

# Sustainable Test Path Generation for Chatbots using Customized Response



Mani Padmanabhan

**Abstract:** In the current researching and Industrial fields have focused on much-attracted technology is chatbots. Informal agents could provide an appropriate and economic environment in online between the users and service provider. Due to a large number of datasets based on the digital tools, the user's queries based satisfying responses providing are critical in the service oriented chatbots. The successful human-chatbots interaction must be apparent and reacted by the user. This paper presents a technique to generate the test path for effective chatbot software testing. Identification of relations among the chatbots, transformations, and boundaries of the existing methodology are described. This experimental results associated with four peak chatbot response methodologies along with technology specifications. An in-depth experimental of proposed chatbot test path generation methodology, provides the experimental results and comparison.

**Index Terms:** Software Engineering, Software Testing, Chatbots, Advances in Natural Language, Deep learning.

## I. INTRODUCTION

The increasing popularity of human-computer interaction has directed many organization releasing If-This-Then-That (IFTTT) frameworks are well suitable. Pankaj R. Telang et.al., describes "The potential applications and popularity of chatbot technology have resulted in leading technology vendors such as IBM, Facebook, Microsoft and Google to releasing IFTTT frameworks to build such chatbots" [1]. The major goals of human-chatbot interaction are expected to ordinary discussions with the human.

Information Technology companies to identifying the natural conversations have been made on the open domain chatbots such as Google dialog flow, Facebook M, Clever-bot to unresolved the natural conversation. In the IFTTT chatbots difficult to provide pleasurable for each individual [4], the satisfying response should be not only sustainable to human quires. Nowadays, the official and unofficial discussion activities are unescapable in social networking tools.

The conversation use the general networks such as WhatsApp, twitter, Facebook, telegram and official networks such as workplace, slack has spread in the digital world. Human can use them universally keep on coupled through the conversion and their simplicity, mobility are major goal for this success.

Pankaj R.Telang et. al., has developed the conceptual

framework for enterprise chatbots. The framework provide the steps to extract intents from an utterance, Extract context from user interactions. The extracted intents, entities and context to generate a response using a dialog model. The author presented the framework with inference engine, knowledge base, planner, and external business services to carry out the user's request.

The interaction manages the dialog manager. In the service all the text and voice services are supported by presented framework. Lisa N. Michaud has describe the experience to develop the "virtual assistant chatbots to hotel guests in London. The author use the IFTTT frameworks. The data collected from the first 1,023 texts sent to human. 53 days of interaction with 491 distinct guest accounts. These texts contain a total of 1,258 different sentences". The IFTTT based chatbots are carried out in the modern of human-computer interaction.

In this paper section II describe the background and current problems in the human-computer interaction. Section III provide the existing chatbots generation methodology. The experiments are demonstrated with the real social media based datasets in the Section IV. Section V provide the comparative analysis. The conclusion in the section VI.

## II. BACKGROUND OF THE PROBLEM

The traditional chatbots response competence is too low. It can answer to the user only if there is a pattern matching between the user query and set of question-answer stored in its knowledge base. The chatbot response has two main responsibilities during the human-chatbot interaction. The initial task to be extract the meaning. The second and major task to identify the response from the datasets.

The major challenge in developing a sustainable chatbots is that generate the user satisfied output based on the input during the real-time with in the time limit.

This experimental based review for the real-time case study is to discover the ability of the customized chatbot interaction to engage in sustainable human conversation. Submit your manuscript electronically for review.

## III. LITERATURE REVIEW

To keep the human-chatbot interaction going, There has been a large amount of work have been made on the chatbots such as Goolge dialogflow, IBM Watson conversation, Microsoft LUIS, Facebook wit.ai, and Amazon lex .

Revised Manuscript Received on October 30, 2019.

\* Correspondence Author

**Dr. Mani Padmanabhan\***, Faculty of Computer Applications, SSL, Vellore Institute of Technology, Vellore, Tamil Nadu, India

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

## Sustainable Test Path Generation for Chatbots using Customized Response

Chatbots could provide timely and cost –effective social support to promote behavioral changes. Chatbots define as “The interaction system software that is designed to compete with human communication”. “A chatbot (also known as a spy, conversational bot, chatterbot, interactive agent, conversational interface, Conversational AI, talkbot or artificial spy entity) is a computer program or an artificial intelligence which conducts a conversation via auditory or textual methods to human”. Natural language processing (NLP) has occurred in the chatbot technology.

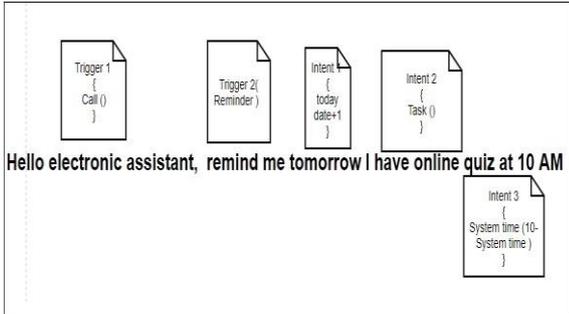


Figure 1 NLP for setting the reminder

The human language input has required for sustainable response in the natural language processing (NLP). The chatbot technology has occurred side by side with, human language input to its intended meaning. Figure 1 shows a sample of the dialog taking place between the college student and the chatbot. The NLP has divide in to trigger and intent. The chatbot response interprets to derive a domain model based on the social networks such as Facebook, WhatsApp, Twitter and Telegram.

The social media communication, involvement, and interaction of people has provide more in the accidental to expose the NLP for the individual’s information for the customized response. Thus, the importance of developing services or mechanisms to gather customized response from domain experts and documents (structured, unstructured, and semi-structured) has increased.

The discussing topics among peers, keeping contact with friends, and organizing all sorts of activities in open communities provide the solution for the customized response in the chatbots. The table 1 describe the list of existing chatbots.

Table 1. Existing Chatbots

Chatbot Name	Developer	Methodology
Elizabot	MIT Lab	Keyword matching
Alicebot	Wallace R.S et al.,	AIML Templates
Elizbet bot	Weizenbaum et al.,	Script command
Mitsuku	Worswick	AIML and keyword matching
Clever bot	Carprnter., R	Datasets matching
Chatfuel	Dumik., Chatfuel.,	Datasets and NLP
IBM Watson	IBM Deep QP project	Information retrieval (IR) and

		QA NLP
Microsoft LUIS	Microsoft	AI Engine
Goolge dialogflow	Goolge	Machine learning and NLP
Amazon Lex	Amazon	Deep learning

The purpose of this study is to explore the capability of the social networks based chatbot agents to involve in human conversation for the sustainable chatbots.

Information Retrieval is the central core topic in the Natural language processing. There have been serval methodology introduced by the AI research community in past years. Figure 2 shows the classification of the chatbot methodology.

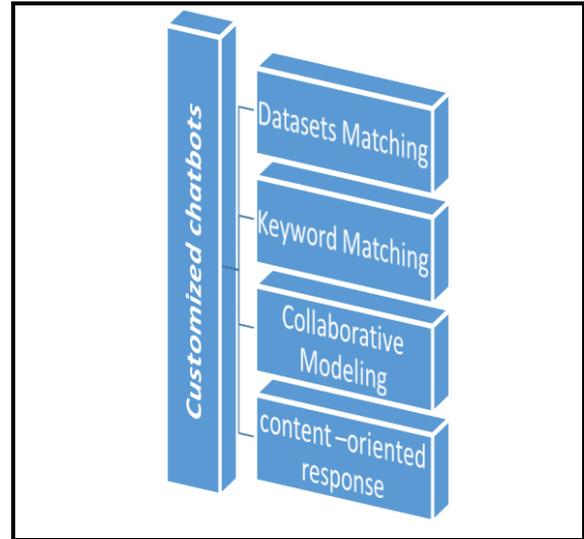


Figure 2 classification of the chatbots.

In the IFTTT methodology more chance to get failure or provide the response as the general. The user expecting customized response from the chatbots. In the below data shows the different between IFFTT and customized response from the electronic chatbot.

**Student:** Hello electronic assistant, remind me tomorrow I have online quiz at 10 AM.

**IFTTT:** Reminder sets

**Customized Response:** Reminder sets at 9:30 AM tomorrow. All the best for the exams.

The methodology has been divided in to four groups such as datasets matching based, Keyword matching based, Collaborative modeling and content –oriented response.

#### IV. PROPOSED METHODOLOGY

Datasets matching technique have been carried out the large amount of work over the past years. Ritter et al. introduced the generation of response based on Statistical Machine Translation (SMT). This SMT technique taking post-response pairs as parallel corpus.

In the SMT based technique have the ranking problem. Ji et al proposed an IR framework by introducing learning to rank strategy with the outputs of SMT, The ranking purpose they used the deep matching model, latent space model and topic-word model for providing ranking during the selection. An investigation has been taken in the datasets matching methodology with a student – faculty interaction during the online exam to draft and organize a new methodology in the chatbot interaction for the exam service domain. Figure 3 shows the interaction between the human-human in the public networks.

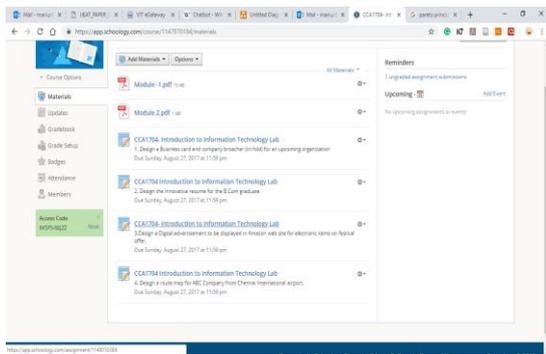


Figure 3 Conversations through URL

Human-to-human conversations over the schoolboy web application (<https://www.schoolboy.com>) are often extremely reduced and concise, holding many unstructured sentences during the official and unofficial chatting.

According to the Pareto’s principle 457 sentences were found in 22 percent of the texts; 45 percent of the customer-typed “sentences” contained only one word and 56 percent contained 8 words or more. The complete list of intent categories used in there analysis such as request, complaint, information, service request, The observations include that 63 percent of the sentences communicated dialogue acts that were in self-service categories and did not require human at all. When including those acts that involved the automated response alerting of human, the students’ knowledge and question asking capability to be change in the different part of the country. The failure rate based on the datasets matching to be 73 percent based on the experimental results with the different IQ students. In this methodology has been strongly suitable during the limited communications such as exam, online quiz. This is clear proof of the potential return in the chatbots based on the large amount of datasets.

The model of results in the datasets matching has be to determine a distribution of Pareto’s principle, so that one might be able to cover 80 percent of exam room conversations by implementing only 20 percent of the different intents.

The maintainability of the keywords is inadequate in terms of the flexibility and maintainability of the chatbots datasets. Figure 4 shows the keyword training phases in the Google dialogflow (<https://dialogflow.com>). The intent has to be fixed to show the response with request. If the 70-100 percentage of the keyword related than response to be selected. If the keyword less than 70 percentage then standard response such as “I can’t get you”, “I unable to understand”. Additionally, the keyword matching rules are constructed

based on the service.

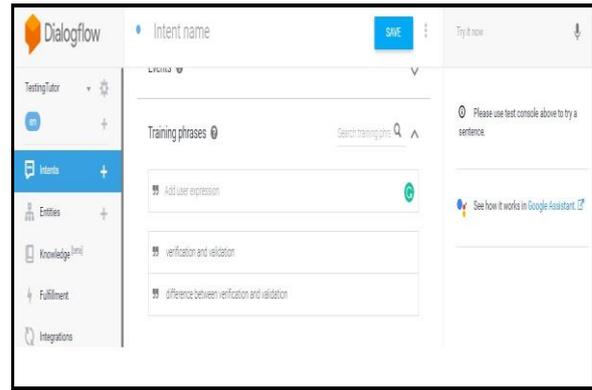


Figure 4. Keyword Matching Training phases

Keyword matching methodology are likely to fail when human discussions are complex. A traditional design approach might result in an implementation that strives for high truth in its responses, the request not understood. For example “set the alarm” is the normal keyword to be used but the same task in the friendly conversation to be the “wake up me tomorrow at 6 AM” both the sustenance shows the task set the alarm but the rejecting to be happen in the second sentence because the keyword missing, sample chatbot interaction shows in below

Set the alarm at 6 AM tomorrow  
Alarm sets at 6 AM tomorrow good night  
But it will incorrectly rejects this input:  
Wake up me at 6 AM tomorrow could be a grate.  
I’m sorry; I didn’t understand

The keyword matching approach deciding whether it is in scope (positive) or out of scope (negative). A negative judgment was counted failure of the keyword matching technique.

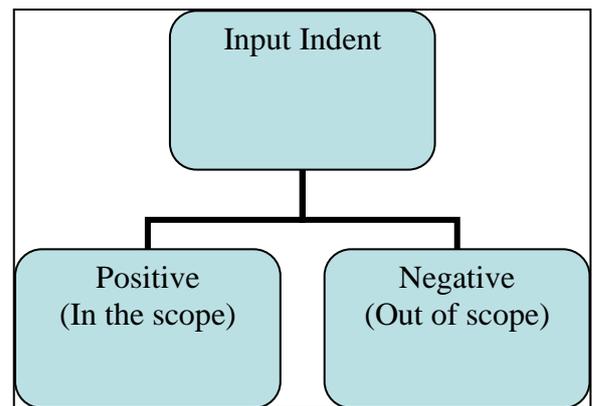
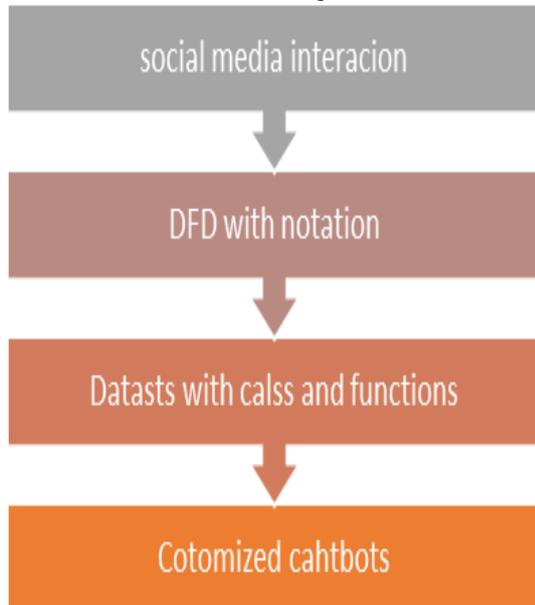


Figure 5. Intent classifier

Collaborative modeling are the suitable model in the datasets organization and selection process during the software development process. The process of chatbot response selection datasets are play the major roles.

## Sustainable Test Path Generation for Chatbots using Customized Response

In the fast growing chatbots management technique the process datasets organization has taken more time. In our proposed methodology DFD based collaborative modeling has been comprised for datasets building. The major source diagrams are acquired from the social networks. In our proposed methodology divided in to two parts. Figure 6 shows the process of generation of customized chatbots based on the Collaborative modeling



**Figure 6. Intent classifier**

Collaborative modeling can occur offline or online. This collaboration mode is more appropriate in our context because it supports early discussions and knowledge building. The social media discussions are converted into DFD. The modeling tool in online such as GenMyModel, MetaEdit+ and SPACE-DESIGN are used to provide the direct manipulation of diagrams from the social media. The large amount of personalized dissociation between the humans are necessary for the customized chatbots datasets. The datasets are the combination of main class and functions for the chatbot intent. In the sample interaction in the online exam as shown below.

### MyExamAPP

Teacher: All the students verify the course code IT5678 in the questing paper.

Students: Yes Sir.

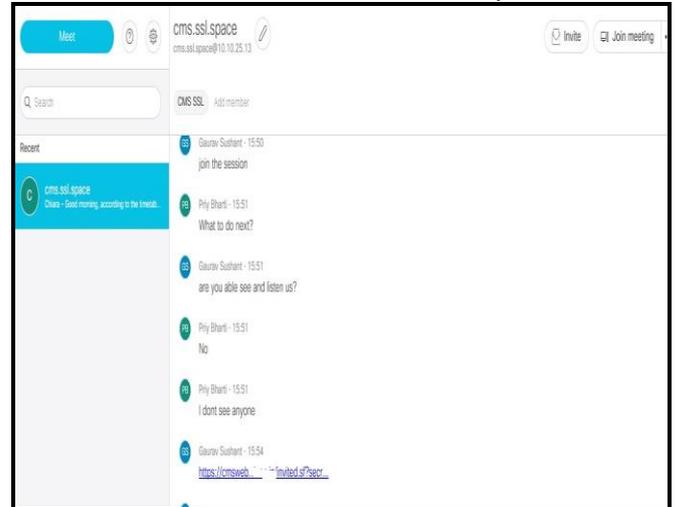
Student: Sir, I am Ram, kindly tell the course code one more time.

Teacher: course code IT5678.

Teacher: Your exam starts now the exam time 3 Hour from now.

Teacher: We will meet tomorrow at 8 AM for the next class.

In the discussion shows the exam has been conducted through the online. The exam starts at 10 AM as per the system time. The keyword ‘now’ to be used for identification of the exam start time. If any of the student in during the exam asking the time left for the exam the chatbot will answer directly to the student. In this real-time intent need the basic information exam start time but according to the interaction exam start time not discussed directly.



**Figure 7. Sample Cisco Social networks discussion**

The keyword matching chatbot are unable to provide the response. The proposed approach DFD collaborative modeling creates possible datasets based on the interaction, like exam start time, end time, course code, and next class timings. Figure 7 shows the traceability for the sample private network message between the university professor and students. Social networks for collaborative modeling brings exciting possibilities, such as involving large groups of people or using it in datasets selection for the customized response in the chatbots. Bingquan Liu et. al., has describe “personal information plays a great role during the procedure of human conversations; the personalized chat is a newly emerging demand in the research on chatbot; thus, little work has been conducted in this field.” Figure 7 shows the interaction between the university professor and students through social media. The dialogs 2055 in 45 days have consider for the DFD conversion from the private online discussion domain Impartus lecture capture system (<https://impartus.com/>). The interaction text contain 3078 different sentence-ending punctuation. 1245 text have found occurrences of non-vocabulary. 675 text complex sentences with unstandardized spelling. The data collected form the interaction are divided in the three categories such as questions, requests, information. For example student inform to the professor tomorrow during the online class “I am unable to attend because my sports event during that time” This type of the information are the customized indent in the chatbot datasets, if same student inform to the processor “We have the sports meeting during the online class timing tomorrow” This type of the interaction applicable to all the students so its under the general indent.

All the categories of response to be divided in general and personalized. This categories has to be useful during the notation formation in the DFD.

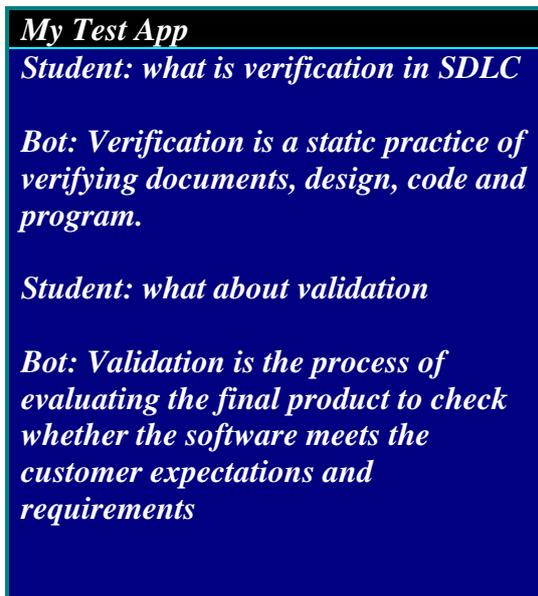
In the content of the chat history shows,  
 “we will meet ooty tomorrow”  
 “I have my son birthday on 3rd July”  
 “My husband working in the IT industry”

The extracted sample data from the social media shows the personalized data for the human like “Female from India with son”. Here ooty that the place in Tamil Nadu, India, my husband shows the chat person is female.

The content- oriented modeling will explore methods to incorporate the user representations learned by the models in this review paper more effectively. There are several studies forum that are trying to provide the customized response in human – chatbot interaction. From the summary of the experimental review the major four methodology has been experimented with the social media content in the following sessions describe the comparative results.

**V. COMPARATIVE RESULTS**

To provide the comparative results, I have consider the true discussion information from the social network such as twitter(<https://twitter.com>), telegram(<https://telegram.org/>) and Impartus lecture capture (<https://impartus.com/>) . Online teaching service management involving a chatbot for effective interaction.



**Figure 8. Sample Chatbot interaction**

Sample Keyword Request and Response

1. "intent": "actions.intent.TEXT",
2. "rawInputs": [
3. {
4. "inputType": "KEYBOARD",
5. "query": "what is the difference between verification and validation"
6. }
7. ],
8. "arguments": [
9. {
10. "name": "text",
11. "rawText": " what is the difference between

- verification and validation ",
12. ],
13. "textValue": " what is the difference between verification and validation "
14. "conversationToken": "[]",
15. "expectUserResponse": true,
16. "expectedInputs": [
17. {
18. "inputPrompt": {
19. "richInitialPrompt": {
20. "items": [
21. {
22. "simpleResponse": {
23. "textToSpeech": "Verification is a static practice of verifying documents, design, code and program"

The training phases of the interaction shows the raw text and the text to speech for the raw text.

The research question considers whether the customized chatbots success and satisfaction level in the human –chatbot interaction. The development time for the public social media based chatbot in the education sector to be consider for the comparative analysis table 2.

**Table 2. Chatbots Comparative Analysis**

	Datasets Matching	Keyword matching	Collaborative modeling	content –oriented response
<b>Development Time</b>	180 minutes	120 minutes	60 minutes	657 minutes
<b>Success Rate</b>	60-75%	40-55%	40-85%	55-95%
<b>Failure Rate</b>	15-45%	20-60%	10-25%	5-22%
<b>User Satisfaction level</b>	Low	Low	High	High

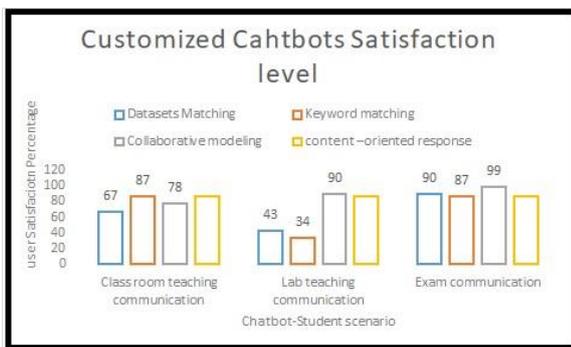
According to the experimental results in the four different methodology the following limitation exist for providing the customized response between the human –chatbot interaction in the higher education courses teaching scenario.

- Indent-Entity based: Simple machine learning approaches based chatbots are set of rule based and template –based. Failure rates are high.
- Language issue: capability of grammatical error analysis is low.
- Relational databases: Rule based templets are not focus the structural relations.



- **Speech Patten:** Natural language processing chatbots are unable to identify the human's current situation like happy, sad or angry Content oriented modeling required the strung deep learning to overcome rule based problems in the chatbots. The development cost and data gathering for the content oriented modeling is high. The customized response needed the well formalized tools for social media content analysis.

Figure 9 shows the comparison results of four methodology for chatbot generation with three different interaction between the University professor and students in the private and public social media during the teaching learning process. The first bar chats shows the 4786 interaction in the class room teaching and the satisfaction level for the students with the chatbots. The exam interaction between the chatbot and students has high (90% -99%) because the interaction level is low in the exam so easily the keyword to be matched the content. Compare with the exam and class room interaction the students has more satisfaction level in the exam room because limited communication. According to our experiments the responses using datasets matching and keyword matching are well suitable for the limited communications.



**Figure 9. Sample Chatbot interaction**

In the collaborative and content –oriented response are provide the sustainable response based on the social media data analysis. In the large number of datasets the customized response are possible based on the proposed collaborative modeling based approach.

## VI. CONCLUSION

The hand – written rules and IFTTT based chatbots shows that nearly 66% of human have experienced dissatisfied with the response of chatbots. The proposed modeling technique based test path identification on social media and content oriented modeling rapidly swapped the end-to-end interaction in the chatbots. More specifically, collaborative modeling tool is a more authoritative and generative-based model to solve the conversational response generation problems. Experimental results of teaching chatbot in the university level teaching showed that nearly 99% of students have experienced good customer service and generation of meaningful, long and informative responses remains a challenging task. Finally it has been observed proposed test path identification methodology well suitable for the sustainable chatbots.

## REFERENCES

1. Pankaj R. Telang et.al., “ A conceptual framework for engineering chatbots”, IEEE internet computing , November 2018. pp. 54–59.
2. A. Ritter, C. Cherry, and W. B. Dolan, “Data-driven response generation in social media,” in Proc. Conf. Empirical Methods Natural Lang. Process. Association for Computational Linguistics, 2011, pp. 583–593
3. Bingquan Liu et. al., “Content-Oriented User Modeling for Personalized Response Ranking in Chatbots”, IEEE/ACM Transactions on Audio, Speech, and Language Processing, Volume: 26 , Issue: 1 , Jan. 2018.
4. Lisa N. Michaud, “Observations of a New Chatbot: Drawing Conclusions from Early Interactions with Users, IT Professional, Sep./Oct. 2018, pp. 40-47, vol. 20.
5. Sara Pérez-Soler ; Esther Guerra ; Juan de Lara, “Collaborative Modeling and Group Decision Making Using Chatbots in Social Networks”, IEEE Software, Volume: 35 , Issue: 6 , November/December 2018, Page(s): 48 – 54.
6. M. Franzago et al., "Collaborative Model-Driven Software Engineering: A Classification Framework and a Research Map", IEEE Trans. Software Eng., [online] Available: 10.1109/TSE.2017.2755039.
7. M.-A. Storey et al., "How Social and Communication Channels Shape and Challenge a Participatory Culture in Software Development", IEEE Trans. Software Eng., vol. 43, no. 2, pp. 185-204, 2017.
8. J. Gallardo, C. Bravo, M.A. Redondo, "A Model-Driven Development Method for Collaborative Modeling Tools", J. Network and Computer Applications, vol. 35, no. 3, pp. 1086-1105, 2012.
9. Mohammad Nuruzzaman ; Omar Khadeer Hussain, “A Survey on Chatbot Implementation in Customer Service Industry through Deep Neural Networks”, 2018 IEEE 15th International Conference on e-Business Engineering (ICEBE), Xi'an, China, 12-14 Oct. 2018.
10. B. Morgan, How Artificial Intelligence will Impact the Insurance Industry in Forbes, July 2017.
11. M.L. Abbate, U. Thiel, T. Kamps, "Can proactive behavior turn chatterbots into conversational agents?", IEEE/WIC/ACM International Conference on Intelligent Agent Technology, 2005.
12. R.S. Wallace, R. Epstein, G. Roberts, G. Beber, The Anatomy of A.L.I.C.E in Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer, Dordrecht:Springer Netherlands, pp. 181-210, 2009.
13. J. Hill, W. Ford, I. Farreras, Real conversations with artificial intelligence: A comparison between human-human online conversations and human-chatbot conversations, vol. 49, 2015.
14. Mani P, Prasanna M: ‘Test Case Generation for Embedded System Software Using UML Interaction Diagram,’ Journal of Engineering Science and Technology, 2017, 12, 4, pp. 860-874
15. I. Sutskever, O. Vinyals, Q.V. Le, "Sequence to sequence learning with neural networks", Proceedings of the 27th International Conference on Neural Information Processing Systems, vol. 2, pp. 3104-3112, 2014
16. Adam, D. Nojan, N. Marios, F and Marin, L.: ‘Chatbots as assistants: an architectural framework’, Proceedings of the 27th Annual International Conference on Computer Science and Software Engineering, Nov 2017, pp. 76-86
17. Wang, H., Xing, J., Yang, Q., Song, W. and Zhang, X. :‘Generating Effective Test Cases Based On Satisfiability Modulo Theory Solvers For Service-Oriented Workflow Applications: Effective Test Cases For Service-Oriented Workflow Applications’, Software Testing, Verification and Reliability, 2016, 26,(2), pp.149–169
18. P. Mani and M. Prasanna.: ‘Validation of automated test cases with specification path,’ Journal of Statistics and Management Systems, Issue on Machine Learning and Software Systems, 2017, 20,4, pp. 535–542
19. V.-K. Tran, L.-M. Nguyen, Semantic Refinement GRU-Based Neural Language Generation for Spoken Dialogue Systems, Singapore:Springer Singapore, 2018.
20. Haolin Wang et. al., “Social Media–based Conversational Agents for Health Management and Interventions”, Computers , August 2018, pp. 26-33, vol. 51.
21. D. Griol and Z. Callejas , “ Mobile Conversational Agents for Context-Aware Care Applications ,” Cognitive Computation , vol. 8 , 2016 , pp. 336 – 356 .

22. Y. Wang , R.E. Kraut , and J.M. Levine , “ Eliciting and Receiving Online Support: Using Computer-Aided Content Analysis to Examine the Dynamics of Online Social Support ,” J. Medical Internet Research , vol. 17 , 2015 , p. 99.
23. H. Wang , Q. Zhang , and J. Yuan , “ Semantically Enhanced Medical Information Retrieval System: A Tensor Factorization Based Approach ,” IEEE Access , vol. 5 , 2017 , pp. 7584 – 7593 .
24. S.A. Rains and V. Young , “ A Meta-Analysis of Research on Formal Computer-Mediated Support Groups: Examining Group Characteristics and Health Outcomes ,” Human Communication Research , vol. 35 , 2009 , pp. 309 – 336
25. N.K. Cobb, A.L. Graham, D.B. Abrams, "Social Network Structure of a Large Online Community for Smoking Cessation", Am. J. Public Health, vol. 100, pp. 1282-1289, 2010.
26. H. Wang, Q. Zhang, J. Yuan, "Semantically Enhanced Medical Information Retrieval System: A Tensor Factorization Based Approach", IEEE Access, vol. 5, pp. 7584-7593, 2017.
27. A. S. Ghiduk,: ‘Automatic generation of basis test paths using variable length genetic algorithm,’ Information Processing Letters, 2014, 114, 6, pp. 304–316

#### AUTHOR PROFILE



**Dr. P. Mani M.S., M.Tech., Ph. D** is an Assistant Professor, Faculty of Computer Applications, SSL, VIT, Vellore, Tamil Nadu, India. He has over 8 years of teaching and research experience. His research interests include Software Engineering, Software Design, UML-based Testing, Software quality, Software Sustainability and Deep learning. [orcid.org/0000-0002-2402-7684](https://orcid.org/0000-0002-2402-7684)