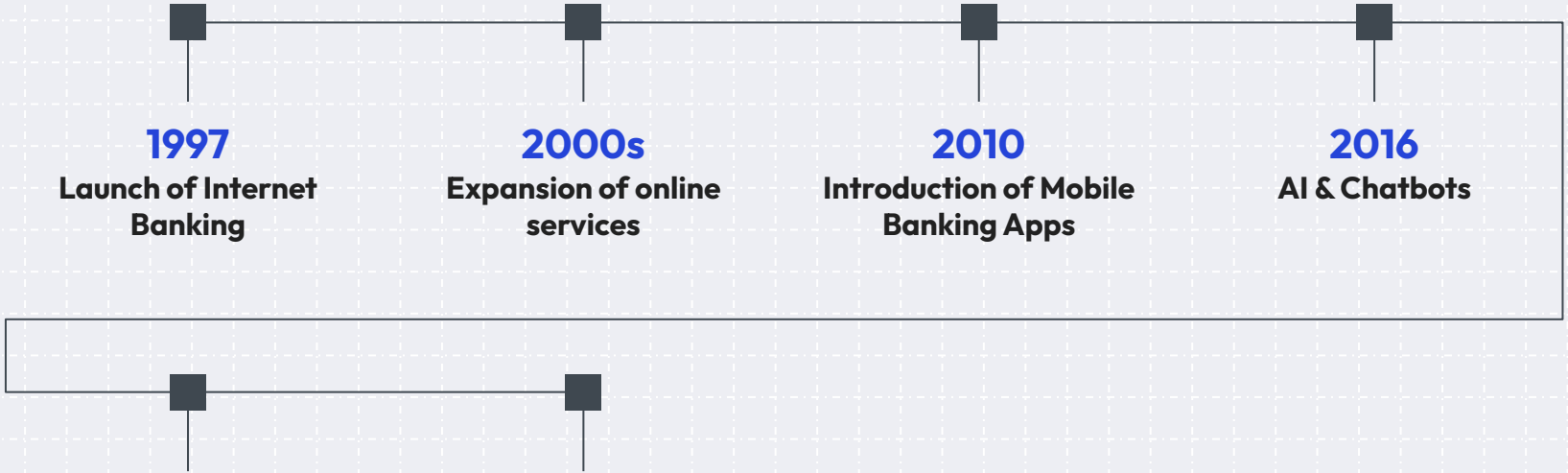
13-05-2024

Table of contents

01

Context & Problem Statement

Introduction



Introduction

cna Top Stories Latest News East Asia Asia Singapore Commentary Sustainability CNA Insider Lifestyle Watch Listen + All Sections

Singapore

The Big Read: As banks go big on digital banking, spare a thought for seniors left behind

As transactions essential to everyday life become increasingly automated in recent years, seniors TODAY spoke with — who were raised in the age of passbooks and real bank tellers — said that they feel increasingly inadequate to perform...see more

mothership

Senior S'poreans share why they find online banking so scary

But less scary with the help of kids and grandkids.

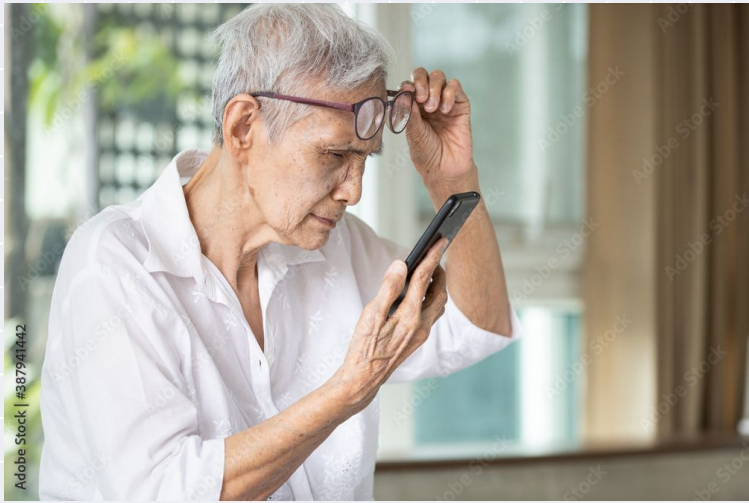
Difficult to find basic banking functions

A 56-year-old parent of a colleague raised a practical concern – that it was difficult to find all the relevant "basic functions" for online banking on the mobile app.

Persona

Cheehong Cao, 68, Retired:

- Declining vision and mobility
- A long-time customer of United Overseas Bank (UOB) and needs to manage his own money as nobody is around to help.
- Motivation: Seeks easy access to banking without the physical need to visit branches.
- Pain Points: He struggles with small text and navigation on standard banking apps due to declining vision and fine motor skills.
- Prefers speaking to a human or human-like interface.



Problem Statement

The rapid digitization of banking has **created accessibility challenges** for elderly retirees with physical impairments, as **traditional digital interfaces demand visual interaction and manual navigation**. This highlights a critical need for an innovative solution that allows these individuals to manage their banking needs independently.

To bridge this gap, this project seeks to develop a speech-to-speech AI chatbot for United Overseas Bank (UOB), **designed to answer FAQs and general queries**. Utilizing Large Language Models for contextual understanding and Retrieval-Augmented Generation for accessing specific knowledge from the bank's FAQs, this chatbot is tailored to **facilitate voice-only interaction, enabling users with physical limitations to engage with bank services informatively.**

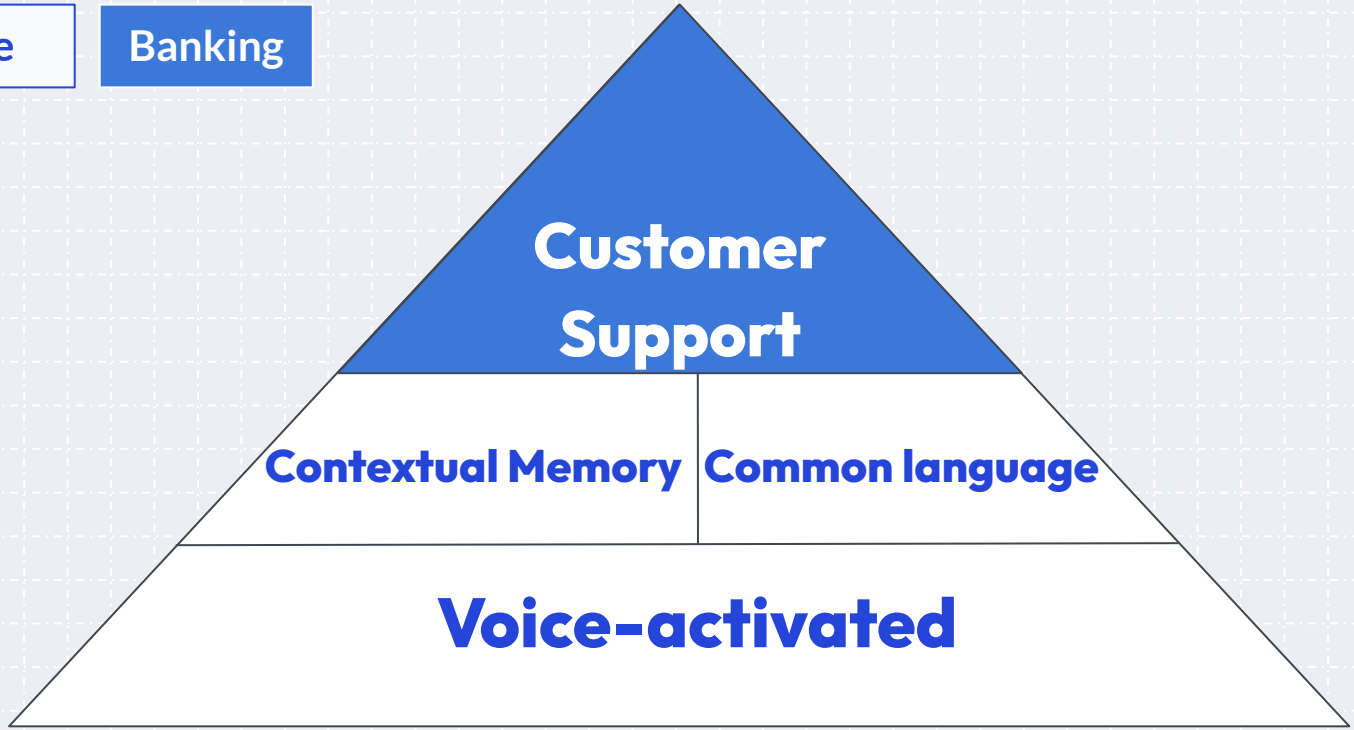
While it does not conduct transactions, the chatbot significantly enhances accessibility by providing reliable information and support, establishing a new standard in inclusive digital banking services.

02

Model Solution, Data Collection & EDA

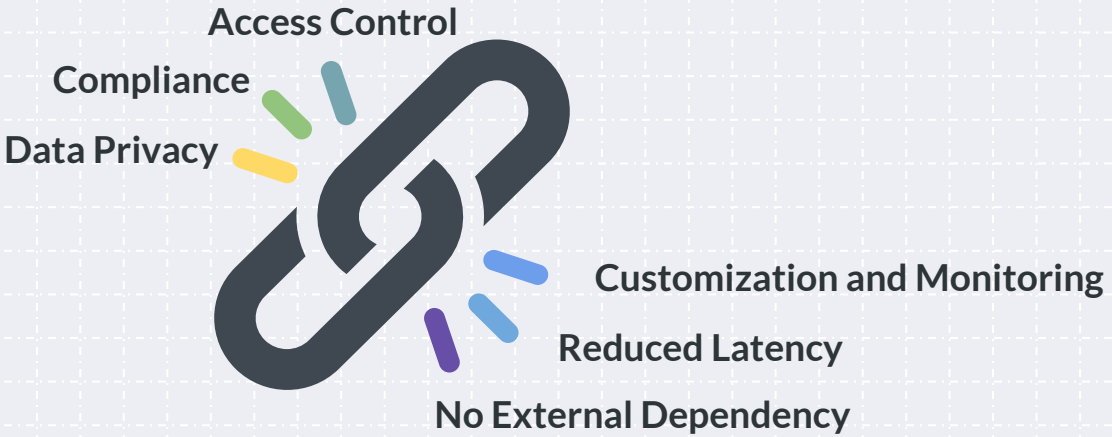
Scope and features of Chatbot

Core Banking



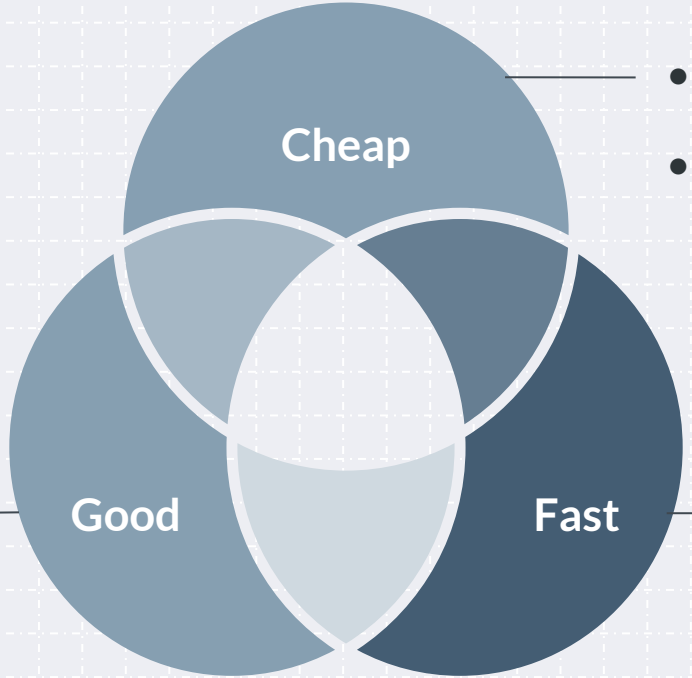
Key Requirement

Running a **local server** for deploying LLM models in banking chatbots is crucial for maintaining strict data security regulations due to several important reasons:



By hosting the chatbot models locally, banks can uphold their responsibility to safeguard sensitive customer data while providing reliable and regulatory-compliant chatbot services.

Key Considerations

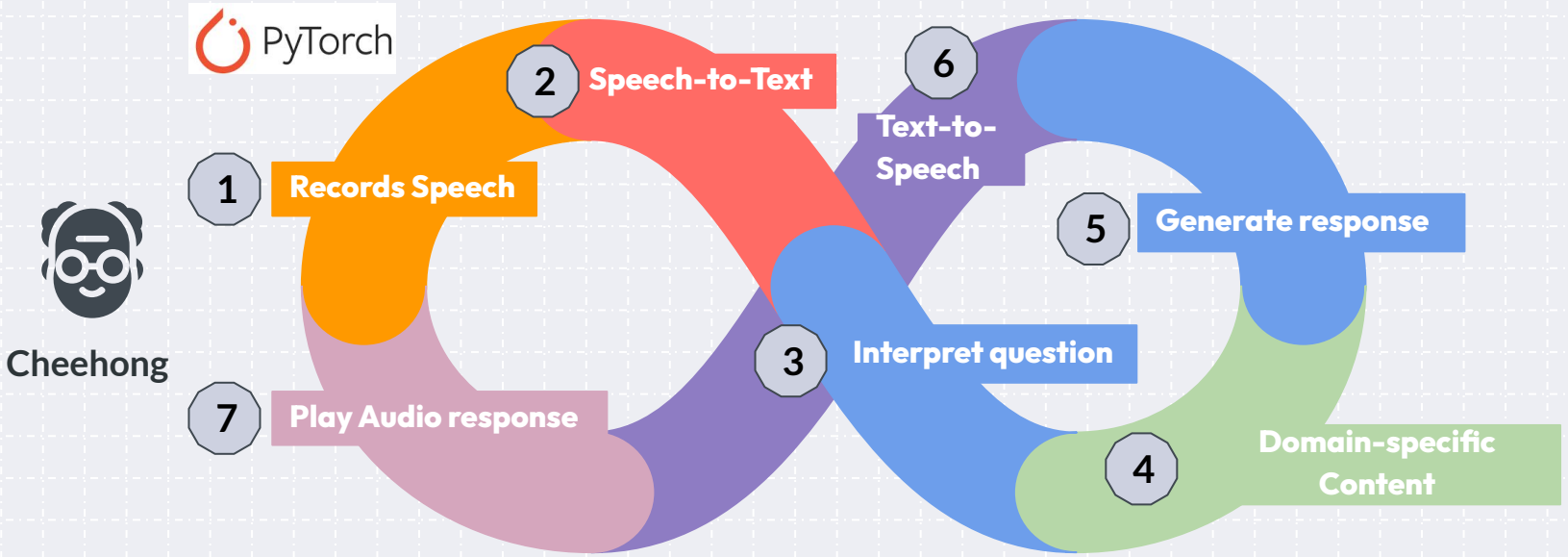


- Open-sourced (Local operation)
- Low consumption of computational resources

- Fit-for-purpose

- Optimized (Low Latency)
- Low consumption of computational resources

Solution Architecture



Speech-To-Text (STT)

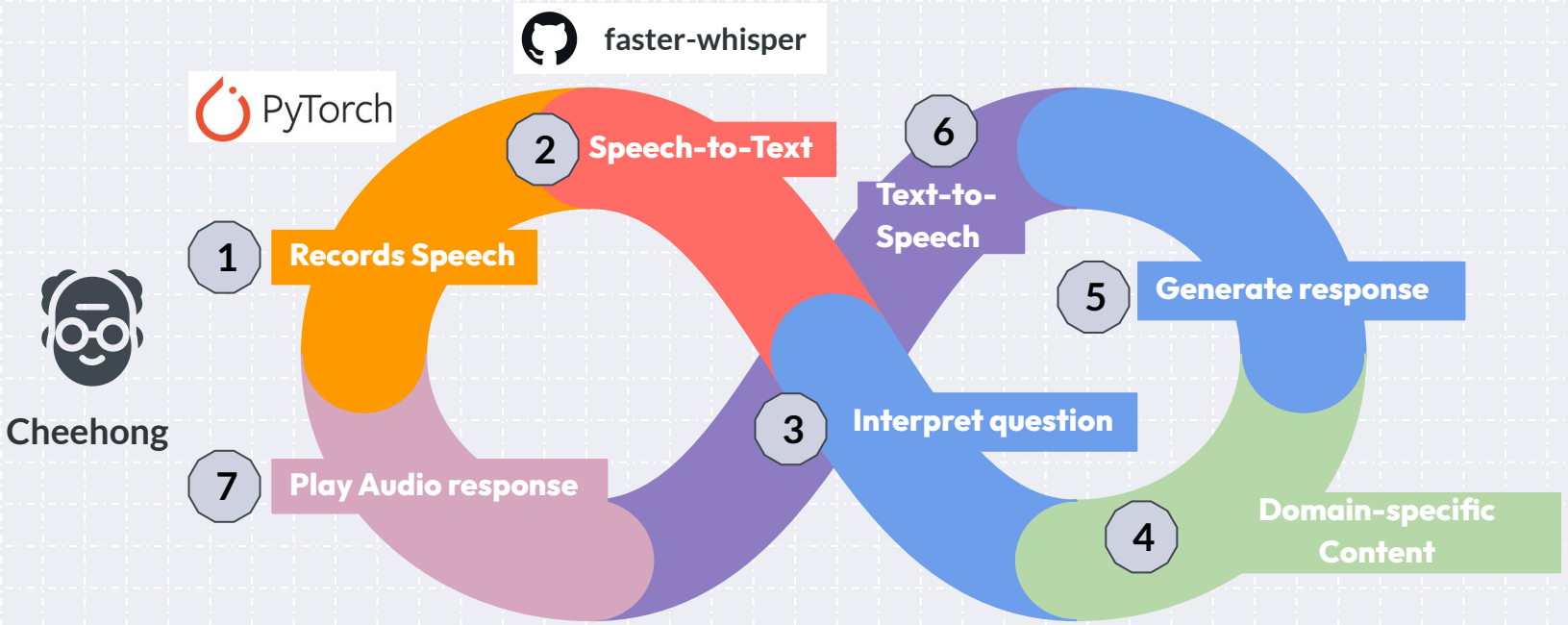
- Voice recognition or speech recognition technology
- Converts spoken language into written text
- Allows digital devices to interpret human speech and transcribe it into text that can be read, stored, or processed.

Comparison & Selection

Model	Device	Audio Length	Transcription Time	VRAM Usage
Original Whisper	Tesla V100S	13 min	4 min 30 sec	11.3 GB
faster-whisper	Tesla V100S	13 min	54 sec	4.7GB

Sources: https://nikolas.blog/making-openai-whisper-faster/; https://pypi.org/project/faster-whisper/

Solution Architecture



Large Language Model

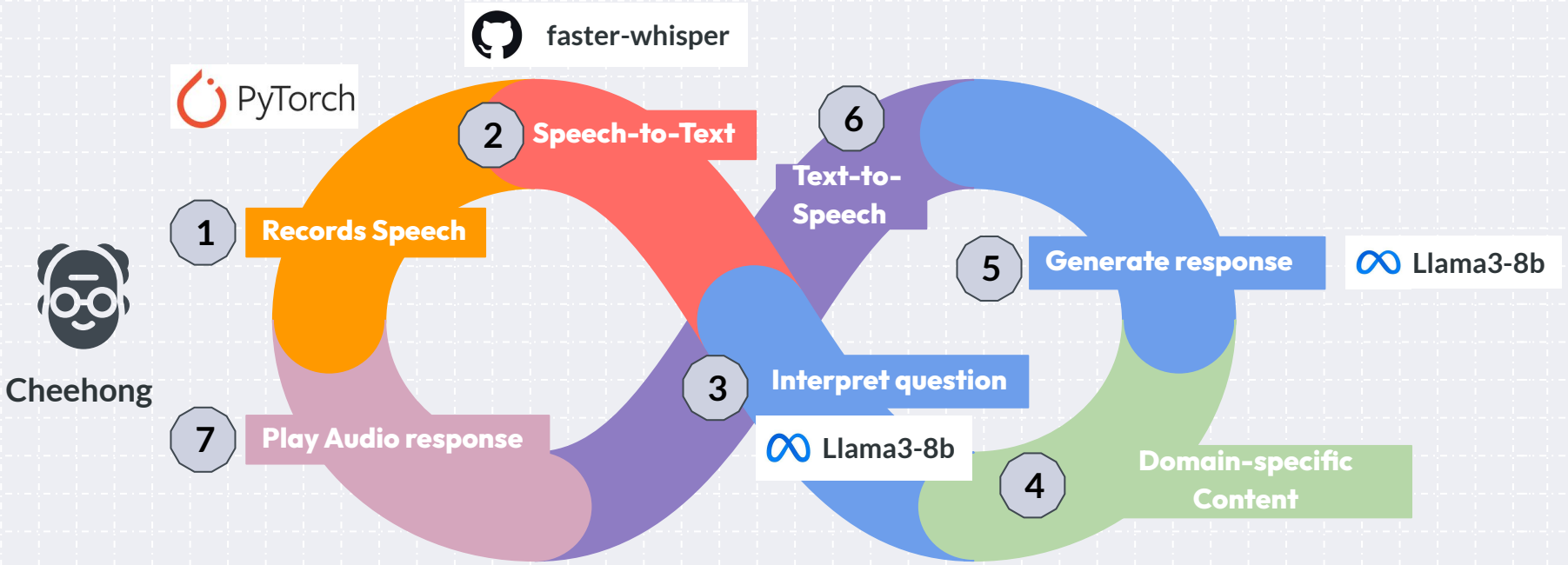
- Designed to understand and generate human-like text by learning from a vast amount of written material.
- Capable in answering questions, having conversations and writing content by predicting what comes next in a sequence of words.

Comparison & Selection

Model	Parameters	Best used for	Languages	Resource requirement	Release Date
Llama3-8B	8 billion	Contextual Tasks	30 languages	Low	Apr 2024
Mistral-7B	7 billion	Contextual Tasks	6 languages	Low	Sep 2023
BLOOM	176 billion	Research/ Coding	46 languages	High	Jul 2022
Falcon-180B	180 billion	Research/ Coding	4 languages	High	Sep 2023

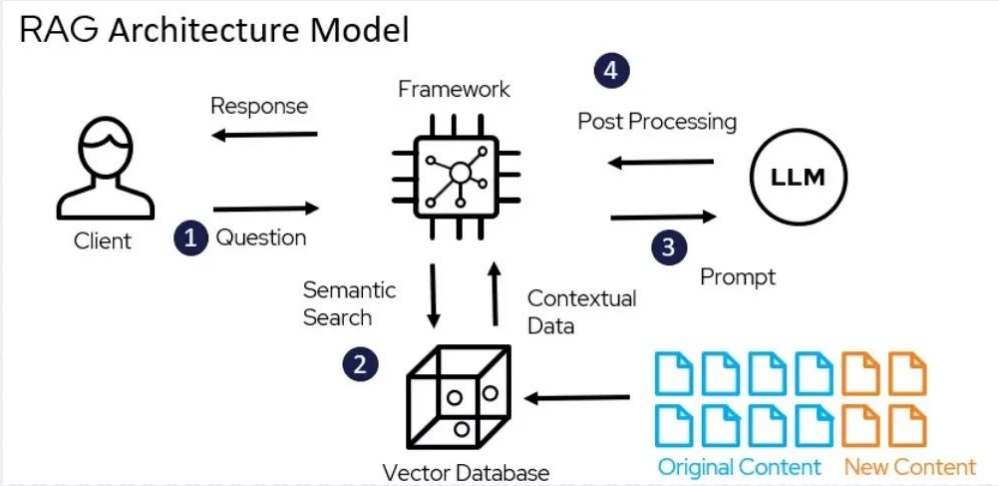
Sources: https://www.scribbledata.io/blog/the-top-10-open-source-llms-2024-edition/; https://www.datacamp.com/blog/top-open-source-llms

Solution Architecture



Retrieval-Augmented Generation (RAG)

- An AI approach that combines the best of both retrieval-based and generative systems
- Introduces new content normally required for domain-specific responses



Source: https://medium.com/@bijit211987/advanced-rag-for-llms-slms-5bcc6fbba411

Data Collection for RAG

- UOB Customer Service FAQ sites
- 7 webpages: General Help, Payment & Transactions, Credit Card, Debit Card, Personal Internet Banking-TMRW, TMRW User Guide, ATM and Loans.
- Selection of web scraping tool:
  - BeautifulSoup
  - Selenium ✅

Exploratory Data Analysis for RAG

Exploratory Data Analysis (EDA) is essential to ensure the dataset's quality, relevance, and diversity. The areas that we will focus on are:

1. Content Quality and Relevance
2. Diversity and Uniqueness
3. Text Statistics
4. Structural Consistency

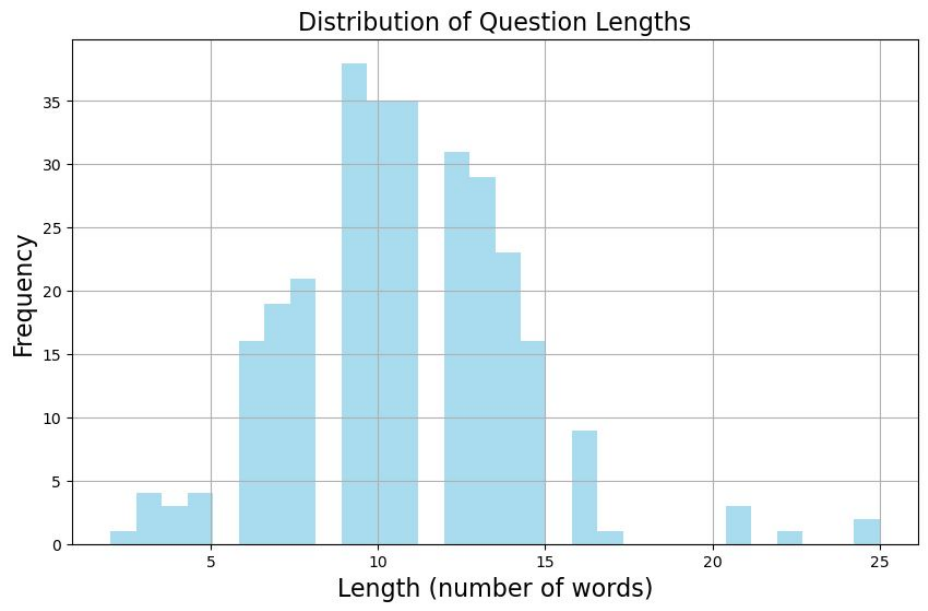
Exploratory Data Analysis for RAG

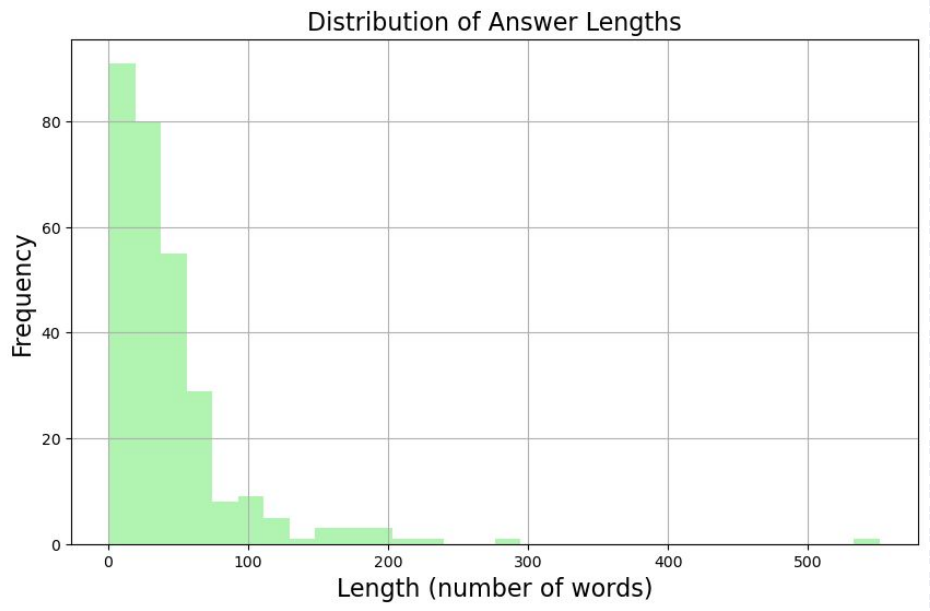
1. **Content Quality and Relevance** ✅
   - Correctness
   - Relevance

2. **Diversity and Uniqueness** ✅
   - Question Uniqueness: 266 out of 291 question are unique
   - Answer Variety: 260 out of 291 answers are unique

Exploratory Data Analysis for RAG

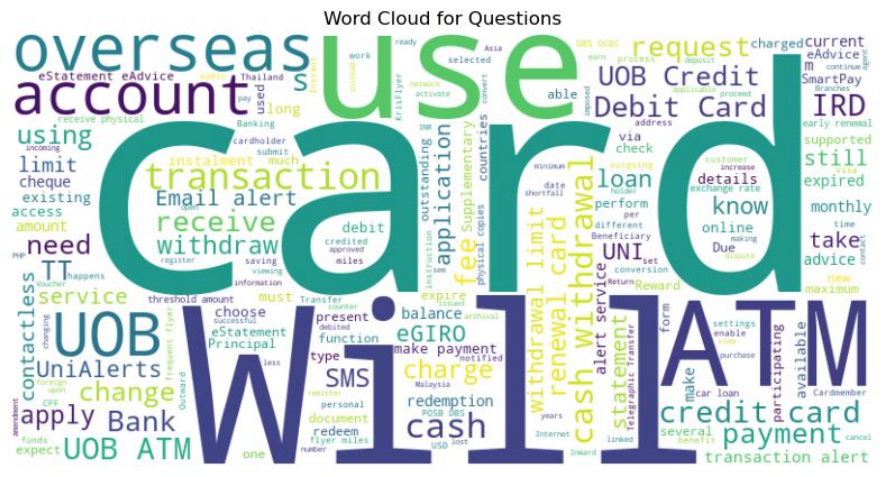
3. Text Statistics - Question and Answer lengths ✅

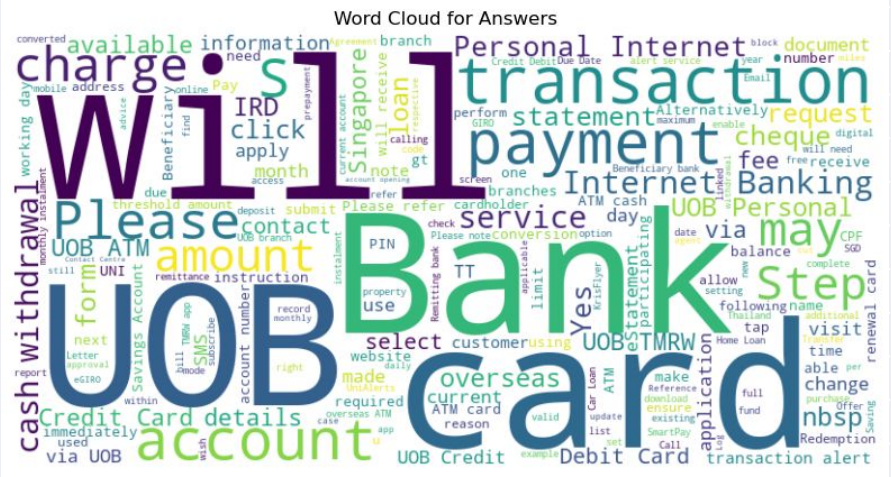




Exploratory Data Analysis for RAG

3. Text Statistics - Word Frequency ✅



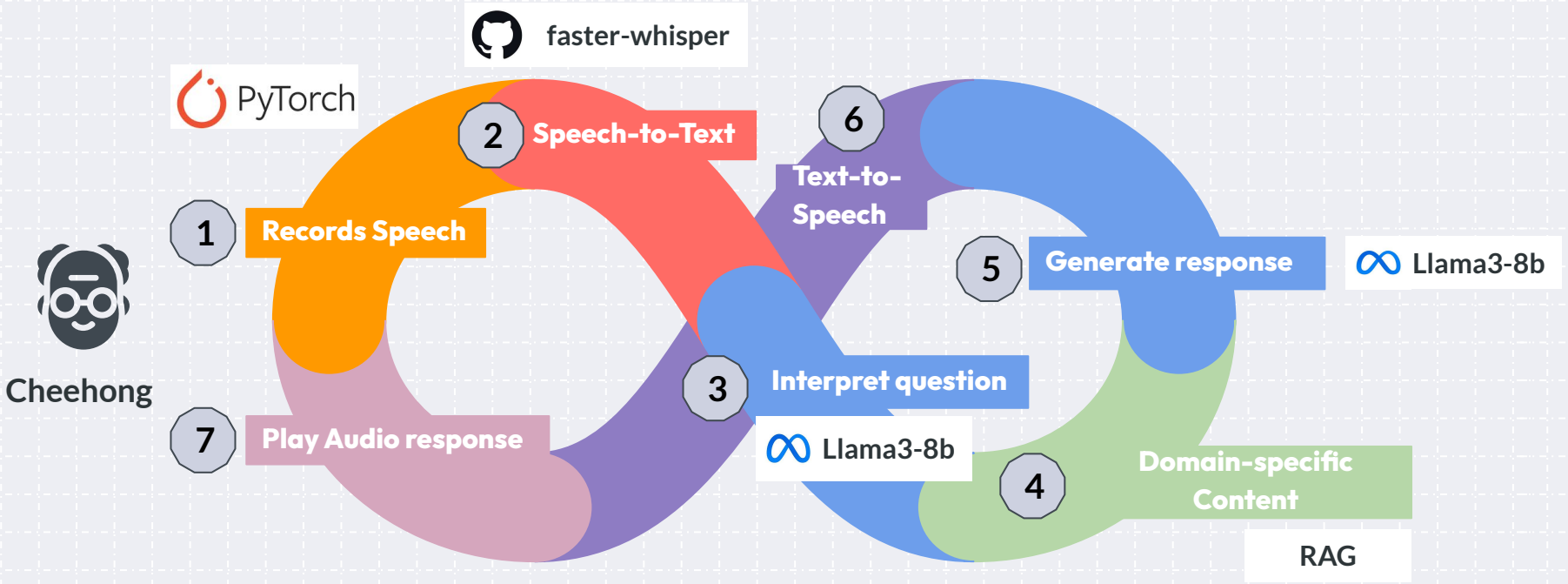


Exploratory Data Analysis for RAG

4. **Structural Consistency** ✅

```
1  Question: I am a Singaporean/PR. How can I open a personal savings account? Answer: For new customers, please bring along your NRIC and initial deposit to any UOB
2
3  Question: I am a Foreigner working in SG. How can I open a personal savings account? Answer: The Bank would require customer to be physically present in Singapore
4
5  Question: What are the service fees associated with my account? Answer: Kindly refer to the following https://www.uob.com.sg/web-resources/personal/pdf/personal/sa
6
7  Question: What types of accounts do you offer? Answer: We offer a myriad of banking accounts to suit your personal needs.Kindly visit the following https://www.uob
```

Solution Architecture



Text-to-Speech (TTS)

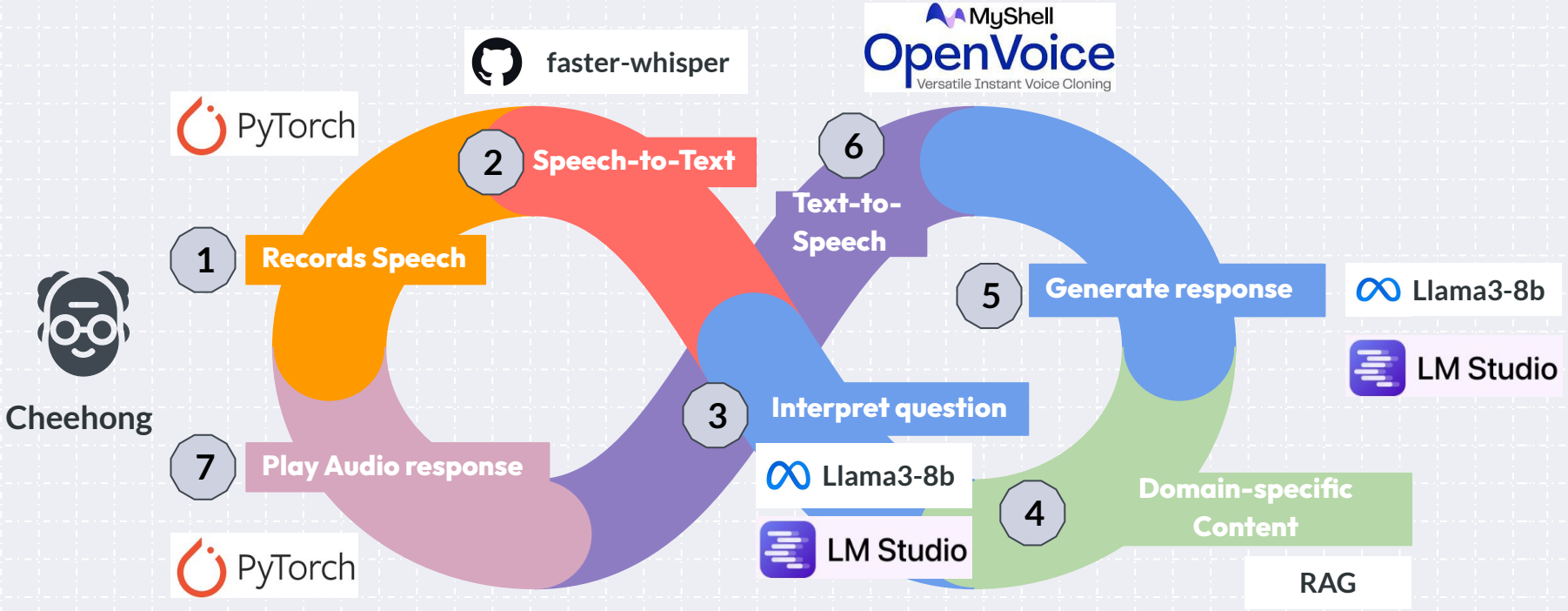
- Converts written text into spoken language
- Natural and human-like voices in multiple languages

Comparison & Selection

Model	Voice Quality	Languages	Customization	Integration Ease
OpenVoice	High, supports emotional nuance and tone control	Multiple, including accents	High, with voice cloning control	Good
MaryTTS	Decent, more robotic	Multiple, limited accents	Moderate, some voice building tools	Moderate
ESpeak	Very robotic, synthetic	Extensive	Low	Good
Coqui TTS	Good	Multiple	High	Good

Sources: https://github.com/myshell-ai/OpenVoice; https://github.com/marytts/marytts; https://espeak.sourceforge.net/; https://github.com/coqui-ai/TTS

Solution Architecture



03

Demonstration

Problem Statement

The rapid digitization of banking has **created accessibility challenges** for elderly retirees with physical impairments, as **traditional digital interfaces demand visual interaction and manual navigation**. This highlights a critical need for an innovative solution that allows these individuals to manage their banking needs independently.

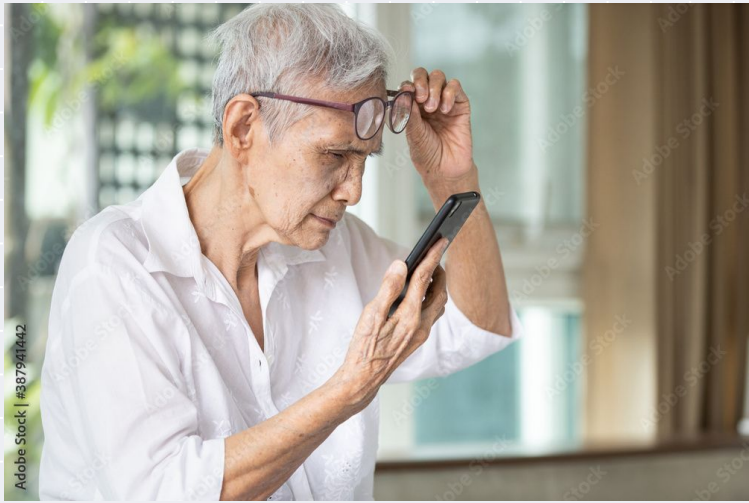
To bridge this gap, this project seeks to develop a speech-to-speech AI chatbot for United Overseas Bank (UOB), **designed to answer FAQs and general queries**. Utilizing Large Language Models for contextual understanding and Retrieval-Augmented Generation for accessing specific knowledge from the bank's FAQs, this chatbot is tailored to **facilitate voice-only interaction, enabling users with physical limitations to engage with bank services informatively.**

While it does not conduct transactions, the chatbot significantly enhances accessibility by providing reliable information and support, establishing a new standard in inclusive digital banking services.

Persona

Cheehong Cao, 68, Retired:

- Declining vision and mobility
- A long-time customer of United Overseas Bank (UOB) and needs to manage his own money as nobody is around to help.
- Motivation: Seeks easy access to banking without the physical need to visit branches.
- Pain Points: He struggles with small text and navigation on standard banking apps due to declining vision and fine motor skills.
- Prefers speaking to a human or human-like interface.

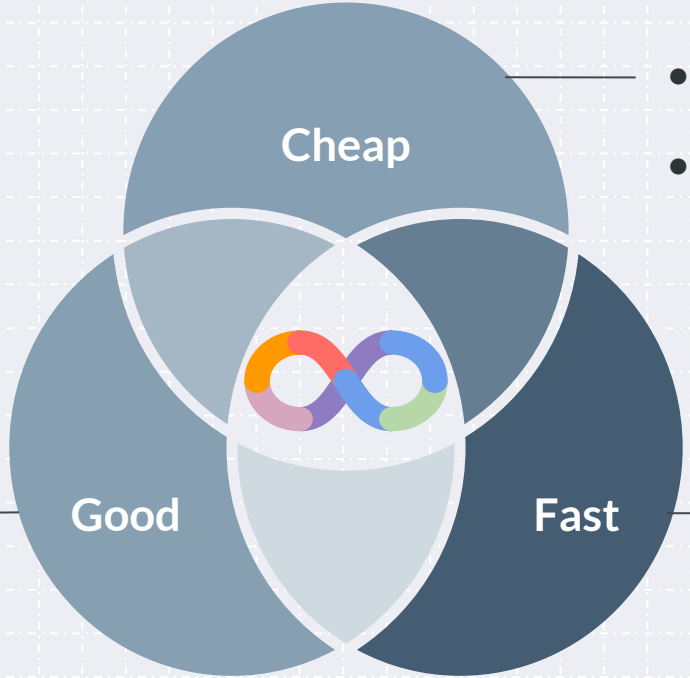


Persona

Cheehong Cao, 68, Retired:

- The speech-to-speech chatbot can provide John with clear, conversational interactions, making banking more accessible and less intimidating.
- A **satisfied** long-time customer of United Overseas Bank (UOB) that is able to seek answers for his banking queries easily.
- Enjoys speaking to the AI Assistant since its free.

Key Considerations



- Open-sourced (Local operation)
- Low consumption of computational resources

- Fit-for-purpose

- Optimized (Low Latency)
- Low consumption of computational resources

04

Limitations & Recommendations

Limitations

- **User Interaction**
  - Currently only caters to English-speaking customers.
  - Voice Authentication
- **Scalability, Performance and Services**
  - Scalability: Handling real-time audio processing and interaction in a scalable way is challenging.
  - Performance: There is still a distinct latency issue for response
  - Scope: Limited to only Q&A and unable to conduct transactions like the traditional digital applications.

Recommendations

- **Multilingual Support**
- **Performance Optimization**
- **Security Enhancements**
- **Integrate with system to provide more services**
- **Expand Test Coverage**

By addressing these areas, the chatbot can be significantly improved and made ready for wider deployment in scenarios demanding high interactivity and user engagement.

Thank you for your attention