Sungkyunkwan University (성균관대학교)

Department of Computer Science and Engineering

Enhancing Workplace Health and Safety in Venezuela: Fine-Tuning AI for Legal Compliance and Queries

Name of Student: Ramon Arias
Student ID: 2020315452
Document: Thesis Proposal
Supervisor Name: 추현승

March 21, 2024

Abstract

This thesis proposes the development and implementation of an artificial intelligence (AI) chatbot to improve Workplace Health and Safety (WHS) in Venezuela. The country does a poor job of ensuring employers comply with WHS regulations, thus risking the safety of the workforce. This is mainly because the relevant information is fragmented, hard to access, and at times outdated. The chatbot's goal is to centralize WHS-related laws, policies, and regulations. By doing this, it offers businesses and employees easy access to said documentation. The desired effect of this AI bot implementation is to reduce workplace accidents and legal complications by simplifying compliance and fostering a safety culture. This project aims to leverage AI to address the challenges faced in the WHS field in Venezuela. The thesis will explore the design and implementation of the AI chatbot, which has as a goal to demonstrate the potential of AI in improving WHS practices in challenging environments.

Keywords - Workplace Health and Safety Venezuela, AI Chatbot for WHS Compliance, Venezuelan Safety Regulations, Legal Framework and Technology Integration

1 Introduction

AI integration in WHS is a hopeful step toward addressing numerous issues that have persisted throughout the nation's history. In order to enhance and promote WHS in Venezuela this thesis suggests developing an AI chatbot. The goal of this chatbot is to give employees a consolidated source of WHS legislation, standards, and guidance, making compliance easier for organizations and assisting workers in their efforts to promote safer workplaces.

It is anticipated that the chatbot will prevent legal problems, promote a culture of safety, and lessen accidents. It might also simplify health and safety administration for companies, resulting in cost savings and improved output. This project not only meets Venezuela's immediate WHS needs but also introduces AI as a solution to broader challenges.

1.1 Background and Motivation

The urgent need to enhance workplace health and safety (WHS) in Venezuela, a nation dealing with serious socioeconomic issues like hyperinflation, crime, and a shortage of skilled labor, is the driving force behind this project. WHS laws are out of date and difficult to follow because they are dispersed over several locations, which causes difficulties for both enterprises and employees in Venezuela. In addition to making it difficult for businesses to follow the law, this circumstance jeopardizes the safety of the workforce.

The proposed solution highlights the value of human life and dignity at work while attempting to make it simpler to comprehend and abide by WHS standards. This initiative, therefore, is not merely about regulatory adherence; it's about fostering a culture of safety and respect, safeguarding the physical and psychological health of the Venezuelan workforce against the backdrop of one of the most challenging economic landscapes in the world.

The motivation behind this thesis is rooted in a deep understanding of the transformative potential of AI in addressing complex societal issues. By bridging the information gap in WHS, the project seeks to catalyze a paradigm shift in how safety is perceived and

implemented across Venezuelan workplaces, potentially setting a precedent for similar interventions in other sectors and countries facing analogous difficulties.

1.2 Objectives

1. **Develop an AI Chatbot:** Create a chatbot that uses AI to interpret and deliver comprehensive WHS information effectively to user.
2. **Launch a User-Friendly Website:** Design and deploy a website that serves as the primary interface for users to interact with the chatbot.
3. **Improve Access to WHS Information:** Make WHS laws, guidelines, and safety practices easily accessible to Venezuelan businesses and employees.
4. **Promote Regulatory Compliance:** Assist businesses in adhering to legal standards by providing tools that simplify the process of finding and understanding WHS requirements.
5. **Enhance Workplace Safety Culture:** Encourage a culture of safety within workplaces by empowering employees with the knowledge they need to advocate for safer working conditions.
6. **Demonstrate AI's Potential in WHS:** Showcase the application of AI and web technologies in solving critical WHS challenges in Venezuela.

1.3 Research Questions

The research is structured around the following refined questions:

1. **How can an AI chatbot be designed to effectively consolidate, update, and communicate complex information on WHS laws and policies in Venezuela to a diverse audience?**
   This primary question focuses on the development process of the AI chatbot, including its ability to aggregate information

from scattered sources and present it in a manner that is both accessible and understandable to all.

2. **What are the broader implications of utilizing AI-driven tools, like the proposed chatbot, for improving WHS practices in contexts similar to Venezuela's?**
   Expanding the scope beyond the immediate project, this question aims to understand how the research could inspire similar initiatives in other developing countries facing analogous challenges in WHS information dissemination and compliance.

3. **What potential challenges and limitations could arise in the implementation of an AI chatbot for WHS in Venezuela, and what strategies can be devised to navigate these hurdles?**
   Acknowledging the complexities involved in the chatbot's deployment, this question seeks to preemptively identify technical, regulatory, cultural, and operational challenges specific to the Venezuelan context. It aims to outline possible solutions or mitigations to ensure the chatbot's effective adoption and impact.

2 Literature Review

2.1 Precedence Research & Review of Related Work

2.1.1 Vlex

Vlex is a comprehensive legal research platform that provides an extensive range of legal documents and resources for Venezuela among other countries. It offers legislation, case law, official journals, books, blogs, and news related to Venezuelan law. One of its products, Vincent AI, enhances the research process by analyzing documents, automating drafting, and answering legal questions, aiming for a significant increase in efficiency for legal practices.

Potential Shortcomings of Vlex:

- **Subscription Cost and Accessibility:** Access to its full features often requires a subscription, which is barrier for Venezuela's crippled economy.

- **Complexity and User-Friendliness:** The vast array of tools and resources can be overwhelming for new users, requiring a steep learning curve.
- **Relevance and Up-to-dateness:** Ensuring that documents are up-to-date and relevant to current legal standards can be challenging.
- **Language and Translation Issues:** Automatic language translation might lose nuances and specificities in legal terminology, potentially leading to misunderstandings.

2.1.2 Scribd

Scribd is a digital document library hosting a vast collection of documents, including over 170 million items across various topics and niches. It facilitates information sharing and inspiration within a global community, offering content not easily accessible elsewhere.

Potential Shortcomings of Scribd:

- **Generalist Focus:** The broad focus on a wide range of document types and topics might mean that highly specialized or up-to-date legal documents specific to Venezuela are less prioritized.
- **User-Generated Content:** The quality, accuracy, and relevance of legal documents can vary, necessitating caution and cross-reference by users.
- **Subscription Model:** Requires a subscription, which might not be feasible for all users.
- **Content Management and Moderation:** Effectively managing and moderating content to ensure accuracy and lawfulness can be challenging.

2.1.3 Acceso a la Justicia

Acceso a la Justicia is a Venezuelan non-profit civil association focused on monitoring the administration of justice and the rule of

law. It aims to defend justice, democracy, freedom, and human rights by disseminating relevant information and promoting a rights-guaranteeing culture.

Potential Areas for Improvement:

- **Accessibility and Reach:** Enhancing resource accessibility to a broader audience, including improving website accessibility features and using more platforms for dissemination.
- **Data Verifiability:** Increasing transparency about research methodologies and sources could help bolster credibility.
- **Engagement with International Community:** Further engagement with international legal communities and human rights organizations could amplify their impact.
- **Response to Rapidly Changing Legal Environment:** Keeping pace with changes and providing timely analysis in a dynamic legal and political landscape.

3 Methodology

This section outlines the methodology employed to develop an AI-powered chatbot capable of answering user queries regarding laws and policies, based on a repository of uploaded documents.

3.1 Dataset Preparation and Preprocessing

The initial phase focuses on collecting, preparing, and preprocessing the dataset, which consists of laws and policies sourced from various official channels. These documents, varying in format (e.g., PDF, Word), will be standardized into a machine-readable format for further processing.

Text Cleaning and Normalization Upon conversion of documents into text format, the next crucial step involves cleaning and normalizing the text. This process includes stripping unnecessary elements such as headers, footers, and metadata that do not contribute

to the chatbot's understanding of the content. Further normalization efforts will ensure consistency throughout the text—standardizing abbreviations, dates, and other formats that might vary across documents. These efforts are critical for reducing noise in the training data and improving the AI model's learning efficiency.

3.2 AI Model Development and Training

With a clean and normalized dataset, the focus shifts to developing and training the AI model that will empower the chatbot with its conversational capabilities.

Training Data Preparation The prepared dataset will be annotated to highlight the structure and key elements of the laws and policies. This structured data serves as input for training the GPT model, enabling it to learn the context and nuances of legal documents.

Model Training Leveraging a pre-trained GPT model, the training process involves fine-tuning this model with the prepared dataset. This fine-tuning adjusts the model's weights and biases to better comprehend and generate responses related to laws and policies. The process iteratively improves the model's ability to parse legal language and provide accurate and contextually relevant answers.

Integration with Website The trained GPT model will then be integrated into a chatbot website. This website is responsible for handling user interactions, processing queries through the AI model, and delivering the model's responses back to the users in a conversational manner. The integration will be designed to ensure smooth communication between the user interface and the underlying AI model, providing a seamless user experience.

3.3 Testing, Evaluation, and Refinement

The final phase ensures the chatbot's effectiveness and readiness for deployment through rigorous testing and evaluation.

Testing Methodology The chatbot will undergo extensive testing, including unit tests for individual components to ensure functional integrity, and user testing to gauge performance in real-world scenarios. User testing, in particular, will provide invaluable insights into the chatbot's conversational capabilities and areas for improvement.

Evaluation Metrics Performance will be evaluated using specific metrics such as the accuracy of responses, response time, and user satisfaction. Accuracy is measured by the chatbot's ability to provide correct information based on the user's queries. Response time evaluates the efficiency of the chatbot in processing queries and returning responses, while user satisfaction gauges the overall user experience and the chatbot's effectiveness in simplifying access to information on laws and policies.

4 Expected Outcomes

4.1 Potential Contributions

This thesis aims to make several contributions to the fields of AI, WHS, and socio-legal studies by showcasing the application of AI to enhance Workplace Health and Safety (WHS) in Venezuela. It seeks to illustrate how technology-driven solutions can foster safety culture in workplaces and offer insights into addressing public health issues in complex socio-economic contexts.

4.2 Limitations of the Study

Despite the anticipated contributions, this study acknowledges its limitations, including potential challenges in user technological literacy, the dynamic socio-economic landscape of Venezuela affecting implementation, and difficulties in data collection due to political and economic instability.

5 Work Plan and Timeline

April 2024: Project Initiation and Dataset Preparation

- Weeks 1-2: Project setup, including defining scope, objectives, and detailed project planning.
- Weeks 3-4: Collection and initial preprocessing of laws and policies documents.

May 2024: Dataset Preprocessing and Initial Model Training

- Weeks 1-2: Text cleaning, normalization, and document indexing completion.
- Weeks 3-4: Start of model training with the prepared dataset.

June 2024: Model Development and Integration

- Weeks 1-2: Continued model training and fine-tuning.
- Weeks 3-4: Integration of the model with the chatbot framework.

July 2024: Testing and Initial Feedback Collection

- Weeks 1-2: Conduct unit and initial user testing.
- Weeks 3-4: Feedback analysis and chatbot adjustments.

August 2024: Refinement and Extended User Testing

- Weeks 1-3: Implement refinements and extend user testing.
- Week 4: Final adjustments and project completion preparations.

September 2024: Project Finalization and Documentation

- Weeks 1-2: Finalize project components and complete documentation.
- Weeks 3-4: Project review, final presentations, and thesis submission.

6 Budget Estimate

The budget for the AI chatbot project is estimated as follows:

6.1 Fixed Costs

- Development Tools and Software Licenses: 500,000

6.2 Variable Costs

- GPT Model Usage (API Calls): With an average of 1,000 API calls per month at a rate of 20 per call, for 5 months the total cost is 100,000.
- Cloud Computing Resources: Approximately 200,000/month for 5 months, totaling 1,000,000.
- Miscellaneous Expenses: 300,000

6.3 Total Budget Estimate

- Total: 1,900,000

A contingency fund of 10-15% is recommended to cover unforeseen expenses, which would adjust the total budget to approximately 2,090,000 - 2,185,000.

References

[1] "(25) Luz Hernandez Gamboa | LinkedIn." Accessed: Mar. 21, 2024. [Online]. Available: https://www.linkedin.com/in/luz-hernandez-gamboa284225a2/

[2] "Ley Organica Sistema Seguridad Social Venezuela | PDF | Seguridad Social | Bancos," Scribd. Accessed: Mar. 21, 2024. [Online]. Available: https://es.scribd.com/doc/58584289/Ley-Organica-Sistema-Seguridad-Social-Venezuela

[3] “Nosotros | Acceso a la Justicia,” Acceso a la Justicia. Accessed: Mar. 21, 2024. [Online]. Available: https://accesoalajusticia.org/quienes-somos/

[4] Y. Caraballo-Arias, “Occupational Safety and Health in Venezuela,” Ann Glob Health, vol. 81, no. 4, pp. 512–521, 2015, doi: 10.1016/j.aogh.2015.08.022.

[5] “vLex Venezuela,” vLex. Accessed: Mar. 21, 2024. [Online]. Available: https://vlexvenezuela.com/

[6] Y. Caraballo-Arias, J. Madrid, and M. Barrios, “Working in Venezuela: How the Crisis has Affected the Labor Conditions,” Ann Glob Health, vol. 84, no. 3, pp. 512–522, Sep. 2018, doi: 10.29024/aogh.2325.