ChatGPT

Enhancing Consular Services in Indonesia through AI-Driven Chatbots and Retrieval-Augmented Generation

Student Name: [Name Placeholder]
Institution: [Institution Name], Department of [Department Name]
Advisor: [Advisor Name Placeholder]
Date: [Month Year]

Abstract

Indonesian consular offices handle a high volume of routine inquiries (e.g. visa requirements, passport renewal, emergency assistance) with limited staff and outdated information channels. This thesis proposes developing a conversational AI system—a retrieval-augmented chatbot powered by deep learning—to improve information access and efficiency in consular services. The research will investigate how large language models (LLMs) augmented with a curated knowledge base (using RAG techniques) can answer citizens' questions about consular procedures accurately and in real time. Key objectives include building a domain-specific dataset from consular FAQs and policy documents, integrating a vector-based retrieval system with an open-source or commercial LLM, and evaluating the system's accuracy and user satisfaction. We will compare the RAG-enabled chatbot against baseline methods (static FAQ search or an LLM without retrieval) using metrics such as answer precision, response time, and user survey ratings. Ethical considerations (data privacy, bias mitigation, and reliability) will be addressed. The expected outcome is a prototype AI assistant that delivers faster, more accurate consular information (24/7), reducing staff workload and improving citizen satisfaction. This work aims to demonstrate the value of AI in streamlining public services and inform broader digitization efforts in Indonesia's foreign affairs sector.

Introduction

Consular services are government functions that assist a country's citizens abroad, including processing visas and passports, providing legal aid, and offering emergency support. These services often involve standard procedures and information (e.g. document requirements, travel advisories, registration) that can overwhelm consular staff when demand is high. In Indonesia, consular affairs are managed by the Ministry of Foreign Affairs and its embassies/consulates worldwide. Currently, information is disseminated through websites (e.g. Portal Peduli WNI, Safe Travel app) and limited call centers. This setup can lead to inefficiencies: citizens may face long waiting times, out-of-date guidance, or language barriers when seeking help. For example, many visa or passport queries are routine and repetitive, yet they require staff time. As one analysis notes, "many consular inquiries—such as visa requirements, document checklists, or passport renewal—are repetitive and straightforward. Chatbots can handle these efficiently, freeing human staff to focus on complex cases" [1] .

The problem addressed in this thesis is the lack of efficient automation and up-to-date information retrieval in Indonesian consular services. Current systems rely heavily on manual processes and static web content, which may not provide timely answers. For instance, government advisories can change rapidly (e.g. travel bans, emergency evacuations), but static FAQs may lag behind. Without automated support, consulates cannot offer 24/7 assistance to citizens in different time zones.

To address this, we propose designing a **RAG-enhanced chatbot** that combines **deep learning (LLMs)** with a **retrieval system** over official consular knowledge. The central research questions are:

- **RQ1:** How can a chatbot combining LLMs and retrieval (RAG) improve the accuracy and relevance of consular information provided to users?
- **RQ2:** What data and design are needed to tailor a RAG chatbot to the Indonesian consular context (languages, topics, official policies)?
- **RQ3:** How does the proposed system compare (in accuracy, response time, and user satisfaction) to existing information channels or simpler chatbots?

The objectives are to (1) construct a comprehensive consular knowledge base (from Indonesian MFA documents, embassy FAQs, etc.), (2) integrate it with a state-of-the-art LLM using RAG techniques, (3) implement a chatbot interface for user queries, and (4) evaluate the system's performance via experiments and user studies. The significance lies in demonstrating a scalable AI solution for government services: improving citizen access to information, reducing operational costs, and showcasing the potential of AI in Indonesian public administration. Successful results could inform wider adoption of AI in other public service domains and support the Indonesian government's digital transformation goals.

Literature Review

AI in Consular and Government Services: Globally, governments are increasingly using AI chatbots to enhance citizen services. For example, the United States Department of State piloted AI chatbots for visa inquiries, allowing applicants to get instant updates on their application status [2] . In the UK, "VisaBot" trials have helped users navigate complex immigration rules. India's consular grievance portal "MADAD" uses a chatbot to assist citizens with complaints or emergencies abroad [3] . Other examples include Australia's "Sophie" chatbot for visa and citizenship queries [4] . At the municipal level, cities like Kansas City and Los Angeles have deployed bots (OpenDataKC, CHIP) to answer civic queries via social platforms [5] [6] . These cases show that conversational agents can handle routine queries 24/7, scale under high demand (e.g. crises), and provide multilingual support [7] [8] .

VFS Global, a global visa outsourcing service, recently launched an AI-powered chatbot for UK visa applicants in 141 countries [9] . This conversational agent uses generative AI (supporting voice/text) to deliver accurate visa information, reducing dependence on support staff [10] . Similarly, private firms like BLS International are exploring AI to automate consular processes (e.g. document verification, risk assessment) as noted in industry reports 【58†】 . These multinational examples indicate a clear trend: both governments and service providers recognize AI's potential to make consular services more efficient.

AI in Indonesian Public Services: Domestically, Indonesian agencies have begun applying AI in public sectors. For instance, the Ministry of Finance's tax directorate (DJP) developed an AI chatbot to help taxpayers find information and file returns more easily [11] . The National Disaster Management Agency (BNPB) deployed "BencanaBot," a chatbot on WhatsApp/Telegram, to allow citizens to report disasters in

real time [12] . These initiatives show feasibility but also highlight challenges: Indonesia faces fragmented data systems and uneven internet access [13] . A survey notes that Indonesia still lags peers in AI adoption due to limited connectivity and data readiness [14] . Thus, while AI chatbots are emerging (e.g. Kominfo hoax detection, education and health superapps), dedicated work is needed to align technology with on-the-ground needs [15] [16] .

RAG and LLMs in Public Sector Applications: Retrieval-Augmented Generation (RAG) is a recent technique where an LLM's responses are grounded in a dynamic knowledge base. Studies and industry sources emphasize that RAG can bridge the "knowledge cutoff" problem of static LLMs by fetching up-to-date facts at query time [17] [18] . For example, an AWS public sector blog describes how RAG integrates an LLM with current data (laws, policies) to provide accurate, context-specific answers [19] [18] . State and local governments in the US are already experimenting with RAG: leaders report it makes chatbot answers "more accurate, and more specific to what [citizens] want to know," by retrieving agency data on the fly [20] . Academic studies also find RAG chatbots improve response relevance in specialized domains (e.g. medical or legal AI assistants) by incorporating domain documents [21] .

Despite these advances, few studies focus on RAG for consular contexts, especially in the Global South. There is a gap in research on customizing RAG chatbots for Indonesian consular needs and local languages. Existing literature highlights potential but notes constraints: data privacy and content sensitivity in consular AI must be carefully managed [22] [23] . Additionally, design must consider user accessibility (mobile usage, local language support). This proposal will build on the global insights (e.g. MFA AI pilots [2] , VFS case [10]) and local efforts (DJP tax bot [11] , SARI mobile bot [24]) to address the specific gap of an Indonesian consular RAG chatbot. The expected contribution is a demonstration of how retrieval-augmented LLMs can be applied to improve public sector information services in a developing country setting.

Methodology

System Overview: We propose a conversational agent combining a **Large Language Model (LLM)** with a **Retrieval system** (RAG architecture) to answer consular queries. Figure 1 illustrates this approach: the user's question triggers a similarity search over a consular knowledge base to fetch relevant documents, which are then supplied as context to the LLM for answer generation [18] [20] .
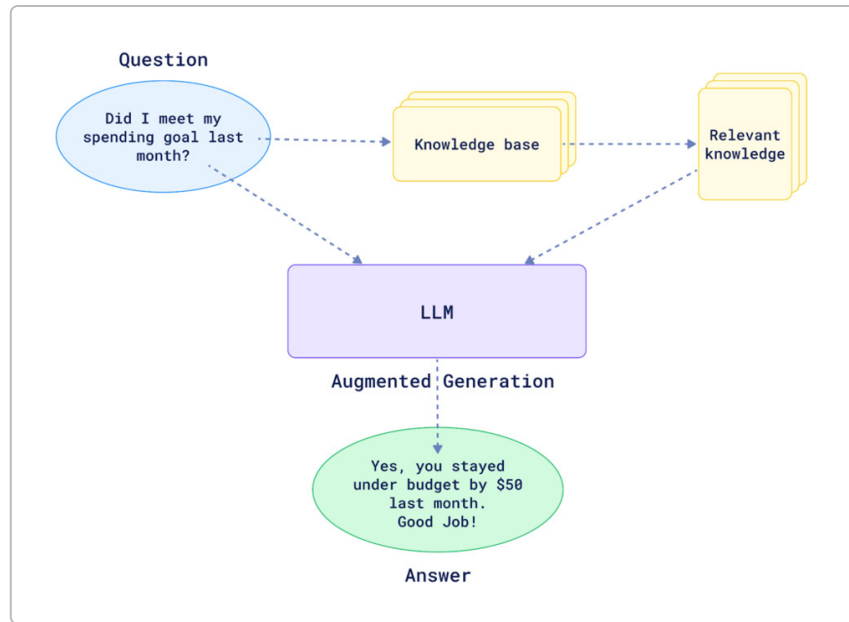
Figure 1. Conceptual flow of a Retrieval-Augmented Generation chatbot: a user question queries a knowledge base, retrieving relevant information that augments the LLM's response (adapted from Qdrant 25).

Data Collection: We will gather official consular content in Indonesian and English, including published FAQ lists, embassy circulars, travel advisories, and regulatory texts (e.g. visa laws, COVID-19 travel rules). Sources include the Indonesian MFA website, embassy/consulate portals, and international guidelines from allied countries (for comparative FAQ content). We will also collect anonymized transcripts of actual consular inquiries (if available) to understand user language. Data will be cleaned and chunked into passages (e.g. 100-300 word segments) for retrieval.

Model and Retrieval: The core model will be a pre-trained LLM (e.g. GPT-3.5/GPT-4 or an open model like LLaMA/BLOOM) with a vector-based retrieval layer. We will use a sentence embedding model (e.g. multilingual BERT) to encode all knowledge base passages into vectors stored in a vector database (e.g. FAISS or Qdrant). At runtime, a user's query is also embedded and compared (via cosine similarity) to retrieve the top-k relevant passages. These passages are then concatenated with the original query to form a prompt for the LLM. Prompt engineering (templates) will ensure the context is used as knowledge and to encourage concise answers.

Prototype Development: We will implement a proof-of-concept chatbot interface (web or messaging app) where users can type questions. The backend will handle the retrieval and LLM querying. If using a closed-source API (e.g. OpenAI), we will manage usage constraints (e.g. token limits) by summarizing or truncating context as needed. We will integrate multilingual support: the system should accept queries in Indonesian or English and respond accordingly. We may fine-tune or prompt-tune the model with a small set of Indonesian-consular QA pairs to improve domain-specific understanding.

Evaluation: The system will be evaluated on (a) **Accuracy** of answers: We will compile a test set of consular questions with gold answers (from MFA docs or expert-written). Metrics include exact match or F1 overlap. (b) **Response Quality**: Human evaluators (language speakers) will rate answers on clarity, correctness, and helpfulness. We will compare the RAG chatbot to a baseline: either an LLM without retrieval or a traditional

keyword FAQ search. (c) **Efficiency**: Measure response time and throughput (e.g. percent queries answered within SLA). (d) **User Study**: A small user study with Indonesian citizens (or consular staff) may be conducted to rate satisfaction and trust. We will analyze improvements over the baseline. Prior medical chatbot research suggests RAG bots can match doctor-given information clarity while reducing consultation time [21]; we will check if similar gains occur for consular info.

Ethical Considerations: Handling consular queries involves privacy-sensitive data (personal travel details, legal status, possibly asylum issues). We will ensure data anonymization and encrypt any personal data. The chatbot will explicitly not store conversation histories. We will include disclaimers in the interface that the AI provides guidance, not legal advice, and provide an option to contact a human officer. To mitigate bias, training data will be checked for fairness (e.g. gender-neutral phrasing) and vetted by domain experts. All responses will be constrained to factual information; we will implement "safe-completion" settings to reduce hallucinations. We will follow Indonesian data protection norms and consult MFA guidelines to ensure the AI system aligns with diplomatic privacy and security standards [22] [26].

Timeline

The 6-month plan is divided into the following phases:

1. **Month 1 – Project Initiation and Data Gathering:** Conduct detailed literature review and finalize research design. Obtain consular documents, FAQs, and any available query logs. Engage with MFA contacts to understand requirements.
2. **Month 2 – Data Processing and Knowledge Base Construction:** Clean and preprocess collected data. Chunk documents and encode into vector embeddings. Set up the retrieval system (vector database).
3. **Month 3 – Prototype Model Development:** Integrate a chosen LLM with the retrieval system. Implement prompt templates and chatbot interface. Perform initial tuning (few-shot examples if needed).
4. **Month 4 – Internal Testing and Refinement:** Test the chatbot internally using a draft Q&A set. Identify failure modes and refine retrieval (e.g. adjust embedding model or chunk size) and prompting strategy.
5. **Month 5 – Evaluation:** Prepare a test suite of consular questions. Conduct quantitative evaluation against baseline (accuracy, response time). Run a small user study or expert review for answer quality. Collect feedback for improvement.
6. **Month 6 – Finalization and Thesis Writing:** Refine the system based on evaluation results. Compile findings, complete thesis writing, and prepare the final defense presentation.

Regular bi-weekly meetings with the advisor will ensure progress alignment. Key deliverables include the compiled data corpus, a working prototype, evaluation report, and the thesis document.

Budget

The estimated budget covers the main costs for this project:

- **Compute Resources (≈\\$1,000):** Cloud GPU instances for training/tuning embeddings and model (e.g. \\$50/month GPU on AWS/GCP) and hosting the chatbot for testing.

- **Software and Tools (≈\$200):** Cloud service fees for LLM API usage or GPU rental (\$150) plus any specialized tools (e.g. vector DB service charges). Most development libraries (PyTorch, Hugging Face, Qdrant) are free.
- **Data Acquisition (≈\$300):** Costs for accessing proprietary consular data if needed (though most MFA documents are public). Potential translation costs if outsourcing Indonesian-English alignment of documents.
- **Miscellaneous (≈\$500):** Token budget for GPT/API calls (\$200), consultation fees or workshops (\$200), and contingency (\$100).

Total projected budget: **~\$2,000**. This is a modest sum, reflecting the reliance on existing open-source frameworks and public data. Funding would ensure adequate compute time and any necessary specialized services without undue constraints.

References

- Global Diplomatic Forum. (2024, December 17). *Chatbots in Consular Affairs: Revolutionising Diplomacy with AI*. LinkedIn. Retrieved from [Global Diplomatic Forum LinkedIn] [27] .
- Nakamura, S. (2025, February 26). *VFS Global pioneers Artificial Intelligence (AI) – launches AI-Powered Chatbot for UK Visa customers*. African Business. Retrieved from [African Business Website] [10] .
- Rahmadini, N. (2025, April 21). *Indonesia, UN Women Launch AI Chatbot SARI to Protect Female Migrant Workers*. JakartaDaily. Retrieved from [JakartaDaily International News] [24] .
- AWS Public Sector Blog. (2024, August). *Demystifying AI: How RAG boosts efficiency in state and local government*. Amazon Web Services. Retrieved from [AWS Public Sector Blog] [19] .
- AWS. (n.d.). *What is Retrieval-Augmented Generation (RAG)?* AWS Learning Blog. Retrieved from [AWS AI Blog] [17] .
- Statetech Magazine. (2025, February). *What Is RAG (Retrieval-Augmented Generation)? How Are State and Local Agencies Using It?* Retrieved from [StateTech] [20] .
- Busch, F., et al. (2025). *Evaluation of a Retrieval-Augmented Generation-Powered Chatbot for Pre-CT Informed Consent: a Prospective Comparative Study*. Journal of Imaging Informatics in Medicine, 12(3). https://doi.org/10.1007/s10278-025-01483-w [21] .
- Nugraha, T., & Hermawan, I. (2024, October). *AI in Indonesian Public Services: A Case Study for Global South Nations*. Modern Diplomacy. Retrieved from [ModernDiplomacy.eu] [16] .
- Consulardiplomacy.com. (2023). *Diplomacy 4.0: How Artificial Intelligence is Changing Diplomacy?* (pp. 15–18). Retrieved from [Diplomacy 4.0 article] [28] [29] .
- V-Soft Consulting. (n.d.). *Chatbots for Government: 12 Global Government Agencies That Use Chatbots*. Retrieved from [V-Soft Blog] [30] [6] .
- GovInsider. (2024, June). *Navigating AI and superapp adoption in Indonesia's public sector*. Retrieved from [GovInsider Asia] [15] .

Appendices

Appendix A. System Architecture Sketch. (Here, we would include a block diagram showing the chatbot's components: user interface, LLM, vector store, and consular document sources.)

Appendix B. Sample Dialogue. (Example Q&A between a user and the proposed chatbot: e.g. a question about visa renewal and the bot's step-by-step answer based on retrieved policy documents.)

Appendix C. Data Schema. (Outline of how the consular documents are indexed and stored in the vector database, and any schema for logging chatbot interactions.)

---

[1] [2] [3] **Chatbots in Consular Affairs: Revolutionising Diplomacy with AI**
[4] [7] [22] https://www.linkedin.com/pulse/chatbots-consular-affairs-revolutionising-diplomacy-n8gbe
[27]

[5] [6] [8] **Chatbots for Government: 12 Global Government Agencies That Use Chatbots**
[30] https://blog.vsoftconsulting.com/blog/15-governments-agencies-that-use-chatbots

[9] **VFS Global pioneers Artificial Intelligence (AI) & digital innovation; Launches Artificial**
[10] **Intelligence (AI)-Powered Chatbot for United Kingdom (UK) visa customers in 141 countries**
https://african.business/2025/02/apo-newsfeed/vfs-global-pioneers-artificial-intelligence-ai-launches-artificial-intelligence-ai-powered-chatbot-for-united-kingdom-uk-visa-customers-in-141-countries

[11] [12] [13] **AI in Indonesian Public Services: A Case Study for Global South Nations - Modern**
[14] [16] **Diplomacy**
https://moderndiplomacy.eu/2024/10/02/ai-in-indonesian-public-services-a-case-study-for-global-south-nations/

[15] **Navigating AI and superapp adoption in Indonesia's public sector**
https://govinsider.asia/intl-en/article/navigating-ai-and-superapp-adoption-in-indonesias-public-sector

[17] [18] **What is RAG? - Retrieval-Augmented Generation AI Explained - AWS**
https://aws.amazon.com/what-is/retrieval-augmented-generation/

[19] **Demystifying AI: How RAG boosts efficiency in state and local government departments | AWS Public Sector Blog**
https://aws.amazon.com/blogs/publicsector/demystifying-ai-how-rag-boosts-efficiency-in-state-and-local-government-departments/

[20] **What Is RAG? | How Are State and Local Agencies Using It? | StateTech**
https://statetechmagazine.com/article/2025/02/what-is-rag-perfcon

[21] **Evaluation of a Retrieval-Augmented Generation-Powered Chatbot for Pre-CT Informed Consent: a Prospective Comparative Study | Journal of Imaging Informatics in Medicine**
https://link.springer.com/article/10.1007/s10278-025-01483-w

[23] [24] **Indonesia, UN Women Launch AI Chatbot SARI to Protect Female Migrant Workers - Jakarta**
[26] **Daily**
https://www.jakartadaily.id/international/16214999370/indonesia-un-women-launch-ai-chatbot-sari-to-protect-female-migrant-workers

[25] **What is RAG: Understanding Retrieval-Augmented Generation - Qdrant**
https://qdrant.tech/articles/what-is-rag-in-ai/

[28] [29] **Diplomacy 4.0: How Artificial Intelligence is Changing Diplomacy?**
https://www.consulardiplomacy.com/home/diplomacy-40-how-artificial-intelligence-is-changing-diplomacy