Conversational AI: *An Overview of Methodologies, Applications & Future Scope*

Pradnya Kulkarni
Dept. of Computer Science
Dr. Vishwanath Karad MIT WPU
Pune, India
pradnya.kulkarni@mitwpu.edu.in

Ameya Mahabaleshwarkar
Dept. of Information Technology
MIT College of Engineering
Pune, India
ameyasm1154@gmail.com

Mrunalini Kulkarni
Dept. of Information Technology
MIT College of Engineering
Pune, India
mrunaliniskulkarni100@gmail.com

Nachiket Sirsikar
Dept. of Information Technology
MIT College of Engineering
Pune, India
nachiket3598@gmail.com

Kunal Gadgil
Dept. of Information Technology
MIT College of Engineering
Pune, India
kunalgadgil98@gmail.com

***Abstract*—Conversational AI is a sub-domain of Artificial Intelligence that deals with speech-based or text-based AI agents that have the capability to simulate and automate conversations and verbal interactions. Conversational AI Agents like chatbots and voice assistants have proliferated due to two main developments. On one hand the methods required to develop highly accurate AI models i.e. Machine Learning, Deep Learning have seen a tremendous amount of advancement due to the increasing research interest in these fields accompanied by the progress in achieving higher computing power with the help of complex hardware architectures like GPUs and TPUs. Secondly, due to the Natural Language interface and the nature of their design, conversational agents have been seen as a natural fit in a wide array of applications like healthcare, customer care, ecommerce and education. This rise in the practical implementation and their demand has in turn made Conversational AI a ripe area for innovation and novel research. Newer and more complex models for the individual core components of a Conversational AI architecture are being introduced at a never before seen rate. This study is intended to shed light on such latest research in Conversational AI architecture development and also to highlight the improvements that these novel innovations have achieved over their traditional counterparts. This paper also provides a comprehensive account of some of the research opportunities in the Conversational AI domain and thus setting up the stage for future research and innovation in this field.**

Keywords—Conversational AI, Dialogue Management, Natural Language Understanding, Natural Language Generation, Entity Recognition, Intent Classification

I. INTRODUCTION

Conversational AI Agents have become mainstream today with the tremendous advancement in methods required to build accurate models, i.e. machine learning and deep learning, and, secondly, due to the fact that they are seen as a natural fit in a wide range of domains, like healthcare, e-commerce, customer care, tourism and education, that heavily depend on natural language conversations in day-to-day operations. This lightning rise in demand has been met with an equally impressive rate of research and development where innovations are now happening every day. However, the meteoric rise in the research interest in this field has brought into spotlight some exciting, yet mercurial, research opportunities. Hence a systematic record of the core concepts of Conversational AI, traditional approaches and current implementations in these areas and the ongoing research, which will act as a platform for future research and innovations, is of paramount importance.

The Conversational AI architecture consists of three main components, each of which is further divided into basic parts that handle more preliminary tasks. The first part deals with the understanding of natural language inputs from users. This activity is essentially a combination of two Natural Language Understanding tasks viz. Intent Classification and Entity Extraction. Intent classification helps the agent understand the Why of the input [1]. Examples of intents in a food ordering chatbot can be - request, inform, place orders, and similarly for a healthcare domain - reporting a symptom, reporting a diagnosis and asking for medicine prescription [2][3]. Entity Extraction deals with the What of the input [4]. It helps the agent identify the discrete pieces of information received from the user, which when combined with the intent allow the agent to fully comprehend the users input. With upcoming better ways of representing natural language in a computer understandable format while also representing relations between different entities ex. Word Embeddings like Word2Vec, NLU has become more accurate than some of its earlier equivalents [5]. After understanding the users input, the agent needs to decide upon its own set of actions which should effectively continue the conversation while preventing states where the Conversational Agent is stuck without necessary information or in an incomplete situation. Ideally the agent should select a set of actions that help in resolving the users request. However, it may happen that the agent does not have the entire information that it needs to conclusively decide upon a specific course of actions. For example, if the user is requesting movie timings for a theatre, but the user has not mentioned which movie, then the agent should recognize the missing piece of information and respond accordingly. Based on this, the agent traverses between different specific states and depending upon the state that the agent finds itself in, the agent decides its next action. The framework module that handles this operation is known as the Dialogue Management System [6]. Some basic concepts of dialogue management include grounding, slot filling, context switching which have been detailed in the following sections of the paper. The last part of a Conversational AI interaction is conveying the current state and results to the other involved interacting entity. The reply should be sent to the user in user understandable format. Natural Language Generation is used for this purpose. Natural Language Generation is the process of converting structured data into user understandable natural language. It works exactly opposite to natural language understanding. It

978-1-7281-4042-1/19/$31.00 ©2019 IEEE

is a complex process which involves stages of content determination, document structuring, aggregation, lexical choice, referring expression generation and realization [7].

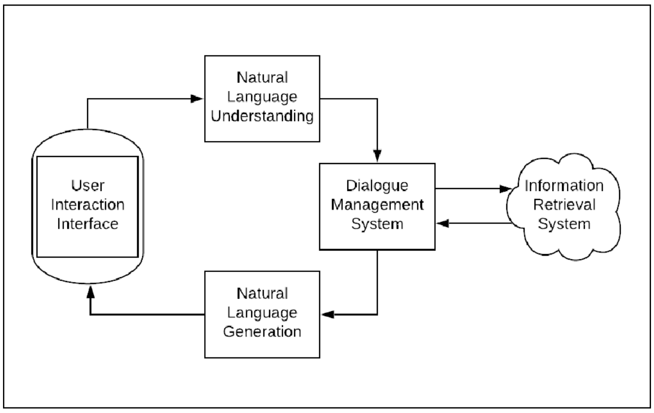

Fig. 1. Conversational AI Architecture

Each of the components explained above is a complex research problem in itself. Various machine learning and deep learning models are used to increase the accuracy of each component. This paper studies the ongoing research related to the components of natural language understanding, dialogue management and natural language generation in conversational AI agents and determines some of the future directions that Conversational AI can take.

II. Natural Language Understanding

Natural language understanding (NLU) is a branch of artificial intelligence (AI) that uses computers to understand input made in the form of unstructured text or speech. The field of NLU is an important and challenging subset of natural language processing (NLP)). NLU is tasked with communicating with untrained individuals and understanding their intent, meaning that NLU goes beyond understanding words and interprets meaning. NLU is even programmed with the ability to understand meaning in spite of common human errors like mispronunciations or transposed letters or words [8]. The NLU provides a direct human-computer interaction. The NLU allows human languages to be understood statically by the computer without the use of if / else. The Natural Language Understanding (NLU) covers one of AI's complex challenges [9]. NLU mainly consists of two tasks - Named Entity Recognition (NER) and Intent Classification (IC). Figure 2 gives an example of Natural Language Understanding in AI agents.

A. Named Entity Recognition (NER)

NER deals with identifying and separating the named entities of a sentence into various predefined classes. These named entities are the particular terms in contain unique information that can help a conversational agent understand the "What" of the user input. Named entities can be classified into classes like people, places, organizations and objects. Earlier works have explored the use of Regular Expressions for NER where in the user inputs are pattern matched against a regular expression to see if they match and are then assigned classes accordingly [10].

This approach is however limited in scope and not suitable for dynamic conversations with wide ranging contexts. Conditional Random Fields (CRFs) are widely used for NER and are extensively found in many applications [11]. However, effective implementation of CRFs requires a tedious amount of feature extraction making this approach less scalable and limited in adaptability [12]. More recently Convolutional Neural Networks (CNNs) have been used for NER and have shown promising results with Micro F scores of 88.64% and 91.13%, which are comparable to the results obtained by CRFs but requires significantly less amount of data pre-processing [12]. A CNN based entity extraction is a two phase process where first the words are represented as numerical vectors using word embeddings and are then fed to the CNN for training on predictions of their labels. Zheng, Suncong, et al. [14] have proposed a Bidirectional Long Short-Term Memory Network (BiLSTM) to capture long term dependencies and a Convolutional Neural Network (CNN) to obtain a feature vector which is given to a sigmoid activation based classifier.

The output of the sigmoid classifier and the BiLSTM is then given to another LSTM to predict the entities. The BiLSTM is also used to capture future and past information as it's beneficial for sequence modelling tasks. The experimental setup consists of a public dataset containing 1441 sentences which was randomly split into training (1153) and test (288) sets. The standard F1 measure is used to evaluate the performance of the model. The performance is compared to two other systems, namely TF and TF-MT. The BiLSTM model got an F1 score of 0.632 as compared to 0.583 (TF-MT) and 0.610 (TF) as an end-to-end model.

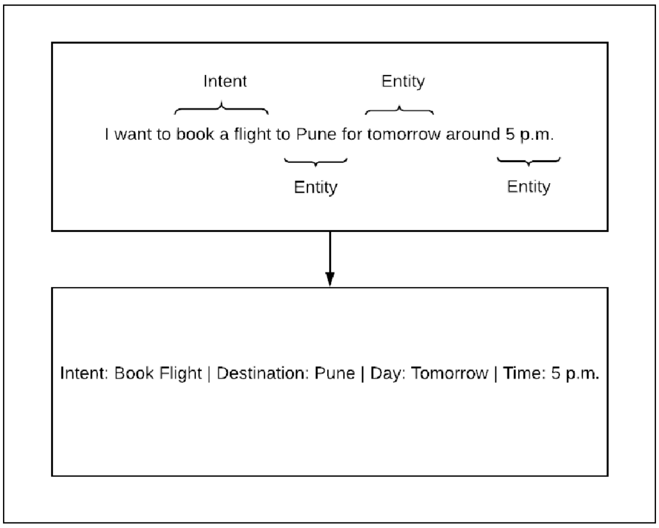

Fig. 2. Natural Language Understanding

B. Intent Classification

IC tries to understand the actual intent of the user as behind providing the input that has been observed. Intent Classification helps the bot understand what exactly is the desired goal or the objective of the user. For example, in healthcare Conversational Agents, user intents can help the bot understand what is the use of the information received from the patient. Some of the intents can be - describing a symptom, pointing out the location of occurrence of a symptom, intensity of a symptom, etc. Various methods have been employed for the task of intent classification. The most notable results have been obtained by Machine Learning and

Deep Learning approaches. Traditional IC models have mainly used supervised learning methods like Hidden Markov Models (HMM) [15] and Decision Trees (DT) [16]. The main drawback of these methods is the requirement of massive training sets. SVMs have been used to classify intents of questions asked online health forums and have shown a precision score of up to 75% [17]. The advent of deep learning has brought forward neural network models that are more successful than earlier and traditional machine learning methods at intent classification. Meng, et al. [18] have proposed two deep hierarchical LSTM models for classifying dialogue intents. The two models are a Hierarchical LSTM (H-LSTM) model and a Memory Augmented Hierarchical LSTM (H-LSTM+Mem) model. The HLSTM relates the sentences and contextual information more closely to make the IC more effective. To make it more effective, a memory unit is added to the H-LSTM to make the H-LSTM+Mem. The memory unit is added to the output of the second LSTM to provide useful contextual information to calculate the sentence vector. The models are evaluated on a dataset provided by an ecommerce company containing 24760 sentences. The dataset is randomly split into training set (80%), validation set (10%) and test set (10%). Based on the top 20 labels, the H-LSTM+Mem has 83.9% accuracy as compared to 81.6% (H-LSTM) and 79.7% (Basic LSTM). More recently, Bidirectional Long Short Term Memory (BiLSTM) Neural Networks have shown state-of-the-art results for classifying multi-class intents with an accuracy up to 94.16% [19].

III. Dialogue Management

Dialogue Management (DM) is an important module in the Conversational AI framework that is responsible for governing the actions of the Conversational Agent and mapping inputs to appropriate outputs and has been an area of tremendous research interest for the past two decades [20]. The DM system is responsible for developing an interaction strategy that can guide the agent in deciding its own actions based on the inputs that have been received from the user. DM systems can be of two types viz. Goal/Task Oriented Systems and Non-Task Oriented Systems [21].

Task Oriented DM Systems are responsible for guiding the user from one state of the conversation to another so as to successfully achieve a predefined or dynamically understood task [22]. Examples of Task Oriented Conversational Agents are widely found today such as conversational agents designed for performing simple tasks like booking a movie ticket, scheduling meetings [23], task management, FAQ answering [24] and can range up to agents designed for performing more complicated tasks like open domain question answering [25], visual question answering and medical diagnosis [26].

The DM system also functions as a state tracker that continuously maintains the state of the conversation and is also responsible for initiating a transfer from one state to another when the control of the conversation is with the Conversational Agent [27]. Table 1 provides details about the different states that a conversation can be in during an interaction between a human and a Conversational Agent.

Following are some of the traditional, current state-of-the-art and promising methods of Dialogue Management System implementation –

TABLE I. States in Dialogue Management

Sr. No.	State	Conversation Control	Description
1.	Grounded	Agent	Acknowledging the users input while deciding upon the agents actions
2.	Slot Filling	Agent	Requesting extra information from human to resolve actions
3.	Initiative	Agent / Human	Steering of the conversation by either entity
4.	Context Switch	Human	Change of the basis or the premise of the conversation

A. *Switch Statements*

The most basic type a DM system can be found in the form of a sizeable Switch Statement. A Switch Statement is a programming control structure that has a set of previously defined actions for each possible input. Every distinct intent that can be expected from the NLU module triggers a unique response or an action [28]. The drawback of this approach arises from the fact that the conversation initiative always lies with the user. This way, the agent becomes merely a reactive interface that cannot provide an engaging conversation.

Another disadvantage with this approach is that the scope of the conversation is limited to the predefined set of actions and responses. Any input that slightly varies from the fixed format, even if valid in context of natural language, will fail to get a satisfactory response from the Conversational Agent.

B. *Finite State Machines*

The goal of a task oriented DM system is to guide the two entities involved in the conversation i.e. the human and the agent through several states of the conversation where the either entity can determine the direction of the conversation. This perspective of viewing DM systems makes Finite State Machines (FSMs) a suitable candidate for implementation [29]. This approach works especially well when the set of possible inputs is limited.

However, the shortfall of this approach is in its lack of flexibility and adaptability. Each conversation that has the same end goal will go through the same set of states even though the current progress and state of the conversation makes some of them redundant.

C. *Machine Learning based Approaches*

A Machine Learning (ML) based approach for DM realization takes the user intents from human input as the input parameters to the ML algorithm and predict what the Conversational Agents actions should be. However, the use of a classic ML algorithm on natural language inputs requires significant pre-processing thus making it an unattractive choice [30]. Deep Learning can help in overcoming this problem with its inherent ability to handle unstructured data. The disadvantage is its lack of ability to improve based on the outcome of the conversation i.e. whether the conversation was successful or not.

D. Deep Reinforcement Learning based Approaches

The aim and functioning of a DM system can be viewed in the form of a sequential decision process. The DM system has to make a series of decisions by observing the inputs to the system and deciding upon the set of actions to take. This paradigm of operation makes Reinforcement Learning a suitable candidate for implementing DM systems. Reinforcement Learning has shown significantly promising results in other tasks that are based on a similar paradigm - an Agent taking Actions based on its interaction with its Environment, like playing Atari games [31]. Deep Reinforcement Learning trained agents have surpassed previous implementations by outperforming heuristic based approaches by over 50% conversation success rate and by outperforming supervised agents with more than 30% conversation success rate [32]. Reinforcement Learning approaches overcome the disadvantage of ML approaches by integrating a conversation success reward and failure penalty which incentivizes the Conversational Agent to try to reach successful conclusions in the conversations. Deep Reinforcement Learning models functioning on top of well-trained prediction models have resulted in agents capable of altering their behaviour in unseen environments and achieving better success rates with the use of a delayed reward signal [33]. Despite their promise, DLR approaches to DM System implementation have their own set of challenges like lack of adjustability in policies, requirement of hand crafted policy and lack of clarity in which dialogues can be modelled and which cannot [34].

E. Belief based DM Systems

A partially observable Markov decision process (POMDP) is a generalization of a Markov decision process (MDP). A POMDP models an agent decision process in which it is assumed that the system dynamics are determined by an MDP, but the agent cannot directly observe the underlying state. Instead, it must maintain a probability distribution over the set of possible states, based on a set of observations and observation probabilities, and the underlying MDP. This is a state-of-the-art approach that is supposed to give highly accurate results but there is little experimental proof due to lack of implementation. The architecture of the model consists of three parts. The bottom sentence-LSTM is developed for slot filling. It is shared for labelling slot values in user utterance at each turn. The middle belief-LSTM is developed for state tracker. It receives previous state and results from the bottom sentence-LSTM to update current states. The upper layer is developed for action selector. It maps states to dialogue actions. A hybrid learning algorithm is used to train the model, where DRL is employed for training the entire network, and supervised learning is used for training sentence LSTM simultaneously [35].

IV. NATURAL LANGUAGE GENERATION

Natural Language Generation (NLG) is a subdomain of NLP that is focused on the methods of response generation in natural language. In the perspective of Conversational AI, NLG plays the important role of making the conversation seem more natural for the human participant which is a critical factor for judging the effectiveness of Conversational Agents.

The NLG module receives input from the Dialogue Management system in a structured format that is based upon the dialogue history and the current context [36]. Thus, the output from the NLG component in a Conversational Agent is a natural language sentence or text which is also the final output of the Conversational AI framework for each dialogue instance. The output of the NLG component is based upon the processing and results of the Natural Language Understanding and Dialogue Management Systems. Following are some of the traditional, current state-of-the-art and promising methods of Natural Language Generation –

A. Template based / Rule based Approach

Template Based systems are NLG systems that map a non-linguistic structured input like a query directly to a natural language representation using predefined templates [37]. The output from such a system often contains gaps, which when appropriately filled after being successfully mapped, give well-formed results as shown by the work of Deemter, K. et al. The disadvantage of template based approaches is their lack of generalization in gender, number and person agreement as well as the fact that handmade rules result in identical outputs that make the conversation tedious [38].

B. N-Gram Generator

As viewed in the work of Adwait R. (2000) the N-Gram approach translates the input as an attribute - value set and generates output by considering the word sequence that has the highest probability of including each input attribute once. A pitfall of N-Gram generator is that it fails to acknowledge that the preceding words are not always the best indicators of the upcoming word and sometimes the overall context of the sentence matters more [39].

C. Neural Network Approach

Tsung-Hsien Wen et al. employed neural networks and deep learning for natural language generation and have tested the use of RNNs for the task [40]. From another one of their experiments it has been found that Semantically Conditioned LSTM (SC-LSTM) gave the best results in BLEU scores as well as evaluation by human judges also confirmed that SC-LSTM model was strongly preferred to the then existing models [41].

D. Seq2Seq

In recent times, Sequence-To-Sequence (Seq2Seq) models have gained a lot of popularity and are supposed to provide state-of-the-art performance in a wide variety of tasks including Conversational Agents [42][43]. Seq2Seq models map an input sequence to a vector representation using LSTM models and then sequentially predicts tokens based on the pre-obtained representation. The model defines a distribution over outputs (Y) and sequentially predicts tokens given inputs (X) using a softmax function. However, such Seq2Seq models suffer from two common problems which are exposure bias and inconsistency between train/test measurement [44].

E. *Seq2Seq + Reinforcement Learning*

Yaser Keneshloo et al. developed a model combining power of reinforcement learning methods with seq2seq models in natural language generation for text summarization. They observed that reinforcement learning model required training time of only a few hours and also a highly superior ROUGE score was recorded in their work. The proposed models used for experimentation were Actor-Critic, Self-Critic, PolicyGenerator, E2E, Argmax. Actor-Critic (Q-Learning) method had ROUGE (1, 2, L) scores of (40.88,17.80,38.54) which is highly superior to the other methods' score average of (38.43,16.64,35.09) [45]. In a combined Seq2Seq and RL approach, the Seq2Seq model produces a set of possible outputs which form the action space for the RL agent and the conversation history combined with the context form the state space or the environment. The agent then learns the optimal response strategy by learning to predict the reward on each action which is defined heuristically by ease of answering, information flow and semantic coherence [46]. Figure 3 shows an architectural overview of a Seq2Seq + RL NLG system.

V. APPLICATIONS

Conversational AI applications have proliferated in the past few decades with an increase in research and development in this domain. Conversational Agents can now be found in a wide range of applications performing a plethora of interesting tasks. Ashay Argal et. al used DNN (Deep Neural Network) and Restricted Boltzmann Machine (RBM) to develop a chatbot in the tourism domain [47]. Kyungyong Chun et al. developed an AI powered conversational agent that provided an online healthcare diagnostic service by using a cloud based knowledge base [48]. Chin-Yuan Huang et al. developed a chatbot that was embedded in a healthcare application for weight management and provided a natural language control system for a wireless healthcare system [49].

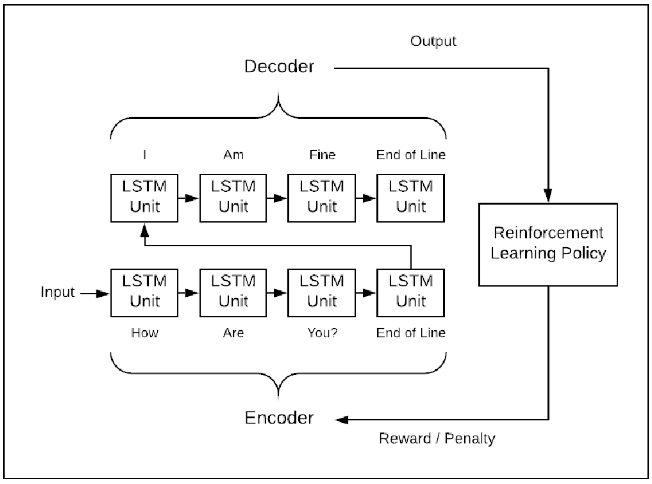

Fig. 3. Seq2Seq + RL

Sharob Sinha et al. have developed a chatbot for the education domain that can answer user queries from education related documents and can function as a virtual teaching assistant [50]. Danilo Arruda et al. developed a KAOS Modelling Assistant Chatbot, known as KAOSBot for assisting requirements engineers to carry out requirement elicitation efficiently [51]. Divya M. et al. demonstrated in their work a medical assistant conversational agent that was capable of disease diagnosis and also providing treatment guidance [52]. Clarizia F. et al. developed a conversational agent that could act as an educational support system to students by detecting students' questions and answering them based on developed ontologies [53]. In a related work, Clarizia F. et al. showed how a conversational agent can be used to provide an engaging tourism experience by automating location based storytelling [54]. Su M. H. et al. demonstrated the use of conversational agents in a sensitive task like edlerly care [55]. The versatility and capability of conversational agents to handle complex tasks that are considered challenging for humans as well was made apparent in the work of Dongkeon Lee et al. who developed an AI agent that could provide a therapeutic counselling service that was also capable of providing emotional responses [56].

VI. CONCLUSION

The study has presented the traditional approaches for Conversational AI implementation in order to gain an insight on this domains evolution. This paper has also reviewed the ongoing research on each of the three basic components of Conversational AI Agents, viz. Natural Language Understanding, Dialogue Management and Natural Language Generation. Based on the accuracy of each model, state of the art methodologies for each component, that give best results have been highlighted. Drawbacks of some of the current widespread methodologies have also been discussed. Some of the cutting edge ongoing research in implementation as well as application of Conversational AI has also been reviewed. This study has thus put the traditional and existing methodologies of Conversational AI into perspective in order to provide a platform for further innovation and exploration.

VII. FUTURE WORK

The work presented in this paper acts as a platform to spearhead further research in Conversational AI which can take multiple directions. This paper has reviewed some of the shortcomings of current implementations in Conversational AI and at the same time has presented some of the current research going on to overcome these shortcomings. This ongoing research can be coupled with parallel implementations which help in widespread adoption of these research works and can also help test them in real-world scenarios. The state-of-the-art works reviewed in this paper have all been results of disparate research endeavours. Future work can be done to bring all these individual component level state-of-the-art works into one single hybrid architecture that is capable of performing exceptionally on all Conversational AI tasks and will also help in determining the compatibility between these disparate research works. In previous sections newer techniques like Reinforcement Learning have been discussed for some of the tasks in Conversational AI which have the potential to expand into novel research problems of their own and efforts can be taken to develop them. Finally, as reviewed in this work, Conversational AI applications in fields like healthcare, education and tourism can be developed further by combining Conversational AI with other subdomains of AI like Computer Vision to explore tasks like visual question-answering and language controlled image segmentation.

REFERENCES

[1] L. Qiu, Y. Chen, H. Jia and Z. Zhang, "Query Intent Recognition Based on Multi-Class Features," in IEEE Access, vol. 6, pp. 52195-52204, 2018.

[2] J. Chen, R. Prasad, S. Stoyanchev, E. Selfridge, S. Bangalore and M. Johnston, "Corpus and Annotation Towards NLU for Customer Ordering Dialogs," 2018 IEEE Spoken Language Technology Workshop (SLT), Athens, Greece, 2018, pp. 707-713.

[3] Razzaq M.A., Khan W.A., Lee S. (2017) Intent-Context Fusioning in Healthcare Dialogue-Based Systems Using JDL Model., Enhanced Quality of Life and Smart Living. ICOST 2017. Lecture Notes in Computer Science, vol 10461. Springer, Cham.

[4] X. Dong, L. Qian, Y. Guan, L. Huang, Q. Yu and J. Yang, "A multiclass classification method based on deep learning for named entity recognition in electronic medical records," 2016 New York Scientific Data Summit (NYSDS), New York, NY, 2016, pp. 1-10.

[5] J.Devlin, M.-W.Chang, K.Lee, K.Toutanova ,Bert: pre-training of deep bidirectional transformers for language understanding,arXiv:1810.04805.

[6] Young, S. (2017). Statistical Spoken Dialogue Systems and the Challenges for Machine Learning. Proceedings of the Tenth ACM International Conference on Web Search and Data Mining – WSDM’17.

[7] T. Young, D. Hazarika, S. Poria and E. Cambria, "Recent Trends in Deep Learning Based Natural Language Processing [Review Article]," in IEEE Computational Intelligence Magazine, vol. 13, no. 3, pp. 55-75, Aug. 2018.

[8] B. H. Juang and Tsuhan Chen, "The past, present, and future of speech processing," in IEEE Signal Processing Magazine, vol. 15, no. 3, pp. 2448, May 1998.

[9] M. C. Surabhi, "Natural language processing future," 2013 International Conference on Optical Imaging Sensor and Security (ICOSS), Coimbatore, 2013, pp. 1-3.

[10] Ek, T., Kirkegaard, C., Jonsson, H., & Nugues, P. (2011). Named Entity Recognition for Short Text Messages. Procedia - Social and Behavioral Sciences, 27, 178–187.

[11] A survey of named entity recognition and classification. (2007). Lingvisticæ Investigationes, 30(1), 3–26.

[12] McCallum, A., & Li, W. (2003). Early results for named entity recognition with conditional random fields, feature induction and web-enhanced lexicons. Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL 2003.

[13] Dong, X., Qian, L., Guan, Y., Huang, L., Yu, Q., & Yang, J. (2016). A multiclass classification method based on deep learning for named entity recognition in electronic medical records. 2016 New York Scientific Data Summit (NYSDS).

[14] Zheng, Suncong, et al. “Joint Entity and Relation Extraction Based on a Hybrid Neural Network.” Neurocomputing, vol. 257, Sept. 2017, pp. 59– 66.

[15] H. Cuayahuitl, S. Renals, O. Lemon and H. Shimodaira, "Humancomputer dialogue simulation using hidden Markov models," IEEE Workshop on Automatic Speech Recognition and Understanding, 2005., San Juan, 2005, pp. 290-295.

[16] Mendoza M., Zamora J. (2009) Building Decision Trees to Identify the Intent of a User Query. In: Velásquez J.D., Ríos S.A., Howlett R.J., Jain L.C. (eds) Knowledge-Based and Intelligent Information and Engineering Systems. KES 2009. Lecture Notes in Computer Science, vol 5711. Springer, Berlin, Heidelberg.

[17] Zhang, T., Cho, J. H. D., & Zhai, C. (2014). Understanding user intents in online health forums. Proceedings of the 5th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics - BCB ’14.

[18] Meng, Lian, and Minlie Huang. “Dialogue Intent Classification with Long Short-Term Memory Networks.” Natural Language Processing and Chinese Computing, edited by Xuanjing Huang et al., Springer International Publishing, 2018, pp. 42–50.

[19] L. Qiu, Y. Chen, H. Jia and Z. Zhang, "Query Intent Recognition Based on Multi-Class Features," in IEEE Access, vol. 6, pp. 52195-52204, 2018.

[20] Ilievski, V., Musat, C., Hossmann, A., and Baeriswyl, M. (2018). Goaloriented chatbot dialog management bootstrapping with transferlearning.arXiv preprint arXiv:1802.00500.

[21] Chen, H., Liu, X., Yin, D., & Tang, J. (2017). A Survey on Dialogue Systems. ACM SIGKDD Explorations Newsletter, 19(2), 25–35.

[22] Hussain, S., Ameri Sianaki, O., & Ababneh, N. (2019). A Survey on Conversational Agents/Chatbots Classification and Design Techniques. Developments in Primatology: Progress and Prospects, 946–956.

[23] Cranshaw, J., Elwany, E., Newman, T., Kocielnik, R.,Yu, B., Soni, S., Teevan, J., and Monroy-Hernandez, A.Calendar. help: Designing a workflow-based schedulingagent with humans in the loop. InProceedings of the2017 CHI Conference on Human Factors in ComputingSystems, ACM (2017), 2382–2393.

[24] Bhavika. R. R., Nidhi. R. and Sanjay. S., "Chatbot for university related FAQs," 2017 International Conference on Advances in Computing Communications and Informatics (ICACCI), Udupi, 2017, pp. 1525-1530Silvia.

[25] Q., Suresh M., A Chatbot-based Interactive Question Answering System, Proceedings of the 11th Workshop on the Semantics and Pragmatics of Dialogue, pages 83–90. Trento, Italy, 30 May – 1 June 2007.

[26] C. Liu et al., "Augmented LSTM Framework to Construct Medical SelfDiagnosis Android," 2016 IEEE 16th International Conference on Data Mining (ICDM), Barcelona, 2016, pp. 251-260.

[27] Traum D.R., Larsson S. (2003) The Information State Approach to Dialogue Management. In: van Kuppevelt J., Smith R.W. (eds) Current and New Directions in Discourse and Dialogue. Text, Speech and Language Technology, vol 22. Springer, Dordrecht.

[28] Zhu W., Chowanda A., Valstar M. (2016) Topic Switch Models for Dialogue Management in Virtual Humans. In: Traum D., Swartout W., Khooshabeh P., Kopp S., Scherer S., Leuski A. (eds) Intelligent Virtual Agents. IVA 2016. Lecture Notes in Computer Science, vol 10011. Springer, China.

[29] C. Hori, Kiyonori Ohtake, Teruhisa Misu, Hideki Kashioka and Satoshi Nakamura, "Statistical dialog management applied to WFST-based dialog systems," 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, Taipei, 2009, pp. 4793-4796.

[30] Olivier Lemon, Olivier Pietquin. Machine Learning for Spoken Dialogue Systems. European Conference on Speech Communication and Technologies (Interspeech'07), Aug 2007, Anvers, Belgium. Pp.26852688.

[31] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., Riedmiller, M.: Playing atari with deep reinforcement learning. In: Proceedings of the NIPS Deep Learning Workshop (2013).

[32] Cuayáhuitl, H., Keizer, S., Lemon, O.: Strategic dialogue management via deep reinforcement learning. In: Proceedings of the NIPS Deep Reinforcement Learning Workshop (2015).

[33] Su, P.-H., Gasic, M., Mrksic, N., Rojas-Barahona, L., Ultes, S., Vandyke, D., Wen, T.-H., and Young,S. (2016a). Continuously Learning Neural Dialogue Management. ArXiv e-prints.

[34] Paek, T., & Pieraccini, R. (2008). Automating spoken dialogue management design using machine learning: An industry perspective. Speech Communication, 50(8-9), 716–729.

[35] S. Lei, X. Wang and C. Yuan, "Word-Based POMDP Dialog Management via Hybrid Learning," in IEEE Access, vol. 7, pp. 39236-39243, 2019.

[36] Albert. G., Emiel. K., Survey of the State of the Art in Natural Language. Generation: Core tasks, applications and evaluation, Journal of Artificial Intelligence Research 61 (2018) 65-170.

[37] Gervás. P., Díaz-Agudo. B., Peinado. F., & Hervás. R. (n.d.). Story Plot Generation based on CBR. Applications and Innovations in Intelligent Systems XII, 33–46.

[38] Deemter, K. van, Theune, M., & Krahmer, E. (2005). Real versus Template-Based Natural Language Generation: A False Opposition? Computational Linguistics, 31(1), 15–24.

[39] Ratnaparkhi. A., (2000) Trainable methods for surface natural language generation. In Proceedings of the 6th Applied Natural Language Processing Conference and the 1st Meeting of the North American Chapter of the Association of Computational Linguistics (ANLPNAACL'00), pp. 194–201.

[40] Wen. T-H., Gasic. M., Kim. D., Mrksic. N., Su. P-H., Vandyke. D., and Young. S. 2015a. Stochastic Language Generation in Dialogue using Recurrent Neural Networks with Convolutional Sentence Re-ranking. In Proceedings of the16th Annual Meeting of

the Special Interest Group on Discourse and Dialogue (SIGDIAL). Association for Computational Linguistics, September.

[41] Wen. T-H., Gasic. M., Kim. D., Mrksic. N., Su. P-H., Vandyke. D., and Young. S., 2015b. Semantically conditioned lstm-based natural language generation for spoken dialogue systems. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP). Association for Computational Linguistics, September.

[42] Ji. Z., Xiong. K., Pang. Y., & Li. X. (2019). Video Summarization with Attention-Based Encoder-Decoder Networks. IEEE Transactions on Circuits and Systems for Video Technology, 1–1.

[43] M. Qiu, F.-L. Li, S. Wang, X. Gao, Y. Chen, W. Zhao,H. Chen, J. Huang, and W. Chu. Alime chat: A se-quence to sequence and rerank based chatbot engine.InProceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2:Short Papers), volume 2, pages 498–503, 2017.

[44] J. Li, X. Chen, E.H. Hovy, D. Jurafsky, Visualizing and understanding neural models in NLP, CoRR abs/1506 (2015) 01066.

[45] Y. Keneshloo, T. Shi, C. K Reddy, and N. Ramakrishnan. 2018. Deep Reinforcement Learning for Sequence to Sequence Models. arXiv preprint arXiv:1805.09461(2018).

[46] Shibata T., Egashira Y., Kurohashi S. (2016) Chat-Like Conversational System Based on Selection of Reply Generating Module with Reinforcement Learning. In: Situated Dialog in Speech-Based HumanComputer Interaction. Signals and Communication Technology. Springer, Cham.

[47] A. Argal, S. Gupta, A. Modi, P. Pandey, S. Shim and C. Choo, "Intelligent travel chatbot for predictive recommendation in echo platform," 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, 2018, pp. 176-183.

[48] Chung, K., & Park, R. C., Chatbot-based healthcare service with a knowledge base for cloud computing. Cluster Computing. (2018).

[49] C. Huang, M. Yang, C. Huang, Y. Chen, M. Wu and K. Chen, "A Chatbotsupported Smart Wireless Interactive Healthcare System for Weight Control and Health Promotion," 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Bangkok, 2018, pp. 1791-1795.

[50] Sinha S., Basak S., Dey Y., Mondal A. (2020) An Educational Chatbot for Answering Queries. In: Emerging Technology in Modelling and Graphics. Advances in Intelligent Systems and Computing, vol 937. Springer, Singapore.

[51] Arruda D., Marinho M., Souza E., Wanderley F. (2019) A Chatbot for Goal-Oriented Requirements Modelling. In: Computational Science and Its Applications – ICCSA 2019. ICCSA 2019. Lecture Notes in Computer Science, vol 11622. Springer, Cham.

[52] D. Madhu, C. J. N. Jain, E. Sebastain, S. Shaji and A. Ajayakumar, "A novel approach for medical assistance using trained chatbot," 2017 International Conference on Inventive Communication and Computational Technologies (ICICCT), Coimbatore, 2017, pp. 243-246.

[53] Clarizia F., Colace F., Lombardi M., Pascale F., Santaniello D. (2018) Chatbot: An Education Support System for Student. In: Cyberspace Safety and Security. CSS 2018. Lecture Notes in Computer Science, vol 11161. Springer, Cham.

[54] Clarizia, F., Lemma, S., Lombardi, M., Pascale, F.: An ontological digital storytelling to enrich tourist destinations and attractions with a mobile tailored story. In: Au, M.H.A., Castiglione, A., Choo, K.-K.R., Palmieri, F., Li, K.-C. (eds.) GPC 2017. LNCS, vol. 10232, pp. 567–581. Springer, Cham (2017).

[55] Su, M.H., Wu, C.H., Huang, K.Y., Hong, Q.B., Wang, H.M.: A chatbot using LSTM-based multi-layer embedding for elderly care. In: International Conference on Orange Technologies (ICOT) (2017).

[56] Dongkeon Lee, Kyo-Joong Oh and Ho-Jin Choi, "The chatbot feels you - a counseling service using emotional response generation," 2017 IEEE International Conference on Big Data and Smart Computing (BigComp), Jeju, 2017, pp. 437-440.