ADVICE: Literature Review

Chatbot System for a Virtual Student Advisor - Literature Review

Avhusaho Ramalala†
Computer Science Department
University of Cape Town
South Africa
RMLAVH001@myuct.ac.za

ABSTRACT

Student advisors offer a great deal of assistance to all students. The University of Cape Town enrolls over 25000 students annually and they all must be registered on time and correctly. A large number of students leads to a large number of queries that Student Advisors may not be able to handle efficiently and effectively and thus leaving dissatisfied students. A Virtual Student Advisor containing a chatbot would offer a good alternative. The chatbot would answer any simple queries that students may have, thus leaving more time for Student Advisors to attend to more complex queries. This paper examines how such a chatbot can be implemented. Literature on previous chatbot implementations and Virtual Student Advisor systems is reviewed. The paper also looks at how chatbot anthropomorphism affects users. Potential security problems that may arise from such an application are also evaluated. It was found that there is relatively sufficient literature on the topic, albeit a few gaps, and that such an application is viable.

CCS CONCEPTS

• Natural Language Processing • Text Classification • Human-Computer Interaction

KEYWORDS

Chatbot, NLP, Student Advisor.

1. INTRODUCTION

One of the most important aspects of any University is its academic department. The department is responsible for offering guidance to students and assisting them with any queries they may have. Their objective is to offer effective and efficient assistance to as many students as possible. This poses a challenge for institutions with a large number of students as they are faced with an overwhelming number of queries from students that Student Advisors cannot effectively and efficiently deal with. Such is the problem faced at the University of Cape Town.

The objective of the project is to create a virtual student advisor that low-income students and students with less access to the university campus can access. The system should be able to do most of the tasks that human student advisors carry out daily. These include, pre-admission guidance, explaining procedures and policies and helping students determine the best fit for their needs. The need for a Virtual Student Advisor has been compounded by the Corona Virus pandemic that has forced most activities to be completed virtually.

This paper will be mainly focused on the chatbot aspect of the system- A chatbot is a software system, which can have a conversation with a human in natural language [1, 2, 6]. The purpose of this paper is to review literature that has been done on implementing a chatbot system for academic institutions. This will include reviewing literature on existing chatbot applications in general, both text and speech. Then literature on current Virtual Student Advisor systems. After this, literature on the UI and anthropomorphism of chatbots will be examined. Finally, testing methods and measurement metrics will be reviewed.

The following sections have been arranged as follows: Section 2 will give the background on the history of chatbot applications and existing text and speech chatbots. Section 3 reviews the various methods of designing a chatbot including the impact of the UI/UX design of the chatbot on people. Section 4 looks into how the effectiveness or success of the chatbot can be measured and tested. Section 5 will discuss the overall implementation of chatbot systems including interesting observations from the literature. Section 6 will provide any conclusions that have been drawn from the reviewed literature.

2. BACKGROUND & RELATED WORK

The Student Advisor plays a vital role in any student's academic timeline and therefore it should be possible for students to get assistance at anytime and anyplace as quick as possible. Student Advisors cannot attend to all queries efficiently due to the large number of students. A website with a virtual Student Advisor customized for each student would be an alternative to a human Student Advisor for less complex problems. A chatbot that is available daily, for 24 hours, would lead to more student satisfaction as their queries would be responded to quickly at any time of the day. More time is also left for human Student Advisors to attend to more complex problems that require more attention.

The first computer program that was capable of Natural Language communication between a machine and a human was created by Joseph Weizenbaum in 1966 called ELIZA [20]. ELIZA analyzed sentences through identifying key words in the input and generating responses by reconstruction rules related to the chosen decomposition rules [4].

2.1. RELATED APPLICATIONS

Language processing programs have since evolved to include both text and voice communication.

2.1.1. TEXT CHATBOT

Two prominent text chatbots, Elizabeth and The Artificial Linguistic Internal Computer Entity (ALICE), will be discussed briefly, although, there are various programs in existence.

Elizabeth was adapted from ELIZA, several improvements in the substitution, selection and storage mechanisms were made [1]. Although Elizabeth still implements the same basic concepts of following decomposition and pattern matching, Elizabeth offers more flexibility and adaptability. It can create an analysis of input sentences and producing a grammatical structure of the sentence. Storage improvements mean the program can store input for further use in the conversation at a later stage [1].

ALICE is a critically acclaimed Artificial Intelligence chatbot created by Dr Richard Wallace in 1995 capable of natural language processing [3]. Like Elizabeth, ALICE is also an adaptation of ELIZA. ALICE is less complicated than other NLP programs, the program uses simple pattern-matching algorithms to match sentence inputs to outputs [2]. The chatbot as won the Loebner prize three times- Loebner prize is a prize awarded to computer programs that have been judged to have the closest resemblance to a human-like conversation using the Turing test.

Both applications are capable of processing complex sentences although their processes are different. Alice uses pattern templates whilst Elizabeth uses a combination of patterns and input and output rules. Both systems have their strengths and weaknesses. For example, ELIZA allows for dynamic adaptability whilst ALICE uses simple patterns which make it easy to adapt to one's own application [1].

2.1.2. VOICE CHATBOT

Language processing agents have evolved to speech-based programs. Speech-based Natural Language Processing (NLP) agents are relatively new compared to text-based language processing agents. The four prominent programs are Amazon's Alexa, Apple's Siri, Microsoft's Cortana, and Google's Google Assistant.

Siri was first released in 2011 making it the oldest of the four agents. Cortana and Alexa were released in 2014 whilst Google Assistant was released in 2016. A comparison of these four programs in 2017 at their current state showed that all four still have room for much improvement in one aspect or the other [23]. The study by Gustavo L. showed that Google Assistant exhibited more natural tone and speech whilst Apple's Siri and Microsoft's Cortana utilized visual aids to enhance answers. All four had their advantages and disadvantages but non was more advanced compared to the other.

2.2. EXISTING STUDENT ADVISOR APPLICATIONS

There are a few applications that have been successfully implemented similarly to what this project aims to achieve- a website including a fully automated chatbot that is available at any time. The chatbot should be able to answer simple queries asked by students in a university. Bhavika R. et al. implemented a chatbot system that answers related Frequently Asked Questions [19]. Herry D. Wijaya et al. designed a chatbot for information management for college [20]. Suvethan N. et al. created a Virtual Student Advisor using NLP, a Schedular and a feedback analyzer [21].

Bhavika R. et al. designed a chatbot using Artificial Intelligence Markup Language (AIML) and Latent Semantic Analysis (LSA). The chatbot was created to effectively answer Frequently Asked Questions in a University [19]. This is very similar to what our project aims to achieve but limited mostly to a database of FAQs. The final product was not a blend of AIML and LSA which would allow for more natural communication. It functions well for basic communications but struggles when it comes to more Complex dialogue or questions. The AIML was implemented to deal with simple general questions and welcome/greetings and that functionality has been achieved well. Implementation of the LSA would improve the system.

Herry D. Wijaya et al. project's aim was to create a virtual assistant to be integrated to the existing website of the Mercu Buana University, Jakarta, Indonesia. The virtual assistant would help with the student traffic and answering questions efficiently. The existing system at the time had challenges in assisting alumni students and undergrad students with queries. The virtual assistant was created with inspiration from ELIZA [23]. They concluded their Chatbot was successful although more research would have been needed. Some downsides of the project would be that there appears to be little research done. The implementation and design could have been improved with more research and more testing.

Suvethan et al. present a mobile and web application students can interact with. The application uses NLP to handle students queries and pattern matching to provide appropriate answers. The system can assist students automatically or give students the option to meet with an actual Student Advisor [21]. It is mainly based on text input.

3. DESIGNING A CHATBOT

3.1. PROCESSING INPUT AND OUTPUT

Jack Cahn from the University of Pennsylvania provides an extensive thesis that outlines the various methods of processing input, extracting the necessary information from the input, and producing the appropriate results [12]. Bayesian or non-Bayesian methods can be used to process input. Bayesian methods process the likelihood of the sentence based on the words in the sentence.

Non-Bayesian methods make use of machine learning techniques like multi-layer perceptron and neural networks.

Several methods can be used to extract information from the input or get the meaning of the passed in sentence. Some of the methods are the "Bag of Words" technique, Latent Semantic Analysis (LSA) and Regular Expressions. "Bag of Words" technique is solely based on the probability of the words relative to the sentence regardless of order, syntax, or sentence structure. LSA also uses probability based on occurrence, the difference being it measures the probability of concepts or meanings. Regular Expressions utilize pattern matching, the input is matched with multiple regular expressions and the most appropriate result is chosen [12].

There are several models which implement response generation strategies. Some of these models are Rule-Based Models, Information Retrieval Models and Reinforcement Learning. Rule-Based Models produce answers based on a pre-defined set of rules created by the programmer, this also means that similar input sentences will have similar responses every time. Information retrieval models are similar to rule-based models with the advantage of being able to handle large data sets such as dialogues on Twitter where there are millions of conversations [12].

Extensive literature has been written on the various methods of processing input and producing outputs for an NLP system. Some systems use a combination of Artificial Intelligence Markup Language (AIML) and Latent Semantic Analysis (LSA) to develop a chatbot [19]. AIML is an XML (Extensible Markup Language) based language that is commonly used to develop conversational agents, with the most famous being A.L.I.C.E. [24]. It was developed Dr. Richard Wallace in 1995. AIML stores the knowledge base of the chatbot in text files. LSA is used for information extraction after the input has been processed.

3.2. UI/UX DESIGN

One of the most important features of a chatbot is its appearance. The appearance is what the user sees when interacting with the chatbot therefore it needs to be customized well. The chatbot's appearance could be generic for all user, or it could be customized for each student based on the profile of the student. Eun Go and S. Shyam Sundar produce literature that gives insight into the advantages and disadvantages of humanizing a chatbot [13]. Research has also been done on the appearance of voice chatbots and anthropomorphism of computers in general. Atieh Poushneh from the California State University-Bakersfield gave insight on the impact of humanizing a voice chatbot [14].

Findings show that a more anthropomorphic chatbot led to users having higher expectations on the functionality of the system. Users attempt to have more complicated or complex dialogue when the system appears to be human-like [13,14]. On the other hand, systems that lack humanness can compensate for it through more interactive messages in the conversation [13].

One could also consider the heuristics of the user interface. The appearance could be made customizable for each student thereby offering the user control, freedom, and flexibility- two of the ten usability heuristics for User Interaction Design [22].

Analysis of speech or voice agents has shown that voice agents that exhibit sincerity, creativity and functional intelligence lead to users interacting with agents more and thus greater user satisfaction [14].

3.3. SECURITY

Protection of student's data is a priority of the system. For the system to be effective, it may have to access sensitive student data. Therefore, the chatbots security will have to abide by data-protection regulations stated in the Protection of Personal Information Act (POPIA) passed in 2013 in South Africa. The POPI Act aims to protect subject's data from theft, discrimination, and any security breaches [25]. The POPI Act affects what kind of data the system can access from the University and how the data that has been accessed can be processed as South African universities need to maintain a certain level of Security infrastructure [16]. This also means third-party software or libraries have to be carefully reviewed before being integrated into the system.

A study done by Asbjorn Folstad et al. from the University of Oslo, Norway, found that customers trusted the chatbot more if its security and privacy measures were perceived to be of a sufficient level [26]. Therefore, it should be made clear by the system to the user that privacy and security of any data shared has been prioritized and the chatbot is secure.

4. TESTING A CHATBOT

Testing is a vital step in any development process. The chatbot must be tested to see how effective and efficient it is in responding to queries submitted by students. The Loebner Prize evaluation methodology is the current testing standard adopted by most chatbot developers [6]. However, evaluation should be adapted for a specific application's needs as they will vary from one system to the other. Minjee C. et al., Bayan Abu Shawar and Eric Atwell, and Aleksandra Przegalinska et al. provide literature on testing or measuring performance of a chatbot [5, 6, 9].

Minjee C. e al. conducted a study on customer satisfaction in luxury brands that use chatbots for customer service [5]. Luxury brands may choose to integrate automatic chat services for several reasons like being able to offer personalized customer service 24 hours daily and thus more customer satisfaction, or at least in theory. A digital presence also improves the connection between customer and company. The study used a five-dimension model; entertainment, interaction, trendiness, problem-solving, and customization, to measure perceptions of customers towards chatbots. The study found that customers were greatly satisfied with the chatbot and suggest that other luxury brands should implement a chatbot system too. Improvements to the study can be made by expanding the age range of participants and getting participants who have not used chatbots before to get more diverse feedback.

Other measurement metrics include dialogue efficiency metric, dialogue quality metric and user satisfaction as tested by Bayan Abu Shawar and Eric Atwell [6]. They also compared their

FAQchat bot with google, students selected their preferred tool for each query. Majority of their participants chose FAQchat bot as their preferred tool. The conclusion was that evaluation of chatbots should not be limited to a standard but rather modified for each application.

Aleksandra Przegalinska et al. argue that the base of a successful chat bot is trust between the human and the bot [9]. After evaluating current performance measurements like those used in finance and other industries they developed three features which make a chatbot more trustworthy and thus more successful. The three features are Transparency, Integrity and explainability.

5. DISCUSSION

It is evident that there is a great deal of research when it comes to implementing a chatbot system. Although, compared to the overall research in natural language systems, implementation in Virtual Student Advisor systems is little, especially in an African setting. Inspiration can be drawn from the research done by Bhavika R. et al. and Suvethan et al. as these applications are closest to what this project aims to achieve [19, 21]. Previous implementations have mostly been catered to their respective universities, as such, this project will have to customize its application for students at the University of Cape Town. This will come with its own set of challenges such us the language of choice as the University of Cape Town is diverse. The chatbot could be implemented to process English initially and then expanded to cater for vernacular languages and various other international languages in the future.

Literature on security of chatbots is minimal, this may be due to the fact that it is difficult to do research on systems that would be handling sensitive information. Companies may be reluctant to give information about their implementation of chatbots for security purposes. The lack of literature could be because chatbots in data sensitive industries like banks and medicine is new. The chatbot will have to implement Network and Internetwork Security services since it is an online system. The system will have to offer an authentication procedure for any data shared or received, access control to make sure only authorized individuals can access the information, confidentiality for student's data and the integrity of the data needs to always remain intact [11].

Jack Cahn presented detailed and extensive literature on how each step of the design process can be executed [12]. When processing input, the choice is between using Bayesian models or non-Bayesian models, each with their advantages and disadvantages. Bayesian models are mainly based on the probability of all dialogue act sequences in the corpora or language base. Non-Bayesian models make use of neural networks, text classification using machine learning, perceptrons and decision trees. Classification and machine learning could be better for a chatbot is it provides context based on past instances and generates better communication. The downside of this may be that classification and machine-learning based algorithms may be more complex and would take longer to implement.

Implementing a speech version of the chatbot would pose a challenge at present as such a complex program would likely lead to third party software or applications (i.e., Google's Google Assistant) being integrated into the system. This may raise security concerns as student data needs to remain confidential and protected.

Student information will have to be shared with the chatbot therefore, security and privacy is one of the main priorities of the system. There are limitations on the amount and type of information the chatbot can have access to. It also must abide by the rules stated in the POPI Act. UCT also contains a significant number of international students, so, the system will have to operate within international regulations and standards.

It should be possible to measure the performance of the chatbot and the satisfaction of students using it. Existing standards for testing could be adapted for our system.

There is little literature on an existing Virtual Student Advising system built for South African institutions. The system would be a valuable contribution to automated student advising systems in Africa.

6. CONCLUSIONS

There is variety of choices when implementing a chatbot, ranging from simple pattern matching models to complex rule based and deep learning models. Sufficient literature on the individual model exists but there is still room for more research to be done on how these models compare against each other in a complex natural language processing system. It is possible to integrate voice into the chatbot using an existing application like Google Assistant. The appearance of the chatbot will affect how students perceive the chatbot and its functionality therefore careful considerations need to be made on the UI design. The system will have to be tested with specified metrics that should be customized for the needs of students at the University of Cape Town.

Creating an interactive text and speech chatbot is a viable objective. Given the time and resources it would be possible to create a chatbot that can be implemented into a Virtual Student Advisor website.

7. ACKNOWLEDGMENTS

Many thanks to the project supervisor, Aslam Safla, for the assistance on the paper.

8. REFERENCES

[1] Bayan A. Shawar, Eric Atwell. 2002. *A Comparison Between Alice and Elizabeth Chatbot Systems*. University of Leeds, School of Computing research report 2002.19. https://eprints.whiterose.ac.uk/81930

[2] Bayan Abu Shawar and Eric Atwell. 2015. ALICE chatbot: Trials and outputs. *Centro de Investigacion en Computación, IPN*, 19 (4). pp. 625-632. https://dx.doi.org/10.13053/CyS-19-4-2326

[3] Bani, B. Singh, and Ajay P. Singh. 2017. College Enquiry Chatbot Using ALICE. *International Journal of New Technology and Research*, Pages 64-65.

[4] Joseph Weizenbaum. 1983. ELIZA - a Computer Program for the Study of Natural Language Communication Between Man and Machine. *Communications of the ACM, New York, NY, USA*, 26.1: Pages 23–28. https://doi-org.ezproxy.uct.ac.za/10.1145/357980.357991

[5] Minjee Chung, Eunju Ko, Heerim Joung, Sang Jin Kim. 2020. Chatbot e-service and customer satisfaction regarding luxury brands. *Journal of Business Research*, Volume 117, Pages 587-595. https://doi.org/10.1016/j.jbusres.2018.10.004

[6] Bayan Abu Shawar, and Eric Atwell. 2007. "Different measurement metrics to evaluate a chatbot system." In *Proceedings of the workshop on bridging the gap: Academic and industrial research in dialog technologies*, pp. 89-96.

[7] Mohit Jain, Pratyush Kumar, Ramachandra Kota, and Shwetak N. Patel. 2018. Evaluating and Informing the Design of Chatbots. *In Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*. Association for Computing Machinery, New York, NY, USA, 895–906. DOI: https://doi.org/10.1145/3196709.3196735

[8] Suhni Abbasi, and Hameedullah Kazi. "Measuring effectiveness of learning chatbot systems on student's learning outcome and memory retention." *Asian Journal of Applied Science and Engineering* 3, no. 2, Pages 251-260. https://www.journals.abc.us.org/index.php/ajase/article/view/251-260

[9] Aleksandra Przegalinska, Leon Ciechanowski, Anna Stroz, Peter Gloor, Grzegorz Mazurek. 2019. In bot we trust: A new methodology of chatbot performance measures. *Business Horizons, Volume 62, Issue 6*, Pages 785-797. https://doi.org/10.1016/j.bushor.2019.08.005

[10] Youjeong Kim, S. Shyam Sundar. 2012. Anthropomorphism of computers: Is it mindful or mindless? *Computers in Human Behavior, Volume 28, Issue 1*, Pages 241-250. https://doi.org/10.1016/j.chb.2011.09.006

[11] Heather Adkins, Betsy Beyer, Paul Blankinship, Piotr Lewandowski, Ana Oprea, and Adam Stubblefield. 2020. Building Secure and Reliable Systems: Best Practices for Designing, Implementing, and Maintaining Systems. *O'Reilly Media*.

[12] Jack Cahn. 2017. CHATBOT: Architecture, design, & development. University of Pennsylvania School of Engineering and Applied Science Department of Computer and Information Science.

[13] Eun Go, S. Shyam Sundar. 2019. Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior, Volume 97*, Pages 304-316. https://doi.org/10.1016/j.chb.2019.01.020

[14] Atieh Poushneh. 2021. Humanizing voice assistant: The impact of voice assistant personality on consumers' attitudes and behaviors. *Journal of Retailing and Consumer Services, Volume 58*. https://doi.org/10.1016/j.jretconser.2020.102283

[15] Kristine L. Nowak. 2004. The Influence of Anthropomorphism and Agency on Social Judgment in Virtual Environments. *Journal of Computer-Mediated Communication, Volume 9, Issue 2*. https://doi.org/10.1111/j.1083-6101.2004.tb00284.x

[16] Dominique Anderson, Oluwafemi Peter Abiodun, Alan Christoffels. 2020. Information security at South African universities—implications for biomedical research. *International Data Privacy Law, Volume 10, Issue 2*, Pages 180–186, https://doi.org/10.1093/idpl/ipaa007

[17] Sameera A. Abdul-Kader and Dr. John Woods. 2015. "Survey on Chatbot Design Techniques in Speech Conversation Systems. *International Journal of Advanced Computer Science and Applications(ijacsa), 6(7)*. http://dx.doi.org/10.14569/IJACSA.2015.060712

[18] Md Shahriare Satu, and Md Hasnat Parvez. 2015. Review of integrated applications with aiml based chatbot. *International Conference on Computer and Information Engineering (ICCIE)*, pp. 87-90. IEEE.

[19] Bhavika R. Ranoliya, Nidhi Raghuwanshi and Sanjay Singh. 2017. Chatbot for university related FAQs. *International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Pages 1525-1530. https://ieeexplore.ieee.org/abstract/document/8126057

[20] Herry Derajad Wijaya, Wawan Gunawan, Reza Avrizal, and Sutan M. Arif. 2020. Designing chatbot for college information management. *IJISCS (International Journal of Information System and Computer Science)* 4, no. 1, Pages 8-13. http://www.ojs.stmikpringsewu.ac.id/index.php/ijiscs/article/view/826

[21] N. Suvethan, K.A.M. Huzaim, R. Mathusagar, M.P.A.W. Gamage and A. Imbulpitiya. 2016. Virtual Student Advisor using NLP and Automatic Appointment Scheduler and Feedback Analyser. *International Journal of Scientific & Engineering Research, Volume 7, Issue 2*, Pages 155-160.

[22] Jacob Nielsen. 1993. Iterative user-interface design. Computer, 26(11), 32-41. doi: https://doi.org/10.1109/2.241424

[23] Gustavo Lopez, Luis Quesada, Luis A. Guerrero. 2018. Alexa vs. Siri vs. Cortana vs. Google Assistant: A Comparison of Speech-Based Natural User Interfaces. Advances in Human Factors and Systems Interaction. *Springer International Publishing*, Pages 241-250. https://link.springer.com/chapter/10.1007/978-3-319-60366-7_23

[24] Fernando A. Mikic, Juan C. Burguillo, Martin Llamas, Daniel A. Rodriguez, Eduardo Rodriguez. 2009. CHARLIE: An AIML-based chatterbot which works as an interface among INES and humans *2009 EAEEIE Annual Conference*, pp. 1-6, doi: https://doi.org/10.1109/EAEEIE.2009.5335493

[25] Johnny Botha, M. M. Grobler, Jade Hahn, and Mariki Eloff. 2017. A high-level comparison between the South African protection of personal information act and international data protection laws. *In ICMLG2017 5th International Conference on Management Leadership and Governance*, p. 57.

[26] Følstad, Asbjørn, Cecilie Bertinussen Nordheim, and Cato Alexander Bjørkli. 2018. What makes users trust a chatbot for customer service? An exploratory interview study. In International conference on internet science, pp. 194-208. Springer, Cham. https://doi.org/10.1007/978-3-030-01437-7_16