

# Using RAG and LangChain to Drive Chatbots for Interactive Economic Data Analysis and Insight\*



Seah Boon Keong, Eilyn Chong, Max Chew  
Central Bank of Malaysia  
Monetary Policy Department

(\*) The views and conclusions presented in this paper are exclusively those of the author(s) and do not necessarily reflect the position of the Central Bank of Malaysia or its Board members.



- Motivation
- Methodology and Approach
- Usage
- Future Work
- Conclusion

## Advancing Economic Data Access via Innovation in Chatbot Technology

- **Problem Statement:** Users often struggle with fragmented and complex economic datasets, requiring specialised knowledge to interpret.
- **Solution:** Our chatbot provides an intuitive interface that simplifies data access and analysis, making it easier for users to gain insights.



Democratize accessibility to vital economic data, extending the benefits to a broader user base including policymakers, researchers, financial analysts.



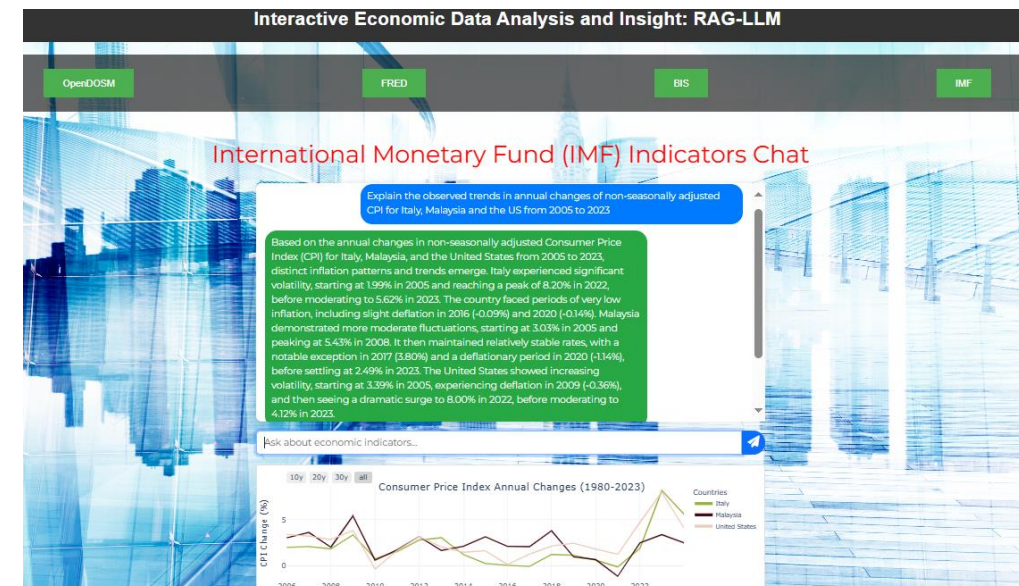
Ingest a wide array of metadata sourced from numerous authoritative institutions.



Offer an interactive and user-friendly interface to explore and analyze intricate economic datasets, thereby making complex data more comprehensible



Utilizes advanced technologies such Retrieval-Augmented Generation (RAG) and Large Language Model (LLM) and Dynamic Charting capabilities



## Data Sources

- We source publicly available economic datasets from authoritative institutions such as BIS, FRED, and IMF, as well as OpenDOSM for Malaysia's data

## Types of Data

- The system ingests from various economic datasets, including macroeconomic indicators and financial statistics.
- It supports: CSV, Parquet, JSON
- The system will retrieve the datasets based on the URLs specified in the metadata, either from the returned API call or from links of CSV or Parquet files.

## 1. Using Retrieval Augmented Generation (RAG) Architecture

Utilizing **relevant information** effectively enhances the output from large language models. Details in the following slide.

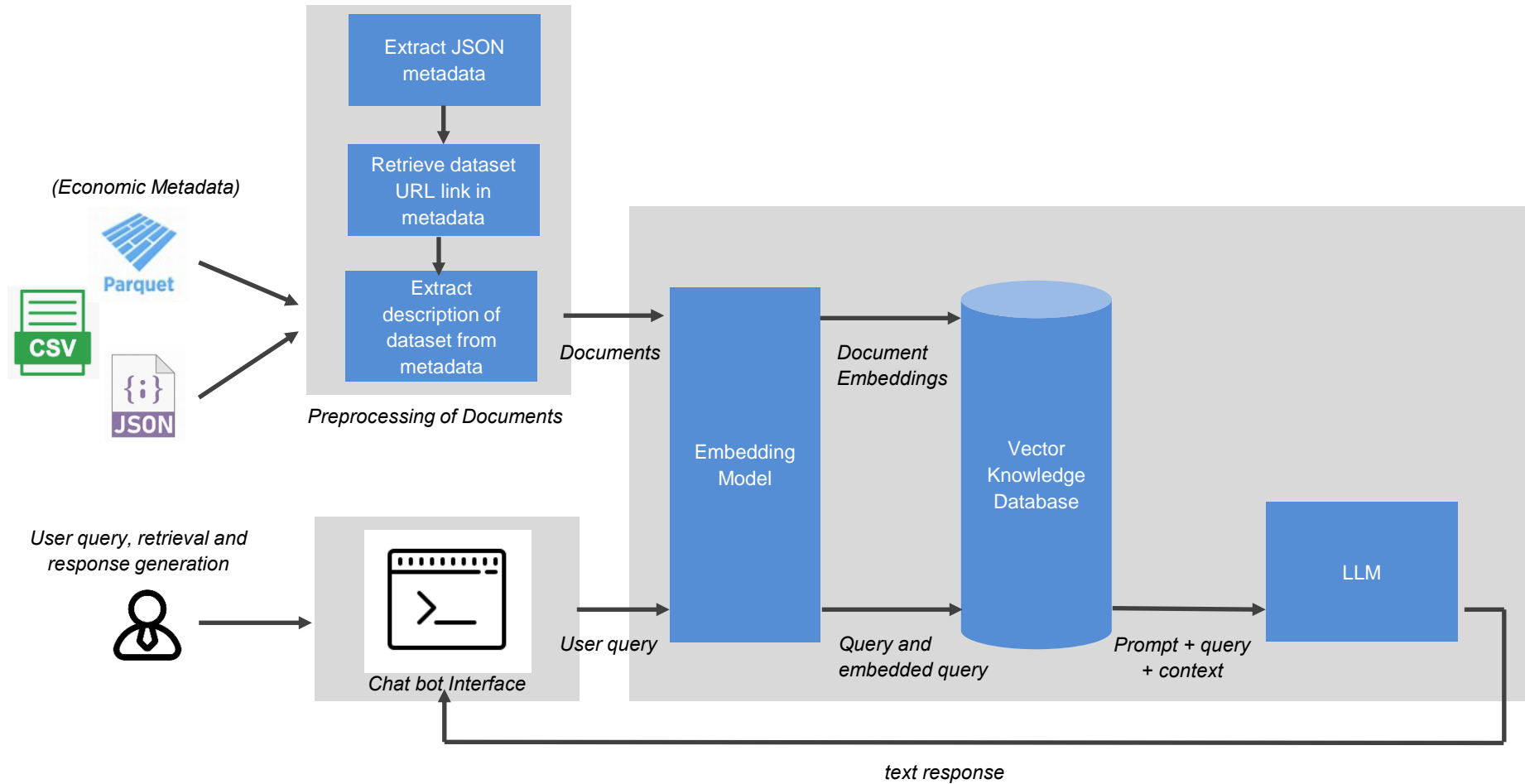
## 2. Data Source Integration

As a proof of concept, this prototype application ingests **datasets** from FRED, BIS, IMF and OpenDOSM for enriched analysis.

## 3. Dynamic Charting Techniques

Employing **interactive visuals** allows for better data interpretation and user engagement.

# Methodology and Approach



# Methodology and Approach

Blend of several AI components in which the key elements are as follow:



# Methodology and Approach

Comparison of FAISS, Chroma, and PineCone based on performance, scalability, ease of use, in-memory storage, API simplicity, and cost efficiency.

Reason for FAISS selection: **Performance and Cost Efficiency**

<i><b>Feature</b></i>	<i><b>FAISS</b></i>	<i><b>Chroma</b></i>	<i><b>PineCone</b></i>
<i><b>Performance</b></i>	<i><b>High</b></i>	<i><b>Moderate</b></i>	<i><b>High</b></i>
<i><b>Scalability</b></i>	<i><b>High</b></i>	<i><b>High</b></i>	<i><b>High</b></i>
<i><b>Ease of Use</b></i>	<i><b>Moderate</b></i>	<i><b>High</b></i>	<i><b>High</b></i>
<i><b>In-Memory Storage</b></i>	<i><b>No</b></i>	<i><b>Yes</b></i>	<i><b>Yes</b></i>
<i><b>API Simplicity</b></i>	<i><b>Moderate</b></i>	<i><b>High</b></i>	<i><b>High</b></i>
<i><b>Cost Efficiency</b></i>	<i><b>High</b></i>	<i><b>Moderate</b></i>	<i><b>Moderate</b></i>

**Source:**

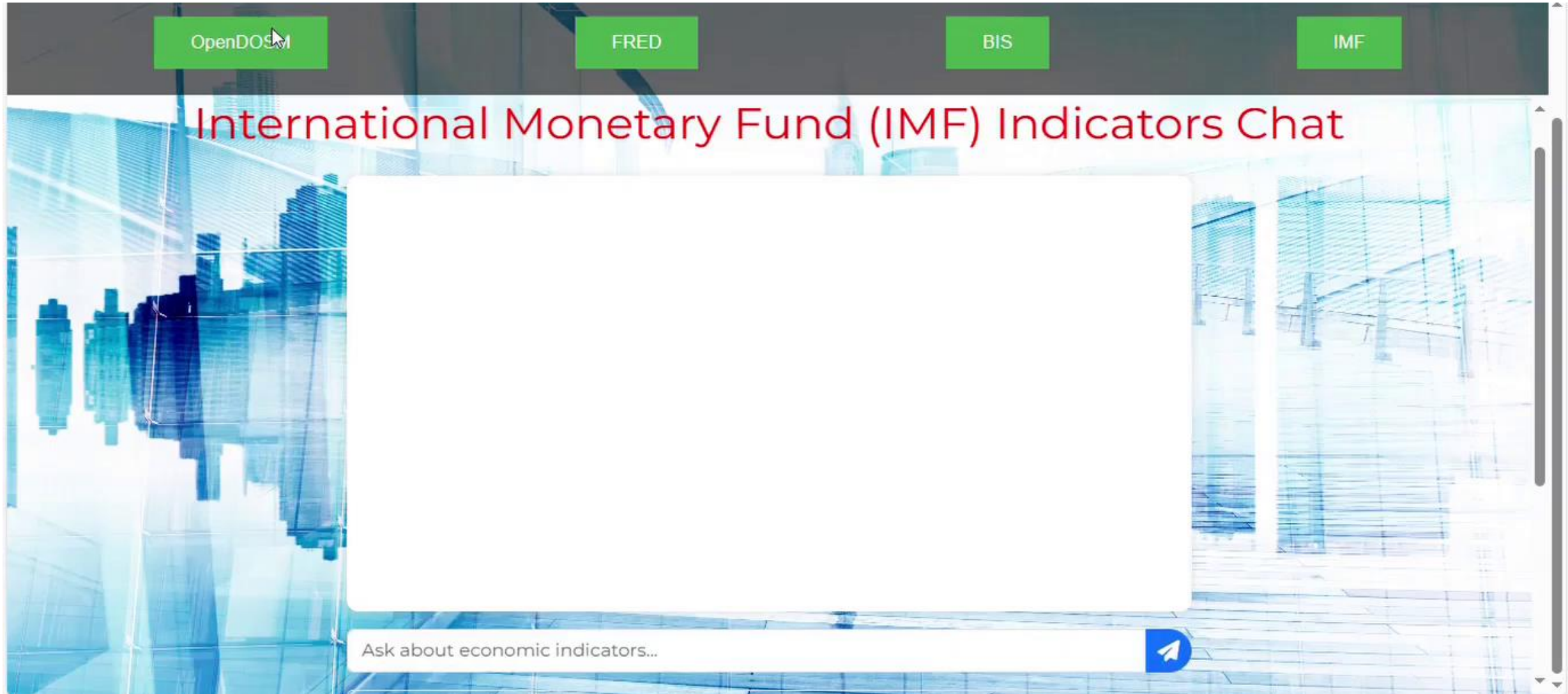
[Pinecone vs FAISS](#), [Pinecone vs Chroma](#)





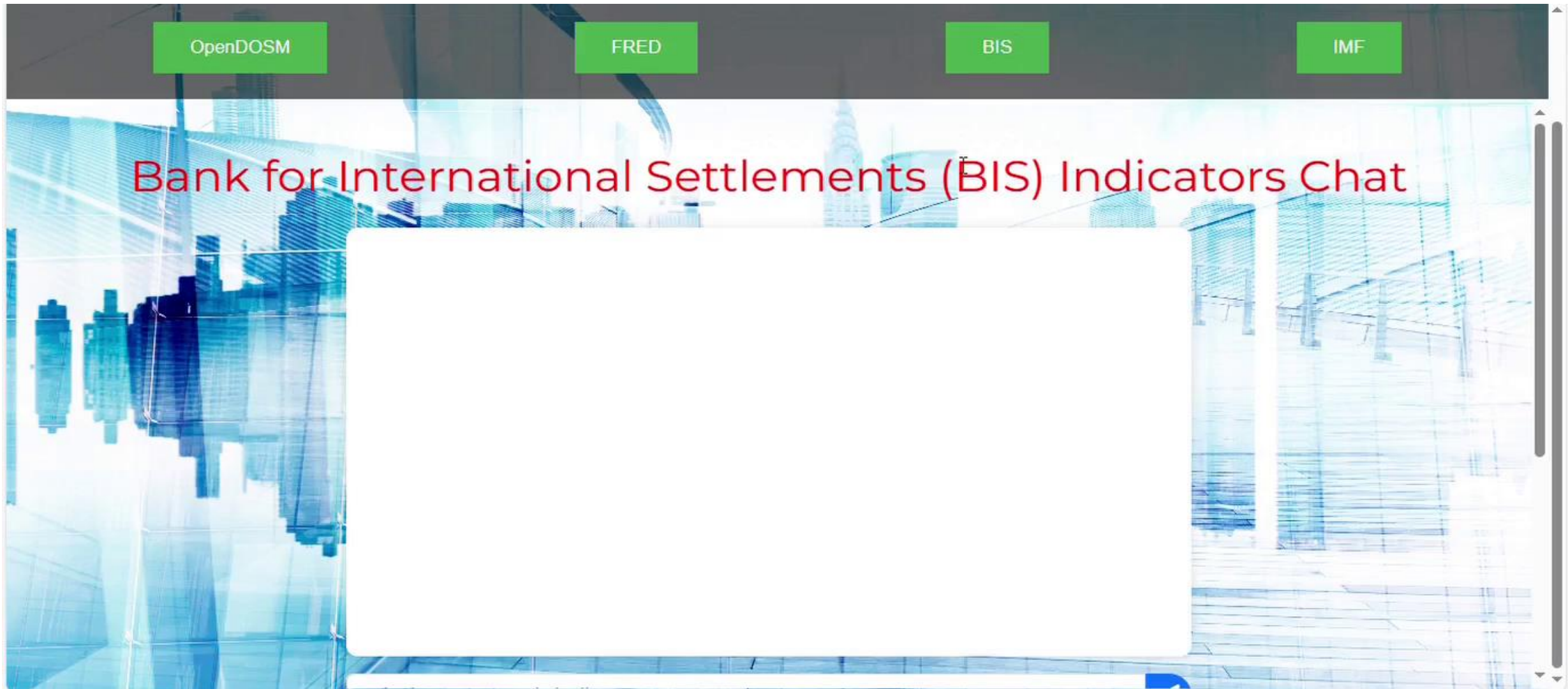
# Usage

An innovative chatbot system that leverages Retrieval-Augmented Generation (RAG) and LangChain to provide an intuitive, interactive interface for exploring and analysing economic datasets. - **Use case sample 1**



# Usage

An innovative chatbot system that leverages Retrieval-Augmented Generation (RAG) and LangChain to provide an intuitive, interactive interface for exploring and analysing economic datasets. - **Use case sample 2**



## Potential Improvements (Prototype - Proof of Concept)

- This prototype (PoC) serves as a foundation for future enhancements. Key areas for improvement include:
  - **Expanding Data Sources**
    - We aim to integrate additional datasets from sources such as BIS, IMF, and FRED. This expansion will significantly enhance the chatbot's ability to deliver comprehensive and insightful analyses.
  - **Enhancing Dynamic Charting**
    - To provide users with a richer and more versatile visualization experience. We plan to expand it by including more chart types, such as bar charts, heatmaps, and scatter plots.

- In summary, our chatbot system democratizes access to economic data, making analysis more accessible and efficient.
- Facilitate faster exploration of economic data to gather data-driven insights and improving efficiency of economic analysis.

# Thank you for your attention!

Seah Boon Keong (Ph.D) [seah.keong@bnm.gov.my](mailto:seah.keong@bnm.gov.my)

Eilyn Chong [eilyn@bnm.gov.my](mailto:eilyn@bnm.gov.my)

Max Chew [kai.xi@bnm.gov.my](mailto:kai.xi@bnm.gov.my)

