An Interactive Voicebot Using Rasa Framework For Migrant Workers

Jyoti Arora[1], Sonalika Roy[2], Vandita Sharan[3], Yogita[4], Yashtika Kakkar[5]

{joy.arora@gmail.com[1], sonalikaroy13@gmail.com[2], vanditasharan14@gmail.com [3]}

Assistant Professor, Dept. of Information Technology, MSIT, Affiliated to GGSIPU, Delhi, India[1], Graduate, Dept. of Information Technology, MSIT, Affiliated to GGSIPU, Delhi, India[2,3]

Abstract. During the pandemic, one thing that came to light and finally received long-overdue attention was the plight of migrant workers in India. This section of society lacks majorly in terms of resources. To help them, this paper focuses upon one such resource called the MiWo Voicebot. This unique voice bot is an AI Voice Bot that focuses on migrant workers and the problems they face. It focuses on areas such as their constitutional rights, labor laws, organizations, and NGOs that can help these workers. The voice bot can be extended to be multilingual which can cater to the needs of a very diverse section of the workers. MiWo aims to increase the quality of life these workers have and give them an option to know how to access their rights or get help when required from the right place.

Keywords: Voice Bots; Chatbot; Human Welfare; Artificial Intelligence, Migrant Workers

1 Introduction

Firstly, let's understand who these migrant workers are. This is the section of society that works seasonally or moves around to different areas looking for employment, which is often temporary. These workers rarely get any benefits and are often financially exploited.[1]

It's important to note that although the actual conditions and minimal resources accessible to migrant workers came to light during the pandemic because suddenly the conditions were so dire that survival became difficult, these workers have been living these underprivileged oppressed lives for a very long duration. With growing technology and awareness, it becomes essential that enough light is shed on their lives and proper mechanisms are established to overcome the difficulties they have been facing for the longest time.[1] MiWo, Migrant Worker's Voice Bot is just one small step towards accommodating their issues in technological advancement. There are a lot of aspects of laws, activism, and opportunities that these workers are not even aware of. The government and various other organizations have taken steps to safeguard their interests, but these workers do not know of it. Through this

ICIDSSD 2022, March 24-25, New Delhi, India

DOI 10.4108/eai.24-3-2022.2319009

voice bot, there is a possibility that these acquired resources can become more accessible if all the information is in one place, stored systematically, and is easily available.

MiWo is designed in a way that caters to these migrant workers directly. The project is based on extensive investigation and actual input from these workers through surveys. Instead of assuming the hardships of these workers, a team went on-ground to gather information about the lives of these migrant workers. Following this, all their queries were studied properly to find optimal solutions to problems they face and who they can contact for resolution. This way these workers can have one place to resolve all their issues. In this paper, the following topics will be covered –

1. Understanding the plight of migrant workers.
2. A study based on voice bot – history, benefits, and drawbacks.
3. MiWo VoiceBot
   3.1 Architecture
   - Speech to Text Conversion
   - Rasa Framework
   - Text to Speech Conversion

   3.2 Benefits

   3.3 Drawbacks

2 The Plight of Migrant Workers

Even though a majority of the migrant workers in the society are regarded as unskilled, they do several important tasks and thus play an important role in the Indian Economy. On 30th January 2020, India officially recorded its first Covid 19 positive case, and ever since people decided to take precautions. In the end, it did not pan out as planned and a lockdown had to be imposed in March of 2020. This resulted in all construction work stopping and all the factories were immediately shut down, directly impacting the lives of millions of migrant workers who still live on daily wages. March 2020 was the first time when India finally understood the reality of these workers' lives when they had nowhere to go, no transportation, and no livelihood. It's not like these migrant workers don't deal with atrocities in their normal lives but it's just that this time their heart-wrenching reality was finally being highlighted under these dire circumstances. People were forced to walk miles to just get back home. Children, the elderly, disabled people, everyone suffered and eventually the people of India witnessed how bound these workers must feel regularly, just to survive. Any luxury is a dream and they have been deprived of the essentials like sanitation and minimum wage for ages with no agency and no help whatsoever. This is the ground reality of these workers and these are the conditions they have been living in thus far. It is high time people start acknowledging the situation and find a way to help them. This has often been overlooked and constructive action is required in this sector urgently.[2]

3 Voice Bot Technology

A voice bot is a voice channel based on natural language performance and understanding (NLU) communication that works by converting audio into text format. AI Technology helps to identify key features in speech and provides the best solution for converting. The Text To Speech Engine (TTS) then converts the response into sound or voice to complete the communication. These robots are trained to complete the whole process of understanding speech and responding to the nearest way that of a man. Voice Assistants are a smart way to communicate and are often called conversational user interfaces because of their qualities and functioning.[3]

As already explained, the key elements of voice bot technology include speech recognition, natural language processing, and speech synthesis. The first thing that requires attention while making a voice bot is to define its functionality, i.e what are the functions that the bot needs to perform successfully to reach the required standard. Planning and creating these rubrics is the first step. Post this, analyzing the basic processes that need to be performed in a specific sequence needs to be figured out. After this, the focus shifts upon implementation. The initial functioning of the voice bot has to be checked before moving forward with it. Once this stage is completed successfully, one can finally start training the bot using various algorithms. The algorithm choice will be dependent on what is the ground functionality of the product in the longer run. It should be verified by multiple people and all errors should be analyzed and subsequently fixed. Lastly, after taking all these things into account, the final step is to deploy the bot.[3][4]

Many major companies provide voice bot services while using various technologies. One such important technology domain is machine learning, which is used by the majority of voice bots. It is an important factor in voice bot development. Through machine learning, various bots can learn how to understand various accents and vocal textures. In the case of machine learning, the most advantageous factor is that the voice bot continues to learn and the data set keeps growing, further amplifying the training of the said machine learning model. Over time, the voice bot can recognize a wider selection of accents and words.

3.1 History of Voice Bots

It was in the year 1961 when IBM launched its first computer which had digital speech recognition. This was called Shoebox. This was relatively very basic compared to the technology we have now and only recognized sixteen words and numbers from 0 to 9. In the 1970s, Carnegie Mellon University, Pittsburg, Pennsylvania developed proper speech recognition which was included in those computers. IBM finally launched its first smartphone named Simon, which in a way created the basis for intelligent conversational technology. Simon was the first commercially available device with speech recognition technology. It could convert speech to text efficiently. Voicebot technology started growing in the 2000s when giants like Google, Apple, and Microsoft started launching technologies like Google

voice search and Siri. Post this, voice technology grew exponentially and has been a part of all modern devices.[5]

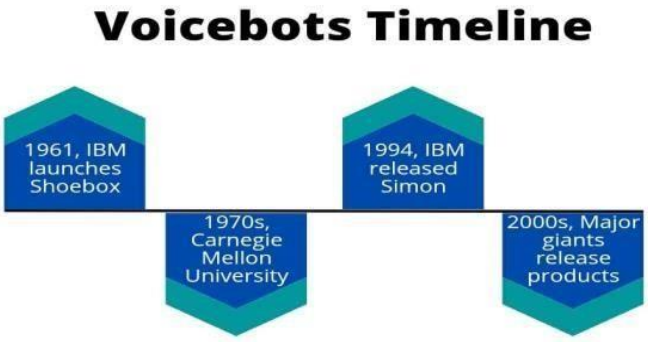

Fig. 1. A brief history of voice bot

3.2 Benefits

1. It's extremely useful for reaching a larger audience. One-on-one interaction is not always plausible, through voice bots, one can reach out to a larger audience. Through this, we can successfully utilize time more efficiently and subsequently increase productivity as it decreases workload. It can easily be used to deliver opportunities and information to large audiences in less time. In the customer service industry, they reduce wait time significantly, thus improving their customer service.[6]
2. These voice bots are accessible at any time 24/7, thus making them more reliable. It can get people the information they require, even on off-hours without increasing any workload. Therefore, they increase the consumer experience by a lot because they don't have to wait.
3. In today's world, a lot of people prefer voice as a mode of communication. It's an extremely convenient and quick way to interact and get the required service.[6]
4. Another important benefit of using voice bots is that they allow a wide range of software that can be integrated with it, like CRM. It in turn makes these voice bots extremely functional and convenient to use.
5. In general, these voice bots save a lot of money as well as time, since they decrease infrastructural costs significantly. All of this can be managed through a remote server through a small team of people. All the repetitive tasks can be taken over by these voice bots.[6][7]

3.3 Drawbacks

1. Initial implementation of a voice bot includes excessive training of the model, which can be expensive and time-consuming.

2. Any model like the means of a voice bot requires a lot of maintenance. Over time there are a lot of bugs that can accumulate and need to be dealt with. Reviewing the product and updates are also equally important.[4]
3. Voice bots lack an emotional quotient. Answers provided by the voice bots are naturally void of emotions and do not take into account the temperament of the user, which can be extremely frustrating in certain situations.
4. A lot of times, voice bots fail to provide the required answer and, in such cases, end up repeating answers in a loop. Repetition is extremely common in voice bots. This again can prove to be annoying and time-consuming for the consumer.
5. In the case of voice bots, data storage, privacy, and security become extremely important for the successful working of any voice bot. These costs can add up to be a lot and thus make it expensive to maintain a functional voice bot.[6]

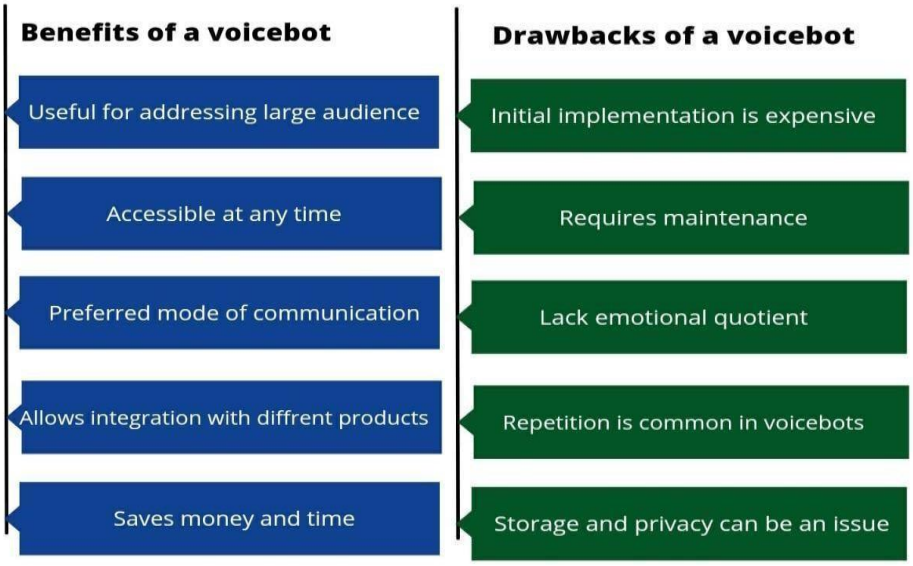

Fig. 2. Benefits of voice bot.

Fig. 3. Drawbacks of voice bot

4 MiWo VoiceBot

MiWo or Migrant Worker's Voicebot is a technology that focuses on helping a marginalized community of migrant workers; it focuses on areas such as living situations of migrant workers, and injustice that takes place in their workplace, human rights violations, and termination laws to just name few. The project is based on an on-ground survey where various migrant workers were interviewed. These workers came forward to tell the heart-wrenching

realities of their lives and explain where they all require help. Surprisingly, the majority of them were illiterate but knew how to use cell phones because of icons and voice technology. Taking this into consideration and the ongoing digital revolution, voice bots seem like a viable option through which these workers can receive help. The bot is divided into sections and tracks various regions in which help can be provided. It keeps them informed about the labor laws of the country, the various labor unions they can join, and connects them to various NGOs, and human rights activists for direct support. All the helpline numbers are also provided to get all required information to these workers. MiWo is also being expanded to be multilingual, to reach a wider number of people in different regions of the country. It can later be associated with government projects like Saksham, to include job search features.

4.1 MiWo Architecture

It Miwo voice bot is based on RASA technology. Let us start by understanding what exactly RASA is. RASA is a recently developed framework that is used for industrial voice and chatbot development. Rasa uses Python as a base and is being widely used for voice bot creation. No prior machine learning knowledge is required for Rasa. It's easily available, open-source, and flexible. In Miwo, there are three major components one needs to grasp to understand the functioning of the bot.

Speech to Text Conversion: The first step in any program is to take input from the user, which in this case will be in the form of speech since it is a voice bot. For the bot to understand what is being said, it is important to convert this speech into text input for further clarification; this is how the bot will make sense of the intent and subsequently extract required entities. With the use of a microphone, the physical sound signal is converted to an electrical one and subsequently, this data becomes digital with the help of an analog-to-digital converter. Once this conversion to digital form is completed, several models can be used to decipher the signal and transcribe the audio to text. In the case of MiWo, PyPI's library for speech recognition is being used for transcribing speech to text. This speech recognition library is used as an envelope that can include a huge number of trending APIs, which makes this framework extremely flexible. An example of such an API is- the Google Web Speech API, which supports a specific default key that's hardcoded in the PyPI library.[8]

Rasa Framework: Two major components of the RASA framework are Natural Language Understanding (NLU) and Dialog Management. This is the base of Rasa through which MiWo is developed.

NLU primarily focuses upon how intents are being classified, which can be explained as grasping what the input intends to say and interpreting the meaning by the bot's functioning. Secondly, it also extracts important information from the user since it extracts entity. Lastly, it also retrieves the response from the bot to complete the process. A trained pipeline is used to process utterances using the NLU model. Based on the context of the conversation, the dialog policies decide what the response should be at the end of it.[9]

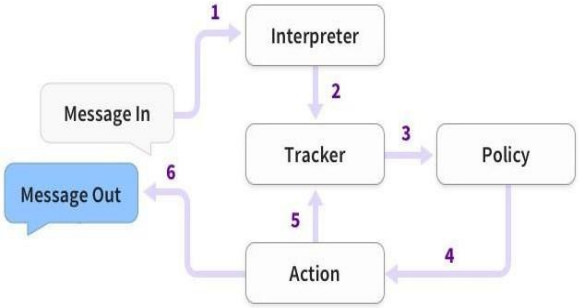

Fig. 4. MiWo Architecture

The following is the flow of the bot:

- The message which is received is forwarded to the available interpreter that in turn converts this information to a dictionary inclusive of the original text sent, its intent, and the extracted entities found. This whole portion of the work is overseen by the NLU.
- This signal is then transferred to a component called a tracker which concentrates upon tracking the state of the conversation.
- Policy receives the current state of interaction through the tracker.
- Every policy decided upon the action it would be taking next.
- These actions previously decided upon by the policies are tracked by the tracker.
- Ultimately, a response is sent to the consumer.

Text to Speech Conversion: The output generated by the voice bot is essentially a text which needs to be converted to speech for the final output, thus full-filling the function of a voice bot. To get this done, MiWo uses the pyttsx library of python. This library is not dependent on the platform being used and can be used independently. One major factor for pyttsx being used is that this library can be used offline as well. Rather than storing the text output as an audio file, this library speaks out the audio output there in reality.[9]

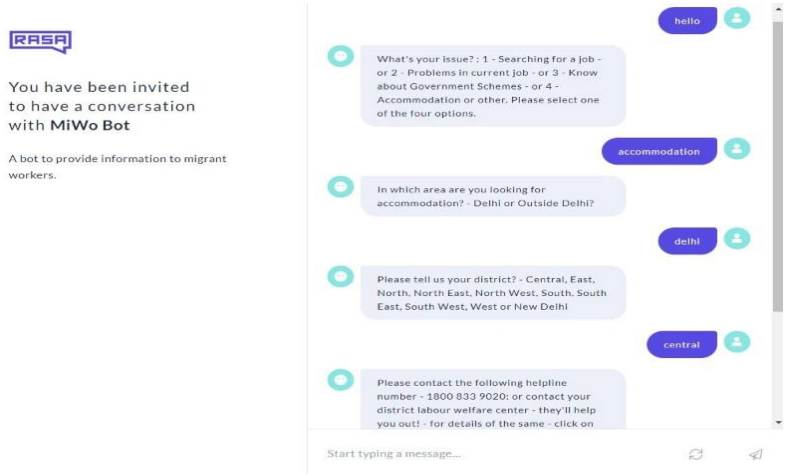

Fig. 5. MiWo Interface

4.2 Benefits

1. It The voice model allows even illiterate workers to get valuable information through this medium. Usually, the vast majority of these workers don't have access to formal education and they rely on word of mouth a lot. This allows them to get useful information through their means.
2. Finally, through this project, this section of society gets the undue attention required. As previously discussed, these people have been marginalized to a very great extent and their living conditions are worrisome. This will help organizations reach out to them easily.
3. The Rasa framework used in MiWo leaves room for expansion since it is extremely flexible. More features can easily be added based on the needs of migrant workers.

4.3 Drawbacks

1. Still, a vast number of migrant workers are not very familiar with these technologies and it will take time for them to get accustomed to it. Real change will take time and groundwork is still faster.
2. The data that is being provided to these migrant workers needs to be verified properly so they don't lose faith in the initiative. Therefore, constantly cross-checking and updating the information is going to be a real task.
3. At times there are going to be failures due to questions being out of bound, voice not being deciphered properly, etc. under such circumstances, there needs to be an admin constantly av ailable to resolve these issues.

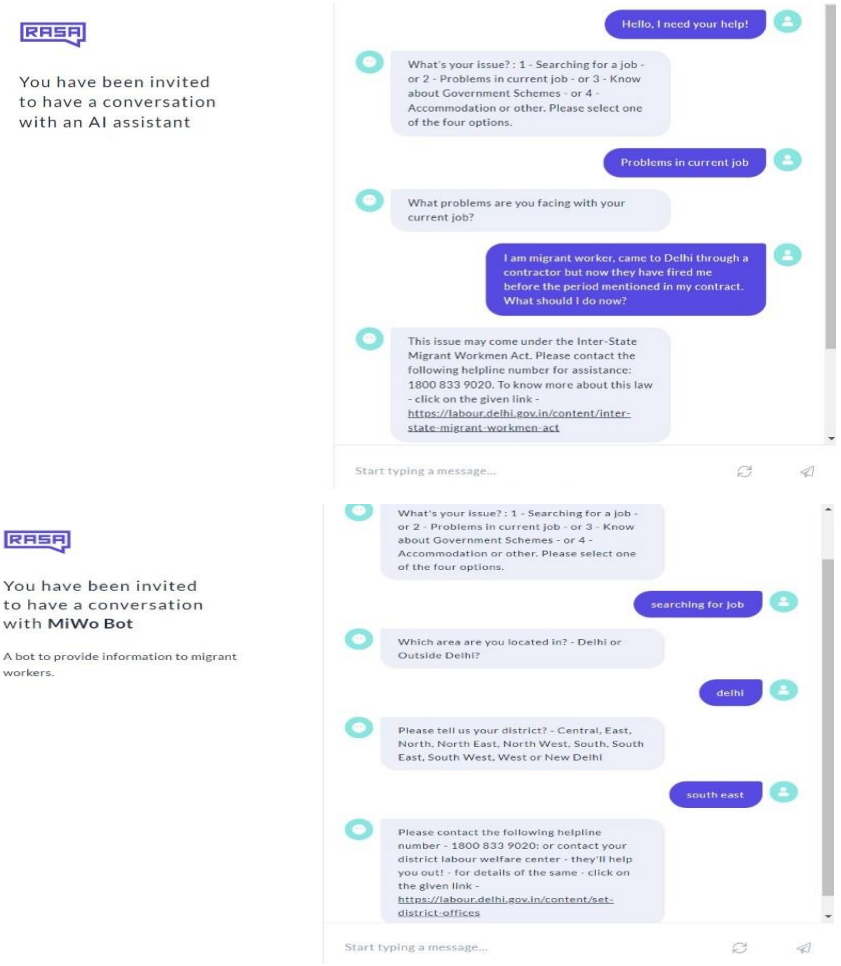

Fig. 6. & 7. Bot Answering questions regarding jobs

5 Conclusion and Future Scope

At the end of it, one gets to understand the gravity of the situation these migrant workers are in and how they have not been getting the required amount of help. Over time the seriousness of their troubles increased and they did not know where to get help from. Under such circumstances, voice bot technology is a boon since it is easily accessible to workers. MiWo is one such initiative in the right direction. It caters to the needs of these workers and constantly adapts to their needs for better performance and outreach. It uses popular frameworks like RASA, which makes coding and maintenance easier. It covers a variety of sections associated with the lives of these workers which will help in the longer run. In the longer run, this bot can be converted into being multilingual, and since it has been made in RASA, this allows widespread use. The bot currently concentrates upon a small geographical area, that can also be expanded eventually with more research.

References

[1] A.Ghosh, A.B.R.Chaudhary, “Need for a holistic perspective towards India’s migrant workers”. URL: https://www.orfonline.org/expert-speak/need-for-a-holistic-perspective-towards-indias-migrant-workers/

[2] J.Jesline, J.Romate, E.Rajkumar, A.J.George, “The plight of migrants during COVID-19 and the impact of circular migration in India: a systematic review”. URL: https://www.nature.com/articles/s41599-021-00915-6

[3] D.Amos, “The Ultimate Guide To Speech Recognition With Python”. URL: https://realpython.com/python-speech-recognition/

[4] N.Jatu, “Python Text to Speech, pyttsx Module”. URL: https://www.geeksforgeeks.org/pythontext- to-speech-pyttsx-module/

[5] A.Tyagi, “Voice AI: The Ultimate Guide”. URL: https://verloop.io/blog/voice-ai-chatbot-guide/

[6] J.Couedel, “What is a voice bot? Explanations, mode of operation and benefits”. URL: https://calldesk.ai/blog/voicebot-definition

[7] “What Are Voicebots And How They Can Help You”. URL: https://voximplant.com/blog/whatare- voice-bots

[8] B.Saji, “An introduction to chatbot development using RASA” URL: https://www.analyticsvidhya.com/blog/2021/11/an-introduction-to-chatbot-development-using-rasa/

[9] J. Petraitytė, "How to build a voice assistant with open source Rasa and Mozilla tools” URL: https://rasa.com/blog/how-to-build-a-voice-assistant-with-open-source-rasa-and-mozilla-tools/