Project Overview:

Addressing the dangers of Immigration

| Heinz Team | [Apoorva Shetty](mailto:ashetty3@andrew.cmu.edu) - Project Manager  [Tasneem Sultana](mailto:tsultana@andrew.cmu.edu) - User Research Specialist  [Abdulaziz Abdulakhadov](mailto:aabdulak@andrew.cmu.edu)- Policy Specialist  [Jhanvi Udani](mailto:judani@andrew.cmu.edu)- Technology Specialist  [Riddhima Singh](mailto:riddhims@andrew.cmu.edu)- Technology Specialist |
| --- | --- |
| Date Last Revised | 11/7/2023 |
| Trello Board |  |

# *Problem Statement*

Human rights violations, conﬂict or poverty have been forcing migrants to ﬂee their country and undertake perilous life-threatening journeys from their departure countries. When the journey is near its end, survival is at many times dependent on successful search and rescue operations conducted by the country that migrants wish to seek asylum in. A successful search and rescue does not always guarantee safety as the country of arrival may not allow migrants to disembark or seek asylum and may even detain the migrants. We believe that a segment of the current migrant population is ending up becoming the victims of either labor or sex trafficking, possibly through job solicitation sites.

The project is in partnership with Eric Schle, who has expressed the need for a tool to identify job solicitation posts linked to the labor market. The identification of such posts would facilitate other broader outcomes such as a geospatial time series to simulate the movement of migrants over time or discovering patterns of trafficking victims at border stops.

# *Project Objectives*

1. Build the below technical artifacts:  
   - Web scrapers to grab data from job solicitation sites  
   - Named Entity Recognition on various text elements
2. Build a report with  
   - User Interview Research

- Immigration Policy Analysis

# *Intended Audience*

Our intended audience is:

* Mr. Eric Schle who is collaborating with a non-profit organization KIND
* The governing body of EU
* Immigration Policy Experts

Other informal, but intended audiences, are the state and federal governments of various countries that can benefit from our analysis and technical artifacts.

Individuals or organizations researching the safety of immigrants and working in the same area could benefit from this project.

# *Key Stakeholders*

* Foreign affairs and immigration departments of governments of the EU
* Eric Schle, researcher and partner
* Non-profit organizations - KIND
* Victims of illegal immigration and trafficking

# *Plan*

1. We are aiming to help immigrants that are wrongfully trafficked through fake jobs available on job solicitation sites:

We have a two-pronged approach:

- Understanding the policy aspect and create a *report*

- Developing web scrapers that can scrape job solicitation websites and use langchain to develop an LLM that can be scaled up for the project

1. To begin our project, we are focusing on user-research : shortlisting the people we talk to, which will mostly be secondary sources since the topic of our project is extremely sensitive.

* Understand the demographic of the victims in this case - more insights be gained from partner
* Document user-interviews and gain insights from them.
* The insights will help us set guidelines for the policy-side of our project - which policies are problematic, which are helping and analyzing them.

1. For the technical aspect : we have received learning objectives from the partner, and as a team we shall spend week 2 developing a learning plan and begin getting acquainted with LangChain.

* Research module for tech development involves finding and shortlisting job solicitation websites that potentially target immigrants and discuss with our partner if we are on the right track

1. After the initial knowledge and skills development through our learning plan, we aim to develop web scrapers and test them out on the job solicitation websites
2. By week 4: we aim to start working on the LLM and understand what is required to build a model

1. The end goal (MVP) will have two parts:

Technical :

~ build a set of web scrapers to grab data from job solicitation sites

~ build open source LLMs to do Named Entity Recognition on various text elements

Policy level :

~ Have a demographic analysis of victims trafficking

~ Analysis of policies aiding/not aiding the safe immigration

Our intention is that our final product will be a tech artifact (LLM) which can be scaled up to recognise fake data from job solicitation websites and prevent trafficking at the root level. Along with that, we want to give the stakeholders a broader understanding of the victim mindset, and assimilate the policy side to connect the dots between immigration and trafficking.

The hopes of this project, if the project goes well, is:

1. The state of PA will have an understanding of the different definitions of blight and their implications on policy and legislation.
2. The state of PA will be better equipped in mapping out their path to addressing blight, which will ultimately trickle down to the municipalities.
3. The state of PA will be able to learn which best practice (or combination of best practices) would prove fruitful if implemented in specific localities in PA.
4. Any funding invested in addressing the issue of blight will be used in an effective manner.

# *Summary of Research*

(To be updated by Sprint 02 / End of Week 3)

# *Communication With Abby*

First email sent 3.23. Team member copied on email. First conference call scheduled for 3.25 11am.

Second conference call scheduled for 4.3 9:30am. Demo Day scheduled for 4.10 in class.