

# Artificial Agents for Humanitarian Assistance

## RefugeeGPT Technical Report

Aarush Gupta

June 2024

### Abstract

The global refugee crisis has reached unprecedented levels, with over 114 million individuals forcibly displaced by the end of 2023 due to conflict, persecution, violence, and human rights violations. In this technical report, we introduce RefugeeGPT, an AI agent designed to assist refugees by providing comprehensive information on laws, regulations, and programs related to asylum in the United States. RefugeeGPT <sup>1</sup> builds upon Google’s Gemini 1.5 Pro to navigate the complexities prevalent in U.S. immigration law, assisted by over 2,500 documents and encompassing more than 40,000 data points on authoritative sources. We explore the potential of AI to enhance decision-making in refugee status determinations, improve access to legal aid, and offer psychological support, while also addressing the significant risks associated with deploying AI in humanitarian contexts.

## 1 Introduction

The global refugee crisis is a significant humanitarian challenge. By the end of 2023, more than 114 million people had been displaced, due to conflict, persecution, violence, and human rights violations. This was the highest number of displaced individuals recorded in recent history, having profound effects to refugees and countries worldwide [9, 4]. Conflicts in the Ukraine, Sudan, the Democratic Republic of Congo (DRC), and Myanmar have been major contributors to this crisis. The ongoing war in Ukraine, for instance, has caused one of the fastest displacement crises since World War II, forcing millions to flee to neighboring countries [9]. Similarly, ongoing conflicts in the DRC and Myanmar have resulted in large numbers of refugees and Internally Displaced Peoples (IDPs), often with minimal international media coverage [12]. Refugees frequently face substantial barriers to accessing aid and assistance. Many low- and middle-income countries, which host the majority of the world’s refugees, have limited resources and infrastructure to support these populations. Over three-quarters of those in need of international protection are hosted by such nations, placing significant strain on resources and exacerbating the challenges refugees face in obtaining essential services such as healthcare, education, and legal protection [4].

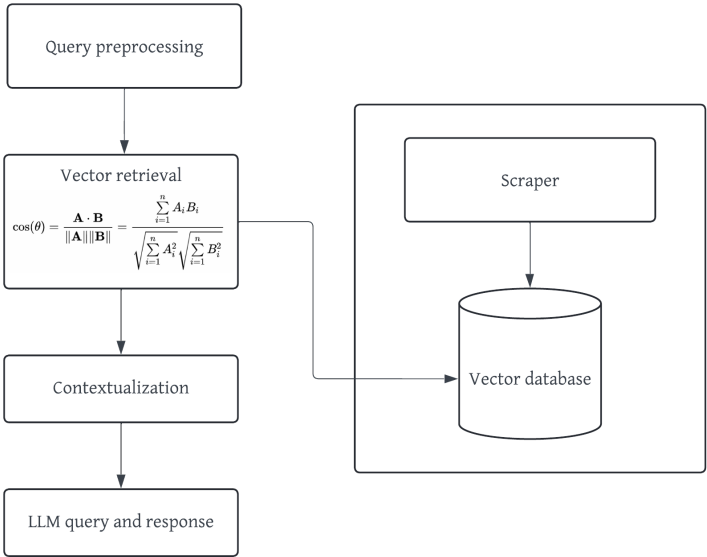
RefugeeGPT aims to mitigate some of these challenges by providing a knowledgeable AI agent to assist refugees. Grounded on a diverse dataset from United States Citizenship and Immigration (USCIS) and United Nations High Commission for Refugees (UNHCR) documents and webpages, RefugeeGPT offers comprehensive guidance on laws, regulations, and programs related to asylum in the United States. Built on Google’s Gemini 1.5 Pro model [7], this agent is designed to handle the complexities and edge cases of U.S. immigration law.

Artificial intelligence has demonstrated potential in the humanitarian sector, offering innovative solutions to support displaced populations. AI technologies can enhance decision-making in refugee status determinations [3, 1], improve access to legal aid, and provide psychological support through

---

<sup>1</sup>Deployed at this [https URL](https://github.com). Source code available on GitHub.

intelligent chatbots [5, 8]. However, the deployment of AI in humanitarian contexts also raises significant concerns, including data quality issues, algorithmic biases, and privacy risks [10]. We discuss RefugeeGPT’s current standing among these concerns and potential future work on improving these.



## 2 Methodology

The foundation of RefugeeGPT’s effectiveness lies in the robustness of its data collection process. To ensure comprehensive and reliable assistance for refugees, we curated a dataset that spans a wide array of authoritative sources. Our dataset includes over 2,500 documents, meticulously selected from organizations and agencies recognized for their expertise in refugee and immigration matters. These sources include:

- **United Nations High Commissioner for Refugees (UNHCR):** Providing global reports, guidelines, and legal texts on refugee protection and rights.
- **U.S. Citizenship and Immigration Services (USCIS):** Offering detailed information on U.S. asylum procedures, forms, and policy updates.
- **Non-Governmental Organizations (NGOs):** Including documents from organizations such as Amnesty International, Human Rights Watch, and the International Rescue Committee, which provide insights into the on-ground realities and support mechanisms for refugees.

To augment this dataset, we developed a web scraper designed to systematically gather documents from the USCIS website. The scraper initiated from the USCIS refugee page as its root and followed all referenced links, recursively forming a tree of sources (depth = 4). This automated process ensured comprehensive coverage of USCIS resources relevant to refugees and asylum seekers. As we remained on the USCIS, UNHCR, and various NGO webdomains, we therefore remained in the public domain and did not infringe on any copyright or private information. To capture the multifaceted nature of refugee experiences and the complexities of U.S. immigration law, we included documents from a diverse range of perspectives. This includes both policy-level guidelines and grassroots reports, providing a holistic view of the challenges and solutions pertinent to refugees.

Our aim was to cover all possible aspects of the asylum process, from initial application to final adjudication, as well as ancillary support services. The comprehensive coverage offered by our scraping mechanism ensured that RefugeeGPT can assist with the various stages of the refugee journey, offering

relevant information at each step. However, given the dynamic nature of immigration policies and refugee issues, our dataset is subject to continuous updates. Therefore, we have established a system for regularly incorporating new documents and updates from our sources by running our scraper regularly. This ensures that RefugeeGPT remains current and responsive to the latest developments in U.S. immigration law and refugee support, as well as consistent with various alerts and waiving of procedures.

The collected documents were pre-processed to remove any redundant or irrelevant information and structured to facilitate efficient retrieval and analysis. This involved tokenization with OpenAI’s Byte Pair Encoding algorithm [6] implementation, correcting case and formatting, and eliminating non-critical words and phrases (stop words). The process mentioned is widely used in Natural Language Processing tasks and enables the LLM to comprehend the semantic meaning of the text rather than spending attention [11] on useless tokens.

Post-processing, we utilized FAISS (Facebook AI Similarity Search) [2] to construct a high-dimensional vector database. Each document was embedded into a 364-dimensional vector space using the `sentence-transformers/all-MiniLM-L6-v2` model from HuggingFace. The embedding process can be formally represented as follows: given a set of documents  $D = \{d_1, d_2, \dots, d_n\}$ , each document  $d_i$  is converted into a vector representation  $e_i$  using the embedding function  $\text{Embed}(d_i)$ , where  $e_i \in \mathbb{R}^{364}$ . We also rely on the FAISS library to facilitate efficient similarity searches within this high-dimensional space with cosine similarity. Recall the cosine similarity algorithm as follows:

$$\text{cosine}(\mathbf{A}, \mathbf{B}) = \frac{\mathbf{A} \cdot \mathbf{B}}{|\mathbf{A}| |\mathbf{B}|}$$

where  $\mathbf{A}$  and  $\mathbf{B}$  are vectors of dimension 384, for us being the embedding vector for the query and  $e_1, e_2, \dots, e_n$ , respectively. This enables efficient search through the vector database, supported by optimizations made by the FAISS authors. Following vector retrieval, we identify the corresponding text and source, offered by FAISS, and contextualize the retrieved text and message in a prompt (see Appendix A). This prompt is fed into the default Google Gemini 1.5 Pro model through the API.

The user interface is designed in Flask and deployed with Gunicorn with 4 workers, served through Cloudflare Tunnel from a self-hosted server running Ubuntu Server 24.04 LTS. The design is aimed to be modern and intuitive and made in pure CSS and requests are streamed with vanilla Javascript and standard web protocols.

### 3 Conclusion

The development of RefugeeGPT represents a significant advancement in leveraging artificial intelligence to support refugees navigating the complex landscape of U.S. asylum laws and processes. By integrating Retrieval-Augmented Generation (RAG) with Google’s Gemini 1.5 Pro model, RefugeeGPT effectively combines the strengths of retrieval-based and generation-based approaches to provide precise, contextually relevant, and comprehensive responses. The prompt guiding the RAG system ensures that the generated responses are informative, free from legal jargon, and include additional resources for further reading. This approach not only enhances the decision-making process for refugees seeking asylum but also improves their access to critical legal information and support services.

To address data quality issues, RefugeeGPT incorporates continuous monitoring and validation protocols to ensure the accuracy and relevance of the information provided. Regular updates to the dataset are performed to reflect the latest developments in U.S. immigration law and refugee support. Algorithmic biases are a significant concern in AI systems, particularly in sensitive areas such as refugee assistance. To mitigate these biases, RefugeeGPT employs a diverse and representative dataset, and undergoes rigorous testing to identify and correct any biases that may arise. Privacy risks are managed through adherence to data protection standards and protocols. User interactions with RefugeeGPT are anonymized and do not require a login, and sensitive data is handled with utmost confidentiality. These measures ensure that refugees can seek assistance without compromising their privacy or security.

In conclusion, RefugeeGPT exemplifies the potential of AI to enhance access to legal aid and support for refugees — the benefits of such a system in providing critical assistance to vulnerable populations are substantial. Future work will focus on expanding the dataset, further refining the AI model, and enhancing the system’s ability to provide personalized support to refugees and asylum seekers.

## References

- [1] Oxford Academic. *What Can Artificial Intelligence Do for Refugee Status Determination?* 2023. URL: <https://academic.oup.com/ijrl/article/34/3-4/373/6839163>.
- [2] Matthijs Douze et al. “The Faiss library”. In: (2024). arXiv: 2401.08281 [cs.LG].
- [3] Madeleine Forster. *Refugee protection in the artificial intelligence era*. 2022. URL: <https://www.chathamhouse.org/2022/09/refugee-protection-artificial-intelligence-era>.
- [4] UN News. *Over 114 million displaced by war, violence worldwide*. 2024. URL: <https://news.un.org/en/story/2024/10/1140743>.
- [5] International Review of the Red Cross. *Harnessing the potential of artificial intelligence for humanitarian action: Opportunities and risks*. 2024. URL: <https://international-review.icrc.org/articles/harnessing-potential-artificial-intelligence-humanitarian-action-opportunities-risks>.
- [6] Rico Sennrich, Barry Haddow, and Alexandra Birch. “Neural Machine Translation of Rare Words with Subword Units.” In: *ACL (1)*. The Association for Computer Linguistics, 2016. ISBN: 978-1-945626-00-5. URL: <http://dblp.uni-trier.de/db/conf/acl/acl2016-1.html#SennrichHB16a>.
- [7] Gemini Team et al. *Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context*. 2024. arXiv: 2403.05530 [cs.CL].
- [8] Dell Technologies. *Four AI-Powered Technologies Aimed at Helping Refugees*. 2023. URL: <https://www.dell.com/en-us/blog/four-ai-powered-technologies-helping-refugees>.
- [9] UNHCR. *Global Trends 2023*. 2023. URL: <https://www.unhcr.org/globaltrends2023>.
- [10] USA for UNHCR. *Artificial intelligence in the Humanitarian Sector*. 2024. URL: <https://www.unrefugees.org/news/artificial-intelligence-in-the-humanitarian-sector>.
- [11] Ashish Vaswani et al. “Attention is all you need”. In: *Advances in Neural Information Processing Systems*. 2017, pp. 5998–6008.
- [12] Concern Worldwide. *The 10 largest refugee crises to know in 2024*. 2024. URL: <https://www.concernusa.org/story/largest-refugee-crises-2024>.

## A Prompt

The following prompt is currently in use. Note that `context` is a formatted result of the vector similarity operation and `message` is the user's query without any processing:

```
You are a helpful and qualified USCIS agent that answers questions on refugee laws, forms,  
and any general information about them.
```

```
Do not worry about anything other than that. You are allowed to give legal advice.
```

```
Given below is some context to help you answer the user's message. You do not have to use it,  
but it may help craft your answer.
```

```
CONTEXT:  
{context}
```

```
Be informative and detailed. Respond WITHOUT legal jargon. Don't be shakespeare. Be helpful.
```

```
MESSAGE: {message}
```