👨‍⚖️ Ethical Considerations for Sprint 2

1. Introduction

In the development of the Furhat Q&A platform, our team has endeavored to align with the highest ethical standards. This document details the specific ethical considerations addressed throughout the various phases of our project. The focus here is on practical, project-specific issues such as data management, user privacy, bias mitigation, and the implications of AI integration with robotic technology.

2. Ethical Framework and Data Management

Data Collection Ethics

The initial phase of our project involves web scraping to gather data from websites such as Melbourne Connect or CIS. It is imperative that this process respects the privacy and intellectual property rights of the data sources. We commit to only scraping data that is publicly available and intended for public use, avoiding any content that requires login credentials or is marked private.

Action Taken: We use advanced filtering techniques to ensure that any personal information inadvertently collected is immediately discarded. Our scraping tools are configured to recognize and skip over any data fields that typically contain personal identifiers.

Data Storage and Usage

Once collected, the data is stored securely with encryption and is accessed only by authorized personnel in our team. The use of this data is strictly confined to the project's scope—powering the language model for the Furhat robot.

Action Taken: We have implemented robust security measures including encrypted databases and strict access controls. Regular audits are conducted to ensure compliance with these security practices.

3. User Privacy and Data Security

User Interaction with Furhat Robot

As the Furhat robot is designed to interact directly with users, special attention has been paid to the privacy and security of user data. The robot is programmed to not store any personal information from its interactions with users.

Action Taken: Each conversation with the Furhat robot is processed in real-time without retaining any personal data post-interaction. Moreover, all interactions are anonymized even during processing to further protect user privacy.

4. Addressing Bias and Ensuring Fairness

Data and Model Bias

Given the diverse user base that the Furhat robot is expected to serve, it is critical to address and minimize any bias in the data and the language model. Bias can distort the robot's responses, leading to unfair or inappropriate interactions.

Action Taken: We conducted a comprehensive audit of the dataset used to train our language model, removing or correcting biased or unrepresentative data. Additionally, ongoing monitoring is set up to identify and correct any emerging biases as the system evolves.

5. Transparency and Accountability in Development

Open Development Practices

Maintaining transparency throughout the development process is essential to build trust with all stakeholders. Detailed documentation of our methodologies and decision-making processes is made available to stakeholders.

Action Taken: All development processes, from data collection methods to model training and robot integration, are documented thoroughly. These documents are reviewed regularly to ensure they reflect current practices and are available for audits.

6. Impact on Stakeholders and Society

Enhancing User Experience

The integration of AI with the Furhat robot aims to create a seamless and efficient user experience. However, the deployment of such technology must consider potential impacts, such as the displacement of traditional roles and the implications of user dependency on robotic technology.

Action Taken: We engaged with our client to understand and mitigate any negative impacts. Adjustments were made to ensure the robot complements rather than replaces human roles, providing assistance that enhances rather than detracts from the human experience.

7. Conclusion

In conclusion, our project's approach to ethical considerations is deeply ingrained in every phase, from data collection to deployment. By maintaining a steadfast focus on privacy, fairness, transparency, and societal impact, we strive to ensure that the Furhat robot not only serves its intended purpose but does so in a manner that is respectful and ethical. This commitment to ethical excellence is fundamental to our project's success and is a cornerstone of our ongoing development efforts. Through continuous ethical evaluation and adaptation, we aim to set a standard for responsible AI development in interactive and robotic technologies.