Chatbot for Elections FAQ - State of South Carolina

BHARATH CHANDRA MUPPASANI, University of South Carolina, USA

Bots, often known as chatbots, are multimodal collaborative assistants that may aid users in carrying out necessary tasks. Chatbots are software applications that mimic human conversation. They can be used to deliver information, respond to queries, make transactions, and have conversations with people. Whenever chatbots are mentioned in relation to elections, it is typically in the context of using bots to propagate false information. But chatbots can also assist individuals in carrying out worthwhile tasks, including registering to vote, locating a polling location, or monitoring the status of their voter registration. With the use of the official FAQ data, we have created a chatbot for the state of South Carolina elections as part of this research project. We built the chatbot using a safe chatbot design architecture to get rid of misleading information from it. Also, the chatbot has been tested and evaluated by a focus group making the chatbot reliable and trustworthy.

CCS Concepts: • **Chatbot**; • **Intent Generation**; • **Paraphrasing**; • **Elections**;

Additional Key Words and Phrases: RASA, FAQ Dataset

ACM Reference Format:
Bharath Chandra Muppasani. 2022. Chatbot for Elections FAQ - State of South Carolina. 1, 1 (November 2022), 6 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

The customer service of the future is already here. While chatbots are frequently used in customer service applications, they can also be used for other things, like giving information about a good or service. Technology is developing quickly, and bots are becoming more clever every day. They are now intelligent enough to answer basic inquiries, and with the right level of escalation, they can even tackle more complicated customer issues. Every industry can benefit from the implementation of chatbots, which will enhance customer service. They are among the best technological tools that firms may utilize to promote customer happiness and operational effectiveness.

2 PROBLEM

Researchers have tried to estimate the effort needed to vote in different states in the United States using COVI, a *cost of voting index* [5]. According to it, Oregon is the easiest state to vote in, while Texas is the most difficult (50th). As the seventh most complicated state for voters, South Carolina is a good location to look into how residents share voting-related information.

In South Carolina, nearly 20% of the population in the state is age 65 and older (Census.gov, 2021). Elections are held in South Carolina to fill a number of municipal, state, and federal positions. These positions include Governor, Lieutenant Governor, Secretary of State, Attorney General, Treasurer, Comptroller General, Superintendent of Education, and Commissioner of Agriculture, as well as State Senate, State House of Representatives, state delegation to the U.S. Senate, and state delegation to the U.S. House of Representatives (Ballotpedia, 2022). The state's election commission oversees

Author's address: Bharath Chandra Muppasani, bharath@email.sc.edu, University of South Carolina, 810 Pulaski St, Columbia, South Carolina, USA, 29201.

XXXX-XXXX/2022/11-ART $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

voter registration, candidate certification, and elections (South Carolina Elections Commission, 2022).

For building the dataset to train the election chatbot, we have made use of the frequently answered questions (FAQs) available on the government website of South Carolina [6]. South Carolina election FAQs [6], consists of 30 question-answer (Q-A) pairs across various topics like early voting, absentee vote, operating hours of the polling places, required documents for voting, election results and questions related to legal actions of the candidate while campaigning. The intuition behind using an official government website to procure QA pairs is to inherently keep the user interaction controlled and secure. A sample data snapshot of internal representation of this FAQ data is presented in Figure 1.

Dataset for South Carolina		
Question	**Answer**	**Intent**
Is the voting system connected to the internet?	No, the voting system is never connected to the internet. Computers used to tabulate votes, BMDs, and ballot scanners used in South Carolina are not even capable of being connected to the internet.	voting_system_connected_internet
Are ballot selfies legal? Can I take a picture of my ballot and share it with others?	No. State law prohibits anyone from showing their ballot to another person (S.C. Code of Laws Section 7-25-100). The use of cameras is not allowed inside the polling place.	ballot_selfies_legal_take

Fig. 1. A snapshot of internal representation for South Carolina data

3 RELATED WORK

Chatbots have been employed for a wide range of functions, including the dissemination of public health information [1], aiding in the treatment of mental health issues [2], and general conversation [3]. [4] has investigated the use of conversational bots in elections for political branding. The emergence of chatbots in elections and their consequences on shifting political intentions are studied by the authors in [4]. One of the studies that comes the closest to the strategy suggested in this study is [3], where the authors created a chatbot to give voting recommendations. Due to the black box nature of the approach, the main drawback of this strategy is that the information presented is frequently inconsistent, inaccurate, and divisive for different users. Therefore, in our approach, we make sure to use a reliable source of information, such as a government website, and present the information gathered, while avoiding inquiries that diverge from the list of official FAQs that is readily available.

4 APPROACH

In this section, we present our approach used to build the election chatbot for South Carolina.

4.1 Dataset Creation

To obtain the Q-A dataset, that we can use to train our chatbot, we have used a web-scraping tool in order to automatically extract the FAQs and save them in the form of comma-seperated values (.csv) files. Figure 1 show a sample snapshot of Q-A pair dataset built for South Carolina. For South

Carolina, we have obtained 30 Q-A pairs. Once the dataset is obtained, we need to tag the questions with their intent in order to capture the core concept the question is answering. This helps the chatbot to better train and match new similar intent questions to the appropriate answer. Due to the presence of several stop words, we have used the n-gram technique to extract the intents for the questions. The intents tagged to the corresponding question can be seen in Figure 1. In the further steps, it is crucial to train the chatbot with similar questions that match an answer, rather than having only a single question associated with an answer. However, in any FAQ wesbite, we only obtain a single question tagged with an answer. Thus, in order to create an efficient chatbot, we made use of a paraphrasing tool to generate similar paraphrased questions as to the one extracted from the election FAQ website.

5 SAFE AND USABLE CHATBOT ARCHITECTURE

We now present a general architecture for building chatbots which try to be safe and useful by providing grounded answers, logging their interaction with users, supporting common conversations and being accessible. This architecture will be specialized to build specific chatbots for supporting FAQs about elections in SC and MS in the next state.

Features that make it safe are:

- Logging of conversations
- Do not answer strategies
- Paraphrasing for better question matching
- Alexa Integration

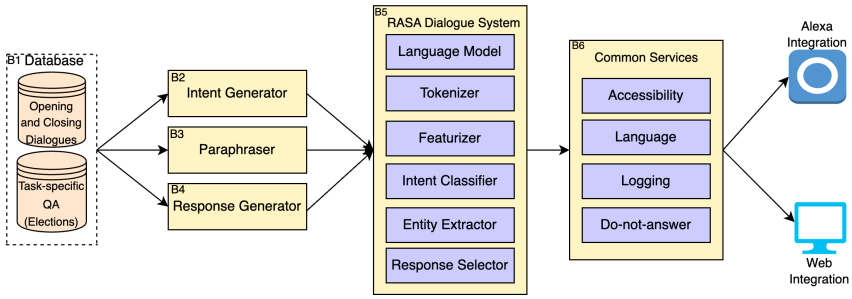

Fig. 2. System Architecture

Figure 2 shows our proposed system architecture. Key components of our system are:

5.0.1 Database (B1). The database is the source from which we extract the training data to train the chatbot. We ensure that the source is reliable and trustworthy. Task-specific QA refers to the data source pertaining to the chosen domain (which is elections, in this case). The opening and closing dialogues are usually generic (like greeting and saying bye).

5.0.2 Intent Generator (B2). Questions in the dataset for South Carolina are small, with an average of 10 words in a question. On further exploration, we identified that most of these words are stop-words, i.e., commonly used words in any language. For example, in English, "the", "is", and "why" qualify as stop-words. In order to obtain important information, it is important to remove stop-words, and when we carried out the similar process on the election FAQ dataset, we identified that there are only a few words left out for every Q-A pair. Thus, in order to obtain the intent for each question, we have used n-gram approach to tag each question with an intent consisting of the important words present in the initially extracted questions.

5.0.3 Paraphraser (B3). In order to train a chatbot, it is crucial to show it similar questions that match an answer, rather than having only a single question associated with an answer. However, in any FAQ wesbite, we only obtain a single question tagged with an answer. Thus, in order to create an efficient chatbot, we made use of a paraphrasing tool to generate similar questions as to the one extracted from the election FAQ website.

5.0.4 Response Generator (B4).

5.0.5 RASA Dialogue System (B5). We use the RASA chatbot framework [**?**] to build the chatbot. The dialogue system has an NLU pipeline with different components for understanding human conversation and responding appropriately. Language models like the Spacy language model can be used if one wants to use pre-trained word vectors. Tokenizer converts sentences to tokens. Featurizer creates a vector representation of the user message and response. The intent classifier classifies the intent of the user message. The entity extractor extracts entities that are specified in the training data. The response selector chooses the appropriate response based on the identified intent and entities.

5.0.6 Common Services (B6). The common services are optional and the user has the flexibility of choosing the services they need. Some of the accessibility options are font settings and Text-to-Speech. The users will be able to converse with the chatbot in the language that is comfortable to them using the language settings. This can be implemented by making use of translators. The conversations can be logged for storage and retrieval using the logging option. This also helps the developers to improve the chatbot conversation by reviewing the stored conversations. We do not want the bot to respond to certain questions (for example, question like 'Do you think the current president is doing a good job?'). These questions are mostly subjective. 'Do-not-answer' option can be used in this case.

5.1 System Integration

Web integration and Alexa integration provide an engaging user interface to converse with the chatbot. We integrated the RASA chatbot with Alexa device as a skill using the Alexa developer console (the skill is still in beta phase). Web integration was done using the RASA webchat package. The system can be accessed from here[1]

Our architecture has been built, generalizing our experience building chatbots across different domains where the usability of the system is essential (for example, education [**?**], networking [**?**], etc.) with little or no modifications. This shows the generalizability of our system. At the same time, common services represent safety-specific services we have implemented to ensure the chatbot's behavior can be audited, controlled, and made more accessible to users. Logging makes the conversations more accessible and helps developers review the conversations to ensure that the interaction flows safely. The 'Do-not-answer' option disables the chatbot from answering inappropriate questions. The blocks B2, B3 and B4 also work in a controlled manner and we only use reliable sources (like Government websites) for B1. Based on our past experience, all these features benefit many other applications where user safety is essential.

5.2 Chatbot Testing

For a user to trust a chatbot, it should be competent and reliable. The performance of chatbots can be evaluated using many methods such as user study, A/B testing and Randomized Controlled Trials (RCTs). Chatbot testing allows us to improve the chatbot design and functionality.

[1]http://casy.cse.sc.edu/Election-Chatbot-SC-main/

5.2.1 Randomized Controlled Trials (RCTs). RCTs are a form of a scientific experiment that is used to test the effectiveness of a new treatment. In this case, it would be a new chatbot that is being tested. It is a study in which the participants are randomly assigned to either a treatment or control group. Participants in the control group receive an alternative treatment (placebo). The two groups are compared to test the effectiveness of the new chatbot (treatment). Though RCTs provide a way to test the chatbots with minimal or no bias and confounding effects, they require time, money, and many participants.

6 EVALUATION

The election chatbot for South Carolina has been hosted on a public server. Figure 3 shows a snapshot of the website. Figure 4 shows a snapshot of the chatbot answering user queries.

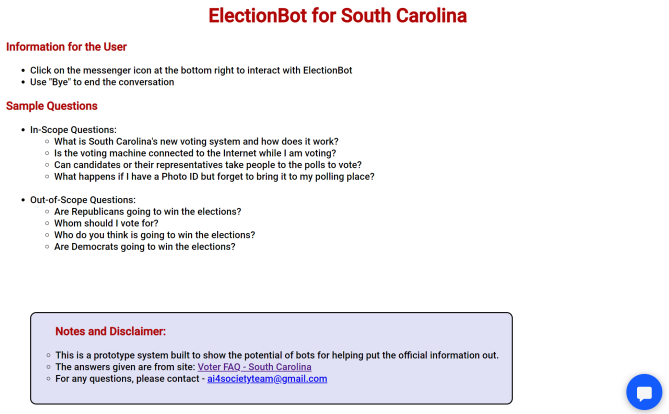

Fig. 3. A snapshot of implemented chatbot hosted on a website

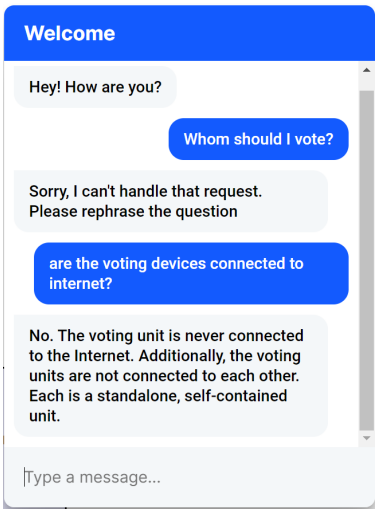

Fig. 4. A snapshot of chatbot answering questions

7 DISCUSSION

About results, What more could be done, anything else interesting

REFERENCES

[1] Parham Amiri and Elena Karahanna. 2022. Chatbot use cases in the Covid-19 public health response. *Journal of the American Medical Informatics Association* 29, 5 (2022), 1000–1010.

[2] Gillian Cameron, David Cameron, Gavin Megaw, Raymond Bond, Maurice Mulvenna, Siobhan O'Neill, Cherie Armour, and Michael McTear. 2018. Assessing the usability of a chatbot for mental health care. In *International Conference on Internet Science*. Springer, 121–132.

[3] Menal Dahiya. 2017. A tool of conversation: Chatbot. *International Journal of Computer Sciences and Engineering* 5, 5 (2017), 158–161.

[4] Marta Mokhnach. 2022. Marketing Approaches in Political Branding during Election Campaigns. (2022).

[5] Scot Schraufnagel, Michael J. Pomante II, and Quan Li. 2020. Cost of Voting in the American States: 2020. *Election Law Journal: Rules, Politics, and Policy* 19, 4 (2020), 503–509. https://doi.org/10.1089/elj.2020.0666 arXiv:https://doi.org/10.1089/elj.2020.0666

[6] South-Carolina. 2022. Frequently Asked Questions. *https://scvotes.gov/voters/voter-faq* (2022).

Received 10 November 2022; revised 10 November 2022; accepted 10 November 2022