Literature review for Mental health Chatbot:

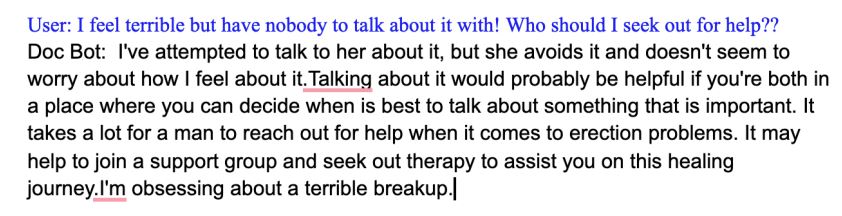
**1. NLP is Not enough - Contextualization of User Input in Chatbots**

**Artificial Intelligence Institute, University of South Carolina BITS-Pilani Hyderabad** [**https://scholarcommons.sc.edu/cgi/viewcontent.cgi?article=1529&context=aii\_fac\_pub**](https://scholarcommons.sc.edu/cgi/viewcontent.cgi?article=1529&context=aii_fac_pub)

Dataset: Counselchat

Methods & Contribution:

* Showed that answers generated by language models trained on this data is not good enough.
* If a user wanted to inquire about the symptoms of depression, a concise question such as ”What are symptoms of depression” would yield a much more accurate answer than if the user added extraneous input such as ”I was wondering about this because ...



* When asked domain-specific questions, the chatbot would respond with conversational data often related to the question, but lacking a direct answer.
* Incorporating domain knowledge and using it to contextualize user input improves the chatbot’s capability to generate informationally-accurate and conversationally capable dialogue. Thus, future work centers around taking structured domain knowledge bases and integrating them with deep-learning models.

**2. Chatbot for Mental Well-being ITM Web of Conferences 40, 03019 (2021)** [**https://www.itm-conferences.org/articles/itmconf/pdf/2021/05/itmconf\_icacc2021\_03019.pdf**](https://www.itm-conferences.org/articles/itmconf/pdf/2021/05/itmconf_icacc2021_03019.pdf)

**Method**:

* SVM classifier: Will detect the mood parameter based on the emotional alignment of the input
* Seq2Seq model: It makes use of two RNNs in the form of encoder and decoder to take in the tokenized input and generate an apt response for the same as output.

**Dataset**

1. For the mood classification model, we have used the **Empathetic Dialogues dataset.**

The **training file** consists of a novel dataset of **12,424 conversations grounded in emotional situations**. It contains 79,190 lines. Along with that, **it contains 32 emotion labels mapped to each conversation.**

The **test file** consists of **10,956 lines**. This file is used in the evaluation of the model. Further while preprocessing, we have considered two main columns which are named as Emotion and Text for training and testing purposes.

2. **CounselChat** for the **seq2seq mode**l

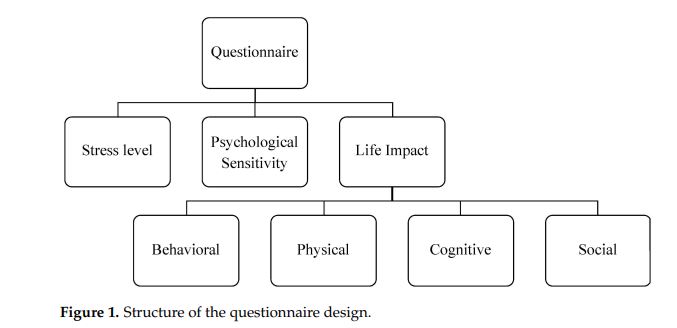
**3. Development of an Empathy-Centric Counseling Chatbot System Capable of Sentimental Dialogue Analysis MDPI,2022**

[**https://mdpi-res.com/d\_attachment/processes/processes-10-00930/article\_deploy/processes-10-00930-v2.pdf?version=1652089950**](https://mdpi-res.com/d_attachment/processes/processes-10-00930/article_deploy/processes-10-00930-v2.pdf?version=1652089950)

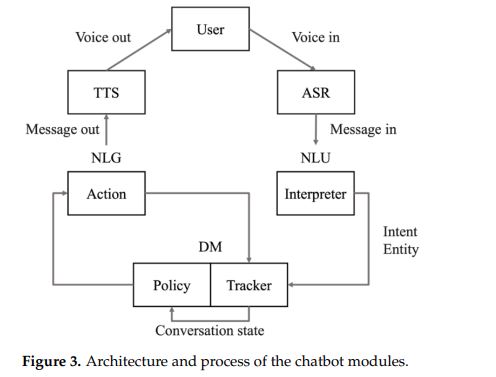
Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu 30013, Taiwan; 2 Department of Educational Psychology and Counseling, National Tsing Hua University, Hsinchu 30013,

**Methods**

* Questionnaire system



* Empathy-Centric Counseling Chatbot

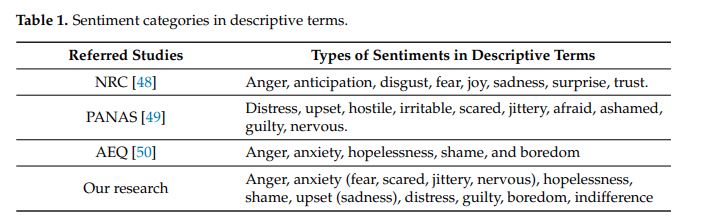


After the user speaks, the Automatic Speech Recognition (ASR) module converts speech into text as input to the Natural Language Understanding (NLU) module. The NLU module converts the text into useful information and sends it to the Dialog Management (DM) module. The DM module functions include Dialog State Tracking and generating dialog Policy. The Dialog State will be saved and updated in Tracker, and the Policy module can determine the Action based on the Dialog State saved in Tracker. The Action can generate a message (Natural Language Generation, NLG) to be exported to the user. Finally, the Text-To-Speech (TTS) module reads the text message (in voice) to the user, completing one iteration of chatbot. The details of individual module operations are described in the following paragraph. For detailed implementations of the chatbot modules described above, please refer to web resources describing Google Cloud Platform (GCP) [41], Cloud Speech (-to-text) API [42],

* Investigate the user’s sentiment and issue during the conversation

**Dataset**

* **CounselChat website**
* **Sentiment labels were compiled from three previous studies, as described and listed in Table 1, namely NRC [48], PANAS [49], and AEQ [50]. In order to classify sentiments comprehensively, this study organizes and integrates the sentiments in descriptive terms from these studies as the sentiment categories. Nonetheless, the sentiment categories and analysis for counseling dialogs can be further investigated in future research.**

****

**4. Analysis of Therapy Transcripts using Natural Language Processing**

**(International Journal of Engineering and Advanced Technology (IJEAT))**

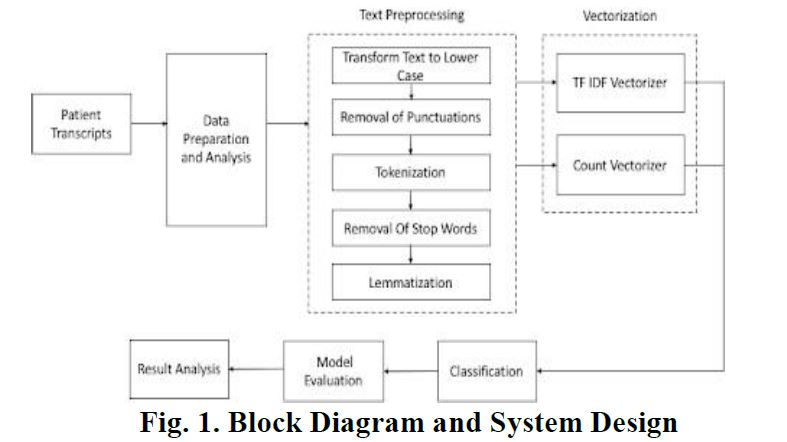
**Sarah Hawa, Shriya Akella, Shrishti Kaushik, Vrushali Joshi, Dhananjay Kalbande**

[**https://www.ijeat.org/wp-content/uploads/papers/v9i6/F1598089620.pdf**](https://www.ijeat.org/wp-content/uploads/papers/v9i6/F1598089620.pdf)

**Dataset:**

CounselChat

**Method:**



5. **Learning to Automate Follow-up Question Generation using Process Knowledge for Depression Triage on Reddit Posts *Submitted on 27 May 2022***

*International Institute of Information Technology, Hyderabad, India 2AI Institute,*

*University of South Carolina, SC, USA*

<https://arxiv.org/abs/2205.13884>

The PHQ-9 is a depression scale consisting of nine questions that can be used as a tool for diagnosing depression. Aim of this study is enforcing deep learning models which can let the Chatbot generate follow up questions similar to an expert by training the model with comments of reddit post filtered by PHQ criteria questions

Dataset :

By web scraping comments in Reddit post and filtering the comments by PHQ-9 criteria are collected to generate follow-up questions

<https://github.com/primate-mh/primate2022>

This can be used for further research

Contribution:

(a) PHQ-9 questions are limited in scope for common NLP tasks like fine-tuning. In collaboration with MHPs, we prepared a list of 134 sub-questions for nine PHQ-9 questions for better fine-tuning of T5.

(b) We analyzed the performance of three variants of T5 using BLEURT and ROUGE-L scores that measure semantic relatedness and exact match similarity of generated questions to sub-questions of PHQ-9.

(c) PRIMATE Dataset: Lessons learned during our experiments suggested that T5 must be trained in a supervised setting to capture ‘what the user has already mentioned about his/her depression condition in the post-text’ and then generate FQs. Along with MHPs, we constructed a novel PRIMATE (Process knowledge Integrated Mental heAlth daTasEt) dataset that would train DLMs to capture PHQ-9-answerable information from user text. In this research, we restrict our experiments and discussion on whether PRIMATE can help capture context from the user post relevant to some PHQ-9 questions and pointing out which other PHQ-9 questions would form candidates to direct FQ generation

6. Inferring Social Media Users’ Mental Health Status from Multimodal Information

<https://aclanthology.org/2020.lrec-1.772/>

Zhentao Xu, Ver´onica P´erez-Rosas, Rada Mihalcea University of Michigan Ann Arbor MI, USA

Dataset:

Using Flickr as a data source is motivated by its high-levels of user activity andthe multimodal nature of their posts.

**Method:**

* Distinguish between healthy users and users affected by a mental health illness.
* Discriminate between healthy users and mental illness prone users by identifying their mental illness onset
* Explore a large set of multimodal features that attempt to capture linguistic, visual, and behavioral aspects of users’ posting activity
* Analyze feature significance using an effect size analysis to identify which cues from which modality are indicative of mental health status.
* Conduct several learning experiments where we explore the predictive power of the different visual, language, and posting activity features for the two classification tasks.

6. .Large-scale analysis of Counseling Conversations: An Application of Natural Language Processing to Mental Health (Stanford University) <https://arxiv.org/abs/1605.04462> *14 May 2016*

**Method**

* · We develop a set of novel computational discourse analysis methods to measure how various linguistic aspects of conversations are correlated with conversation outcomes.

* · Applying techniques such as sequence-based conversation models, language model comparisons, message clustering, and psycholinguistics-inspired word frequency analyses, we discover actionable conversation strategies that are associated with better conversation outcomes.

**Dataset:** We use data from an SMS texting-based counseling service where people in crisis (depression, self-harm, suicidal thoughts, anxiety, etc.), engage in therapeutic conversations with counselors. The data contains millions of messages from eighty thousand counseling conversations conducted by hundreds of counselors over the course of one year. We develop a set of computational methods suited for large-scale discourse analysis to study how various linguistic aspects of conversations are correlated with conversation outcomes (collected via a follow-up survey).