EE/CprE/SE 492 WEEKLY REPORT 5

March 14 - April 3

Group number: 35

Project title: Universal Response Engine: LLMs for Good

Client &/Advisor: Ahmed Nazar and Mohamed Selim

Team Members/Role:

Abrahim Toutoungi - Stakeholder Liaison

Gabriel Carlson - Communications Manager

Halle Northway - Meeting Coordinator

Brianna Norman - Project Deliverables Manager

Ellery Sabado - Timeline Coordinator

Emma Zatkalik - Assignment Manager

Weekly Summary
We worked on putting all of the Iowa data for every one of our chosen topics into our RAG. We also created csv files for all of our data for each of the fifty states. Answer evaluation questions were created and compared to what our LLM produces to generate metrics on how well our LLM answers questions based on the data we fed it.

Past Week accomplishments
-   Finalizing Iowa data in our RAG
-   Evaluation of RAG answers
-   Finalizing our list of topics the users can choose from

Pending Issues
-

Individual Contributions

Name	Individual Contributions	Hours this week	Hours cumulative
Abrahim Toutoungi	- Transferred data to CSVs - Worked on gathering more data for sexual health - Filled gaps in resource list in Addiction, Sexual Health, Mental Health, - Worked on updating our project directory to accommodate data for all 50 states	12	105

Gabriel Carlson	- Tested and collected LLM response time and output parsing data using different chunking sizes - Updated error handling styling and display message on frontend - Updated landing page and logo click handling - Updated font to more accessible Lexend - Added SSL/TLS certificates to nginx and Flask API - Configured firewall rules on VM for port 22, 443, 80, and 5000	14	90
Halle Northway	- Completed data collection - Structured data for suicide and homelessness topics into CSVs - Reworking output parser for model, which involves converting model response to HTML format and modifying chat prompt	7	104
Brianna Norman	- Formatted data into csvs - Finishing data collection - Reworked brand issues in Git	10	92
Ellery Sabado	- WRoked on finished the Chronic Iliness and Finanacial topic dataset and converting to csv - Helping out with Drug addiction dataset collection and csv fiels	11	103
Emma Zatkalik	- Structuring the data for my topics into csv files - Editing the prompts for conciseness - Answer evaluation using langsmith - Reworked branch issues in Git	12	105

Comments and extended discussion (optional)
N/A

Plans for upcoming week
- Work on organizing data and finalizing project
- Work on setting up our research paper

Summary of weekly advisor meeting

TLDR for next week
- Senior Design Presentation Draft
- Draft paper sections
- More evaluations (Talk to Emma)

We went over the structure/format of the paper
- The limit is 6 pages
- Writing will be more technical than other academic papers, Ahmed might reach out asking for improvements or to rephrase things
- https://www.overleaf.com/project/67eca2a637461eb02a8ad858
- Only the abstract and conclusion go in the main file
- Cite using the BibTex export option
    - If it doesn't exist on the page, look for it with google scholar
    - Add to bibliography.bib
    - Dataset
        - Write a sentence description for each important ones, Ahmed will sift through
    - Data preprocessing
        - Talk about how we structured our data
    - Retrieval and Intent-Aware Response Generation
        - Whoever feels most confident
        - Make a diagram similar to the system design first before writing a description

LLM Issues
- Check to see if websites are being scraped every time
- Check to see if a FAISS database exists, if it does don't reparse
- Check to see if the document sources have double links, emails, and phone numbers
-