



User

What are the asylum procedures in Turkey?

Chatbot

According to international law, in order to make an asylum application in a country, you need to have reached the borders of that country. If you left your country for reasons of war or persecution and you are currently in Turkey, the responsibility to evaluate your asylum application and offer you legal protection lies solely with the Government of Turkey.

Persons who were forced or compelled to flee their home countries because of war or persecution and who are therefore unable to go back have the right to seek asylum in Turkey.

Navigating Humanitarian AI: Lessons Learned from Building a Chatbot Proof- of-Concept

Kevin Cole, Refugee Solidarity Network

ABOUT THE AUTHOR

Kevin Cole is Head of Partnerships and Research at the New York-based non-profit **Refugee Solidarity Network** (RSN). The orientATion project was led by Kevin Cole with the generous contributions of student volunteer Ziyad Rahman.

ACKNOWLEDGEMENTS

This research initiative was supported by a grant by the AWS Disaster Response Team, providing RSN with \$2,000 USD in AWS credits used to cover the costs of using AWS resources to conduct this research. More broadly, the underlying legal rights information was produced in the course of joint programs carried out by Refugee Solidarity Network and Refugee Rights Turkey since 2014 with funding from the US Department of State's Bureau of Population, Refugees, and Migration (BRPM).

Table of Contents

Introduction	4
Background	4
Motivations	5
Core Concerns with the Implementation of 'Humanitarian AI'	6
Factual Inaccuracies and the Principle of Accuracy	6
Algorithmic Bias & Performance Disparities	6
Accessibility and Connectivity Issues	7
Privacy and Informed Consent	7
Costs and Resource Allocation	8
Environmental Impact	8
Labor Market and Economic Displacement	8
Community Involvement and Decision-Making	9

Research Design and Implementation	9
Document Extraction from FAQ Pages	9
Creating and Saving Embeddings	10
Retrieval Augmented Generation	11
Preliminary Findings	12
LLMs Can Facilitate Information Retrieval, Albeit With Limitations	13
Currently Available ‘Guardrails’ to Prevent Misuse are of Limited Effectiveness	15
Limitations in Support for Less-Represented Languages	16
Developing and Maintaining LLM Applications is Resource Intensive	16
AI Technology Driven Data Commodification May Intensify ‘Data Extractivism’	17
Conclusions	17
Recommendations for Humanitarian Actors	18
Center Humanitarian Principles and Standards in Approaches to AI Adoption	18
Enable Genuine Community Participation & Informed Consent of Participants	19
Critically Assess Public-Private Partnership Opportunities	19
Consider Concerns of Both Self-Hosted & Hosted LLM Infrastructures	19
Embed Chatbot Applications Among Accessible, Multimodal Support Infrastructure	20
Monitor & Evaluate AI Implementations Regularly	20
Establish Robust Guardrails Against Misuse	20
Prioritize Adoption of Internal AI Tooling for Lower-Risk Applications	20
Advocate for Responsible & Equitable AI Development	21
Glossary of Terms	22

Introduction

This initiative, dubbed the “orientATIon Project,” derives its name from the term Access To Information (ATI), highlighting the core mission of facilitating access to critical legal information for asylum-seeking communities. ATI is a foundational right, serving as a crucial element of empowerment and a precursor to enabling communities to understand and enact their legal rights effectively.

Between October 2023 - January 2024, Refugee Solidarity Network (RSN) conducted an exploratory research project designed to assess the ethical and practical feasibility of leveraging **Large Language Models (LLMs)** to power a **chatbot** offering individuals seeking protection with general guidance on asylum procedures from a perspective grounded in the applied experience of developing a proof of concept.

During the course of the orientATIon project, RSN tested the current bounds of LLM-powered chatbots, applying various techniques to refine the ability of the prototype chatbot's capacities to provide accurate, specific and relevant guidance on asylum-related questions in multiple languages.

We hope that this praxis-grounded exploration will contribute to discussions surrounding the significant challenges humanitarian actors can expect to face when implementing **Artificial Intelligence (AI) applications** in sensitive contexts, while also highlighting some of the potential positive impacts of adopting these technologies in our work. The report concludes with a set of recommendations drawing on both our own reflections as well as commentary from other members of the “Humanitarian AI” community.

Background

RSN understands refugees' access to information to be an essential, precursor component to enabling their legal empowerment. As a result, an important pillar of RSN's Theory of Change has been to assist in the development of informational materials covering the rights and obligations of refugees and asylum-seekers in the countries which host them.

Since March 2015, the US-based non-profit, Refugee Solidarity Network (RSN), and Turkish non-governmental organization, Refugee Rights Turkey (RRT), have been engaged in a comprehensive partnership with the primary goal of improving legal protection of refugees in Turkey, which hosts one of the largest populations of refugees in the world. The nearly decade-long partnership between RSN and RRT has significantly strengthened the availability of reliable legal information and specialized legal assistance services for refugees in Turkey while also strengthening refugee protection capacities among Turkey's community of legal professionals.

Only having promulgated a domestic legal framework for asylum procedures in 2014, the protection landscape in Turkey was initially quite weak and catered almost exclusively to (Arabic-speaking) Syrians forcibly displaced by the intense violence of the Syrian civil war. In 2015, RSN and RRT collaborated to produce and distribute a series of “Know Your Rights”-style informational booklets in multiple languages to better address the needs of the diverse and growing refugee and asylum-seeker community in Turkey. Recognizing the diverse needs and informational preferences within refugee communities in Turkey, both

organizations committed to continuous, iterative expansion and improvement of their legal information ecosystem. In 2019, RSN and RRT launched the [Informational Portal for Refugees in Turkey](#), a dedicated digital information platform designed with input from refugee communities in Turkey to provide critical know-your-rights information to asylum-seekers in accessible and effective formats.

The launch of the [Informational Portal for Refugees](#) not only expanded access to legal information into multiple formats, but also enabled RSN and RRT to make more frequent updates in response to legislative changes and other critical developments, enhancing the timeliness and relevance of information. Currently, the platform provides information for both Syrians (who are eligible for group-based temporary protection status in Turkey) and non-Syrians (who are eligible for individually-assessed conditional refugee status, or for refugee status for individuals from European countries of origin, in Turkey).

In total, a body of more than 120 question-answer pairs covering the core rights, obligations and procedures relevant to asylum-seeking communities in Turkey have been collected and published on the platform in a total of ten languages: Arabic, English, Farsi, French, Pashto, Russian, Somali, Sorani Kurdish, and Urdu.

Motivations

A significant motivation behind the orientATIOn research project was one of the key findings of user experience research carried out with users of the Informational Portal for Refugees in Turkey; namely, that the information needs and preferences of asylum-seekers in Turkey are quite diverse, and users express a desire to receive information tailored to their particular situations.¹

A few of the leading questions behind this research were:

- Can LLM technology be utilized to improve refugee and asylum-seekers' access to critical information regarding their rights and obligations?
- Can LLM technology help refugees and asylum-seekers to receive more personally-applicable information?

It is our hope that with positive answers to these questions, we can explore further integration of these technologies into future programming, complementing and improving the efficiency of ATI interventions and activities that are typically resource intensive.

Our exploratory investigation was also underpinned by RSN's mandate to uphold humanitarian ethical principles: while there is much ongoing discussion about the use of AI in humanitarian contexts, incorporating a range of viewpoints,² we hope that reflecting on our practical experiences during this project can contribute actionable guidance to

¹ For example, in an evaluation activity held with RSN and RRT's beneficiaries between July - November of 2022, 41% of respondents surveyed found informational materials to be too specific in detail, while 28% found them too general in nature.

² Among other resources, the "Humanitarian AI Today" podcast brings together a wide range of perspectives and actors working on this topic. "Podcast on Humanitarian Applications of Artificial Intelligence," *Humanitarian AI Today*, accessed May 13, 2024, <https://humanitarianaitoday.org/>.

conversations tailored to assist humanitarian organizations, donors, and private sector partners to effectively navigate the opportunities and challenges presented by these technologies.

Core Concerns with the Implementation of ‘Humanitarian AI’

Humanitarian organizations have been cautious in their adoption of AI tools, opting for limited proof of concept initiatives.³ This caution reflects the complexity of using new technologies safely and responsibly, rather than a lack of engagement. Substantially accelerated by the exponential growth in public uptake of AI technologies over the past two years, a subspecialty of "Humanitarian AI" has begun to crystallize as both organizations and individuals seek to create a community of practice and guidance around the adoption of AI tools towards humanitarian ends.⁴

In discussions taking place under the broad umbrella of Humanitarian AI, a number of ethical and practical concerns associated with, but certainly not all unique to, the application of AI systems to humanitarian concerns, have emerged.

Factual Inaccuracies and the Principle of Accuracy

A persistent concern is the propensity for LLMs and other generative AI techniques to generate outputs which contain factual inaccuracies, often referred to as "hallucinations."⁵ While the massive, primarily human-generated bodies of text that LLMs are trained on do enable these models to generate outputs that mimic natural language and often seem remarkably thoughtful, the inherently probabilistic algorithm they employ to do so is not capable of ensuring the factual accuracy of the contents they produce.⁶ The challenge of ensuring accuracy is particularly acute in humanitarian contexts, where incorrect advice or information can have serious, potentially fatal, consequences.⁷

Algorithmic Bias & Performance Disparities

Algorithmic bias in AI systems is a significant ethical concern,⁸ particularly elevated in sensitive areas such as healthcare, justice, and humanitarian aid. Bias frequently stems from the reproduction of human bias in data used to train models, the use of unrepresentative data and flawed prediction models. It is important to highlight that the potential negative impacts of algorithmic bias are not

³ Lydia Tanner, "Learning from Innovations Using Artificial Intelligence" (Creating Hope in Conflict: A Humanitarian Grand Challenge, 2023), <https://humanitariangrandchallenge.org/research-learning-reports/>.

⁴ An illustrative example can be found in the call to consider the formulation and adoption of a "humanitarian AI manifesto": Abeba Birhane et al., "Do We Need a Humanitarian Manifesto for AI?" (AI UK Fringe, London, UK, March 18, 2024), <https://www.cdacnetwork.org/news/do-we-need-a-humanitarian-manifesto-for-ai-join-cdac-network-at-the-alan-turing-institutes-ai-uk-fringe>.

⁵ Jessica Fjeld et al., "Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI," SSRN Scholarly Paper (Rochester, NY, January 15, 2020), 31, <https://doi.org/10.2139/ssrn.3518482>.

⁶ Emily M. Bender et al., "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜," in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '21 (New York, NY, USA: Association for Computing Machinery, 2021), 610–23, <https://doi.org/10.1145/3442188.3445922>.

⁷ Andrew Deck, "AI Translation Is Jeopardizing Afghan Asylum Claims," *Rest of World*, April 19, 2023, <https://restofworld.org/2023/ai-translation-errors-afghan-refugees-asylum/>.

⁸ Leon Yin, Davey Alba, and Leonardo Nicoletti Technology + Equality, "OpenAI's GPT Is a Recruiter's Dream Tool. Tests Show There's Racial Bias," *Bloomberg*, March 8, 2024, <https://www.bloomberg.com/graphics/2024-openai-gpt-hiring-racial-discrimination/>.

limited to applications directly employing algorithmic decision-making (e.g., automated evaluation of the credibility of a written asylum claim), but may also be reflected in a lower quality of service when using an AI-powered system.

One issue of particular salience for the application of AI, and particularly **Natural Language Processing (NLP)** technologies, in the humanitarian sector is that these technologies exhibit significantly reduced performance for lesser-represented languages,⁹ which may lead to exclusion from services that leverage these technologies, further marginalizing speakers of these languages.

In this scenario, humanitarian adoption of AI tools could potentially exacerbate existing linguistic injustice, underscoring the difficulty of creating inclusive AI tools without perpetuating and even amplifying existing biases. The fact that the scarcity of diverse datasets to train AI models underlies this issue reflects the reality that complex issues of linguistic and cultural representation cannot be effectively overcome through technological advancements alone.¹⁰

Accessibility and Connectivity Issues

The dependence of LLM-based applications on continuous internet connectivity limits their usability in regions with low internet access as well as by individuals with limited or no access to a suitable device with which to access the application. Moreover, chatbot interfaces that are dependent on text input may present accessibility barriers to individuals who experience difficulty in using written language to communicate or who do not have access to assistive technologies.

Privacy and Informed Consent

Users' interactions with AI systems, and particularly LLM-powered chatbot applications, can present complex privacy concerns and increase the likelihood of personally identifying information (PII) being collected, including unintentionally, and transmitted to third parties (such as an LLM model provider). This raises significant privacy concerns and highlights the importance of ensuring that users provide and maintain informed consent,¹¹ a principle requiring that individuals be fully aware of the risks, benefits, and alternatives of sharing their data.¹²

Establishing and maintaining informed consent can be particularly challenging in the context of AI-driven systems, both due to limited understanding of these novel technologies by the general public, as well as the multiple use cases for user-provided data, such as model evaluation and fine-tuning, throughout the AI product lifecycle.

⁹ Becky Kazansky et al., "Chatbots in Humanitarian Contexts: Learning from Practitioner Experiences" (International Federation of Red Cross and Red Crescent Societies, 2023), 8, https://communityengagementhub.org/wp-content/uploads/sites/2/2023/06/20230623_CEA_Chatbots.pdf.

¹⁰ Fjeld et al., "Principled Artificial Intelligence," 27; Matteo Wong, "The AI Revolution Is Crushing Thousands of Languages," *The Atlantic*, April 12, 2024, <https://www.theatlantic.com/technology/archive/2024/04/generative-ai-low-resource-languages/678042/>.

¹¹ Tanner, "Learning from Innovations Using Artificial Intelligence," 8: "Consent should be fully informed, specific, unambiguous, and freely given. This requires particular consideration where there are language barriers, low literacy levels or people may be confused or afraid. Often, the specific uses of data may not be known at the time of data collection - for example, if it may be used as training data for multiple systems."

¹² Fjeld et al., "Principled Artificial Intelligence," 14.

Costs and Resource Allocation

Because the development and deployment of AI-driven applications frequently depends on contracting the services of third party providers (whether to supply the computational power to train or run LLM models, the infrastructure to transmit and store data, or other services critical to AI product architectures), the adoption of AI-driven applications in humanitarian efforts poses the risk of substantially redirecting aid funding towards commercial technology companies. Despite promises to lower costs and improve efficiency, the development, monitoring, and maintenance resources these applications require over their lifespan may negate anticipated gains.¹³

The following paraphrase of a discussion held with managers and directors of humanitarian organizations brought together by MERL Tech succinctly encapsulates the implications of humanitarian funding prerogatives on the safety of humanitarian AI initiatives:

NGOs should start by integrating AI into internal business processes rather than going straight to programmatic use of AI, suggested one person, because we don't have the budget or capacity to use AI responsibly in programs. Another person agreed. 'Our overhead is restricted by charity watch dogs. We lose our rating on Charity Navigator if we go above a certain percentage. So just 2-3% of our budget goes to cover our entire IT and data strategy.' This means that NGOs are severely handicapped when it comes to safely exploring the use of AI for programs.¹⁴

Environmental Impact

The environmental footprint of training and operating large-scale AI models, including substantial resource consumption necessary to house the data centers that form their operational infrastructure, presents a significant ethical dilemma.¹⁵ Concerns center around the potential for the environmental impacts of adopting AI-powered technologies to result in further heightening the risks related to climate change which are already producing vulnerability across the globe.¹⁶

Labor Market and Economic Displacement

AI applications promise cost savings but have also been observed contributing to a "race-to-the-bottom" in wages and conditions for workers,¹⁷ exacerbating economic inequalities. Moreover, the

¹³ Kazansky et al., "Chatbots in Humanitarian Contexts: Learning from Practitioner Experiences," 12.

¹⁴ Linda Raftree, "10 Ways to Help Your Organization Manage the Pain of AI Adoption," *MERL Tech* (blog), April 1, 2024, <https://merltech.org/10-ways-to-help-your-organization-manage-the-pain-of-ai-adoption/>.

¹⁵ Jude Coleman, "AI's Climate Impact Goes beyond Its Emissions," *Scientific American*, accessed May 14, 2024, <https://www.scientificamerican.com/article/ais-climate-impact-goes-beyond-its-emissions/>.

¹⁶ "The US Must Balance Climate Justice Challenges in the Era of Artificial Intelligence," *Brookings*, accessed May 14, 2024, <https://www.brookings.edu/articles/the-us-must-balance-climate-justice-challenges-in-the-era-of-artificial-intelligence/>.

¹⁷ Billy Perrigo, "OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic," *TIME*, January 18, 2023, <https://time.com/6247678/openai-chatgpt-kenya-workers/>; AI Intersections Database, "The 'Ghost Work' of AI Is Done by Low-Paid Workers across Developing Countries," *AI Intersections Database*, accessed April 22, 2024, <https://aidatabase.mozilla.org/ai-intersections-database/ai-0272/>; Milagros Miceli and Julian Posada, "The Data-Production Dispositif" (arXiv, May 24, 2022), <http://arxiv.org/abs/2205.11963>.

reality that most LLMs have been trained on unlicensed materials scraped from the internet raises questions about the ethical use of AI, given its propensity for reproducing copyrighted content.¹⁸

Community Involvement and Decision-Making

Although participatory design is frequently underlined as an ethical imperative for the development of humanitarian services, engaging affected communities in the development and implementation of AI solutions remains a challenge.¹⁹

Community participation practices that are ultimately tokenistic or extractive in their design or implementation may undermine the potential for meaningful involvement in the creation, application and refinement of AI-driven tools, highlighting the need for genuinely participatory design. The risk of dedicating resources towards the implementation of solutions that do not align with the needs or preferences of those they aim to serve underscores the importance of context-sensitive, user-centered design in humanitarian AI projects.

Research Design and Implementation

The orientATlon project was structured around the development of a proof of concept, LLM-powered chatbot capable of receiving free-form (text) user input and responding to inquiries regarding asylum-seekers' rights and navigation of the asylum system in Turkey.

The project's architecture consisted of three primary components:

1. Document extraction from RSN & RRT's Informational Portal for Refugees in Turkey,
2. Creation and saving of vector embeddings from extracted documents,
3. Generating responses to user queries using retrieval augmented generation (RAG) methods.

Because this research initiative was supported by a grant from the Amazon Web Services (AWS) Disaster Response Team, the technical implementation of the project was restricted to using resources available via the AWS Bedrock platform,²⁰ including the foundational models this platform supported at the time of research.

Document Extraction from FAQ Pages

To ensure that the orientATlon chatbot was capable of providing up-to-date and verified information pertaining to asylum-seekers' rights and obligations under the Turkish asylum system and other relevant law, it was first necessary to collect and format the question-answer pairs from the Informational Portal for Refugees in Turkey that would serve as the chatbot's "knowledge base."

¹⁸ Harry H. Jiang et al., "AI Art and Its Impact on Artists," in *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, AIES '23 (New York, NY, USA: Association for Computing Machinery, 2023), 363–74, <https://doi.org/10.1145/3600211.3604681>.

¹⁹ Kazansky et al., "Chatbots in Humanitarian Contexts: Learning from Practitioner Experiences," 20.

²⁰ "What Is Amazon Bedrock? - Amazon Bedrock," accessed May 17, 2024, <https://docs.aws.amazon.com/bedrock/latest/userguide/what-is-bedrock.html>.

Document Extraction from FAQ Pages

```
{  
  "section": "Registration & Status",  
  "question": "What is temporary protection?",  
  "answer": "Temporary protection is the protection status established by the Temporary Protection Regulation dated 22 October 2014. According to this regulation, the status is applied to foreigners who were forced to leave their country, who cannot return to the country they left, who arrived at or crossed Turkish borders en masse or individually, and who will not have their international protection needs adjudicated under an individual procedure.",  
  "language": "English",  
  "deep_link": "https://multecihaklari.info/services/registration-status/?section=questions&question=0"  
}
```

```
{  
  "section": "Marriage & Divorce",  
  "question": "What are the conditions for civil marriages in Turkey?",  
  "answer": "In Turkey, a marriage ceremony can be performed between a freely and fully consenting man and woman. Persons wishing to marry must be bachelor at the time they apply for the marriage. In Turkey, the legal age for marriage is 18. ....",  
  "language": "English",  
  "deep_link": "https://multecihaklari.info/services/marriage-divorce/?section=questions&question=2"  
}
```

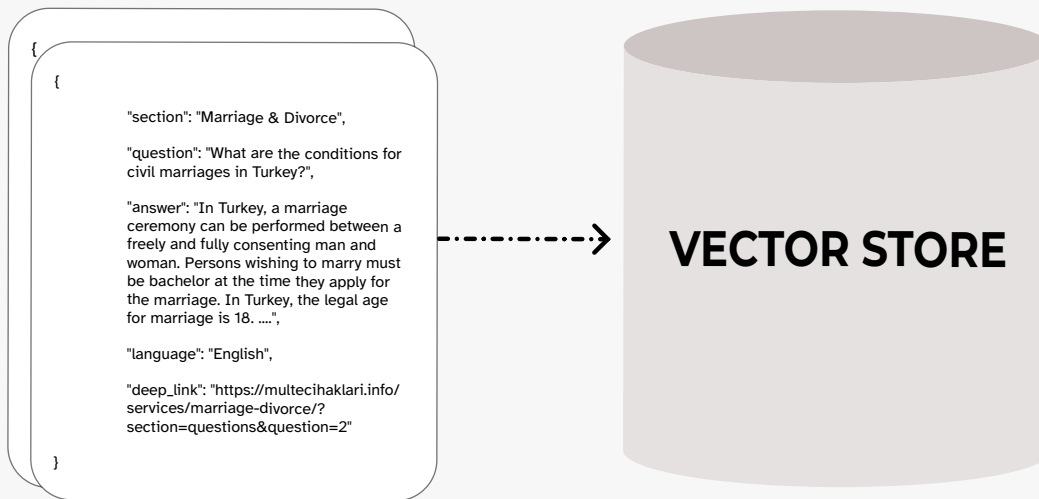
To accomplish this, we created a web scraping tool to extract the information directly from the Informational Portal's website, saving the question-answer pairs in a machine-readable format. This process was done separately for information applicable for Syrians and for information applicable to non-Syrians in order to ensure that the chatbot could tailor its responses to the chatbot user's situation under Turkish law.

The extracted data includes both the full text of each question-answer pair as well as additional metadata such as the thematic section which the information was sourced from and a direct link to the source, enhancing the chatbot's ability to provide precise answers and to cite sources it uses in generating a response.

Creating and Saving Embeddings

In this phase, information previously extracted from FAQ pages and saved as machine-readable files are subsequently transformed into **vector embeddings** using an **embedding model**. These embeddings offer a means of capturing the semantic contents of specific words and sentences, serving as a detailed mathematical representation of the input texts. By producing this mathematical representation of underlying texts, the system is later able to retrieve relevant contents saved in the **vector store**.

Creation and Saving of Embeddings



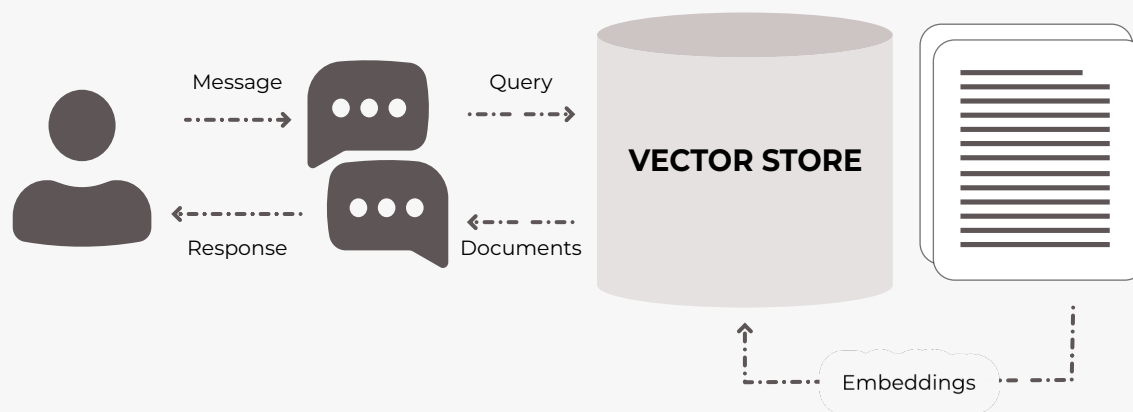
Vector embeddings are generated using an embedding model, which produce a mathematical representation of words/sentences' semantic relatedness and are stored in a searchable database known as a vector store.

The vector store acts as a searchable database where these embeddings are saved. By producing and saving embeddings in a vector store, the system is later able to perform **semantic searches** and retrieve information that is most relevant to the user's inquiry.

Retrieval Augmented Generation

The final component in the project's architecture utilizes the previously stored embeddings to improve the chatbot's responses through a process called **Retrieval Augmented Generation (RAG)**. This method involves an intermediate step where the chatbot uses its embeddings to perform semantic searches, retrieving relevant documents or snippets of information.

Retrieval Augmented Generation (RAG)



RAG is a method that leverages semantic search over a set of provided documents to dynamically embed context into prompts processed by an LLM, enabling the LLM to draw on this contextual information in formulating a response.

These retrieved snippets provide context for the chatbot, enabling it to generate more accurate and contextually relevant responses. This advanced approach allows the chatbot to produce better-informed answers than would otherwise be generated when relying solely on the LLM's initial (unsupervised) training data.

This general project architecture was implemented and tested on researcher-generated “user” queries in Arabic, English, Farsi, Pashto, Somali, Sorani Kurdish, Turkish and Urdu.

Preliminary Findings

After implementing the project architecture described above, we performed an informal evaluation of the orientATlon chatbot's performance in fielding a set of researcher generated inquiries on a range of topics related to navigating the asylum system in Turkey. This evaluation, while not exhaustive, was helpful in identifying a number of preliminary findings which we believe to demonstrate key concerns with the application of LLM-based informational products in humanitarian settings.

It is important to note that the limited scope of this research pilot, as well as the incredibly fast pace of advancement in LLM technologies and corresponding implementation techniques, meant it was not possible to apply many of the advanced and emergent approaches to the field-wide challenges of optimizing the processing of natural language towards specific operational goals. As such, this report **does not intend to draw conclusive statements with regards to the capabilities of LLM-powered applications**, but instead to reflect on the significant challenges that organizations are faced with when attempting to apply these technologies to complex, real world use cases, particularly within the heightened realm of concerns relevant to humanitarian interventions.

LLMs Can Facilitate Information Retrieval, Albeit With Limitations

In our testing, we found that the orientATlon chatbot was capable of synthesizing relevant and accurate responses in many cases, especially in response to user inquiries that were reflective of the types of questions included in the Refugee Info Portal's existing QA pairs.

The following are illustrative examples of responses to researcher-generated "user inquiries" provided by the orientATlon chatbot:²¹

User

What do I need to study at a university in Turkey?

Chatbot

Based on the information from the Information Portal for Refugees in Turkey, the main requirements to study at a university in Turkey are:

1. Graduate from high school
2. Meet other entrance conditions such as taking the Foreigner Student Exam (Yabancı Öğrenci Sınavı – YÖS) and obtaining a qualifying score for the university you want to attend
3. For those who graduated high school outside of Turkey or cannot demonstrate previous school attendance, you may need to register at Open Education High School (Açık Öğretim Lisesi)
4. Take and pass the Examination for Foreign Students (YÖS) organized by universities
5. Pay any required application fees
6. Provide documentation related to your previous education as requested by the university
7. Demonstrate language proficiency if required by the university or department

I would suggest consulting the Foreign Student Departments or university websites directly for the full and latest requirements. The Information Portal also has more details in its Education section. Let me know if you need any clarification or have additional questions!

Sources

<https://multecihaklari.info/services/rights-and-procedures-for-unaccompanied-minors-2/?section=questions&question=16>
<https://multecihaklari.info/services/rights-and-procedures-for-unaccompanied-minors/?section=questions&question=16>
<https://multecihaklari.info/services/education/?section=questions&question=9>

However, more nuanced inquiries requiring recalled source documents' contents to undergo complex manipulation (e.g., discarding some retrieved information as contextually irrelevant, synthesizing information from across multiple document sources) frequently resulted in responses

²¹ Due to space and interactivity limitations when working with the PDF format, a limited selection of English-language chatbot interactions is included here. An interactive exploration tool can be found on the author's website: <https://kevin-cole.com/blog/orientation-access-to-information/>.

that were significantly less precise and/or relevant than would be expected from human-generated responses.

User

What should I do if my asylum application is rejected?

Chatbot

This means that your international protection application is rejected. The authorities are under the obligation to communicate this decision, together with its reasons, to you or your legal representative or lawyer. You have the right to appeal against this decision.

As international protection applications made by persons under administrative detention are usually assessed under the accelerated procedure, the key legal remedy to challenge negative decisions is judicial. You or your legal representative or your lawyer, if any, must appeal to the administrative court against the negative decision latest within fifteen days from the notification of the decision to you.

For detailed information on how to access legal support while under administrative detention, you can review the relevant sections. You can also get information by calling the Detention Hotline of Refugee Rights Turkey, at 0507 218 62 85.

Sources

<https://multecihaklari.info/services/registration-status-2/?section=questions&question=30>
<https://multecihaklari.info/services/rights-and-procedures-for-persons-in-administrative-detention-2/?section=questions&question=15>
<https://multecihaklari.info/services/rights-and-procedures-for-persons-in-administrative-detention/?section=questions&question=15>

While the information provided in the above response is factually correct, it may be considered out of context as the user did not specify being under administrative detention. The inclusion of this potentially confusing information would likely not be made by a human expert, who would instead prefer to provide more complete contextualizing information in their response.

The following example demonstrates the chatbot's behavior when the retrieved context information is determined to be insufficient to answer the user's inquiry. This behavior can be influenced by modifying the underlying LLM's system prompt; however, adherence to these instructions cannot be guaranteed. Correct handling of this type of scenario was less common when working in less-represented languages such as Somali and Sorani Kurdish.

User

How long does the asylum process take in Turkey?

Chatbot

Unfortunately, my training materials do not include specific information on how long the asylum process takes in Turkey. However, I would suggest reaching out to Refugee Rights Turkey for more details on the asylum procedures and timelines in Turkey. As an organization providing legal assistance to asylum seekers, they are likely to have helpful information to answer your question. I apologize that I do not have more details to provide. Please let me know if you have any other questions!

Sources

<https://multecihaklari.info/services/registration-status-2/?section=questions&question=5>
<https://multecihaklari.info/services/registration-status-2/?section=questions&question=49>
<https://multecihaklari.info/services/registration-status-2/?section=questions&question=42>
<https://multecihaklari.info/services/registration-status-2/?section=questions&question=34>

While a number of the performance issues related to the accuracy, completeness and relevance of the responses generated by the orientATlon chatbot can likely be improved through the application of more advanced techniques, these observations demonstrate a core misapprehension regarding the state of ‘artificial intelligence’ technologies, particularly as concerns LLMs. While these models are certainly capable of generating authoritative sounding natural language texts, their algorithmic function remains *predicative*.

This raises important considerations for their deployment in sensitive humanitarian settings: can AI tools be effectively ‘de-risked’? Who is responsible for (mis-) information produced by an AI interface?

Currently Available ‘Guardrails’ to Prevent Misuse are of Limited Effectiveness

The freeform nature of LLMs' text generation capabilities pose complicated security concerns when used in production settings. These security risks are cause for concern not only within the humanitarian sector; they are operational risks being actively confronted by actors across all levels, from individual product developers to the largest, most well-equipped AI technology companies like OpenAI and Alphabet.

Primary concerns include adversarial misuse, such as applying techniques like prompt injection or jailbreaking to redirect chatbots to perform other tasks, essentially ‘hijacking’ the chatbot's underlying resources to complete irrelevant tasks, with the associated costs being borne by the product's owners, or enabling the chatbot to return undesirable responses which reflect poorly on the host organization. While some mitigation techniques can reduce the potential harms or lower the ‘rewards’ of this type of misuse, the operating nature of LLMs makes them inherently vulnerable to adversarial redirection.²²

²² Kai Greshake et al., “Not What You’ve Signed up for: Compromising Real-World LLM-Integrated Applications with Indirect Prompt Injection” (arXiv, May 5, 2023), 5, <https://doi.org/10.48550/arXiv.2302.12173>.

Limitations in Support for Less-Represented Languages

We also anecdotally observed that the orientATlon chatbot's performance, even on relatively simple user inquiries, was degraded when working with less-represented languages, as well as when users prefer less-represented dialects, registers or non-standard writing methods otherwise poorly represented in the model's training data.²³

In our limited testing, the lowest levels of performance were observed when working with user inquiries and source documents written in Somali and Sorani Kurdish. Performance degradation was primarily seen in the generation of grammatically incorrect response texts, although response latency was also higher when working with these languages.

This may be attributable to lower-quality translations of source FAQ materials into these languages, rather than reflecting directly on the performance of embedding/completion LLMs utilized in the orientATlon chatbot. This underlines the important and practical challenges that humanitarian organizations already face in attempts to ensure the highest possible quality of data (including translations), especially in languages not spoken by humanitarian organizations' own staff members. Moreover, these observations illustrate the importance of engaging directly with affected communities to proactively identify and resolve issues with quality control and cultural appropriateness when localizing contents into languages spoken by affected communities.

While enhancing the performance of models in specific languages using techniques such as custom LLM model creation or model fine-tuning offer potential solutions to these challenges,²⁴ these approaches require the collection of large, high-quality datasets in targeted languages, which may be impractical or even impossible in many scenarios. This limitation also raises concerns about whether AI-powered tool performance degradation and/or lack of support for some languages could lead to the exclusion of specific, and typically already underrepresented, communities from the benefits touted by these technologies.

Developing and Maintaining LLM Applications is Resource Intensive

Another key takeaway from this research pilot project has been the understanding that while the resources required for effectively and safely developing, deploying, monitoring and refining LLM-based applications (in other words, ensuring such applications are "production-ready") are becoming more accessible, they still remain out of reach for many non-profit and humanitarian actors. The rapid pace of advancements in LLM technologies poses both an opportunity as well as significant risk for adoption by actors operating in the humanitarian sector, which is prototypically resource-limited in nature.

Given the cutting-edge nature of these technologies there is a general shortage of AI knowledge and expertise in the humanitarian sector and to combat this shortage, it is unclear whether concentrating expertise in-house or leveraging external partnerships will provide the strongest long-term advantages while also enabling humanitarian organizations to maintain rigorous adherence to their guiding ethical principles while advancing their core mandates.²⁵

²³ This observation echoes findings reported from the experience of other practitioners, c.f. Kazansky et al., "Chatbots in Humanitarian Contexts: Learning from Practitioner Experiences," 8.

²⁴ Prabha Kannan, "Improving Equity and Access to Non-English Large Language Models," Stanford University Human-Centered Artificial Intelligence, April 22, 2024, <https://hai.stanford.edu/news/improving-equity-and-access-non-english-large-language-models>.

²⁵ Tanner, "Learning from Innovations Using Artificial Intelligence," 7.

Observations published in The Engine Room's 2023 report titled "Chatbots in Humanitarian Contexts: Learning from Practitioner Experiences" underline the challenges that resource limitations pose for humanitarian organizations seeking to implement more advanced chatbot systems:

Some humanitarian staff members we spoke to wonder whether humanitarian organisations should rather build and host their own communications infrastructures; others argue that having humanitarian organisations create and manage their own internal technological platforms also creates risks, as the in-house resources necessary for their technical development and maintenance often don't match what's available.²⁶

AI Technology Driven Data Commodification May Intensify 'Data Extractivism'

The fact that data, whether in the form of training inputs, user-generated refinement, or otherwise, holds a privileged and pivotal role in the operationalization and advancement of AI and LLM-based technologies means that data itself functions as one of the most valuable commodities in the AI economy. A 2023 report "Learning from Innovations Using Artificial Intelligence" commissioned by Grand Challenges Canada captures the centrality of sourcing data for the successful implementation of AI-driven interventions:

"Access to sufficient high-quality data was the single most common issue preventing success. More than half of the projects encountered data as an issue and it was the primary issue for almost half of the projects. Access to data was often constrained or managed by on-the-ground third-party organizations, a very common situation in humanitarian settings."²⁷

This heavy reliance on large and industry-specific datasets could further encourage an already documented form of 'humanitarian extractivism' whereby humanitarian actors are driven to collect extensive data, oftentimes in exchange for providing communities in situations of humanitarian need with the support that they require.²⁸ Concerns have been rightly raised that public-private partnerships between humanitarian actors and privately-held tech companies, frequently billed as "tech for good" initiatives, may be particularly susceptible to such arrangements whereby technical support for the adoption and deployment of AI driven products is offered in exchange for access to valuable data produced by humanitarian actors.²⁹

Conclusions

In our research pilot, the orientATIon chatbot demonstrated some promising capabilities in terms of retrieving and synthesizing information on the rights of asylum-seekers in Turkey, responding to natural language questions in multiple languages. If integrated into RSN and RRT's existing ecosystem of informational services, this type of application could help to transform the experience

²⁶ Kazansky et al., "Chatbots in Humanitarian Contexts: Learning from Practitioner Experiences," 19.

²⁷ Tanner, "Learning from Innovations Using Artificial Intelligence," 5.

²⁸ The concept of "humanitarian extractivism" has been popularized by Kristin Bergtora Sandvik, who published a book on the topic in 2023: *Humanitarian Extractivism: The Digital Transformation of Aid, Humanitarianism* (Manchester: Manchester University Press, 2023).

²⁹ Sofie Elbæk Henriksen and Lisa Ann Richey, "Google's Tech Philanthropy: Capitalism and Humanitarianism in the Digital Age," *Public Anthropologist* 4, no. 1 (March 31, 2022): 21–50, <https://doi.org/10.1163/25891715-bja10030>; Mirca Madianou, "Nonhuman Humanitarianism: When 'AI for Good' Can Be Harmful," *Information, Communication & Society* 24, no. 6 (April 26, 2021): 850–68, <https://doi.org/10.1080/1369118X.2021.1909100>.

of accessing information, providing an experience which bridges the gap between the online [Informational Portal for Refugees in Turkey](#) and traditional assistance formats such as telephone hotlines and in-person counseling.

For example, interacting with this type of chatbot could help to provide users with shorter and more personally-relevant information to answer their inquiries, or enable informational contents to be dynamically rephrased or simplified in accordance with a user's individual preferences and needs. However, despite the technology's potential, our experience has also uncovered a number of practical, ethical and security concerns that the adoption of LLM-based applications pose for humanitarian organizations. The rapid pace of technological advancements in AI necessitates a cautious approach that considers related multifaceted and intersectional concerns.³⁰

Recommendations for Humanitarian Actors

Reflecting on the project's preliminary findings as well as broader discussions taking place under the banner of "humanitarian AI," we have elaborated the following list of recommendations for humanitarian actors evaluating the potential adoption of LLM-powered chatbot information dissemination tools, as well as for donor organizations seeking to enable their responsible development and deployment:

- **Center Humanitarian Principles and Standards in Approaches to AI Adoption**

Whether while adopting or considering the adoption of AI technologies, humanitarian organizations must always ensure that their efforts and approach are firmly anchored in humanitarian principles and reflect humanitarian standards. In particular, the humanitarian standards of appropriate and relevant response; communication, participation and feedback; and the proper management of resources for their intended purpose.³¹

AI tools must be designed and deployed in a manner that prioritizes the well-being and rights of the individuals they aim to assist, without causing harm or exacerbating vulnerabilities, and with the informed consent of all involved parties. It is crucial for organizations to critically evaluate how and if AI solutions align with their core missions, and to probe the potential ethical implications of their use, ensuring that these technologies serve as a means to enhance human dignity and aid delivery rather than merely as 'cost-saving' mechanisms.

One of the most important steps that humanitarian actors can take towards reducing their risk of causing harm through the adoption of AI tools is to recognize the basic truth that intersectional, socially-rooted issues cannot find solutions based solely in the application of any particular technology, no matter how advanced. In other words, rejecting 'technosolutionism'.³²

³⁰ Marjolein Lanzing and Katrin Schulz, "Harmful NLP: Towards a Systemic Injustice Approach" (*Journée d'étude journée éthique et TAL*, Nancy, France, 2024), <https://www.atala.org/content/ethique-tal>. Lanzing and Schulz use the term "systemic injustice approach" to describe a mindset which includes both technological and social-political challenges in the application of AI.

³¹ "What Are Humanitarian Standards? | Sphere Standards," Sphere, accessed May 16, 2024, <https://www.spherestandards.org/humanitarian-standards/>.

³² The term "technosolutionism" was coined and popularized by Evgeny Morozov in his book *To Save Everything, Click Here: The Folly of Technological Solutionism* (New York: PublicAffairs, 2013); Broussard has described a similar phenomenon under the name of "technochauvinism": Meredith Broussard, *Artificial Unintelligence: How Computers Misunderstand the World* (Cambridge, Massachusetts: The MIT Press, 2019).

Assessments of the potential risk for human rights impacts related to the adoption of AI-driven applications should not only consider purely technical, security or financial risks associated, but also take into account the potential for causing or exacerbating other forms of injustice and harm. Towards this end, humanitarian actors may contribute to calls for the creation of a 'Humanitarian AI Manifesto'.³³

- **Enable Genuine Community Participation & Informed Consent of Participants**

To achieve effective community engagement, broader efforts to provide affected communities with the resources to understand the potential risks and benefits of AI technologies must be undertaken. Ideally, affected communities should benefit from general awareness-raising initiatives on the potential risks and benefits of AI, decoupled from individual project/initiative-centered community engagement processes. It is also important for donor organizations to recognize the importance of these activities and, where possible, to provide financial support for their inclusion in humanitarian responses.

Human Rights Impact Assessment (HRIA) tools such as the "Framework for Meaningful Engagement: Human Rights Impact Assessments of AI" can assist humanitarian organizations to engage in more genuine and participatory conversations with affected communities.³⁴

In order for the consent provided by users from affected communities to be effectively informed, a clear and comprehensive overview of planned data collection and usage should be shared with potential application users.

- **Critically Assess Public-Private Partnership Opportunities**

In line with humanitarian organizations' responsibility to uphold humanitarian principles and standards, organizations and actors considering engaging in partnerships with private, for-profit entities such as technology providers have a duty to critically assess such arrangements and their potential negative impacts on affected communities, for example through improper and unethical data governance arrangements. This means evaluating the terms of engagement to guard against potential conflicts of interest and ensuring that such collaborations do not compromise the autonomy or data privacy of affected communities.

- **Consider Concerns of Both Self-Hosted & Hosted LLM Infrastructures**

Whether choosing to host LLMs and their supporting infrastructure internally or through external providers, humanitarian organizations must consider the complex implications of these decisions for user privacy, data protection, and long-term sustainability. While self-hosting this infrastructure may offer greater control over data governance, this approach also requires significant long-term investment and resources in order to maintain secure and sustainable operations. Given the pace of change in the AI field, there is no such thing as a "one and done" LLM application.

Conversely, deploying LLM-based applications using externally-hosted infrastructure and models offers to reduce operational burdens while increasing risks related to data sovereignty and privacy. For example, cloud storage providers facing pressure from judicial bodies of countries in which they provide service may feel compelled to hand over sensitive data stored via their services. Moreover,

³³ Birhane et al., "Do We Need a Humanitarian Manifesto for AI?"

³⁴ "Framework for Meaningful Engagement: Human Rights Impact Assessments of AI | ECNL," March 8, 2023, <https://ecnl.org/publications/framework-meaningful-engagement-human-rights-impact-assessments-ai>.

the processing of PII by third-party tools in some jurisdictions may violate local personal privacy laws and expose organizations to legal liabilities.

Decisions should be based on a thorough risk assessment, including a HRIA, while also considering long-term sustainability. Donor organizations also play an important role in these decisions. While many donors are keen to support innovation, it is important for those making funding decisions to recognize that responsible AI product development requires resources that stretch beyond direct product outputs: participative design, data lifecycle management, and building internal staff capacities are all important examples. Funders should consider the sustainability impacts of AI-related funding decisions, including resourcing for humanitarian organizations to build internal capacities.

- **Embed Chatbot Applications Among Accessible, Multimodal Support Infrastructure**

Echoing the findings of other researchers,³⁵ the importance of integrating chatbot and other AI applications within a broader, multimodal communications and support infrastructure, such as phone-based hotlines and face-to-face consultations, is vital to ensuring that AI tools complement, rather than replace, traditional forms of support cannot be overstated.

By treating AI products as complementary components of a larger ecosystem, affected communities are enabled to utilize their preferred form of communication and information acquisition. Moreover, this approach mitigates some of the primary risks of over-reliance on digital solutions, which inherently exclude those with limited or no access to digital communications technology and services.

- **Monitor & Evaluate AI Implementations Regularly**

In the context of releasing an AI tool to production, continuous monitoring and evaluation is essential to ensuring that these tools are effective and responsive. It is important to note that directly evaluating chatbot performance while also respecting the privacy and security of these tools' users can be particularly challenging, especially given the freeform nature of chat-based inputs. Due care must be taken to collect anonymous data to be used for evaluation purposes.

- **Establish Robust Guardrails Against Misuse**

To prevent the misuse of LLM-based applications and the resources that power them, robust guardrails should be implemented and frequently tested. Humanitarian organizations should also dedicate resources towards cross-organizational training of staff regarding the potential risks and ethical considerations of using AI tools.

- **Prioritize Adoption of Internal AI Tooling for Lower-Risk Applications**

While building up organizational expertise on AI, humanitarian organizations may prioritize the adoption of internal-facing tools which pose more limited direct humanitarian risks. This approach can enable humanitarian organizations and their staff to build a stronger understanding of the use cases and limitations of AI-powered tools, while still delivering valuable cost/efficiency savings that contribute to more robust and sustainable humanitarian responses.

³⁵ Kazansky et al., "Chatbots in Humanitarian Contexts: Learning from Practitioner Experiences," 5.

It is important to note that while internal tools can pose fewer humanitarian risks than external/public-facing applications, this is not necessarily the case. When considering the adoption of AI-driven applications, a thorough risk assessment should always be conducted.

- **Advocate for Responsible & Equitable AI Development**

Humanitarian organizations and actors should take an active role in advocating for policies and practices that promote responsible and equitable AI development, with particular focus on the intersection of AI technologies and human rights. At the same time, humanitarian organizations must also advocate for and enable the direct participation of affected communities in discussions regarding the regulation of AI technologies. Such advocacy is crucial in shaping a global framework for the responsible and ethical government of AI.

Humanitarian organizations and individual actors should consider contributing to humanitarian AI “communities of practice” in order to benefit from the advantages of co-learning, share experience and best practices, and encourage the further expansion and democratization of AI knowledge and expertise.

Glossary of Terms

Artificial Intelligence (AI) Application

Artificial intelligence is a broad term with multiple context-specific meanings. Within the context of this report, an artificial intelligence application denotes any technology which applies concepts from machine learning, natural language processing, and other computational fields in order to solve complex problems which require human-like capacities for perception and reasoning.

Natural Language Processing (NLP)

Natural Language Processing refers to a set of computational techniques applied towards enabling machines to recognize, understand and generate human text and/or speech.

Large Language Model (LLM)

A large language model is a type of artificial intelligence model designed to interpret and generate human-like text based on massive datasets of existing text.

Chatbot

Chatbots are software applications used to conduct human-like online conversations via text or text-to-speech, instead of providing direct contact with a live human agent.

Vector Embeddings

Vector embeddings are representations or mappings of text data wherein words, sentences or documents are converted into mathematical vector representations where words that share similar meanings are grouped closer together. Vector embeddings are used in NLP to help understand and analyze the semantic meaning of texts.

Vector Store

A vector store is a database or other repository where vector embeddings are stored for later retrieval. They allow for the systematic retrieval of stored embeddings on the basis of semantic similarity to the provided query.

Embedding Model

An embedding model is an algorithm trained on massive amounts of information (usually, but not exclusively, natural language texts) to classify input information, mapping this data onto a multi-dimensional mathematical representation.

Retrieval Augmented Generation (RAG)

Retrieval augmented generation is a technique in AI where a system retrieves information relevant to a query from a vector store or other information source, subsequently using this information to generate responses. RAG combines the retrieval of specific informational sources with the text summary/generation capabilities of LLMs to improve the relevance and accuracy of model-provided responses.

Semantic Search

Semantic searches are an advanced type of search which seeks to understand the contextual meaning of the search query, rather than relying purely on keyword matching, in order to find relevant matches.