Let's Chat about Brexit! A Politically-Sensitive Dialog System based on Twitter Data

Aparup Khatua
Temasek Laboratories
Nanyang Technological University
Singapore
aparupkhatua@gmail.com

Erik Cambria
School of Comp. Sc. and Engineering, Nanyang Technological University
Singapore
cambria@ntu.edu.sg

Apalak Khatua
Strategic Management Area
XLRI Xavier School of Management Jamshedpur
India
apalak@xlri.ac.in

Iti Chaturvedi
Temasek Laboratories
Nanyang Technological University
Singapore
iti@ntu.edu.sg

***Abstract*—Data scientists are exploring various semi-supervised learning methods to build conversational agents - commonly known as chatterbot. This paper investigates various issues related to a political chatterbot where human agents are politically opinionated. Here, understanding the latent intent of human agent is crucial for developing an efficient political chatterbot. We set our study in the context of 2016 Brexit referendum. We argue that employing a subjectivity detector and an emotion analyzer, in addition to the keyword based topic detector, enhances the intent detection process. Next, we discuss the importance of maintaining political neutrality. To maintain its neutrality, a chatterbot needs to disassociate itself from a politically opinionated response. This can be achieved by associating a response with a user or a set of users. Nowadays, the Twitter platform provides an enormous amount of user-generated contents for various socio-economic events. Hence, we have considered tweet feeds for developing the overall chatterbot architecture in the political domain.**

Keywords—Chatterbot; Twitter; 2016 Brexit Referendum

I. INTRODUCTION

Recent developments in the domain of artificial intelligence (AI) and natural language generation have enabled a machine to engage in a conversation with a human agent. This conversation agent or chatterbot is an old concept in the discipline of natural language generation. Conversation capability (regarding the scope of the discussion) and response mechanism are two important aspects of a dialog system (Fig. 1).

It is worth noting that prior attempts to develop open-domain conversational agents failed to generate a meaningful conversation. In the last decade, rule-based closed-domain chatterbots, which are less challenging from a computational perspective, became quite popular for various business applications. For instance, automated agents can complement a business organization's efforts on the online platform by assisting a human agent in choosing travel/accommodation options or by giving healthcare supports to patients residing in remote locations.

These rule-based chatterbots are fine for short conversations, but this approach fails to get engaged in a long conversation with the human agent due to the lack of training data or rules. This breakdown, due to inappropriate utterances by the chatterbot, impedes the process of long conversations between machine and human agents. This is why today preventing conversation breakdown is the key issue in designing dialog systems.

To this end, researchers are trying to move from symbolic AI to sub-symbolic AI to enhance the efficiency and broaden the applications of dialog systems. This paper proposes a framework for developing a politically-sensitive dialog system and employs unstructured text data, extracted from the Twitter platform, for the overall chatterbot architecture. We probe various challenges associated with the existing approaches in a political domain and, subsequently, attempt to address them. We consider the 2016 Brexit referendum related discussions on Twitter for our study. To the best of our knowledge, none of the prior studies have investigated the challenges associated with developing a dialog system in the political domain.

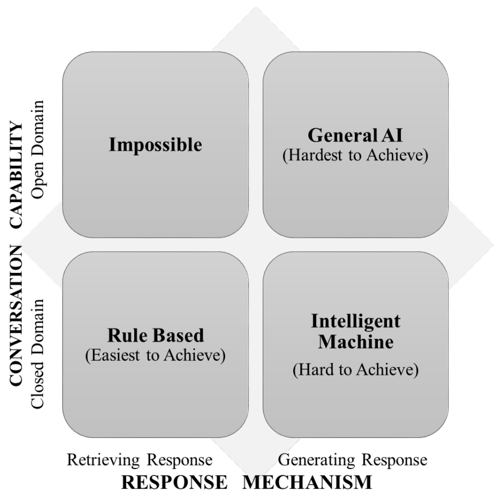

Fig. 1. Categories of dialog system research

The nature of a political chatterbot is different from other commercial chatterbots. For a political conversation agent, it is important to understand the political 'intent' or leaning of the human agent for generating an appropriate unbiased response, which is not crucial for a chatterbot providing IT supports. In other words, there are no ideological differences between the machine and the human agent in the context of an IT helpdesk-related or travel-related queries. Thus, in our proposed framework, even if a politically-opinionated human agent says something in a biased manner on a particular topic, then our chatterbot agent considers the political intent of the human agent for 'generating response'. In brief, we are trying to incorporate emotional intelligence in the chatterbot architecture to mimic human-like behavior during the conversation.

II. RELATED WORKS

Dialog system research is classified into two major areas, namely: non-goal oriented and goal-oriented. Developing a non-goal oriented generic dialog system is a challenging task in comparison to a goal-oriented system. However, recent research is employing deep learning algorithms and the huge volume of training data to focus on non-goal oriented system design. Recurrent neural network (RNN) based sequence-to-sequence framework is gaining popularity for designing a neural conversion model. For training purpose, researchers are using input and corresponding output sequence similar to a conversation between two humans [27].

In dialog system research, corpus-based learning is a prevalent method to train the chatterbot. Large text corpora, such as tweet feeds, movie subtitles, IT Helpdesk, and Ubuntu dialogs, were used by prior studies [20], [27], [16] to train the open as well as closed-domain systems with a certain level of accuracy. Prior studies have also attempted to train the dialog system by using live interactions with human agents [23].

In 2015, a research team from Google had attempted to design an end-to-end open-domain dialogue system using sequence-to-sequence model and considered movie subtitles for training purpose [27]. Google's dataset comprises of 62 million sentences for training purpose and 26 million sentences for validation purpose. Google's model has attempted to predict the next sentence on the basis of previous sentence/sentences in the conversation. Thus, this approach has eliminated the need for the hand-crafted rule to design a domain specific dialogue system. Google's study considered both closed-domain IT helpdesk data set as well as open-domain subtitles dataset. In both closed- and open-domain, this approach performed decently, but their proposed dialog system failed to pass the Turing test [27].

In 2016, researchers from McGill and Montreal University have attempted to design a dialog system by using Ubuntu Dialogue Corpus (which comprises of 1 million multi-turn dialogue) to train their neural conversion language model [16]. This project has considered both RNN and long short-term memory (LSTM) to generate the next best possible response in a conversation and has observed that LSTM performs better than RNN [16].

In 2017, a team of Microsoft researchers has also attempted to design a fully data driven knowledge grounded neural conversation model to produce a more content-rich response without slot filling [9]. In this project, they used two large datasets namely, 23 million tweet conversations and 1.1 million tips from Foursquare to train their model. Their simple entity matching approach for incorporating the external information makes the model generic and this model is also applicable to open-domain systems [9]. However, this approach also failed to match the desired performance level.

The goal-driven approach relies on hand-crafted rules to generate the response in a dialog system. In the goal-driven approach, communication logs of business sectors, in formats such as chat records, SMS, and email conversations, are considered as an input training data. These models have some inherent limitations, but this goal-driven approach is commercially more viable in comparison to non-goal oriented dialog system. Nowadays, researchers are employing machine learning techniques for classifying the 'intention' [22] to precisely understand a user's requirement, and subsequently, to generate a corresponding response for the said user. Prior studies also considered the emotion of input message to generate a human-like response to developing an emotional chatting machine [29].

An end-to-end dialogue system comprises of many sub-components such as speech recognizer, natural language interpreter, state tracker, response generator, natural language generator and speech synthesizer [22]. However, exploring all these components is beyond the scope of a single paper. Hence, this paper concentrates on developing a chatterbot architecture. For generating an 'appropriate' response, our proposed framework also considers the intent/emotion associated with the input message to understand the cognitive behaviors of the human agent [11]. Employing emotional intelligence for generating a response is essential to mimic human-like conversations.

Currently, researchers are employing voluminous unstructured text data, from social media platforms, to train the dialog system. Twitter is becoming popular not only among commercial organizations to promote their new products or services but also among social activists, as well as other social media users, to express their opinion about various socio political issues [14]. This is why we chose to extract context-specific political conversation from the Twitter platform. However, the failure of Microsoft Tay project highlights the inherent limitations of the Twitter corpus as a training dataset for developing a dialog system. [12]. It is worth noting that Microsoft Tay project considered unfiltered tweets for the training purpose. On the contrary, we are considering filtered context-specific tweet data for the training purpose. We have generated intent-specific datasets from our twitter corpus to train our proposed model.

III. RESEARCH CONTEXT AND DATA

UK electorates voted on June 23, 2016, to decide whether Britain should exit from the European Union or not. This referendum was commonly mentioned as Brexit, an abbreviation of Britain's exit, in the media [10], [13].

We have considered this Brexit referendum, as a context, for designing a political chatterbot. We have extracted around 2.7 million tweet feeds during June 18, 2016, to June 28, 2016. To extract our data, we have considered four hashtags/keywords as follows: *Brexit*, *EUref*, *VoteLeave* or *VoteRemain*. The last two keywords can aptly summarise the political discourse of the 2016 Brexit referendum. On the one hand, the *VoteLeave* lobby wanted to leave the European Union "to reap the benefits of free trade with the rest of the world" [13]. Influential leaders of this camp emphasized the need to 'takeback control' of their country to ensure a 'real freedom' from immigration problems and 'a dysfunctional bureaucracy' to guarantee 'safety, education, health and happiness' for the next generation [13].

On the other hand, the *VoteRemain* lobby highlighted the benefits of being in the European Union and argued that the other European countries might impose high import tariff barriers to UK industrialists as a consequence of Britain's exit from the European Union. In brief, the remain lobby argued that the UK will be 'better off' as a member of European Union – otherwise 'a country isolated from its neighbours' will 'struggle' 'to protect jobs' [13]. Hence, two dominant political ideologies or *intents* of our research context were 'leave' and 'remain'. Table 1 reports a few sample tweets and their intents.

TABLE I: SAMPLE TWEETS AND INTENTS

Intents	**Sample Tweets** [13]
Leave	£10billion per year to support a dysfunctional beauraucracy! #VoteLeave #LeaveEUOfficial #IvoteLeave #LeaveEU
Leave	£17bn, the true cost of immigration to the UK every year #VoteLeave #Brexit #LeaveEU
Leave	@David_Cameron if we leave we'll still be part of a vibrant, connected world, but free to shape our own destiny #voteleave #TakeBackControl
Remain	#Brexit fallacy: UK great enough to stand alone in world but not big enough to influence EU #VoteRemain : UK big influence in EU #StrongerIn
Remain	"A country isolated from its neighbours is one that struggles to progress." #StrongerIN #GreenerIN #EUref
Remain	#VoteRemain for an equal britain. for a strong economy. for a diverse and interesting place to live. #LabourInForBritain

IV. PROPOSED CHATTERBOT ARCHITECTURE

Figure 2 graphically depicts our proposed chatterbot architecture. This chatterbot is comprised of various components, such as response handler, response analyzer, tweet and topic database. A crucial component of our proposed framework is the *response handler*, the first layer of our framework. The response handler will analyze the input message from the human agent. Following prior studies [25], we propose a keyword-based topic detector –as the first layer. The topic detection process, in our response handler, tries to map an input message with our set of pre-identified topics (which we have prepared through manual annotation by using our contextual understanding about the political event).

If our response handler can map an input message with a predefined topic from our database, then the input message will get tagged accordingly, and our chatterbot will continue with that topic until a new topic is identified during the conversation [25]. However, a keyword-based topic identification has some inherent limitations. For instance, let us consider the following input messages:

Human: *What is European Union?*

Human: *David Cameron has misled people about immigration.*

A keyword-based topic identifier will tag this input message on the basis of keywords such as the 'European Union' or 'David Cameron.'

Incidentally, UK voters desperately Googled the first question immediately after the polling day [13]. The first question (or the input message) is *objective*. However, the David Cameron related second input message is *subjective*. The keyword(s) 'David Cameron' was mentioned by both the leave and remain camp [13], and our conversational agent needs to know the intent of the human agent. The keyword-based topic identifier is fine for objective input message but not for the subjective input message. Thus, in the next step, we apply *subjectivity detection* on the input message. In the context of chatterbot, subjectivity detection is of utmost importance, because it allows to identify politically opinionnated input messages.

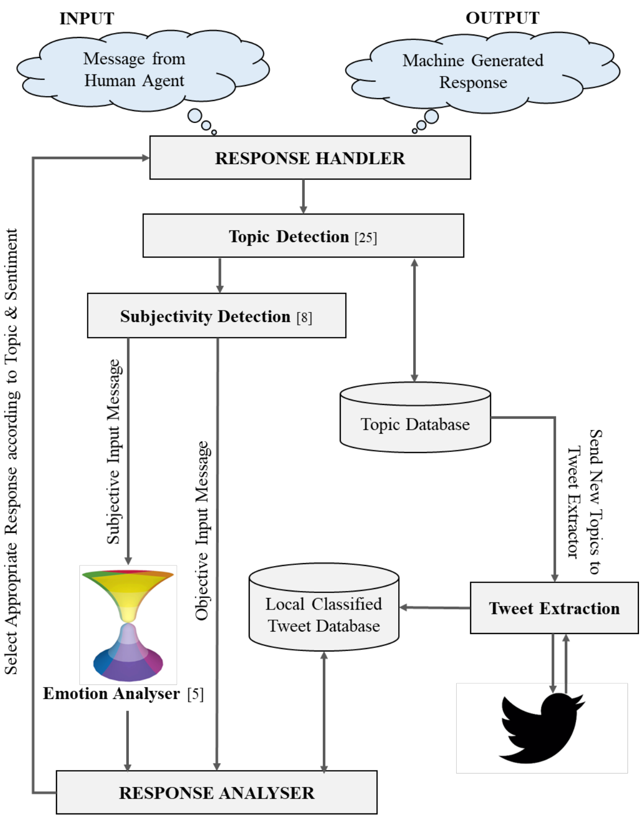

Fig. 2. Proposed Framework

For example, objective input messages such as 'What is the purpose of EU?' are not politically opinionated. Thus, polarity detection is not required. Subjectivity detection can be a complex task in natural language processing For instance, in the sentence 'Supporting Brexit today will ruin the trade prospects of UK tomorrow', he [a user] said while citing the example of past regional trading agreements' there is clearly an objective frame for the writer and a direct subjective frame for the user with the text anchor 'said'. However, it is ambiguous whether the texts anchor 'citing' is objective or subjective in nature.

Prior studies observe that subjective sentences can be of wishful nature such as indicating purchasing interest [26]. In [1], the authors show that subjective sentences in online forums can be identified by 'Dialog Acts' such as 'Question', 'Repeated Question', 'Clarification', etc. They also show that subjective sentences are longer than objective sentences and often contain inappropriate content such as abusive language. In the context of chatterbot, mostly politically opinionated users will use 'abusive language'.

In breif, probing the subjectivity/objectivity of an input message helps us to understand the political leaning of a user better. In other words, an objective message indicates that the user is probably not opinionated. On the contrary, for a subjective message, it is important to identify the user's political leaning. However, only subjectivity detection might not be sufficient. For instance, let us consider a few input messages as follows:

Human: *The argument of #<u>VoteLeave</u> lobby sounds more logical than #<u>VoteRemain</u>.*

Human: *The argument of #<u>VoteRemain</u> lobby sounds more logical than #<u>VoteLeave</u>.*

Human: *<u>DavidCameron</u> has misled people about immigration. The truth is, we can only control it if we <u>#VoteLeave.</u>*

Here, it is difficult to interpret the intent of the human agent due to the presence of hashtags/keywords from both the leave and remain lobby within the above input messages. A bag-of-words perspective will treat the first two input messages uniformly. In reality, these two messages are conveying two different, rather opposite, political intents. Similarly, the last input message is containing keywords such as *David Cameron* (related to VoteReamin lobby) as well as *VoteLeave*. Thus, it will be difficult for the chatterbot to interpret the intent of human in such situations correctly.

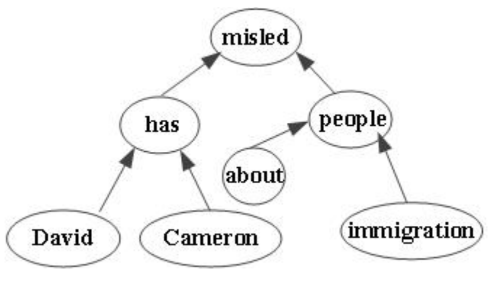

Fig 3. Bayesian network for a sentence

To address this, we need to move from statistical aspect to linguistic aspects of an input message, and this can be achieved by considering the the dependency relation between clauses. In [8], the authors used Bayesian networks to learn the weights of a neural network and detect subjectivity. An example of a Bayesian network representing inter-dependencies between the words of the sentence "David Cameron has misled the people about immigration" is illustrated in Figure 3. Once the structure of the Bayesian network is determine heuristically using the training data, the context (that is, the parents for each word) can be established. For example, the context of the word `people' is 'immigration' and 'about'. Hence, [8] argues that structurally related words, among all the words within the sentence, provide the best contextual information for polarity detection.

Furthermore, two sentences as follows - "iPhoneX is expensive but nice" and "iPhoneX is nice but expensive" would look similar from a bag-of-words perspective. However, they bear exactly opposite polarity because "the former is positive as the user seems to be willing to make the effort to buy the product despite its high price, the latter is negative as the user complains about the price of iPhoneX although he/she likes it" [4], [8]. Thus, interpreting the emotion of a user is extremely crucial – especially in the domain of a political chatterbot. For instance, a keyword-based topic detection will simply fail to interpret the actual intent of the following message:

Human: *Think twice before supporting the #VoteRemain campaign.*

Thus, the next layer, after the subjectivity/objectivity detection, is emotion analyzer. Here, we propose to incorporate sentic computing [4], a commonsense-based framework for concept-level sentiment/emotion analysis, which aims to bridge the gap between statistical NLP and many other disciplines that are necessary for understanding human language, such as linguistics, commonsense reasoning, and affective computing [3], [6], [18]. Sentic computing, whose term derives from the Latin sensus (as in commonsense) and sentire (root of words such as sentiment and sentience), enables the analysis of text not only at document, page or paragraph level, but also at sentence, clause, and concept level [19]. This is possible thanks to an approach to NLP that is both top-down and bottom-up: top-down for the fact that sentic computing leverages symbolic models such as semantic networks and conceptual dependency representations to encode meaning; bottom-up because it uses sub-symbolic methods such as deep neural networks and multiple kernel learning to infer syntactic patterns from data.

Coupling symbolic and sub-symbolic AI is key for stepping forward in the path from NLP to natural language understanding. Relying solely on machine learning, in fact, is simply useful to make a 'good guess' based on past experience, because sub-symbolic methods only encode correlation and their decision-making process is merely probabilistic. Natural language understanding, however, requires much more than that.

To use Noam Chomsky's words, 'you do not get discoveries in the sciences by taking huge amounts of data, throwing them into a computer and doing statistical analysis of them: that's not the way you understand things, you have to have theoretical insights'.

Following prior studies [25], if our response handler fails to identify the intent, due to lack of relevant data, then our chatterbot will create "its own meta-dialogue with the user" to identify the intent. For instance,

Human: *What is the prediction about Brexit?*

Machine: *It is difficult to predict. What do you think?*

Human: *UK is not with David Cameron.*

Machine: *So, you are supporting VoteLeave.*

Human: *What other option do we have?*

After this brief meta-dialog, the chatterbot will be able to identify the political intent of the user and then it can efficiently continue the chat. However, if our chatterbot repeatedly fails to tag input messages with a relevant topic/intent, even after a brief meta-dialog, then we need manual intervention. Generating the set of predefined topic/intents will be an iterative process in a political domain. In other words, we need to regularly update and adapt our set of topics/intents in a supervised manner. This supervised manual intervention is essential for a political chatterbot where the trending topics/hashtags (or emotions of users about a particular topic) are transient. On the contrary, the intents of an IT helpdesk chatterbot will be relatively stable.

Moreover, if we notice that an input message is associated with a new topic (which is not present in our predefined set of topics) then we immediately need to incorporate that topic into our tweet extraction algorithm to collect tweets on that particular topic. This component of our framework would help us to extract tweet feeds from the Twitter platform using search API, and subsequently, we would classify these newly extracted tweets on the basis of topic and intent before storing them in the local tweet database. The richness and contemporaneous of our training data heavily depends on this Tweet Extraction module. Similarly, the Local Tweet Database is the main backbone of our response analyzer for generating a response. In brief, our proposed framework would consider the emotion of a user using emotion analyzer module and reciprocate with appropriate intent while generating the response.

V. NEUTRALITY OF POLITICAL CHATTERBOT

As discussed above in a political chatterbot, it is important to recognize the correct intention of the human agent for generating the most appropriate response. In other words, if the human agents ask a question $\mathbf{Q_i}$ then our neural model will tag this question $\mathbf{Q_i}$ with an intent $\mathbf{I_i}$. Intent $\mathbf{I_i}$ can be associated with a set of appropriate response $\{ \mathbf{r_1}, \mathbf{r_2}, \ldots, \mathbf{r_n} \}$. Thus, in response to $\mathbf{Q_i}$, our chatterbot_will select an appropriate response from the above set. The efficiency of our model broadly depends on proper tagging of $\mathbf{Q_i}$ with $\mathbf{I_i}$.

As we noted earlier, our proposed chatterbot model in the political domain is different from other chatterbots such as IT helpdesk, tourism and travel, restaurants and so on. For instance, [9] designed a chatterbot to respond 'queries about a specific restaurant.' They have considered relevant segments of user reviews from Zomato and Yelp to address queries related to service or atmosphere of a restaurant. There might be differences in opinions among users, but broadly the reviews will converge. However, in a political domain, there can be a plurality of ideologies and convergence might not occur at all. In other words, one particular political view, as well as the counter view of the same, might be equally relevant for different strata of the social media users. For instance,

Human: *Why we will vote for united EU?*

Machine: *#VoteRemain today and secure a stronger future for our country. We're stronger, safer and better off in Europe*

Here, the chatterbot response is biased in favor of *#VoteRemain*. Consequently, the creditworthiness of this chatterbot, from the perspective of *#VoteLeave* supporter, would get severely compromised. Ideally, a political chatterbot should not have its ideology. Hence, our chatterbot response should satisfy the need/expectation of a user (by answering the question - *Why we will vote for united EU?*), and at the same time, our chatterbot should also maintain its neutrality about the topic. Our proposed model attempts to maintain its neutrality by associating a response $\mathbf{r_i}$ with a user $\mathbf{u_i}$ as follows:

Human: *Why we will vote for united EU?*

Machine: <u>*@David_Cameron* says</u> "*#VoteRemain today and secure a stronger future for our country. We're stronger, safer and better off in Europe*"

In other words, our political conversation is as follows: a question $\boldsymbol{Q_i}$ related to intent (which is *VoteReamin* in the above example) $\boldsymbol{I_i}$ generates a response $\boldsymbol{(u_i + r_i)}$. Associating the response $\boldsymbol{r_i}$ with user $\boldsymbol{u_i}$ allows our chatterbot to maintain its neutrality for the intent/context $\boldsymbol{I_i}$. However, this approach might not be appealing to a politically active human agent. Thus, we can also explore the option of text summarization of all tweets by David Cameron (who supported the VoteRemain campaign) or summarization of a few appropriate tweets which supports the VoteRemain ideology as follows:

Human: *Why we will vote for united EU?*

Machine: <u>*VoteRemain supporters are saying*</u> *that united EU will secure a stronger future for our country.*

The follow-up question is how to select a set of appropriate response/tweets for the above text summarization. Following prior studies in social media analytics, we can consider retweet counts which are a crude proxy for the appropriateness or acceptance among other social media users. However, it is also worth noting that political discussions are transient.

In other words, trending topics keep on changing in a vibrant social media platform such as Twitter. Thus, appropriateness of a response not only depends on acceptance by other social media users but also depends on time dimension. Hence, in our proposed model appropriate responses can not be static. Thus, our local tweet database will regularly update the tweet corpus to maintain the time relevance of our response generator.

VI. Conclusion

This paper has explored various issues and challenges of a political chatterbot. Unlike commercial chatterbots for IT helpdesk or restaurant services, we noted that the interpretation of the intent of human agents is extremely crucial in political domain because human agents might differ in their political ideology. We demonstrate, by using various anecdotal evidence from the Brexit related deliberations, that the initial layers of a political chatterbot need to employ both subjectivity and emotion analyzer in addition to the keyword based topic detector.

Next, we also demonstrate the need to maintain neutrality for a political chatterbot. Our Twitter data from Brexit discussion portrays that input messages from politically active human agents might be highly opinionated. However, the chatterbot should maintain its neutrality while generating a response. Thus, we propose to associate the response with a user or a set of users. In other words, our chatterbot should disassociate itself from a politically opinionated response. We have also made a humble attempt to develop the overall chatterbot architecture in this paper.

References

[1] Biyani, P., Bhatia, S., Caragea, C., Mitra, P. (2014) Using non-lexical features for identifying factual and opinionative threads in online forums, *Knowledge-Based Systems* 69, pp. 170–178.

[2] Cambria, E. (2016). Affective computing and sentiment analysis. *IEEE Intelligent Systems*, 31(2), 102-107.

[3] Cambria, E. Hussain, A., Havasi, C., Eckl, C. (2010) SenticSpace: Visualizing opinions and sentiments in a multi-dimensional vector space. In: LNAI, vol. 6279, pp. 385–393, Springer

[4] Cambria, E., & Hussain, A. (2015). *Sentic computing: a common-sense-based framework for concept-level sentiment analysis*. Springer.

[5] Cambria, E., Livingstone, A., & Hussain, A. (2012). The hourglass of emotions. *Cognitive behavioural systems*, 144-157.

[6] Cambria, E., Mazzocco, T., Hussain, A., Eckl, C. (2011) Sentic medoids: Organizing affective common sense knowledge in a multi-dimensional vector space. In: LNCS, vol. 6677, pp. 601-610, Springer

[7] Cambria, E., Poria, S., Bajpai, R., & Schuller, B. (2016). SenticNet 4: A Semantic Resource for Sentiment Analysis Based on Conceptual Primitives. In *Proceedings of COLING*, pp. 2666-2677.

[8] Chaturvedi, I., Ragusa, E., Gastaldo, P., Zunino, R., & Cambria, E. (2017). Bayesian network based extreme learning machine for subjectivity detection. *Journal of The Franklin Institute* doi.org/10.1016/j.jfranklin.2017.06.007

[9] Ghazvininejad, M., Brockett, C., Chang, M. W., Dolan, B., Gao, J., Yih, W. T., & Galley, M. (2017). A Knowledge-Grounded Neural Conversation Model. *arXiv preprint* arXiv:1702.01932.

[10] Grčar, M., Cherepnalkoski, D., Mozetič, I., & Novak, P. K. (2017). Stance and influence of Twitter users regarding the Brexit referendum. *Computational Social Networks*, 4(1).

[11] Howard, N., Cambria, E. (2013) Intention awareness: Improving upon situation awareness in human-centric environments. *Human-centric Computing and Information Sciences* 3(9)

[12] https://qz.com/653084/microsofts-disastrous-tay-experiment-shows-the-hidden-dangers-of-ai/ Accessed on September 5, 2017.

[13] Khatua, A., & Khatua, A. (2016). Leave or Remain? Deciphering Brexit Deliberations on Twitter. In 2016 *IEEE Procedings of 16th International Conference on Data Mining Workshops* (ICDMW), pp. 428-433.

[14] Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is Twitter, a social network or a news media?. In *Proceedings of the 19th international conference on World wide web* (pp. 591-600). ACM.

[15] Kyusong Lee, Tiancheng Zhao, Yulun Du, Edward Cai, Allen Lu, Eli Pincus, David Traum,Stefan Ultes, Lina M. Rojas-Barahona, Milica Gasic, Steve Young and Maxine Eskenazi DialPort, Gone Live: An Update After A Year of Development, Proceedings of the SIGDIAL 2017 Conference.

[16] Lowe, R. T., Pow, N., Serban, I. V., Charlin, L., Liu, C. W., & Pineau, J. (2017). Training end-to-end dialogue systems with the ubuntu dialogue corpus. *Dialogue & Discourse*, 8(1), 31-65.

[17] Lowe, R., Pow, N., Serban, I., & Pineau, J. (2015). The ubuntu dialogue corpus: A large dataset for research in unstructured multi-turn dialogue systems. *arXiv preprint* arXiv:1506.08909.

[18] Poria, S., Gelbukh, A., Agarwal, B., Cambria, E., Howard, N. (2013) Common sense knowledge based personality recognition from text Advances in Soft Computing and Its Applications, pp. 484-496

[19] Poria, S., Gelbukh, A., Cambria, E., Yang, P., Hussain, A., Durrani, T. (2012) Merging SenticNet and WordNet-Affect emotion lists for sentiment analysis. Proceedings of ICSP, pp. 1251-1255

[20] Ritter, A., Cherry, C., & Dolan, B. (2010). Unsupervised modeling of twitter conversations. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 172-180).

[21] Ritter, A., Cherry, C., & Dolan, W. B. (2011). Data-driven response generation in social media. In *Proceedings of the conference on empirical methods in natural language processing* (pp. 583-593). Association for Computational Linguistics.

[22] Serban, I. V., Lowe, R., Charlin, L., & Pineau, J. (2015). A survey of available corpora for building data-driven dialogue systems. arXiv preprint arXiv:1512.05742.

[23] Sordoni, A., Galley, M., Auli, M., Brockett, C., Ji, Y., Mitchell, M., Nie, JN, Gao, J & Dolan, B. (2015). A neural network approach to context-sensitive generation of conversational responses. Paper presented at the meeting of the HLT-NAACL, 2015. *arXiv preprint* arXiv:506.06714.

[24] Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. In *Advances in neural information processing systems*, pp. 3104-3112.

[25] Ultes, S., Rojas-Barahona, L., Su, P. H., Vandyke, D., Kim, D., Casanueva, I., Budzianowski, P., Mrkšić, N., Wen, T.H., Gasic, M. & Young, S. (2017). Pydial: A multi-domain statistical dialogue system toolkit. *Proceedings of ACL 2017, System Demonstrations, 73-78.*

[26] Vazquez, S., Munoz-Garcıa, O., Campanella, I., Poch, M., Fisas, B., Bel, N., Andreu, G. (2014) A classification of user-generated content into consumer decision journey stages, *Neural Networks* 58, pp. 68–81.

[27] Vinyals, O., & Le, Q. (2015). A neural conversational model. *arXiv preprint* arXiv:1506.05869.

[28] Xu, A., Liu, Z., Guo, Y., Sinha, V., & Akkiraju, R. (2017). A New Chatbot for Customer Service on Social Media. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 3506-3510). ACM

[29] Zhou, H., Huang, M., Zhang, T., Zhu, X., & Liu, B. (2017). Emotional Chatting Machine: Emotional Conversation Generation with Internal and External Memory. *arXiv preprint* arXiv:1704.01074.