Advancing Indonesian Consular Services through SahabatAI and Retrieval-Augmented Generation: Alternative Thesis Proposals

I. Introduction: Identifying Advanced AI Solutions for Indonesian Consular Services

A. Reiteration of User Objective and Context

This report aims to support the identification and development of robust, alternative thesis proposal topics centered on the application of Artificial Intelligence (AI) within the consular services of the Ministry of Foreign Affairs of the Republic of Indonesia (Kemlu). The exploration emphasizes leveraging advanced Indonesian-centric AI technologies, notably the SahabatAI Large Language Model (LLM) ecosystem and Retrieval-Augmented Generation (RAG) techniques. Furthermore, it considers existing AI implementations, such as the SARI (Sahabat Artifisial Migran Indonesia) chatbot, as valuable reference points and learning opportunities for future innovations. The objective is to furnish a postgraduate researcher with well-grounded options for impactful academic inquiry in this domain.

B. The Transformative Potential of AI in Consular Services

Artificial Intelligence, particularly the synergy between sophisticated LLMs like SahabatAI and the information-grounding capabilities of RAG, holds significant transformative potential for consular operations. Such technologies can revolutionize how Kemlu delivers services by enhancing the accuracy and accessibility of information for Indonesian citizens abroad, streamlining complex administrative procedures like document legalization, and enabling more personalized and timely assistance. The Ministry of Foreign Affairs of the Republic of Indonesia has demonstrated a strategic inclination towards digital transformation and AI adoption. This is evidenced by initiatives such as the SARI chatbot, designed to protect and inform Indonesian migrant workers [1], and the national effort to develop sovereign AI capabilities through the SahabatAI project.[3] Deputy Minister of Foreign Affairs, Arrmanatha Nasir, has explicitly stated that Kemlu will continue to prioritize the use of technology, including AI, to support the services and protection of Indonesian citizens abroad.[5] This commitment suggests a receptive environment for AI-driven innovations within consular services.

The development and deployment of digital tools are not isolated technological upgrades but rather reflect a broader strategic imperative. Kemlu's existing digital infrastructure, including the Portal Peduli WNI for online consular services [5] and the AI-powered SARI chatbot [1], alongside the high-level commitment to AI [5], indicates a comprehensive strategy to modernize

public service delivery. This modernization is likely driven by the complexities of serving a large and diverse Indonesian diaspora and aligns with overarching national ambitions, such as the "Golden Indonesia 2045" vision, which SahabatAI is also linked to.[3] Consequently, proposing AI research for Kemlu is not merely about technical novelty; it is about contributing to a demonstrable strategic direction and national aspiration, thereby enhancing the potential impact and support for such research endeavors.

C. Report Structure and Aims

This document will present several meticulously researched alternative thesis topics. Each topic will be detailed with potential research questions, proposed methodologies, data acquisition and management strategies, and relevant performance evaluation metrics, all structured to align with the rigorous standards of an academic thesis proposal. The overarching aim is to provide a comprehensive and actionable foundation from which a researcher can select and further develop a compelling and academically sound thesis proposal.

D. AI Sovereignty and Consular Services

The deliberate focus on SahabatAI and RAG as core technologies for the proposed thesis topics extends beyond mere technical preference. It inherently involves an exploration of "AI sovereignty" within the critical context of government functions.[3] SahabatAI is explicitly framed as an Indonesian LLM ecosystem designed to foster technological independence and address "critical context and cultural reference gaps left by the global large language models".[3] Consular services are deeply embedded in national legal frameworks, require nuanced communication, and demand an understanding of Indonesian cultural sensitivities. Therefore, a thesis that investigates the application of SahabatAI to consular tasks will, by its nature, assess the practical efficacy of a sovereign AI in addressing specific national requirements. The outcomes of such research would have broader implications for Indonesia's technological autonomy and its capacity to develop culturally attuned AI solutions for a wide range of public services.

II. Contextual Framework: AI, SahabatAI, RAG, and Indonesian Consular Operations

A. The Landscape of Indonesian Consular Services

The Ministry of Foreign Affairs of the Republic of Indonesia (Kemlu) provides a wide array of consular services aimed at protecting its citizens abroad and facilitating international legal and administrative processes. Understanding these functions, existing digital tools, and inherent challenges is crucial for identifying impactful AI applications.

- Key Functions:
  The core responsibilities of Kemlu's Directorate General of Protocol and Consular Affairs are multifaceted. A primary mandate is the protection of Indonesian citizens (Perlindungan WNI) and Indonesian legal entities overseas.8 This includes assistance to Indonesian migrant workers, a particularly vulnerable group for whom the SARI chatbot was developed.1 Document services are another significant component, encompassing

the legalization of documents for international use 10, and the issuance of passports and visas.11 Furthermore, consular services involve providing general information and assistance to Indonesian citizens 12, and facilitating administrative processes such as the registration of births, marriages, and deaths occurring abroad.6 The Directorate General of Protocol and Consular Affairs is tasked with the formulation and execution of policies in these areas.8

- Existing Digital Initiatives:
  Kemlu has already embarked on digital transformation through several key initiatives:
  - **Portal Peduli WNI:** This web-based platform serves as a central hub for Indonesian citizens abroad, offering features for self-reporting (Lapor Diri), online application for various consular services (including document legalization, passport services, and issuance of various certificates), and a system for reporting cases or lodging complaints.[5] The portal's content, such as FAQs, user guides, and forms, represents a valuable, structured knowledge source that could be leveraged for a RAG system.[6]
  - **SARI Chatbot (Sahabat Artifisial Migran Indonesia):** Launched in collaboration with UN Women, SARI is an AI-powered chatbot specifically designed to support Indonesian migrant workers, with a focus on women.[1] It is integrated into the Safe Travel mobile application and aims to provide quick, empathetic responses based on comprehensive and credible data, developed through a human-centered design process.[1] SARI signifies a concrete step by Kemlu in utilizing AI for citizen protection.
  - **"STEMPEL ASLI" and "SITPROTKONS":** Kemlu's performance reports mention the development and implementation of digital applications for consular services, including "STEMPEL ASLI" for document legalization and "SITPROTKONS".[9] The "STEMPEL ASLI" application is available on the Google Play Store and facilitates online document legalization requests.[15] These systems indicate an ongoing effort to digitize core consular workflows.
- Identified Challenges and Areas for AI Innovation:
  Despite these advancements, several challenges persist where AI could offer innovative solutions:
  - The procedures for document legalization can be complex, particularly with the introduction of the Apostille system alongside traditional legalization methods, creating potential confusion for users.[17]
  - Consular offices often handle a high volume of inquiries, many of which are repetitive and concern standard procedures. AI can automate responses to such queries.
  - There is a continuous need for accurate, up-to-date, and easily accessible information for citizens dispersed globally across different time zones.
  - Ensuring consistency of information across numerous platforms (websites, apps, individual representative offices) is a significant challenge.
  - Data collection ("Pendataan") for effective citizen protection has been identified

as a major hurdle [19], an area where AI-driven data analysis might offer improvements.

The presence of multiple digital platforms within Kemlu for consular services—Portal Peduli WNI, SARI, STEMPEL ASLI, and SITPROTKONS—while indicative of progress, also suggests a potential for fragmented user experiences and information silos. Users might find it confusing to determine the appropriate platform for their specific needs, and information could be inconsistently updated across these systems. An integrated AI solution, particularly one based on RAG, could address this by acting as a unified, intelligent interface. Such a system could draw knowledge from all these disparate sources, providing a coherent and user-friendly "single window" to Kemlu's digital consular services, thereby streamlining access and improving service quality.

B. SahabatAI: An Indonesian LLM Ecosystem for Sovereign AI

SahabatAI represents a significant national initiative to develop an indigenous Large Language Model ecosystem tailored to Indonesian needs, languages, and cultural contexts.

- Capabilities and Development:
  SahabatAI is an open-source LLM ecosystem co-initiated by prominent Indonesian technology and telecommunication companies, Indosat Ooredoo Hutchison and GoTo Group, with support from global technology partners like NVIDIA and research institutions such as AI Singapore and various Indonesian universities.3 A core objective of SahabatAI is to empower Indonesians by providing AI models built by Indonesians, for Indonesians. These models are specifically designed to understand local contexts, Bahasa Indonesia, and several local dialects, including Javanese and Sundanese, thereby addressing linguistic and cultural nuances often missed by global LLMs.3
- Technical Details:
  The SahabatAI ecosystem includes models based on established architectures like Llama3 8B and Gemma2 9B.21 These base models undergo continued pre-training on extensive datasets rich in Indonesian language content, amounting to approximately 50 billion tokens. This data includes sources like the SEA-LION Pile (a large Southeast Asian language dataset), Indonesian news articles, and Wikipedia content in Indonesian.21 Furthermore, instruction-tuned versions, such as gemma2-9b-cpt-sahabatai-v1-instruct, are available. These models are fine-tuned using a substantial number of instruction-completion pairs in Bahasa Indonesia, Javanese, Sundanese, and English, making them particularly suitable for conversational AI, question-answering, and other instruction-following tasks.23 These models typically feature a context length of 8192 tokens.21
- Relevance for Public Sector:
  SahabatAI is explicitly intended to enhance government services and facilitate seamless business-to-government (B2G) and citizen-to-government interactions.3 Its application in Indosat Ooredoo Hutchison's AI chatbot for various citizen and resident services demonstrates its practical utility in the public domain.4 The Indonesian government has expressed commitment to leveraging SahabatAI, aligning with the "Golden Indonesia 2045" vision and supporting national leadership in technology.7

The development of SahabatAI, with its specific and extensive fine-tuning for Bahasa

Indonesia and other local languages [21], directly confronts the "low-resource language" challenge often encountered in Natural Language Processing (NLP) for languages other than English.[24] This focused effort significantly enhances the feasibility and potential quality of an AI system designed for Kemlu. By utilizing models deeply trained on relevant Indonesian linguistic and cultural data, a thesis project can anticipate better performance and contextual understanding for consular queries compared to relying solely on global LLMs that may lack such specialized training. This inherently reduces a major technical risk and promises more accurate and culturally appropriate AI interactions.

C. Retrieval-Augmented Generation (RAG): Principles and Advantages

Retrieval-Augmented Generation (RAG) is an AI architecture that enhances the capabilities of LLMs by grounding their responses in external, verifiable knowledge sources.

- Core Concept:
  RAG is a hybrid model that integrates two primary components: a retrieval mechanism and a generation module.25 The retriever is responsible for searching and fetching relevant documents or information snippets from an external knowledge base (e.g., a collection of official regulations, FAQs, or policy documents). The generator, typically a powerful LLM, then processes this retrieved information along with the original user query to produce a coherent, contextually appropriate, and factually grounded textual response.25
- Key Advantages for Government Services:
  The RAG architecture offers several compelling advantages for applications in government services, particularly consular affairs:
  - **Addresses LLM Limitations:** RAG significantly mitigates common LLM issues such as "hallucination" (generating plausible but incorrect information) by compelling the model to base its output on retrieved evidence. This improves factual accuracy and ensures that responses are rooted in real-world, up-to-date data.[25] This is paramount for consular services that dispense legal and procedural information where errors can have serious consequences.
  - **Knowledge-Intensive Tasks:** RAG excels in domains requiring access to specific, verifiable information. It is well-suited for tasks like open-domain question answering, providing legal advisory services, and powering sophisticated conversational AI systems.[25]
  - **Dynamic Knowledge Integration:** RAG systems can be connected to knowledge bases that are regularly updated. This allows the AI to provide information that reflects the latest changes in regulations, policies, or emergency situations, a critical feature for dynamic environments like consular services.[27]
  - **Enhanced Verifiability and Transparency:** Because RAG responses are grounded in specific retrieved documents, it is often possible to cite the sources used by the AI. This increases the verifiability of the information provided and enhances user trust and transparency in the AI system.
  - **Reduced Need for Frequent Retraining:** By externalizing knowledge, RAG systems can often be updated by simply modifying the knowledge base, rather

than requiring extensive and costly retraining of the LLM itself.

- Architectural Components:
  A typical RAG system comprises several key components. The retriever often uses dense vector representations (embeddings) of documents and queries to find semantically similar information. The generator is usually a transformer-based LLM. Beyond these core elements, RAG systems also involve upstream processes like document preprocessing (cleaning, structuring) and chunking (dividing documents into smaller, manageable pieces for retrieval), and embedding generation (converting text into numerical vectors). Downstream elements can include mechanisms for ensuring response safety and credibility.25

The global trend of foreign ministries adopting AI and chatbot technologies for consular services underscores the relevance and timeliness of research in this area for Kemlu. For instance, the Italian Ministry of Foreign Affairs and International Cooperation (MAECI) has deployed a virtual assistant (chatbot) across its website and numerous diplomatic-consular offices to guide users on consular and visa services.[29] Similarly, Kenya's Ministry of Foreign Affairs is actively integrating AI into its diplomatic practices, including enhancing consular services and crisis prediction.[32] The United Arab Emirates' Ministry of Foreign Affairs (MoFA) is also leveraging AI as part of its "Smart Mission" initiative to achieve bureaucracy-free consular services.[33] These international precedents demonstrate a clear trajectory towards AI adoption in this sector. Kemlu's exploration of AI, including the SARI chatbot and considerations for SahabatAI, positions it within this global movement. A thesis focused on AI for Indonesian consular services can, therefore, draw valuable insights, best practices, and potential comparative analyses from these international examples, adding significant depth and global context to the research.

III. Proposed Alternative Thesis Topics in AI for Enhanced Consular Services

This section outlines three distinct thesis proposal topics, each leveraging SahabatAI and RAG to address specific challenges and opportunities within Indonesian consular services. Each proposal includes a problem statement, the proposed AI solution, key research questions, and a methodological blueprint.

A. Thesis Topic 1: "Developing an Intelligent Q&A System using SahabatAI and RAG for Indonesian Consular Document Legalization Procedures"

- 1. **Problem Statement and Research Significance:** The process of document legalization for use abroad is a critical consular service that often presents considerable complexity for both Indonesian citizens and foreign nationals. Users must navigate a multi-layered system that includes traditional legalization pathways involving Kemlu and the Ministry of Law and Human Rights (Kemenkumham), as well as the more recent Apostille certification system, applicable to member countries of the Hague Apostille Convention.[17]

Understanding which pathway applies, the specific requirements for different document types, and the correct sequence of procedures can be daunting. Official regulations, such as the Minister of Foreign Affairs Regulation No. 14 of 2022 concerning Procedures for Legalization of Documents at the Ministry of Foreign Affairs (Permenlu No. 14 Tahun 2022) [18], provide comprehensive details but may not be easily digestible for the general public. This information complexity can lead to errors in applications, processing delays, and user frustration. The significance of this research lies in its potential to vastly improve service efficiency and user satisfaction. An AI-powered Question & Answering (Q&A) system can offer immediate, accurate, and personalized guidance on document legalization, tailored to the user's specific document and destination country. This would not only empower users but also reduce the repetitive query load on consular staff, allowing them to focus on more complex cases.

2. **Proposed AI-Driven Solution and Novelty:** The proposed solution is a RAG-based Q&A system. The **retriever** component will access a curated knowledge base composed of official Indonesian government documents. This includes, centrally, Permenlu No. 14 Tahun 2022 [18], along with related regulations from Kemlu's Legal Documentation and Information Network (JDIH Kemlu) [35] and Kemenkumham's Directorate General of General Legal Administration (AHU) concerning Apostille services.[36] The knowledge base will also incorporate FAQs and procedural guides scraped from official Kemlu websites (e.g., general FAQs [11], information from embassy/consulate sites [17]), Kemenkumham's Apostille portal, and informational content from Kemlu's "STEMPEL ASLI" mobile application.[15] The **generator** component will be an advanced SahabatAI model, such as gemma2-9b-cpt-sahabatai-v1-instruct.[23] This model, fine-tuned or expertly prompted, will generate clear, concise, and actionable answers in Bahasa Indonesia, ensuring that all responses are strictly grounded in the information retrieved from the knowledge base. The novelty of this solution manifests in several aspects:
   - The specific application of SahabatAI, an Indonesian-centric LLM, to the complex domain of consular document legalization, leveraging its nuanced understanding of Bahasa Indonesia and local legal contexts.
   - The creation of a specialized RAG system tailored to interpret and synthesize information from Indonesian regulatory texts and procedural guidelines related to both traditional legalization and the Apostille Convention.
   - The potential incorporation of a decision-tree-like logic within the RAG framework. This would enable the system to guide users more effectively by asking clarifying questions (e.g., document type, country of origin, destination country's Apostille status) to determine the correct legalization pathway.

3. **Key Research Questions:**

- How effectively can a SahabatAI-RAG system interpret user queries formulated in natural Bahasa Indonesia concerning document legalization and retrieve the most pertinent information segments from a diverse knowledge base of legal texts and FAQs?
- What is the optimal RAG architecture—including chunking strategies for legal documents, choice of embedding models proficient in Indonesian, and retriever configurations (e.g., dense, sparse, or hybrid)—for processing Indonesian regulatory documents like Permenlu No. 14 Tahun 2022 and associated guidance materials?
- Can the system accurately distinguish between the procedural requirements for Apostille certification and traditional multi-step legalization based on the user's query context and document details?
- How do end-users perceive the accuracy, completeness, clarity, and overall usability of the AI-generated responses when compared to seeking information through traditional channels (e.g., website navigation, direct inquiry)?
- What are the primary challenges and best practices associated with curating, maintaining, and updating a knowledge base for a dynamic legal procedure like document legalization, ensuring the AI system consistently provides current information?

4. **Methodological Blueprint:**
   - *AI Models & Techniques:*
     - **Primary LLM:** A SahabatAI instruction-tuned model, such as gemma2-9b-cpt-sahabatai-v1-instruct [23] or llama3-8b-cpt-sahabatai-v1-instruct (pending comparative performance analysis on Indonesian Q&A tasks). Justification for SahabatAI includes its Indonesian-centric design, demonstrated proficiency in Bahasa Indonesia, open-source nature, and alignment with national AI development goals.[3]
     - **RAG Framework:** Implementation of a standard RAG pipeline [25], with consideration for advanced RAG techniques such as query expansion or re-ranking modules to enhance retrieval precision if initial results warrant.[39]
     - **Embedding Models:** Evaluation of leading sentence transformer models with strong Indonesian language support, or leveraging embedding capabilities inherent in the chosen SahabatAI base model if suitable for retrieval tasks.
     - **Comparative Analysis:** The performance of the SahabatAI-RAG system will be benchmarked against a baseline SahabatAI model operating without RAG. Optionally, a comparison with a global LLM (e.g., GPT-3.5) integrated with a similar RAG setup could be conducted to highlight the specific advantages of SahabatAI's linguistic and contextual specialization for this Indonesian use case.
   - *Data Strategy:*

        - **Knowledge Base Creation:**
            - Systematic collection of official documents: Permenlu No. 14 Tahun 2022 concerning Kemlu legalization procedures [18], related ministerial regulations and decrees from JDIH Kemlu [35], and regulations/guides concerning Apostille from Kemenkumham AHU's official portal (e.g., legalisasi.ahu.go.id, panduan.ahu.go.id).[36]
            - Web scraping of FAQs, procedural guides, and service descriptions from:
                - The official Kemlu website (kemlu.go.id), including its general FAQ section [11] and specific consular service pages (e.g., document legalization information on main site or embassy sub-sites [17]). Efforts should be made to access key pages like kemlu.go.id/portal/id/read/16/konsuler/layanan-legalisasi-dokumen which were previously inaccessible, potentially through official requests or archival services.
                - The Portal Peduli WNI, which contains service information, FAQs, and guides relevant to citizens abroad.[6]
                - Informational content related to Kemlu's "STEMPEL ASLI" mobile application for document legalization.[15]
            - Explicit inclusion of materials that delineate the differences, procedures, and applicable contexts for Apostille versus traditional legalization.[17]
        - **Data Preprocessing:**
            - Conversion of PDF documents (the likely format for official regulations [35]) into clean text.
            - Structuring and normalizing the extracted text.
            - Document segmentation into appropriate chunk sizes (e.g., paragraphs, logical sections) for effective RAG retrieval, with experimentation to determine optimal chunking strategies.[39]
        - **Q&A Pair Generation (Optional, for Fine-tuning/Evaluation):** Development of a dataset comprising question-answer pairs based on the curated knowledge base. This can be created manually by domain experts or semi-automatically using LLM-assisted methods, followed by human verification.
    - *System Development & Training:*
        - Implementation of the RAG pipeline, including setting up a retriever (e.g., using a vector database like FAISS or ChromaDB) and integrating it with the SahabatAI generator model.
        - If a sufficiently large and high-quality Q&A dataset is created, consider fine-tuning the chosen SahabatAI instruct model specifically on consular legalization interactions. This can enhance its domain adaptation and response quality, even though models like

gemma2-9b-cpt-sahabatai-v1-instruct are already fine-tuned on general Indonesian instructions.[23]

- *Performance Evaluation:*
  - **Retrieval Metrics:** Assess the retriever's performance using metrics such as Context Precision, Context Recall, and Context Relevance, which measure the accuracy and appropriateness of the documents fetched in response to a query.[28]
  - **Generation Metrics (e.g., using RAGAS framework [41]):** Evaluate the generator's output based on:
    - *Faithfulness:* The degree to which the generated answer accurately reflects the information in the retrieved context, minimizing hallucination.
    - *Answer Relevancy:* The pertinence of the generated answer to the user's original question.
    - *Answer Correctness:* The factual accuracy of the information provided, benchmarked against ground truth documents.
  - **NLP Quality Metrics:** Standard metrics like Perplexity (for language model fluency) and BLEU scores (if high-quality reference answers are available for comparison).
  - **Task-Specific Metrics:** Develop metrics to evaluate the system's accuracy in correctly identifying the appropriate legalization path (Apostille vs. traditional Kemlu/Kemenkumham route) and the correctness of the procedural steps outlined in its responses.
  - **User Study:** Conduct qualitative evaluations with potential end-users (e.g., students planning to study abroad, business professionals) to gather feedback on the system's usability, the clarity of its responses, and their perceived trustworthiness.
  - **Handling of Ambiguity:** Evaluate the system's ability to handle unanswerable or ambiguous queries gracefully, for instance, by asking for clarification or indicating when information is not available in its knowledge base.[42]

The successful development of such a Q&A system for document legalization would yield more than just a practical tool. It would inherently establish a reusable methodology and a domain-specific knowledge base architecture. The processes refined for structuring complex legal and procedural information (like Permenlu No. 14 Tahun 2022 and Apostille rules), along with the optimized RAG components (chunking, retrieval, generation strategies), could be readily adapted for other complex consular services, such as visa applications or emergency assistance protocols. Furthermore, the experience gained in curating, updating, and managing the knowledge base for legalization would establish best practices for maintaining the currency and accuracy of information in other AI-driven consular tools. This positions the thesis not just as a solution to a specific problem, but as a foundational contribution to Kemlu's broader digital transformation strategy, potentially setting a precedent for how the Ministry digitizes and makes accessible other intricate consular information.Moreover, this

research endeavor would inevitably highlight critical data governance challenges within Kemlu. A RAG system's efficacy is fundamentally tied to the quality and currency of its underlying knowledge base.[25] Legal and procedural landscapes, like document legalization, are dynamic; regulations change (as exemplified by the relatively recent introduction of the Apostille system in Indonesia [17]). If an AI system, particularly one providing official government information, relies on outdated source documents, it can lead to significant errors and misguidance for users. Therefore, the research would need to consider and potentially propose mechanisms for ensuring that the source documents feeding the RAG system are consistently updated. This involves establishing clear pipelines and responsibilities for reflecting any changes in laws, regulations, or internal procedures within the AI's knowledge base. This focus on information management and update protocols is a crucial operational aspect that extends beyond the AI model itself, touching upon the internal administrative processes necessary to support such advanced digital services sustainably.

- **B. Thesis Topic 2: "Enhancing Crisis Response and Citizen Safety: A SahabatAI and RAG-based System for Real-time Emergency Information Dissemination to Indonesian Citizens Abroad"**
    - 1. **Problem Statement and Research Significance:** Indonesian citizens residing or traveling overseas can encounter a spectrum of emergencies, including natural disasters (earthquakes, tsunamis, volcanic eruptions), political instability and civil unrest, public health crises, and personal accidents. In such critical situations, the timely provision of accurate, actionable, and easily understandable information from Kemlu and its representatives is paramount for ensuring their safety and well-being. Current methods for disseminating emergency information may rely on website updates, social media posts, or direct communication, which can sometimes be fragmented, slow to update across all channels, or difficult for affected individuals to find amidst the chaos of a crisis. While the SARI chatbot specifically targets the needs of Indonesian migrant workers [1], a broader, more dynamic system designed for all citizens facing acute emergency situations could significantly enhance Kemlu's crisis response capabilities. The significance of this research lies in its potential to create an AI-driven system that can rapidly process incoming situational updates, official consular advisories, and local emergency contact information. Such a system could provide tailored, immediate, and life-saving guidance to Indonesian citizens in affected areas, directly supporting Kemlu's core mandate of citizen protection (Perlindungan WNI).[9]
    - 2. **Proposed AI-Driven Solution and Novelty:** The proposed solution is a RAG-based system designed for real-time emergency information dissemination. The **retriever** component would be engineered to ingest and process information from a variety of dynamic sources. These would include official Kemlu travel advisories, emergency announcements published on

Portal Peduli WNI [6] or other designated Kemlu communication channels, verified information from Indonesian embassies and consulates in affected regions [12], and potentially, carefully curated and validated reputable news feeds providing on-the-ground updates. The **generator** component will be a SahabatAI model (e.g., gemma2-9b-cpt-sahabatai-v1-instruct [23]), which would be responsible for generating concise, clear, and actionable alerts, summaries, and responses to user queries in Bahasa Indonesia. These outputs would be tailored to the specific nature of the emergency and, where ethically permissible and technically feasible with user consent, to the user's reported location. The novelty of this approach is multi-faceted:

- A primary focus on dynamic, real-time information retrieval, processing, and synthesis specifically for crisis communication in a consular context.
- The application of SahabatAI, with its strong Indonesian language capabilities, to a high-stakes, safety-critical consular function where clarity and empathy are crucial.
- An exploration of how RAG architectures can be adapted to handle rapidly evolving information streams while maintaining trustworthiness and providing reliable summaries.
- The potential for integration with location-based services (always with explicit user consent and robust privacy safeguards) to deliver highly contextualized and relevant alerts and assistance information.

3. **Key Research Questions:**

- How can a RAG system be architected to effectively retrieve, prioritize, and synthesize information from diverse, dynamic, and potentially conflicting sources (e.g., official advisories, news reports, local alerts) for the purpose of emergency consular communication?
- What are the most effective prompting strategies and fine-tuning approaches for SahabatAI to generate clear, empathetic, unambiguous, and actionable emergency guidance in Bahasa Indonesia, based on the context retrieved by the RAG system?
- How can the system ensure the reliability, accuracy, and timeliness of the information it disseminates, and what mechanisms can be implemented to manage or flag potential misinformation from external, less controlled feeds?
- What performance metrics are most appropriate for evaluating an AI system designed for a crisis communication context? This would include speed of information updates, accuracy of advice provided, clarity of instructions, and the system's ability to avoid inducing panic.
- What are the significant ethical considerations in deploying AI for emergency communication, particularly concerning data privacy (e.g., handling of user location data), the potential for AI-induced anxiety or

misdirection, and ensuring equitable access to information for all affected citizens?

4. **Methodological Blueprint:**

- *AI Models & Techniques:*
  - **Primary LLM:** SahabatAI, specifically an instruction-tuned model like gemma2-9b-cpt-sahabatai-v1-instruct [23], chosen for its proficiency in Bahasa Indonesia and its potential for generating empathetic responses, a quality highlighted as a goal for the SARI chatbot.[1]
  - **RAG Framework:** An advanced RAG architecture may be necessary to handle the diversity of data types (structured advisories, unstructured news) and to implement logic for prioritizing information based on urgency, source credibility, and timeliness. This could involve developing source weighting mechanisms or integrating credibility assessment modules.
  - **Information Extraction (IE) / Event Detection:** Consideration should be given to using IE techniques to automatically identify key information (e.g., type of emergency, location, affected numbers, official instructions) from unstructured news feeds or social media (if used) before this information is passed to the RAG system for contextualization and response generation.
- *Data Strategy:*
  - **Knowledge Base Sources:**
    - Kemlu's official travel advisories, emergency protocols, and contingency plans (sourced from kemlu.go.id, Portal Peduli WNI [6], and internal documents if ethically accessible for research).
    - Historical data on past consular responses to various crises (if available and properly de-identified to protect privacy).
    - Comprehensive contact information for all Indonesian embassies, consulates general, and honorary consulates.[44]
    - General safety guidelines and standard operating procedures for various types of emergencies (e.g., earthquake preparedness, what to do during civil unrest, health pandemic advice).
    - A carefully curated and continuously validated list of trusted news APIs, public safety information feeds (e.g., from international organizations like the WHO, or national disaster management agencies of host countries). This requires a robust source selection and validation strategy.
  - **Data Preprocessing:**
    - Development of real-time or near real-time data ingestion pipelines for dynamic sources.

        - Implementation of techniques for summarizing lengthy reports or news articles and extracting key entities and events relevant to the crisis.
        - Chunking strategies suitable for rapidly changing information, possibly with timestamping and versioning of information chunks.
- *System Development & Training:*
    - Emphasis on building a robust retrieval system capable of querying dynamic and diverse information sources efficiently.
    - Development of sophisticated prompting strategies for SahabatAI that prioritize clarity, conciseness, authoritativeness, and reassurance in emergency communications.
    - Creation of simulated crisis scenarios to rigorously test the system's responsiveness, information accuracy under pressure, and ability to handle incomplete or conflicting data.
- *Performance Evaluation:*
    - **Information Accuracy & Timeliness:** Compare AI-generated advice and alerts against official information releases from Kemlu and other trusted sources. Assess the lag time in the AI system incorporating new critical updates.
    - **Relevance of Retrieved Information:** Utilize metrics like Context Precision and Context Recall for the RAG component [28] to ensure the information provided to the LLM is pertinent to the query or situation.
    - **Clarity and Actionability of Generated Responses:** Conduct user studies involving scenario-based evaluations where participants assess the comprehensibility, usefulness, and trustworthiness of the AI's emergency guidance.
    - **Robustness and Resilience:** Test the system against noisy, incomplete, or deliberately conflicting information sources to evaluate its ability to maintain accuracy and avoid propagating misinformation.
    - **Ethical Audit and Bias Assessment:** Systematically review the system for potential privacy infringements (especially if location data is used), biases in information prioritization or advice given to different user groups, and the risk of generating advice that could be harmful or counterproductive in a crisis.

This line of research could significantly advance the development of "trustworthy AI" for government crisis communication. Establishing robust frameworks for validating dynamic information sources and ensuring that AI-generated responses are not only factually accurate but also ethically sound and responsible is a critical challenge. The thesis could pioneer methodologies in this area, particularly for high-stakes scenarios where misinformation can have severe consequences. The RAG architecture's inherent link to source data [25] provides a foundation for verifiability, but in a crisis, the sources themselves can be volatile. Therefore, the research would contribute to the emerging

field of AI safety and ethics in public sector AI applications, addressing how to build systems that remain reliable even when drawing from a fluid information environment.Furthermore, a successful AI system for Kemlu's crisis response could serve as a foundational model for broader inter-agency crisis information sharing and coordination within Indonesia. Consular crises experienced by citizens abroad often have domestic implications, requiring coordinated efforts with national agencies such as the National Disaster Management Agency (BNPB) or health authorities. An AI system designed to process and synthesize crisis information could potentially be architected to interface with or cross-reference data from these other Indonesian emergency response entities. Given that SahabatAI is envisioned as an "ecosystem" [3], this implies a design philosophy that could support interoperability. Consequently, the thesis could explore the potential for a networked AI approach to national crisis management, where Kemlu's system acts as a vital node, extending its utility beyond the direct consular role to contribute to a more holistic national preparedness and response capability.

- **(Optional) C. Thesis Topic 3: "Evaluating and Enhancing the SARI Chatbot: A Comparative Study of RAG Integration with SahabatAI for Improved Support to Indonesian Migrant Workers"**
    - 1. **Problem Statement and Research Significance:** The SARI (Sahabat Artifisial Migran Indonesia) chatbot, developed by Kemlu in collaboration with UN Women, represents a significant and commendable initiative to provide information and protection services to Indonesian migrant workers, particularly women.[1] These workers often face complex challenges, including legal and contractual issues, unsafe working conditions, potential for abuse and exploitation, and difficulties accessing support services. While SARI aims to address these needs by offering quick, empathetic, and credible information [1], any AI system, especially in its early iterations, has the potential for enhancement as AI technology evolves and user needs become better understood. The knowledge base of such a system and its natural language understanding and response generation capabilities can always be improved. The significance of this research lies in its direct engagement with an existing Kemlu AI system. It offers the opportunity to provide data-driven, actionable recommendations for SARI's improvement, thereby tangibly enhancing the support available to a vulnerable segment of Indonesian citizens abroad. This research would specifically explore how advanced RAG techniques, coupled with the latest SahabatAI models, could elevate SARI's effectiveness in delivering comprehensive and nuanced assistance.
    - 2. **Proposed AI-Driven Solution and Novelty:** The core of this thesis would be a multi-stage approach:
        - **Evaluation of Current SARI:** Conduct a thorough evaluation of the existing SARI chatbot's capabilities. As direct access to SARI's backend or internal

design documents may be limited for an external researcher, this evaluation would likely rely on publicly available information about its features and goals [1], analysis of its interaction patterns (if accessible as a user), and potentially the development of a framework to simulate user interactions and assess responses against defined criteria.

- **Prototyping an Enhanced SARI (SARI 2.0):** Propose and develop a prototype of an enhanced version of SARI by:
  - Integrating a RAG architecture to significantly expand and deepen its knowledge base. This would involve incorporating a broader range of relevant documents, such as detailed Indonesian and host-country labor laws, regulations concerning migrant worker protection, anonymized case studies of common issues, and comprehensive information on available support services (legal aid, shelters, counseling).
  - Utilizing a more advanced and instruction-tuned SahabatAI model (e.g., gemma2-9b-cpt-sahabatai-v1-instruct [23]) as the generative component. This aims to achieve more nuanced, contextually aware, and empathetic responses, aligning with SARI's stated design philosophy. The novelty of this research includes:
- Potentially being the first in-depth academic study focused on proposing concrete RAG-based enhancements for an existing Indonesian government-deployed chatbot.
- A comparative analysis of SARI's current (or inferred) architectural approach versus a more advanced RAG-SahabatAI architecture, highlighting specific areas of improvement.
- A dedicated focus on improving AI-driven support for a vulnerable demographic, ensuring that technological advancements directly contribute to social welfare and align with SARI's established human-centered and bias-free design principles.[1]

3. **Key Research Questions:**

- Based on available descriptions and simulated interactions, what are the current strengths and limitations of the SARI chatbot in comprehensively addressing the diverse and complex information needs of Indonesian migrant workers?
- How can a RAG architecture, designed to source information from an expanded knowledge base (including labor laws, protection agency guidelines, and potentially anonymized case precedents), demonstrably improve the depth, accuracy, and specificity of SARI's responses to migrant worker queries?
- Can newer, more powerful SahabatAI models (like Gemma2-based versions) provide more empathetic, contextually sensitive, and less biased responses

for potentially distressing migrant worker issues, compared to the AI capabilities likely underpinning the current SARI?
- What specific RAG components and configurations (e.g., retrieval strategies for legal vs. narrative case documents, optimal chunking methods for multilingual or jargon-heavy texts) are most effective for the domain of migrant worker support and protection?
- How can the established "human-centered," "female-friendly," and "bias-free" design principles of the SARI initiative [1] be rigorously maintained and even enhanced through the integration of a RAG architecture and more advanced LLMs?

4. **Methodological Blueprint:**
- *AI Models & Techniques:*
  - **LLM:** An advanced instruction-tuned SahabatAI model.[23]
  - **RAG Framework:** The RAG architecture would need to be tailored for handling legal documents, case-based information, and potentially emotionally charged queries. This might involve drawing inspiration from RAG applications in legal domains, which often require sophisticated retrieval and reasoning over complex texts.[26]
  - **Comparative Analysis:** The baseline for comparison would be the current SARI chatbot's capabilities (as understood from public information or direct interaction). If specific details about SARI's underlying technology (e.g., if it uses an earlier SahabatAI model or a different NLU engine) become known, this would allow for a more direct technical comparison. Otherwise, the comparison might be against a simpler RAG setup or a non-RAG LLM to demonstrate the added value of the proposed enhancements.
- *Data Strategy:*
  - **Knowledge Base Augmentation for SARI 2.0:**
    - Attempt to understand the scope of the existing SARI knowledge base from public descriptions.[1]
    - Collect relevant Indonesian labor laws, international conventions on migrant workers' rights, and specific regulations from destination countries concerning migrant workers. Sources include JDIH Kemlu [35], the Ministry of Manpower, and BP2MI (The Indonesian Migrant Worker Protection Board).[14]
    - Incorporate publications and guidelines from UN Women and other international organizations involved in migrant worker protection.[1]
    - Crucially, explore ethical ways to include anonymized case studies or common scenarios faced by migrant workers to enable the AI to provide more practical and

experience-informed advice. This requires careful attention to data privacy and anonymization protocols.
        - Gather FAQs, guides, and contact information from relevant NGOs and support organizations that assist migrant workers.
    - **Data Preprocessing:** Similar to Topic 1, but with an added emphasis on handling potentially sensitive personal narratives (from case studies, even if anonymized) and diverse document types (legal texts, reports, guides).
- *System Development & Training (SARI 2.0 Prototype):*
    - Develop the RAG-enhanced SARI prototype.
    - Place significant focus on prompt engineering to elicit empathetic, clear, and supportive responses from SahabatAI, building upon SARI's stated goals of being "female-friendly" and "non-stigmatizing".[1]
    - Develop a comprehensive evaluation dataset that accurately reflects the range and complexity of real-world queries and situations encountered by Indonesian migrant workers.
- *Performance Evaluation:*
    - Utilize metrics from the RAGAS framework (e.g., faithfulness, answer relevancy, context relevancy, answer correctness) [41] to assess the core performance of the RAG system.
    - Conduct qualitative evaluations, possibly through user studies with individuals familiar with migrant worker issues (or simulated user personas), to assess the empathetic tone, clarity of advice, and overall helpfulness of the AI's responses.
    - Measure the accuracy of the system in providing correct information regarding migrant workers' rights, available legal procedures, and appropriate support channels.
    - Benchmark the performance of the SARI 2.0 prototype against the (hypothesized or known/observed) performance of the current SARI system to quantify improvements.
    - Specifically assess the enhanced prototype against SARI's original design goals: providing quick responses, empathetic interaction, comprehensive and credible data, being free from gender bias, and offering non-stigmatizing and non-prejudicial support.[1]

This research offers a valuable model for the iterative improvement and lifecycle management of AI systems deployed in the public sector. SARI is an operational AI system [1], and AI technology, including the SahabatAI ecosystem itself [3], is in a constant state of evolution. Governments require sustainable strategies to upgrade their AI tools, not merely deploy them as static solutions. This thesis would provide a practical case study in "AI lifecycle management" for public services, demonstrating how newer AI advancements—such as more powerful locally-developed LLMs and refined RAG techniques—can be systematically integrated to enhance existing services. This is a crucial aspect for ensuring the long-term viability, relevance, and effectiveness of AI in government.Furthermore, by concentrating on a

vulnerable group (Indonesian migrant workers) and aiming to enhance an AI system specifically designed for their protection and support (SARI), this thesis can make a significant contribution to the discourse on "AI for Good" and the ethical development of AI. This is particularly relevant in the context of a developing nation like Indonesia addressing pressing social challenges. The research would inherently need to grapple with profound ethical considerations, such as the data privacy of migrant workers, mitigating bias in the information provided by the AI, and ensuring that the technology empowers rather than disempowers its users. Consequently, the thesis can serve as a compelling case study demonstrating how advanced AI can be responsibly and effectively applied to address specific Sustainable Development Goals (SDGs), such as those related to decent work and economic growth, gender equality, and reducing inequalities, using Indonesia's proactive efforts as a tangible example.

IV. Foundational Elements for Thesis Development

Successful development of any of the proposed thesis topics requires a solid understanding of the available AI tools, data resources, and lessons from existing implementations. This section provides a framework for these foundational elements.

A. Comparative Framework for AI Models and Techniques

The selection of the most appropriate LLM and RAG configuration is a critical early step for any thesis. A structured comparison can guide this choice.

- **Purpose:** To establish an evidence-based rationale for selecting specific AI models and techniques, particularly focusing on SahabatAI variants and RAG components, tailored to the demands of Indonesian consular service applications.
- **Criteria for Comparison:**
    - **Bahasa Indonesia Proficiency:** This is paramount. Evaluation should consider performance on established Indonesian language benchmarks such as SEA HELM (also known as BHASA) and IndoMMLU, for which SahabatAI models have reported scores.[21] The ability to handle formal Indonesian (as found in legal documents), colloquial expressions (in user queries), and domain-specific terminology (e.g., legal terms in document legalization, crisis-specific vocabulary) is crucial.
    - **Instruction Following Capability:** For instruct-tuned models like SahabatAI gemma2-9b-cpt-sahabatai-v1-instruct [23], their capacity to accurately adhere to complex prompts and generate responses in desired formats (e.g., summaries, step-by-step instructions, empathetic dialogue) is key. IFEval benchmark scores can provide insights here.[23]
    - **Context Length:** The effective context window of the LLM dictates how much retrieved information can be processed by the RAG generator. SahabatAI models generally support a context length of 8192 tokens [21], which is usually adequate for RAG applications.
    - **Data Requirements for Further Fine-tuning:** If the thesis proposes additional domain-specific fine-tuning of a SahabatAI model, the data volume and quality required, as well as the complexity of the fine-tuning process, must be assessed.

    - **Computational Resources:** The computational demands for both training (if fine-tuning) and inference are important practical considerations. SahabatAI benefits from support via Indosat's GPU Merdeka sovereign AI cloud infrastructure [4], which might offer accessibility advantages for Indonesian researchers.
    - **Open Source vs. Proprietary Nature:** SahabatAI's open-source nature [3] is highly beneficial for academic research, allowing for greater transparency, customizability, and reproducibility.
    - **RAG Component Variations:** The RAG framework itself is modular. The thesis should consider and potentially compare different retrieval strategies (e.g., sparse like BM25, dense vector-based, or hybrid approaches), the use of re-rankers to improve the relevance of retrieved documents [39], and various document chunking methods.[28]
    - **Hallucination Rate / Factual Accuracy (when combined with RAG):** A critical metric for government services, where misinformation can have serious repercussions. RAG is designed to improve this, but the extent of improvement needs evaluation.
- **Table 1: Comparative Analysis of Candidate LLMs and RAG Techniques for Indonesian Consular Services**

Model/Technique	Key Features	Bahasa Indonesia Performance (Selected Benchmarks)	RAG Suitability	Availability/License	Potential for Consular Tasks
SahabatAI gemma2-9b-cpt-sahabatai-v1-instruct	Gemma2 9B architecture; CPT on ~50B tokens (Indonesian focus); Fine-tuned on ~448K Indonesian, 96K Javanese, 98K Sundanese, 129K English instructions.[22] Context: 8192 tokens.	IndoMMLU Overall: 62.6%.[23] SEA HELM (Bahasa Indonesia - Base model sahabatai-v1-9B): 60.040.[22] IFEval (English - Instruct): High score indicates good instruction following.[23]	Excellent: Designed for instruction following, good context length, strong Indonesian language grounding.	Gemma Community License.[22] Open Source.	High potential for Q&A, summarization, empathetic dialogue, procedural guidance in Bahasa Indonesia.
SahabatAI	Llama3 8B	SEA HELM	Good (Instruct	Llama3	Good potential

llama3-8b-cpt-sahabatai-v1-base	architecture; CPT on ~50B tokens (Indonesian focus from AI Singapore-Llama-3-8B-Sea-Lion v2.1-Instruct).[21] Context: 8192 tokens.	(Bahasa Indonesia - Base model sahabatai-v1-8B): 53.454.[21] (Instruct version would be needed for direct RAG generation).	version): Good context length, strong Indonesian pre-training. Performance of instruct version to be verified.	Community License.[21] Open Source.	for similar tasks, especially if an instruct-tuned version performs comparably or offers specific advantages (e.g., different reasoning patterns).
Generic Global LLM (e.g., GPT-3.5 via API)	Very large general-purpose model.	May have good general Bahasa Indonesia, but likely less specialized than SahabatAI. No specific Indonesian public benchmarks typically reported.	Good context length, strong general instruction following.	Proprietary API access.	Can serve as a benchmark for SahabatAI's performance, especially regarding nuanced Indonesian understanding and cultural context. May require more prompt engineering for specific tasks.
Advanced RAG Techniques	Query expansion, re-ranking, hybrid retrieval, sentence-level context retrieval (Focus Mode).[39]	N/A (Technique, not model specific)	Can significantly improve retrieval accuracy and relevance, reducing load on the generator.	Open-source libraries available.	Highly relevant for all proposed topics, especially those dealing with complex or ambiguous queries, or large, diverse knowledge bases.

This table serves as a crucial tool for justifying methodological choices within the thesis

proposal. The user query explicitly requested a "comparison of different models/techniques" and "justification for choices." By systematically laying out the characteristics of SahabatAI models against relevant criteria—such as their specific pre-training on Indonesian data [21, 22], performance on Indonesian language benchmarks [21, 23], and suitability for instruction-driven tasks—the table provides a concise and evidence-based foundation for selecting the most appropriate AI tools. This structured comparison helps in de-risking the technical aspects of the thesis and demonstrates a thorough understanding of the available technologies.

B. Data Resources and Management for Consular AI

The effectiveness of any RAG-based system is fundamentally dependent on the quality, comprehensiveness, and currency of its knowledge base.

- Overview of Key Data Sources:
  A wealth of publicly accessible information can form the core of the knowledge base for AI-driven consular services:
  - **JDIH Kemlu (jdih.kemlu.go.id):** This is the primary repository for official legal documents from the Ministry of Foreign Affairs, including *Peraturan Menteri* (Ministerial Regulations) and *Keputusan Menteri* (Ministerial Decrees). A key document is Permenlu No. 14 Tahun 2022 concerning Document Legalization.[18] These documents are typically available in PDF format.[35]
  - **Kemlu Main Website (kemlu.go.id):** This portal contains extensive information, including:
    - General FAQs covering a wide range of consular services, such as document legalization, passport and visa procedures.[11]
    - Specific service pages detailing procedures, although consistent access to a centralized "Layanan Legalisasi Dokumen" page proved problematic during preliminary research.[49] Information is often distributed across main portal sections and the websites of individual Indonesian embassies and consulates.[17]
    - Official announcements, press releases, and news articles, which can provide updates on service changes, such as those following the implementation of the Apostille Convention.[53]
    - Contact information for various directorates, including the Directorate of Consular Affairs, which handles legalization.[54]
  - **Portal Peduli WNI (peduliwni.kemlu.go.id):** A critical resource, this portal offers online application for services (including legalization, passport issuance), FAQs, user guides, and announcements for Indonesian citizens abroad.[5] Its structured content is highly suitable for ingestion into a RAG knowledge base.[6]
  - **Kemenkumham Apostille Portal (legalisasi.ahu.go.id / apostille.ahu.go.id):** The official source for information, FAQs, and application procedures related to Apostille certificates issued by the Ministry of Law and Human Rights.[36] This is crucial for any system needing to differentiate between Apostille and traditional

legalization.
    - **"STEMPEL ASLI" Application:** This Android mobile application facilitates online document legalization requests with Kemlu.[15] Its informational content, user interface flow, and any embedded guidance could inform the design of a Q&A system.
    - **SARI Chatbot Documentation/Information:** Publicly available descriptions of SARI's functionalities, objectives, and development process provide insights into Kemlu's approach to AI in citizen services.[1]
    - **Kemlu Performance Reports (Laporan Kinerja):** These official reports from the Directorate General of Protocol and Consular Affairs or specific directorates (e.g., Directorate of Consular Affairs) may contain valuable data on service volumes, identified challenges, and details about ongoing digital initiatives like SITPROTKONS and STEMPEL ASLI.[9]
- **Data Collection and Preparation:**
    - **Collection:** Employ web scraping techniques for public textual data from websites. For PDF documents (e.g., regulations), use PDF-to-text conversion tools.
    - **Preparation:** Clean the extracted text to remove irrelevant elements (HTML tags, navigation menus). Structure the data where possible (e.g., identify Q&A pairs from FAQs). Segment large documents into smaller, semantically coherent chunks suitable for RAG retrieval. This may involve strategies based on paragraphs, sections, or fixed token counts.
    - **NLP Preprocessing:** Apply Bahasa Indonesia-specific NLP preprocessing steps if necessary, such as stemming or lemmatization, although modern embedding models often handle raw text well.
- **Ethical Considerations:**
    - Primarily focus on the use of publicly available data from official government sources.
    - If any non-public data were to be considered (which is generally beyond the scope of a typical thesis but could be relevant for future extensions), strict adherence to data privacy principles, anonymization techniques, and compliance with Indonesian data protection laws (e.g., UU PDP) would be absolutely paramount.
    - Actively work to ensure that the AI system does not perpetuate or amplify any biases that might be present in the source data. This requires careful curation and ongoing monitoring.
- **Table 2: Potential Data Sources for AI-driven Consular Service Thesis Research**

Data Source	Document/Information Types	Specific Content Example	Relevance to Consular Task	Notes on Accessibility/Format
JDIH Kemlu	Ministerial	Permenlu No. 14	Core legal	Publicly

(jdih.kemlu.go.id)	Regulations (Permenlu), Decrees (Kepmenlu)	Tahun 2022 (Document Legalization Procedures) [18]	framework for document legalization Q&A, procedural guidance.	accessible; primarily PDF format.[35]
Kemlu Main Website (kemlu.go.id)	FAQs, service descriptions, announcements, contact lists	FAQ on legalization types, fees [11]; Embassy contact details [44]; News on Apostille impact.[53]	General Q&A, specific procedure details, emergency contacts, policy updates.	Publicly accessible; HTML, PDF. Some specific service pages may require targeted searching or alternative access methods.
Portal Peduli WNI (peduliwni.kemlu.go.id)	Online service applications, FAQs, user guides, announcements	Guidebook for self-registration [6]; Legalization service application portal [40]; FAQ section.[6]	Rich source for Q&A on various consular services, user journey insights, procedural steps.	Publicly accessible; HTML, PDF for guides. Structured data within the portal.
Kemenkumham Apostille Portal (apostille.ahu.go.id)	Apostille application procedures, FAQs, list of accepted documents, user guides (Panduan)	User guide for online Apostille application [36]; FAQ on Apostille process.[38]	Essential for distinguishing Apostille from Kemlu legalization; detailed procedural steps for Apostille.	Publicly accessible; HTML, PDF for guides. Structured application process.
"STEMPEL ASLI" Mobile App (Google Play Store)	Application for Kemlu document legalization	App description, user interface elements, potential embedded help text.[15]	Understanding current online legalization process, user interaction patterns.	App publicly available; content may need to be extracted via app usage or description.
SARI Chatbot Public Information	Descriptions of features, goals, development approach	News articles on SARI launch, features (empathy, bias-free data).[1]	Insights into Kemlu's AI strategy, human-centered design principles, target user considerations for vulnerable groups.	Publicly available news and reports.
Kemlu	Annual reports,	Mentions of digital	Context on service	Available on

Performance Reports (Laporan Kinerja)	service statistics, project updates	initiatives like SITPROTKONS [9]; Data on service delivery.[57]	challenges, existing digital tools, quantitative data for background.	Kemlu's repository [9]; PDF format.

This systematic mapping of data resources is invaluable for planning the research. It provides a practical inventory of the raw materials available for constructing the RAG knowledge base, helping to assess the richness of available data for each potential thesis topic and identify any potential gaps. This directly addresses the user's need to understand data collection and preparation strategies. The existence of online application systems for both Apostille at Kemenkumham [36, 37] and traditional legalization at Kemlu via the "STEMPEL ASLI" app [15, 16] is particularly noteworthy. These systems imply a significant degree of digitalization and data structuration already in place for these complex procedures. Backend databases and digital document handling are likely components of these services. This existing digital infrastructure is a strong enabler for a RAG system, as structured or semi-structured data is generally easier to ingest, process, and manage within a knowledge base compared to dealing purely with unstructured documents like scanned PDFs. This groundwork can make the development of a RAG-based Q&A system more feasible and potentially more robust, as it might be possible to leverage or be informed by these existing structured data environments.

C. Leveraging Insights from SARI Chatbot Implementation

The SARI chatbot, Kemlu's existing AI application for migrant workers, offers valuable lessons for any new AI project in the consular domain.

- **Key Learnings from SARI:**
    - **Human-Centered Design:** A standout feature of SARI's development was its participatory and human-centered design approach. This involved direct consultations with the primary target users—Indonesian migrant women—as well as with violence service providers, civil society organizations, and youth groups.[1] This methodology ensures that the AI solution is genuinely aligned with user needs and contexts.
    - **Focus on Empathy and Trust:** SARI was designed to provide empathetic conversation and relies on a complete and credible knowledge base, with an emphasis on integrating data that is free from gender bias and providing responses without stigma or prejudice.[1] These qualities are crucial for any consular AI, especially those dealing with users who may be stressed, vulnerable, or facing sensitive situations.
    - **Technology Choice and Evolution:** SARI is described as an "AI chatbot".[1] The fact that Kemlu is also exploring the use of the more advanced SahabatAI for other citizen services [5] indicates an institutional openness to adopting and

evolving with AI technologies.
    - **Collaborative Development:** The SARI initiative was a joint effort between Kemlu and UN Women, with support from the Migration Multi-Partner Trust Fund.[1] This highlights the potential and often the necessity for multi-stakeholder collaboration in developing effective public service AI solutions.
    - **Integration with Existing Platforms:** SARI is integrated as a feature within the Safe Travel mobile application.[1] This demonstrates a practical model for embedding AI tools into broader digital service delivery platforms, rather than creating standalone, isolated applications.
- Applicability to New Consular AI Projects:
  The lessons from SARI are directly applicable to the proposed thesis topics:
    - Any new system should adopt a human-centered design methodology, involving potential users in the design and testing phases.
    - Building user trust is paramount. This can be achieved by ensuring the AI provides accurate, verifiable information (a key strength of RAG) and by designing interactions that are respectful and supportive.
    - The AI's communication tone should be carefully crafted to be appropriate for the service context—whether it's providing procedural information, emergency guidance, or support for sensitive issues. Empathy, clarity, and cultural appropriateness are key.
    - Consideration should be given to how any new AI tool might integrate with or complement existing Kemlu platforms like Portal Peduli WNI or other relevant applications to create a more seamless user experience.

The explicit call for collaboration in the ongoing development of the SahabatAI ecosystem [20] presents a unique opportunity. The SahabatAI team invites contributions from researchers, developers, and language enthusiasts in areas such as identifying technical issues, sharing data, improving documentation, and proposing new evaluation tasks and metrics. Academic thesis work, which involves in-depth research, rigorous model evaluation, and potentially the creation of new datasets or fine-tuning methodologies, aligns perfectly with these collaborative goals. The findings, datasets, or even refined models resulting from a thesis project could be valuable contributions back to the SahabatAI community. This transforms the thesis from a one-way consumption of technology into a potential avenue for scholarly contribution to a significant national AI initiative, thereby enhancing the academic value, impact, and potential visibility of the research.

V. Conclusion and Strategic Recommendations for Thesis Topic Selection

This report has presented a contextual framework and several alternative thesis topics for leveraging SahabatAI and Retrieval-Augmented Generation (RAG) to enhance the consular services of the Indonesian Ministry of Foreign Affairs (Kemlu). The aim is to provide a strong foundation for a postgraduate researcher to select and develop an impactful thesis proposal.

A. Recap of Proposed Alternative Thesis Topics

Three primary alternative thesis topics have been detailed:

1. **Developing an Intelligent Q&A System using SahabatAI and RAG for Indonesian Consular Document Legalization Procedures:** This topic focuses on creating an AI assistant to help users navigate the complex rules and procedures for document legalization, including the distinction between traditional methods and the Apostille Convention. Its strength lies in addressing a clear, persistent user need with a well-documented, albeit complex, process, and relatively straightforward access to official data sources like Permenlu No. 14 Tahun 2022.[18]
2. **Enhancing Crisis Response and Citizen Safety: A SahabatAI and RAG-based System for Real-time Emergency Information Dissemination to Indonesian Citizens Abroad:** This topic proposes an AI system for rapidly processing and disseminating critical information during emergencies. Its strength is its high potential impact on citizen safety and its exploration of dynamic RAG capabilities. However, sourcing and validating real-time information presents a more significant data challenge.
3. **(Optional) Evaluating and Enhancing the SARI Chatbot: A Comparative Study of RAG Integration with SahabatAI for Improved Support to Indonesian Migrant Workers:** This topic involves analyzing the existing SARI chatbot and proposing RAG-based enhancements using newer SahabatAI models to better serve Indonesian migrant workers. Its strength is its direct relevance to an existing Kemlu AI initiative and a vulnerable user group. However, it may depend on the availability of more detailed information about the current SARI system's architecture and knowledge base.

B. Guidance on Selecting a Topic

The choice of thesis topic should be guided by several factors:

- **User Interest and Expertise:** The researcher should select a topic that aligns with their specific interests within AI (e.g., NLP, RAG architectures, ethical AI), consular affairs, or public administration. Passion for the subject matter is a key driver for sustained research.
- **Data Accessibility and Feasibility:** This is a critical practical consideration.
    - Topic 1 (Legalization Q&A) appears to offer the most straightforward data access, primarily relying on publicly available legal documents, official FAQs, and website content from Kemlu and Kemenkumham. The existence of Permenlu No. 14 Tahun 2022 [18] as a core, accessible document is a significant advantage.
    - Topic 2 (Emergency Information) requires more complex data strategies for handling dynamic, real-time information and robust source validation protocols.
    - Topic 3 (SARI Enhancement) might present challenges in obtaining detailed internal information about the current SARI system, potentially requiring more assumptions or a focus on publicly described features and goals.
- **Methodological Complexity:** All proposed topics involve the sophisticated application of RAG and SahabatAI. However, Topic 2 (with its emphasis on dynamic RAG and the complexities of crisis contexts) and Topic 3 (which may involve a comparative study against an existing system with potentially limited public technical details) could present

higher methodological hurdles than Topic 1.

- **Potential for Novel Contribution:** All three topics offer substantial avenues for novel academic contribution. Topic 1 allows for innovation in applying SahabatAI-RAG to a specific Indonesian bureaucratic and legal process. Topic 2 can break new ground in real-time crisis AI using a national LLM. Topic 3 offers a unique opportunity to contribute to the iterative improvement of an existing public sector AI system and address AI for social good.
- **Alignment with Academic Rigor:** The chosen topic must allow for the formulation of clear, researchable questions, the development of a robust and defensible methodology, and the definition of measurable evaluation criteria, all of which are essential for fulfilling the requirements of a postgraduate thesis.

C. Suggestions for Refining the Chosen Topic into a Comprehensive Thesis Proposal

Once a broad topic area is selected, the following steps are recommended for developing it into a full thesis proposal:

- **Narrow the Scope:** Define highly specific research questions and precise objectives. For instance, if choosing Topic 1, one might focus on a particular type of document or a specific user journey within the legalization process.
- **Conduct a Deeper Literature Review:** Expand the review beyond the initial sources. This should include a thorough examination of the latest advancements in RAG variants, LLM fine-tuning techniques, comparative studies of AI applications in government services globally [29], and documented challenges in public sector AI deployment (e.g., ethics, data governance, adoption barriers).
- **Develop a Detailed Data Plan:** Specify the exact data sources to be used, the methods for their collection (e.g., targeted web scraping, API access if available, official requests for documents), detailed preprocessing steps, and robust ethical handling protocols. If certain key Kemlu web pages (e.g., the central consular legalization page [49]) remain inaccessible, the plan must outline alternative strategies to obtain the necessary information. This could involve using internet archive services, focusing on information available through embassy sub-sites, or formally requesting the information from Kemlu.
- **Elaborate on the Evaluation Strategy:** Define precise quantitative and qualitative metrics for assessing the AI system's performance. Plan the creation of evaluation datasets (e.g., curated Q&A pairs, scenario-based test cases). If user studies are proposed, outline the participant recruitment strategy, task design, and data analysis methods.
- **Consider Limitations and Future Work:** Acknowledge any potential limitations of the proposed research (e.g., scope constraints, data access issues, limitations of current AI models) and suggest concrete avenues for future research that could build upon the thesis findings.

D. Final Encouragement

The application of advanced AI like SahabatAI and RAG to Indonesian consular services is a rich and impactful area for research. Any of the proposed topics, if pursued with academic rigor and a commitment to addressing real-world needs, has the potential to result in a

significant thesis. Such work can contribute not only to the academic body of knowledge in AI and public administration but also to the practical enhancement of services provided by the Ministry of Foreign Affairs to Indonesian citizens and other stakeholders.

The selection of a thesis topic is a strategic decision that can significantly influence a researcher's future career trajectory and areas of specialization. For example, focusing on Topic 1 (Legalization Q&A) would cultivate deep expertise in AI applications for regulatory compliance, legal information systems, and the design of user-centric government e-services. Opting for Topic 2 (Emergency Information Dissemination) would build specialized skills in real-time data processing, AI for crisis management, and the ethical considerations of AI in high-stakes environments. Choosing Topic 3 (SARI Enhancement) would offer valuable experience in evaluating and iteratively improving existing AI systems, with a strong focus on AI for social good and supporting vulnerable populations. Therefore, the researcher should consider not only the immediate academic requirements but also how the chosen topic aligns with their long-term professional aspirations and research interests.

Regardless of the specific topic chosen, the research will invariably engage with the broader and increasingly critical challenges of AI governance in the public sector. These include ensuring accountability for AI-driven decisions, maintaining transparency in how AI systems operate, safeguarding data security and citizen privacy, and promoting the ethical use of these powerful technologies.[60] Concerns about data privacy, algorithmic bias, and potential societal impacts like job displacement are inherent in the deployment of AI, especially in government.[32] It is also important to note that current SahabatAI models, while powerful, have not yet undergone specific safety alignment, placing the onus on developers and researchers to implement their own safety fine-tuning and related security measures.[22] Consequently, any thesis undertaken in this domain must, either explicitly or implicitly, address these governance dimensions. This ensures that the research contributes not only to the advancement of AI technology but also to the vital discourse on the responsible and ethical adoption of AI in public administration, a field of growing importance both in Indonesia and globally.

Works cited

1. Kemlu RI dan UN Women Luncurkan Chatbot SARI, Inovasi AI untuk ..., accessed May 10, 2025, https://www.liputan6.com/global/read/6001356/kemlu-ri-dan-un-women-luncurkan-chatbot-sari-inovasi-ai-untuk-pelindungan-perempuan-pekerja-migran
2. Portal Kemlu, accessed May 10, 2025, https://kemlu.go.id/berita/kementerian-luar-negeri-dan-un-women-memperkuat-pelindungan-perempuan-pekerja-migran-indonesia-melalui-inovasi-chatbot-ai-sari?type=publication
3. The Evolving Journey of Sahabat-AI - Twimbit, accessed May 10, 2025, https://twimbit.com/about/blogs/the-evolving-journey-of-sahabat-ai
4. Indonesia Tech Leaders Team With NVIDIA, Partners to Launch Nation's AI, accessed May 10, 2025, https://blogs.nvidia.com/blog/indonesia-tech-leaders-sovereign-ai/

5. Foreign Ministry to Utilize AI-Based Services for Indonesian Citizens Abroad, accessed May 10, 2025, https://en.tempo.co/read/1975389/foreign-ministry-to-utilize-ai-based-services-for-indonesian-citizens-abroad
6. Peduli WNI - Kemenlu, accessed May 10, 2025, https://peduliwni.kemlu.go.id
7. Indosat and GoTo launch Sahabat-AI, Indonesia's open-source LLM for digital sovereignty, accessed May 10, 2025, https://ciosea.economictimes.indiatimes.com/news/next-gen-technologies/indosat-and-goto-launch-sahabat-ai-indonesias-open-source-llm-for-digital-sovereignty/115286783
8. peraturan presiden republik indonesia - Badan Pembinaan Hukum Nasional, accessed May 10, 2025, https://bphn.go.id/data/documents/15pr056.pdf
9. 1 LKj Ditjen Protokol dan Konsuler TA 2022 - Portal Kemlu, accessed May 10, 2025, https://kemlu.go.id/files/repositori/65397/Laporan%20Kinerja%20Direktorat%20Jenderal%20Protokol%20dan%20Konsuler%20TA%202022.pdf
10. Kekonsuleran - Portal Kemlu, accessed May 10, 2025, https://fe-non-production.apps.opppd2-dev.layanan.go.id/layanan/layanan-konsuler-fasilitas-diplomatik-dan-pelindungan-wni/kekonsuleran
11. Portal Kemlu, accessed May 10, 2025, https://kemlu.go.id/faq
12. About Us – Indonesia, accessed May 10, 2025, https://indonesia.nl/en/about-us
13. New Clarity for Indonesian Citizens Abroad: Understanding Ministry of Law Regulation 6/2025 - Assegaf Hamzah & Partners, accessed May 10, 2025, https://www.ahp.id/new-clarity-for-indonesian-citizens-abroad-understanding-ministry-of-law-regulation-6-2025/
14. Migration MPTF Final Report, accessed May 10, 2025, https://mptf.undp.org/sites/default/files/documents/2025-02/final_report_2024_migration_mptf_indonesia.pdf
15. Layanan Legalisasi Dokumen Kemenlu - Indonesia.go.id, accessed May 10, 2025, https://indonesia.go.id/kategori/kependudukan/807/layanan-legalisasi-dokumen-kemenlu
16. STEMPEL ASLI - Aplikasi di Google Play, accessed May 10, 2025, https://play.google.com/store/apps/details?id=id.go.kemlu.protkons.stempelasli2&hl=id
17. Legalisasi Dokumen - Newyork, accessed May 10, 2025, https://www.kemlu.go.id/newyork/pelayanan-perwakilan/legalisasi-dokumen
18. peraturan.bpk.go.id, accessed May 10, 2025, https://peraturan.bpk.go.id/Download/310940/PERMENLU-14-2022.pdf
19. Pendataan Jadi Tantangan Besar dalam Perlindungan WNI di Luar Negeri - VOA Indonesia, accessed May 10, 2025, https://www.voaindonesia.com/a/pendataan-jadi-tantangan-besar-dalam-perlindungan-wni-di-luar-negeri/6016183.html
20. Sahabat-AI, accessed May 10, 2025, https://sahabat-ai.com/
21. GoToCompany/llama3-8b-cpt-sahabatai-v1-base · Hugging Face, accessed May 10, 2025, https://huggingface.co/GoToCompany/llama3-8b-cpt-sahabatai-v1-base

22. GoToCompany/gemma2-9b-cpt-sahabatai-v1-base - Hugging Face, accessed May 10, 2025, https://huggingface.co/GoToCompany/gemma2-9b-cpt-sahabatai-v1-base
23. GoToCompany/gemma2-9b-cpt-sahabatai-v1-instruct - Hugging Face, accessed May 10, 2025, https://huggingface.co/GoToCompany/gemma2-9b-cpt-sahabatai-v1-instruct
24. IndoGovBERT: A Domain-Specific Language Model for Processing Indonesian Government SDG Documents - MDPI, accessed May 10, 2025, https://www.mdpi.com/2504-2289/8/11/153
25. A Comprehensive Review of Retrieval-Augmented Generation (RAG): Key Challenges and Future Directions - arXiv, accessed May 10, 2025, https://arxiv.org/pdf/2410.12837
26. Retrieval-Augmented Generation with Vector Stores, Knowledge Graphs, and Hierarchical Non-negative Matrix Factorization - arXiv, accessed May 10, 2025, https://arxiv.org/html/2502.20364v1
27. How RAG Systems Improve Public Sector Management - - Datategy, accessed May 10, 2025, https://www.datategy.net/2025/04/07/how-rag-systems-improve-public-sector-management/
28. Retrieval Augmented Generation Evaluation in the Era of Large Language Models: A Comprehensive Survey - arXiv, accessed May 10, 2025, https://arxiv.org/html/2504.14891v1
29. The virtual assistant of the Embassy's website is now active - Ambasciata d'Italia Yangon, accessed May 10, 2025, https://ambyangon.esteri.it/en/news/dall_ambasciata/2024/02/da-oggi-e-attivo-lassistente-virtuale-del-sito-dellambasciata-a-yangon/
30. The virtual assistant of the website of the Consulate General of Italy in Los Angeles is online, accessed May 10, 2025, https://conslosangeles.esteri.it/en/news/dal_consolato/2024/02/da-oggi-e-attivo-l assistente-virtuale-del-sito-del-consolato-generale-ditalia-a-los-angeles/
31. The virtual consular assistant “Chatbot“ is online - Consolato Generale d'Italia Melbourne, accessed May 10, 2025, https://consmelbourne.esteri.it/en/news/dal_consolato/2024/02/the-virtual-consular-assistant-chatbot-is-online/
32. Mungai Onyango Integration Of Artificial Intelligence On Diplomatic Practice In Kenya 2000-2024 - Journal of African Interdisciplinary Studies (JAIS): ISSN 2523-6725 (online), accessed May 10, 2025, https://cedred.or.ke/jais/images/nov-2024/15PDF_Mungai__Onyango_Integration_Of_Artificial_Intelligence_On_Diplomatic_Practice_In_Kenya_2000-2024.pdf
33. The Ministry of Foreign Affairs Showcases Digital Transformation and the Future of Consular Services at GITEX Global 2024, accessed May 10, 2025, https://www.mofa.gov.ae/en/mediahub/news/2024/10/17/17-10-2024-uae-uae
34. Aturan Terbaru Legalisasi Apostille di Indonesia - CPT Corporate, accessed May 10, 2025, https://cptcorporate.com/dokumen-apostille-di-indonesia/
35. Website JDIH Kemlu, accessed May 10, 2025, https://jdih.kemlu.go.id

36. panduan_firma [AHU ONLINE], accessed May 10, 2025, https://panduan.ahu.go.id/doku.php?id=panduan_firma
37. Panduan Pengguna Aplikasi AHU Legalisasi - Apostille Untuk Permohonan Apostille, accessed May 10, 2025, https://panduan.ahu.go.id/lib/exe/fetch.php?media=panduan_pengguna_aplikasi_apostille.pdf
38. FAQ | AHU - LEGALISASI APOSTILLE, accessed May 10, 2025, https://apostille.ahu.go.id/faq
39. [2501.07391] Enhancing Retrieval-Augmented Generation: A Study of Best Practices - arXiv, accessed May 10, 2025, https://arxiv.org/abs/2501.07391
40. Peduli WNI - Kemenlu, accessed May 10, 2025, https://peduliwni.kemlu.go.id/pelayanan.html
41. How can I measure the response quality of my RAG? - Reddit, accessed May 10, 2025, https://www.reddit.com/r/Rag/comments/1hzgazk/how_can_i_measure_the_response_quality_of_my_rag/
42. Unanswerability Evaluation for Retrieval Augmented Generation - arXiv, accessed May 10, 2025, https://arxiv.org/html/2412.12300v1
43. Direktorat Jenderal Protokol dan Konsuler - Wikipedia bahasa Indonesia, ensiklopedia bebas, accessed May 10, 2025, https://id.wikipedia.org/wiki/Direktorat_Jenderal_Protokol_dan_Konsuler
44. The Embassy of the Republic of Indonesia - Washington, accessed May 10, 2025, https://kemlu.go.id/washington/tentang-perwakilan/kontak-kami
45. Indonesia International Travel Information, accessed May 10, 2025, https://travel.state.gov/content/travel/en/international-travel/International-Travel-Country-Information-Pages/Indonesia.html
46. (PDF) Legal Query RAG - ResearchGate, accessed May 10, 2025, https://www.researchgate.net/publication/389012050_Legal_Query_RAG
47. Bantu Pekerja Migran, Wamen Christina Apresiasi Peluncuran Fitur Chatbot SARI dari Kemlu - BP2MI, accessed May 10, 2025, https://www.bp2mi.go.id/berita-detail/bantu-pekerja-migran-wamen-christina-apresiasi-peluncuran-fitur-chatbot-sari-dari-kemlu
48. Peraturan Menteri Luar Negeri Nomor 14 Tahun 2022 ... - JDIH Kemlu, accessed May 10, 2025, https://jdih.kemlu.go.id/dokumen/109450
49. accessed January 1, 1970, https://kemlu.go.id/portal/id/read/16/konsuler/layanan-legalisasi-dokumen
50. Mengenal Lebih Jauh KBRI Belarusia di Moskow: Lokasi, Layanan, dan Kegiatannya, accessed May 10, 2025, https://www.traveloka.com/id-id/explore/destination/kbri-belarusia-di-moskow-trp/380373
51. Tentang Pelayanan Perwakilan - Brasilia, accessed May 10, 2025, https://kemlu.go.id/brasilia/tentang-perwakilan/tentang-pelayanan-perwakilan
52. mekanisme pelayanan kekonsuleran - Frankfurt, accessed May 10, 2025, https://kemlu.go.id/frankfurt/pelayanan-perwakilan/legalisasi
53. UNTUK PERIODE YANG BERAKHIR 31 DESEMBER 2023 AUDITED - PPID, accessed

May 10, 2025, https://e-ppid.kemlu.go.id/storage/2220/Laporan-Keuangan-Tahun-2023-Audited_.pdf
54. Kementerian Luar Negeri RI - Portal Kemlu, accessed May 10, 2025, https://kemlu.go.id/kontak
55. Kontak Kami - | Portal Kementerian Luar Negeri Republik Indonesia, accessed May 10, 2025, https://arsipportal.kemlu.go.id/portal/id/kontak-kami
56. Apostille Indonesia | PDF - Scribd, accessed May 10, 2025, https://id.scribd.com/document/663270573/apostille-indonesia
57. Laporan Kinerja Direktorat Konsuler tahun 2023, accessed May 10, 2025, https://kemlu.go.id/files/repositori/65397/LKj%20Dit.%20Konsuler%202023%20final.pdf
58. Peduli WNI - Kemenlu, accessed May 10, 2025, https://peduliwni.kemlu.go.id/faq.html
59. How Governments are Using AI: 8 Real-World Case Studies, accessed May 10, 2025, https://blog.govnet.co.uk/technology/ai-in-government-case-studies
60. Security Concerns Surround AI Vision for US Government - Baker Institute, accessed May 10, 2025, https://www.bakerinstitute.org/research/security-concerns-surround-ai-vision-us-government
61. Tackling AI Challenges in Public Services with Solutions Designed for the Complexity | F5, accessed May 10, 2025, https://www.f5.com/company/blog/tackling-ai-challenges-in-public-services-with-solutions-designed-for-the-complexity
62. AI in Government: A Strategic Framework for Digital Transformation - REI Systems, accessed May 10, 2025, https://www.reisystems.com/ai-in-government-a-strategic-framework-for-digital-transformation/