

TruthBot: An Automated Conversational Tool for Intent Learning, Curated Information Presenting, and Fake News Alerting

Ankur Gupta^{*1}, Yash Varun^{*1}, Prarthana Das^{*1}, Nithya Muttineni^{*1}, Parth Srivastava^{*1}, Hamim Zafar¹, Swaprava Nath¹, and Tanmoy Chakraborty²

¹Indian Institute of Technology Kanpur, ^{*}equal contribution,
{ankugupt,yashyv,prdas,mnithya,parthsri,hamim,swaprava}@iitk.ac.in
²Indian Institute of Information Technology Delhi, tanmoy@iiitd.ac.in

Abstract

We present TruthBot (name is anonymized), an all-in-one multilingual conversational chatbot designed for seeking truth (trustworthy and verified information) on specific topics. It helps users to obtain information specific to certain topics, fact-check information, and get recent news. The chatbot learns the intent of a query by training a deep neural network from the data of the previous intents and responds appropriately when it classifies the intent in one of the classes above. Each class is implemented as a separate module which uses either its own curated knowledge-base or searches the web to obtain the correct information. The topic of the chatbot is currently set to COVID-19. However, the bot can be easily customized to any topic-specific responses. Our experimental results show that each module performs significantly better than its closest competitor, which is verified both quantitatively and through several user-based surveys in multiple languages. TruthBot has been deployed in June 2020 and is currently running.

1 Introduction

Social media has endowed us with a massive volume of information, which has both useful, not-so-useful, and misleading content (Vraga and Tully, 2019; Shu et al., 2017). This is worrisome since a usual netizen’s social media time is significant. According to eMarketer¹, US adults use social media for an average of 12 hours per day. During the crisis situations like a global pandemic, people desperately look for solutions and end up consuming unverified information from various online resources including social media.

It is observed that the users often trust the messages they see on social media and instant messengers. A study by Bridgman et al. (2020) shows that those who receive most of the news from social media are more likely to believe falsehoods about COVID-19. It is also observed that the social media has taken over the role of a news platform and now ranks just behind television (Kohut et al., 2010) which may mask true information for certain users. To help such users reach to the correct information, technology needs to be developed within the social media and instant messengers itself.

¹www.emarketer.com

On the other hand, popularity of chatbots among the Internet users has increased significantly in the past few years² (GrandViewResearch, 2017). Studies on US-based users have shown that 86% of them prefer chatting with a chatbot than a human agent (Forbes.com, 2019). A study by Brandtzaeg and Følstad (2017) shows that the major motivations for using chatbots are their timely and efficient assistance for information which improves productivity. Hence, the prospect of a chatbot to truth-seek information is bright, and serves as the motivation of this paper.

1.0.1 State-of-the-art and limitations.

The current fact-checking apps require the users to enter the suspicious messages in those apps to fact-check or ask the users to read the current fact-checking articles to find out if their answers lie within the articles. Such apps are not very *useful*, since suspicious messages typically come via social media platforms or instant messengers and the cost of switching apps or read verified messages to find out the truth is quite high. The other kind of solutions that involve forwarding service (e.g., to WhatsApp business accounts) for fact-checking are typically manually verified and responded by a team of journalists. The chatbot approach for fact-checking is very limited as we discuss in Section 2. However, those approaches also do *not* (a) consider a complete *truth-seeking* design that provides holistic information on a topic, rather take care of certain frequently asked questions on a topic (e.g., COVID-19), and (b) handle low-resource languages.

Note that we consider the term *truth-seek* to be more general than *fact-check*. While fact-checking tries to classify a piece of news to be true or not, truth-seeking provides a complete information against a query. For instance, the query can be a general question about a topic or a news article about which a user is only partially aware or an area-based infection statistics of a disease. The scope of truth-seeking, therefore, subsumes fact-checking as we discuss in the following section.

1.1 Proposed Solution: TruthBot.

We introduce **TruthBot**, a truth-seeking chatbot that (a) provides answers to topic-specific frequently asked questions (FAQ), (b) returns articles after searching on the web, (c) responses to custom queries (e.g., area-wise infection statistics for COVID-19), in addition to (d) fact-checking news articles. **TruthBot** uses a deep neural network to identify the intent of the user’s query. However, if the confidence of the identification is not significant, it consults with the user through a conversation to make sure that the right intent is conveyed to the chatbot. If the query topic falls in any of the four classes discussed above, it triggers the corresponding module and retrieves the answer. The modules have dependencies, e.g., if the FAQ module cannot respond to the query, it automatically searches for the query in the fact-check or Google search modules. The details are in Section 3.

1.2 Evaluation.

We evaluate **TruthBot** in three aspects:

- ▷ We measure the *response accuracy* by finding the relevance of the content of the response with the content of the original query.

²<https://g.co/trends/7tkFb>

Query: monkeys are being tortured in laboratories to find COVID-19 vaccine
IFCN chatbot (click https://wa.me/17272912606?text=hi)
Response: (1) Claim rated False: A vaccine for coronavirus was available since 2001. Fact checked by Newsmeter.in on 04/26/2020 This is a dog vaccine wrongly claimed as being a vaccine for COVID-19. (provides the link)
(2) Claim rated False: A WhatsApp forward claims that the COVID-19 vaccine is ready, was announced by U.S. President Donald Trump, and will be launched next Sunday ... Fact checked by Newschecker on 03/23/2020 (provides the link)
TruthBot
Response: (1) (provides link to news which reads) Fact check: This disturbing viral video of animal cruelty is not of COVID-19 vaccine test A disturbing video purportedly showing animal cruelty circulating on social media claims: this is how monkeys are being tortured in laboratories to test the COVID-19 vaccine. Viral video showing animal torture is from last year reportedly from a German laboratory ...
(2) (provides link to news which reads) Ferrets are among the animals researchers are infecting with the Covid-19 virus as they search for a suitable species to use in tests ...

Table 1: An example query and response obtained from two chatbots – IFCN (a baseline) and TruthBot (our method).

- ▷ We conduct a survey on the *satisfaction* of the chatbot regarding the topic it handles (COVID-19 in our example).
- ▷ We also conduct another survey with a different population on the *user interface* about the usefulness, ease of use, credibility, precision, and value of the chatbot.

Table 1 shows a typical example of the responses returned by the *international fact-checking network (IFCN)* chatbot on WhatsApp³ vis-a-vis TruthBot, where the latter is much more precise against the query.

1.3 Our contributions.

The contributions of this paper can be summarized as follows:

- ▷ Unlike other fact-checking chatbots, TruthBot is a truth-seeking bot, which performs four tasks (Section 3.3 to Section 3.6) via a conversational AI engine (Section 3.2) which identifies the intent of a query and guides the user through a piece of curated information or gets the intent clarified from the user.
- ▷ The chatbot identifies the query language and responds in that language (Section 3.7). This is supported for the 108 languages that Google translate is capable of handling (we use the Google translate service for this task), which is a major advantage of using TruthBot, particularly for low-resource languages.

³<https://www.poynter.org/fact-checking/2020/power-in-your-pocket-how-to-use-the-ifcns-new-whatsapp-chatbot/>

- ▷ Experiments (Section 4) show that the response accuracy for all types of queries is much better than that of the IFCN chatbot, which also does a query-based fact-checking and is the closest in its design to our chatbot.
- ▷ **TruthBot** has been operational since June 2020 and, it is currently accepting queries related to COVID-19.

1.4 Reproducibility

We provide all data, code, and survey results separately.

2 Related work

We discuss the related studies briefly in two aspects – fake news detection and task-specific chatbots.

Detecting malicious content on social media such as fake news, misinformation, disinformation, rumors, cyberbullying has been one of the primary research agenda in natural language processing and social computing. Due to the abundance of literature, we would like the readers to point to two recent survey articles by [Zhou and Zafarani \(2018\)](#) and [Bondielli and Marcelloni \(2019\)](#). Existing approaches can be broadly divided into four categories – (i) *knowledge-based*, which verifies whether the given news is consistent with a knowledge base (true knowledge) [Pan et al. \(2018\)](#); (ii) *style-based*, which employs various stylometry analysis modules to check the similarity of the writing style of a given news with existing fake news [Zhou et al. \(2019\)](#); (iii) *propagation-based*, which explores the propagation of a given news on social media to check how it correlates with the usual fake information diffusion [Bian et al. \(2020\)](#); and (iv) *source-based*, which examines the trustworthiness of the source (the account) from which the news has started propagating [Zhou and Zafarani \(2019\)](#).

Another direction of research relevant to the current study deals with the use of chatbots as an assistant for different applications ([Lin et al., 2020](#); [Radziwill and Benton, 2017](#); [Følstad et al., 2018](#)). Though chatbots originated in the 1960s, its use as an AI assistant and widespread acceptance is a recent phenomenon ([Shum et al., 2018](#)). The use of chatbots for fake news detection is quite a recent development. The approach closest to this paper is by the *international fact-checking network (IFCN)* ([IFCN, 2020](#)), that has tied up with WhatsApp to develop a fact-checking chatbot on that platform for people to fact-check news articles. In this paper, we provide case-studies and a comparison of outputs of various queries of our chatbot vis-a-vis the IFCN chatbot.

TruthBot is quite different from all chatbot approaches to true information retrieval, since it is not limited to fact-checking alone, rather provides a complete information on a topic in multiple languages.

3 Architecture of TruthBot

The objective of **TruthBot** is to bring all relevant truth-seeking activities on a topic within the scope of a single chatbot. Users can interact with **TruthBot** to get (a) general information regarding the topic, (b) fact-check potentially fake news, (c) general information on search queries from Google search, and (d) response from other custom information modules. The current setup of **TruthBot** is tuned to provide all such information regarding

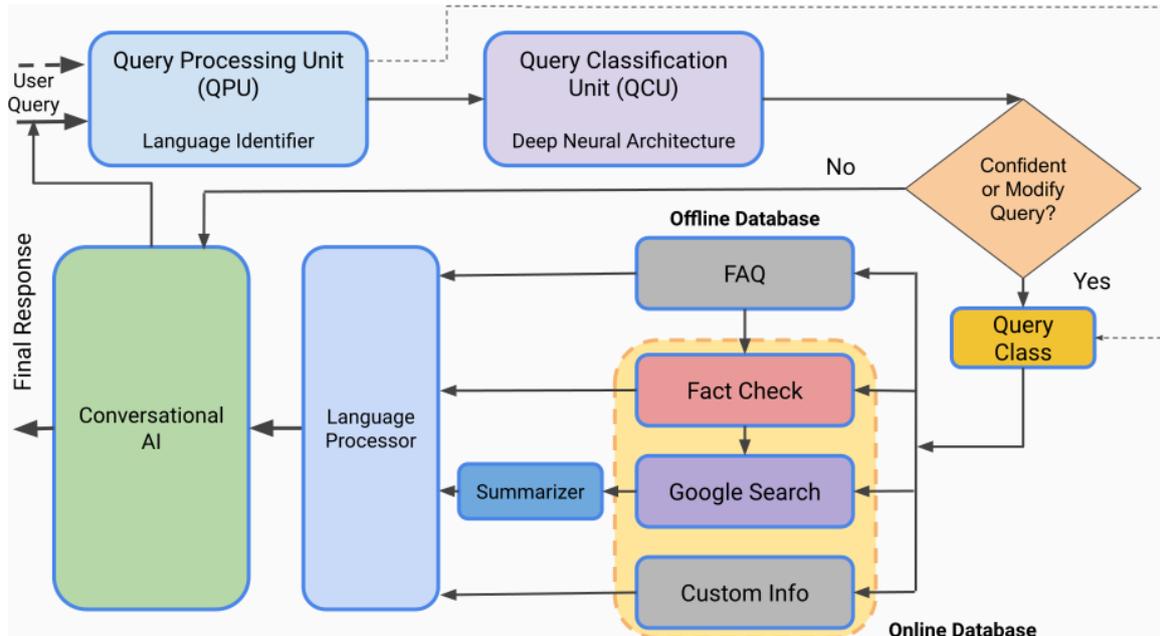


Figure 1: Schematic workflow of TruthBot.

COVID-19 pandemic, with the custom information module being the infection and death statistics of a city, district, state, or country. However, it is perfectly general to be tuned to any other topic with a little customization. We present the basic architecture of TruthBot in this section. Fig. 1 shows the graphical outline of the working units of TruthBot.

3.1 User queries

The beta-version of the bot has been operational since June, 2020, and has attracted 696 unique users (who either have directly used or are informed about the bot till August, 2020) on the three platforms – WhatsApp, Facebook Messenger, and Telegram. For the analyses presented in this paper, we used about 1300 queries which were received till July, 2020.⁴ We manually classified these queries into *six* classes – *frequently asked questions* (FAQ) on a specialized topic (COVID-19 in our case), *fake information verification* (FAKE), *general search for a news* (GEN), *some custom information* (area-wise infection, death statistics–AREASTAT–in our case of COVID-19), *greeting*, and *spam*. Some of the queries, e.g., “*warm saline water gargling can cure coronavirus*”, were classified into multiple classes, as this example can be classified as FAQ, FAKE, and GEN. This manual labeling task took about 30 person-hours since one needed to read every query and use human justification to classify them. This process generated the training set (the JSON file with these classification is available in the supplementary material) which we used to classify new queries.

3.2 Conversational AI approach

TruthBot is designed to help users check the truth of a news or a message. The focus is to provide the least cognitive load to the user and understand the intent of her queries before proceeding on to searching the information. This is important since the trajectory of the

⁴Collectively we got 2270 different queries sent to our bot till September 9, 2020. The performance of the bot will improve as this training set becomes richer.

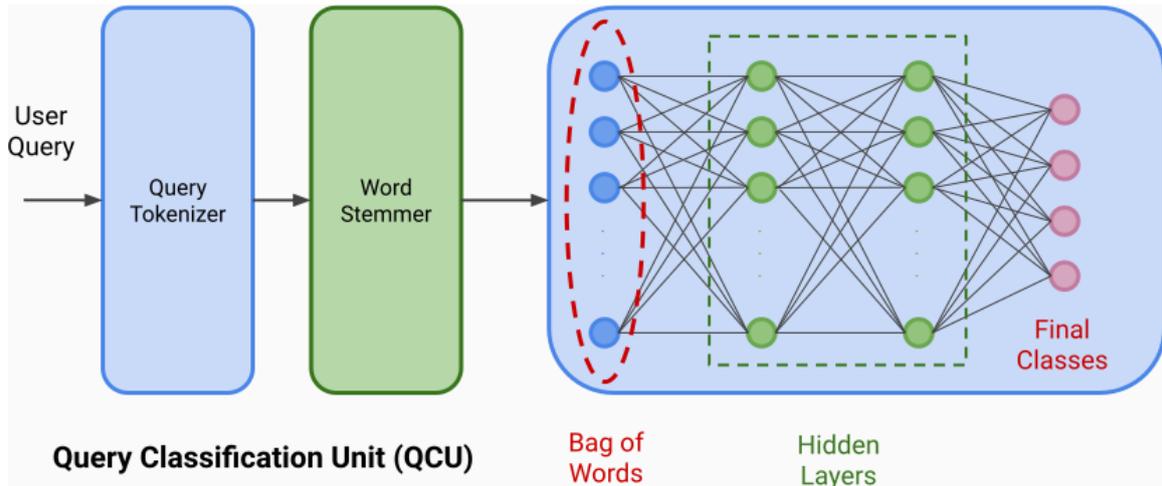


Figure 2: Deep neural network for query classification (number of nodes is representative).

search for the different tasks **TruthBot** is capable of doing is quite different. The query is tokenized and stemmed into root words to create a bag of words. The collection of such words and the query classes are used to train the Query Classification Unit (QCU). This unit classifies the intents of the queries into the six classes discussed in Section 3.1 using a *deep neural network* (DNN). Any new query is soft-classified into these classes using the softmax scores. If the scores are significant—the threshold has been tuned based on manual inspection of the classification—then the action corresponding to that class is invoked (e.g., it calls the module Fact Check if it is classified as a fake information verification query). If the scores of classification are not significant, then it returns to the user and provides a selection choice of the intent to make a focused truth-search of the query. A couple of examples of how the conversation unfolds in this method are provided in Fig. 3. Fig. 2 shows the schematic representation of QCU. The performance metrics of this classification for different classes are shown in Table 2.

Metric (Shu et al., 2017)	FAQ	FAKE	GEN
Precision	0.44	0.22	0.58
Recall	0.61	0.08	0.64
F1 score	0.51	0.12	0.61
Accuracy	0.67	0.602	0.68

Table 2: Performance metrics of the DNN classification.

3.3 Topic-specific frequently asked questions (FAQs)

The purpose of this module is to answer some standard questions users may have on the topic **TruthBot** is tuned to. Since standard answers to these FAQs are already available, it is possible to create a knowledge-base from which these questions can be directly answered and reduce the latency of the responses. In the following subsection, we discuss the case study of the FAQs on COVID-19 that this bot is currently tuned to.

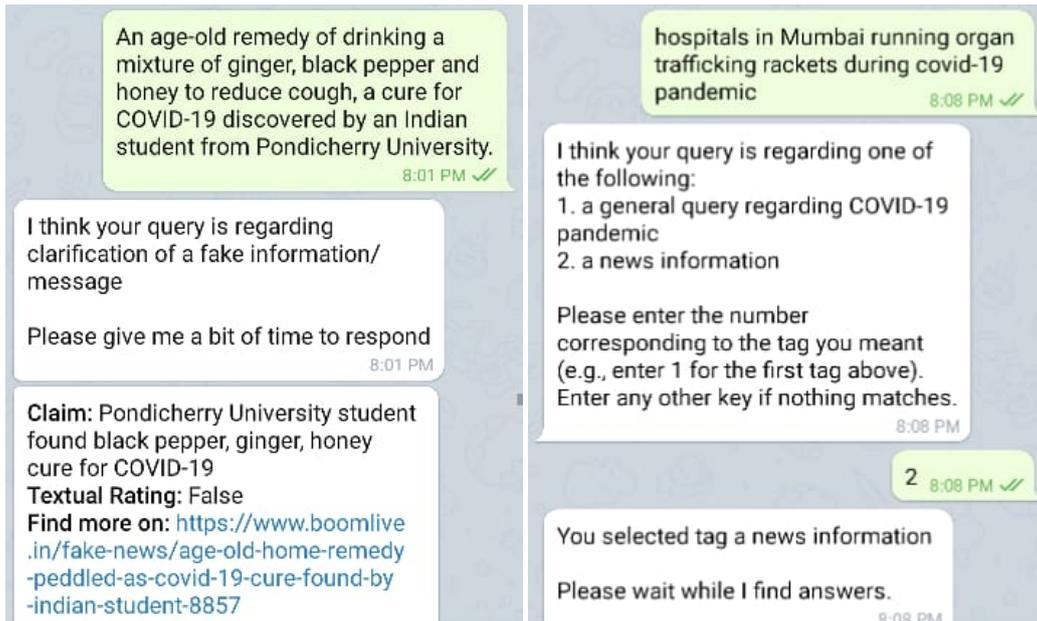


Figure 3: Examples of the conversation with the user when TruthBot is confident (left) and not completely confident (right).

3.3.1 Case study: COVID-19.

For COVID-19, websites like WHO (<https://covid19.who.int/>), CDC (<https://www.cdc.gov/>), several government health departments (e.g., the Indian Ministry of Health and Family Welfare (<https://www.mohfw.gov.in/>)) provide a large collection of FAQs and myth-busters. We first create a knowledge-base by *web-scraping* those articles from these websites. Fig. 4 shows the flowchart of this module.

The Text Scraping Unit (TSU) consists of several Python libraries. The TSU downloads the HTML content of the target webpage (e.g., the FAQ page of CDC⁵) by sending an HTTP request to the associated URL and uses it to construct a parse tree using `html.parser`.

⁵<https://www.cdc.gov/coronavirus/2019-ncov/faq.html>

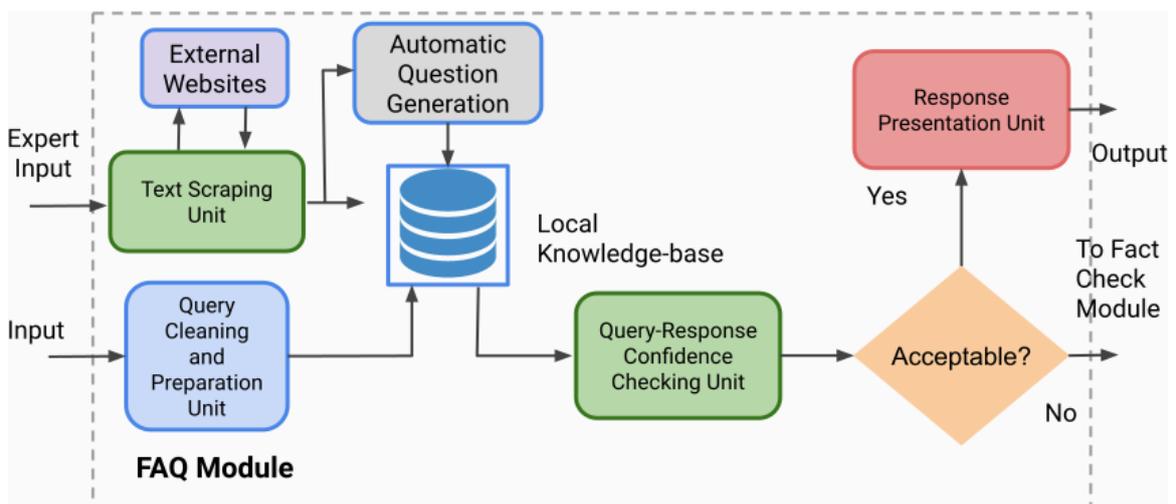


Figure 4: Architecture of the FAQ layer.

The useful FAQs and the paragraphs answering each FAQ are extracted from the parse tree and stored in the local knowledge-base.

To augment the local knowledge-base, we also employ automatic question generation from the answer texts. This is done such that a slightly rephrased question gets a better similarity with one of the questions in our knowledge-base. The answer, however, remains same for all the questions generated from that text. The most straightforward means for this task is answer-aware question generation Lopez et al. (2020). In answer-aware question generation, the model is presented with the answer and a passage to generate a question for that answer considering the passage as the context. As the answer-aware models require answers for generating questions, we first split the passage into sentences and select a subset based on named-entity recognition and noun-phrase extraction. We have used a text-to-text transformer model Raffel et al. (2019) for question generation. The model is fine-tuned in a multitask way using task prefixes with the help of SQuADv1 dataset Rajpurkar et al. (2016) along with pre-trained transformers (specifically seq-2-seq models) on paragraphs that provide answers to COVID-19 related FAQs extracted from online databases.

When a user sends the query, we need to find the appropriate question that matches this query. We use BERT Devlin et al. (2018) fine-tuned on COVID-19 dataset⁶ to generate the contextual embedding for each *question* sentence in the corpus. Then, we use cosine similarity to match a query with one of the questions in the local knowledge-base. A cosine similarity score of 0.85 is used as an empirically determined threshold for the correctness of the matchings. If a query scores above this threshold, we return the answer corresponding to the matched question in the corpus. Otherwise, the query is transferred to Fact Check module.

3.4 Custom information query

Several topic-specific custom modules can be added to TruthBot. The purpose of these modules varies with the kind of topic we tune the chatbot to. As an example, we explain one such custom information module for COVID-19 below.

3.4.1 Case study: COVID-19.

Currently, TruthBot is running a custom information module, called AREASTAT. The goal of this module is to provide the user with the infection, death, recovery statistics of a city, state, or country. Currently, the bot responds to the finer AREASTAT queries on Indian states and cities. For other countries, it provides country-level aggregate information. However, this can be easily extended with an appropriate database providing the finer information for the cities and states of other countries.

After the intent of the query is classified as AREASTAT, it queries the databases with the name of the city, state, or country extracted from the query. The following workflow is repeated at most twice until the result is found.

- ▷ The name of the place is extracted from the query.
- ▷ The place is looked up in three databases – state wise, district wise, and country wise in order.

⁶<https://allenai.org/data/cord-19>

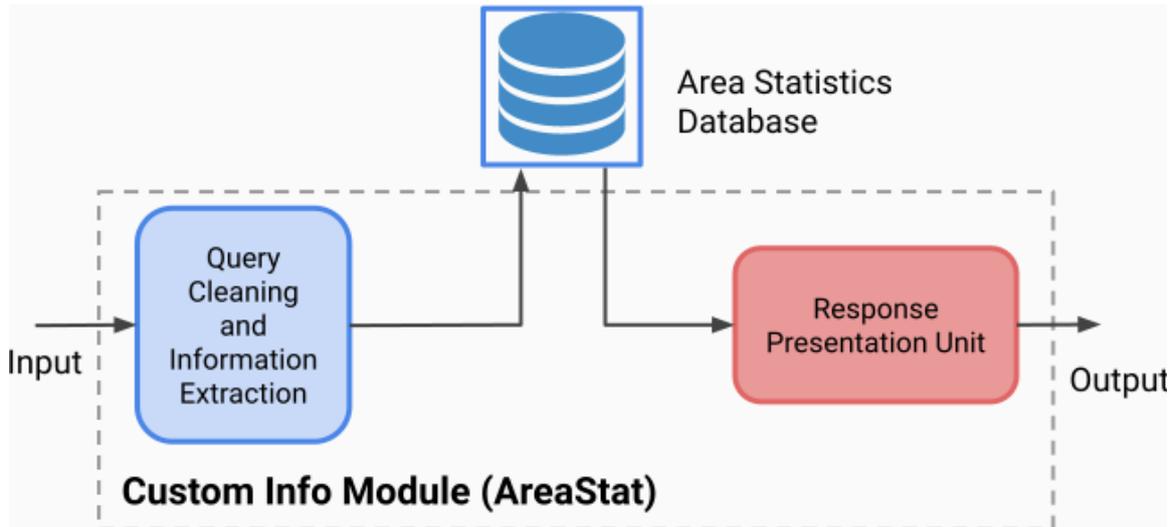


Figure 5: Custom information query (AREASTAT) layer.

- ▷ If a match is found in the database, the results of the number of COVID-19 cases, partitioned into *confirmed*, *recovered*, *active*, and *deceased* are sent to the user.
- ▷ If no match is found in the database, the user is sent a message asking to enter only the name of the place to ensure that there is no spelling mistake.

3.4.2 Verifying if the query is about AREASTAT.

The AREASTAT queries follow a pattern, e.g., “what are the number of cases in ⟨place name⟩”, “number of COVID-19 cases in ⟨place name⟩”, “how many people are infected with covid in ⟨place name⟩” etc. The query is first processed for the common phrases such as “number of cases” or “how many” followed by the place name.

However, since the database for area-wise statistics available to us is only for COVID-19, we need to distinguish similar queries that ask the statistics for a different topic, e.g., the death in a cyclone. We first identify if the query has any COVID-19 related keywords along with the phrases discussed above. Once the query is confirmed to be a COVID-19 statistics query, we detect the place name using the usual patterns of such AREASTAT questions.

There are two ways in which the statistics query may not be responded to: (a) if the statistics asked is not related to COVID-19, or (b) if the place name is not available in the database. In the former case, the query is transferred to the GEN module for a Google search of the query. In the latter, the user is responded with a message asking to re-enter the place correctly.

Statistics databases. The covid19india API⁷ crowdsources the latest number of COVID-19 cases from various sources pertaining to each state and city of India. For the countries other than India, the data is scraped from the Worldometers website⁸ at the time when the query is received.

⁷<https://api.covid19india.org/>

⁸<https://www.worldometers.info/coronavirus/>

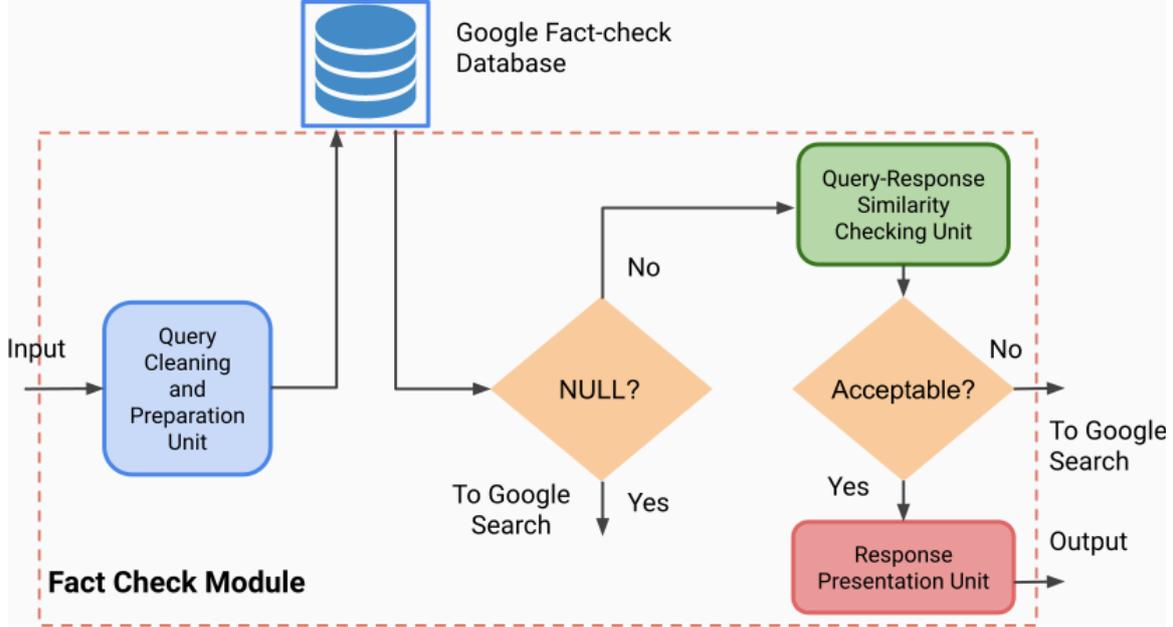


Figure 6: Architecture of Fact Check module.

3.5 Fake news alerting

Several responsible news agencies bust fake news and post such analysis on their dedicated fact-check section of their websites. If a query is classified as FAKE, it is processed through the Fact Check module that uses ‘Google Fact Check Claim Search API’ to search for the already fact-checked claims by sending an HTTP request. A typical url for calling the API consists of the query, language code, page size and an API key. An API call can return with a nested dictionary (denoted by FCR) of multiple objects corresponding to the number of fact-checked claims available on different fact-checker websites (e.g., Alt News⁹). For each of the retrieved results, we store the TEXT, TEXTUAL RATING, and URL attributes. For each result FCR_i , we assign a relevance score rs_i that computes the semantically matching similarity (based on cosine similarity) between the contextual embedding of the TEXT attribute of FCR_i and the contextual embedding of the query. The results in FCR sorted based on their relevance scores and the top- k ($k \leq 3$) results are displayed to the user if their relevance scores exceed a predefined (empirically determined) cutoff. If the query language is a language other than English, the results are displayed after translating to the original language using the Language Processor module. If the API call does not return any result or the relevance scores of all the returned results are below the cutoff, the query is forwarded to the google-search module. A regional language query is forwarded to the Google search module in its original language. The flow-chart for the Fact Check module is shown in Fig. 6.

3.6 Google search articles

When a query is classified as GEN or the FAQ and Fact Check module does not produce any response to a query in FAQ or FAKE category respectively, the query is forwarded to the Google Search module that retrieves results by performing Google search using the google

⁹<https://www.altnews.in/>

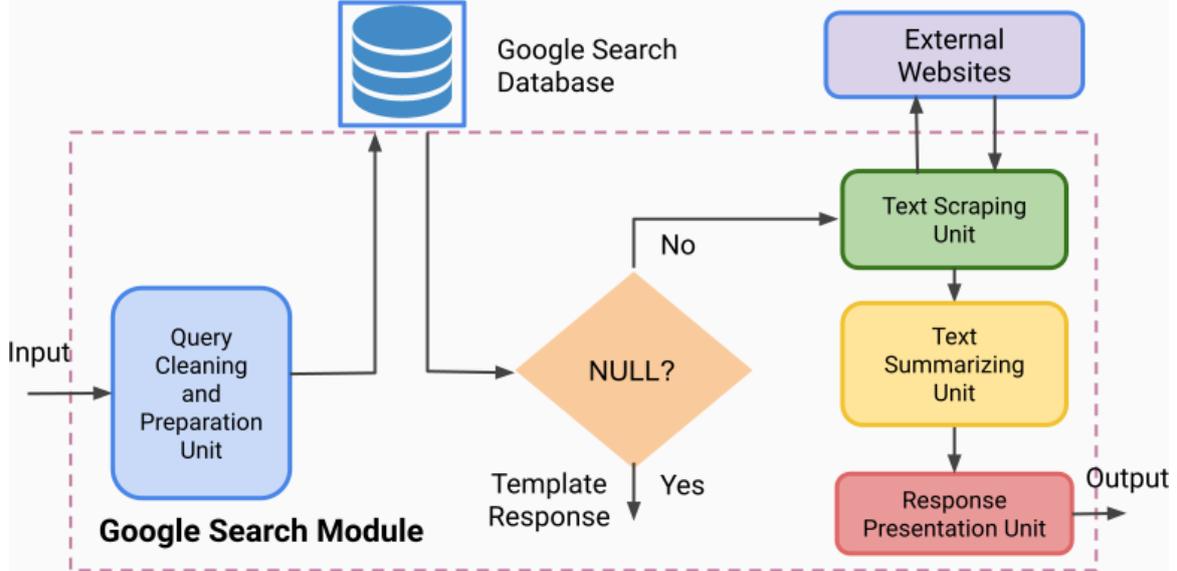


Figure 7: Consolidated architecture of Google search and summarizer layers.

search library. The query is fed to the `search()` method of Google search library and the top two urls (already sorted based on their relevance to the query) are retrieved. For each of the two URLs obtained in the previous step, we scrape the webpage so that we can perform summarization using that scraped data. After scraping the url, the content is summarized using TextRank Mihalcea and Tarau (2004), a graph-based summarization algorithm. To compute the edge weights of the graph used in TextRank, we use three different similarity metrics

1. Cosine similarity between sentences using TF-IDF model,
2. Longest common substring between any two sentences,
3. BM25 / Okapi-BM25 (Robertson and Zaragoza, 2009), a bag-of-words ranking function used for information retrieval tasks.

The obtained summary and the url are displayed to the user. If the original query is in a low-resource language and the scraped data of the webpage is in the same language, no translation is required. However, if the obtained webpages are in English, then the language processor (LP) module (see Fig. 1) translates the summary into the language of the original query before displaying it to the user.

3.7 Extension to low-resource languages

One distinguishing factor of **TruthBot** is that it is capable of responding in multiple languages. This is particularly helpful for the low-resource language users since the social media and instant messengers have significant user-base in these languages but technologies for various applications (including true information finding applications) are not well-developed for them. We use the language detection and translation service of Google translate¹⁰ in the QPU and language processor modules in Fig. 1. In QPU, if the detected language is different from English, then that information is stored, the query is translated into English

¹⁰<https://translate.google.com>



Figure 8: Response of IFCN chatbot (left) and TruthBot (right) in Hindi.

and sent to the later modules. At the language processor module, the response of the bot is translated back into the original language of the query and passed through the conversational AI module to be displayed on user's screen. Fig. 8 shows the responses of a typical query in *Hindi* in IFCN chatbot and in TruthBot.

4 Experimental Results

The experiments are performed to validate the efficacy of TruthBot with its closest competitor, the IFCN chatbot. The comparison considers three metrics. First, we perform a quantitative comparison on the response accuracy, which compares the textual similarity of the query and the response provided by the two chatbots. Second, we run a user survey to capture their qualitative satisfaction with the responses of the two chatbots. Finally, we conduct a user survey about the user interface and the presentation of the curated information in these two chatbots.

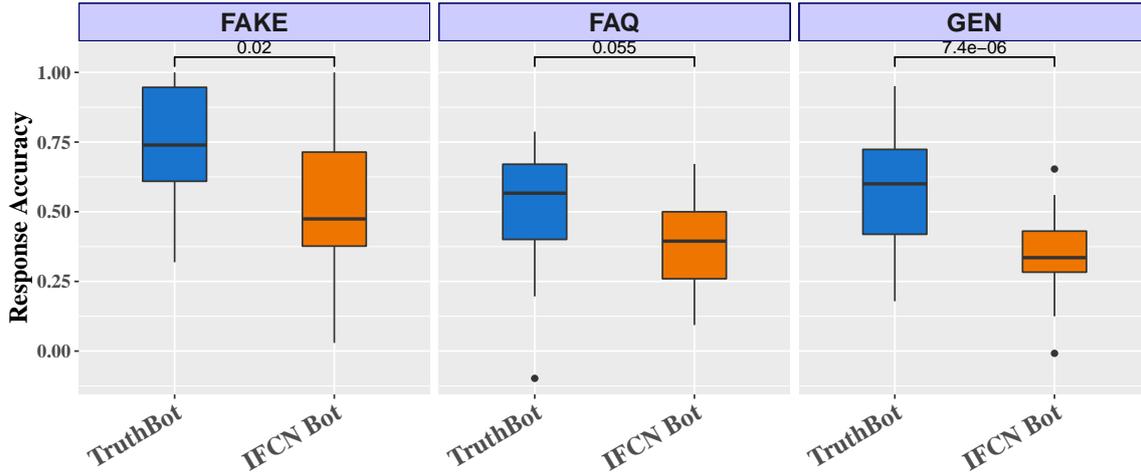


Figure 9: Comparison of Response Accuracy for FAKE, FAQ, and GEN queries. The numbers on the square brackets on the top of the plots denote the p -values of the t -test for statistical significance.

4.1 Evaluation dataset

To evaluate **TruthBot**, we curated a dataset consisting of 79 queries. The queries belonged to four categories - FAQ, FAKE, GEN, and AREASTAT. The FAQ queries were manually curated from WHO and CDC which provide general information on COVID-19. For FAKE and GEN queries, we collected most recent (as of August 10, 2020) fake information and true news articles on COVID-19 from different media outlets (e.g., <https://www.indiatoday.in/>). Single sentence queries that conveyed the main message were curated from the collected articles. We curated the AREASTAT queries manually. Supplementary materials accompany the list of queries and the responses of the two chatbots. In the quantitative comparison of the response accuracy, we used the first three categories, since the textual similarity between AREASTAT queries and responses is not informative about accuracy.

4.2 Evaluating response accuracy

First, we wanted to evaluate the accuracy of **TruthBot** in retrieving the correct article available in the online or offline databases in response to the specific query. For this, we designed a metric called ‘response accuracy’, which measures the similarity of the retrieved response against the actual query. We used cosine similarity between the main text of the retrieved response and the actual query for computing response accuracy. Results of **TruthBot** were compared against that of the chatbot by IFCN. For IFCN chatbot, the sentence following the phrase “Claim rated false” or “Claim rated misleading” was selected as the main text of the response. For **TruthBot**, for FAKE queries, the sentence followed by the phrase “Claim” was selected as the main text, whereas for FAQ and GEN queries, the complete response was compared against that of the actual query. Fig. 9 compares the response accuracy of **TruthBot** against that of IFCN chatbot for different classes of queries. For FAKE queries, average response accuracy of **TruthBot** was 0.74 as compared to 0.56 for IFCN chatbot. Since IFCN is only meant for fake news alerting, the response

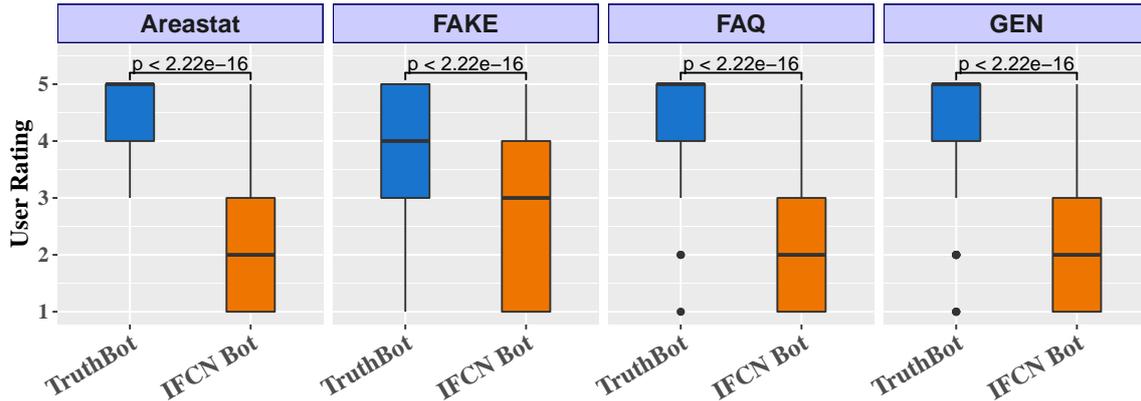


Figure 10: Comparison of user satisfaction while using TruthBot and IFCN chatbot based for AREASTAT, FAKE, FAQ and GEN queries. The results are based on a user survey. The numbers on the square brackets on the top of the plots denote the p -values of the t -test for statistical significance.

that it produces for FAQ and GEN queries may not be meaningful. However, since the response accuracy measures the textual similarity between the response and the query, it also indicates whether relevant articles were retrieved by the chatbot. For FAQ and GEN queries, TruthBot’s average response accuracy was 35.64% and 68.2% better than that of IFCN chatbot.

4.3 Chatbot satisfaction survey

We conducted a human-centric evaluation to determine how satisfied a user would be with the responses received from TruthBot as compared to IFCN chatbot. The subjects were chosen through an open call sent via email and 64 users participated in this survey. The age-group of the participants were: 10-30 years (80.3%), 31-50 years (16.3%), and above 50 years (3.4%). Among them, 20.2% were female and 79.2% were male participants. Each user was assigned 10 queries (2 AREASTAT, 3 FAKE, 2 FAQ and 3 GEN) and the corresponding responses from TruthBot and IFCN chatbot. The queries were randomly chosen from the set of queries described in Section 4.1 so that we obtain multiple ratings (by multiple users) for a single query. Each user was instructed to rate the response of each bot based on how satisfied (1 for least and 5 for most) they were with the response obtained from a bot for each of their assigned queries. The result of the user survey is shown in Fig. 10.

For each of the query classes, TruthBot achieved much better user rating as compared to IFCN chatbot. For AREASTAT, FAQ and GEN, median user rating of TruthBot was 5 as compared to the median rating of 2 for IFCN chatbot. Even for FAKE queries (for which IFCN chatbot is specialized), TruthBot achieved better rating (median rating 4 as compared to 3 for IFCN chatbot).

4.4 User interface study

Finally, we also compared TruthBot against the IFCN bot based on the feedback from the users regarding different aspects of the user interface - Usefulness, Ease of Use, Credibility, Precision, and Value (interaction.design.org, 2020). We obtained responses from 51 users

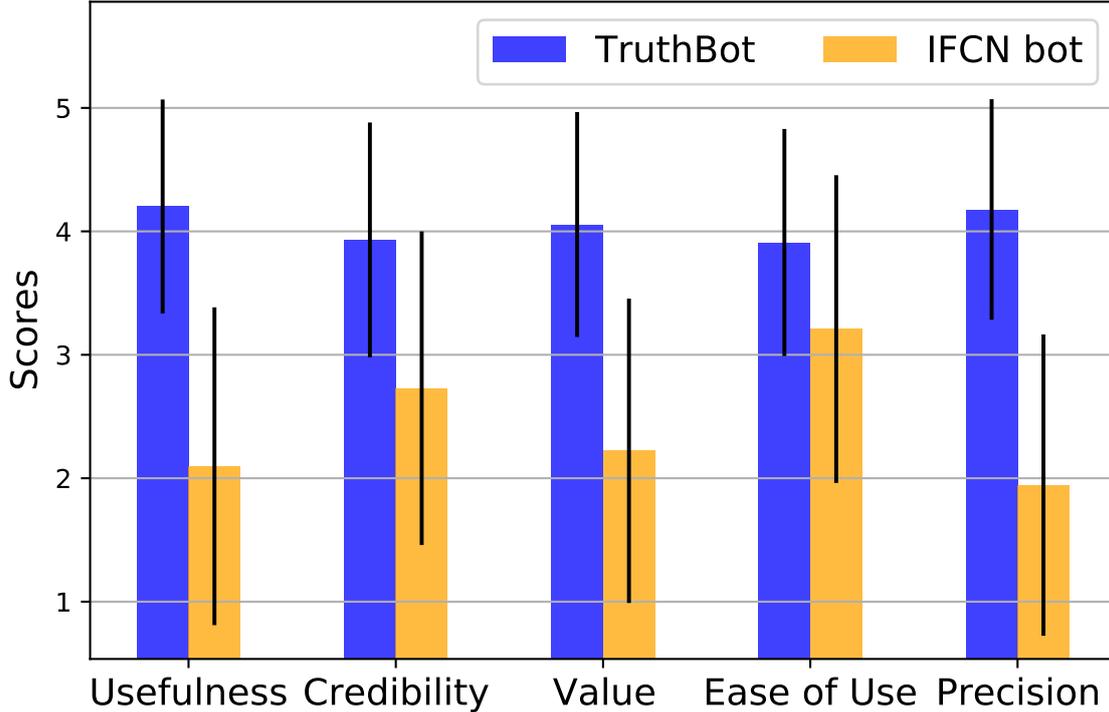


Figure 11: Comparison of the GUIs of the two chatbots along different metrics.

for this survey. In the survey, the users had to rate the bots for each of the above aspects. Possible ratings were from 1 (worst) to 5 (best). The result of the user survey is shown in Fig. 11. For each of the above aspects, TruthBot obtained better ratings compared to that of IFCN bot. Specifically, for Usefulness and Precision, mean rating for TruthBot was 2 times better than that IFCN bot on an average. The p -values of the t -test for statistical significance on all the aspects were < 0.00001 .

5 Summary and Future Work

We presented a multilingual and multipurpose chatbot that provides almost all aspects of information on a topic a user can be interested in. Such chatbots will be a single-point destination for users during situations such as COVID-19. Performance evaluation and user survey (as exemplified for COVID-19) also demonstrate the superiority of the bot over the existing options (IFCN chatbot) in providing multi-faceted well-searched information on a specific topic. The bot can be customized for other topics too. A primary goal of this paper is to introduce and exhibit the concept of such holistic information dissemination through a chatbot and provide evaluations to show its efficacy.

The chatbot still depends on a number of external databases. While this may not be completely avoidable, but in our future work, we plan to make some of these information to be cached in our local databases to reduce the dependencies on such external factors, which helps both in the response latency and reliability. We also plan to expand the scope of the chatbot to image and multimedia messages, which are often sources of misinformation.

References

- Tian Bian, Xi Xiao, Tingyang Xu, Peilin Zhao, Wenbing Huang, Yu Rong, and Junzhou Huang. Rumor detection on social media with bi-directional graph convolutional networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 549–556, 2020.
- Alessandro Bondielli and Francesco Marcelloni. A survey on fake news and rumour detection techniques. *Information Sciences*, 497:38–55, 2019.
- Petter Bae Brandtzaeg and Asbjørn Følstad. Why people use chatbots. In *International Conference on Internet Science*, pages 377–392. Springer, 2017.
- Aengus Bridgman, Eric Merkley, Peter John Loewen, Taylor Owen, Derek Ruths, Lisa Teichmann, and Oleg Zhilin. The causes and consequences of covid-19 misperceptions: Understanding the role of news and social media. 2020.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.
- Asbjørn Følstad, Petter Bae Brandtzaeg, Tom Feltwell, Effie L-C Law, Manfred Tscheligi, and Ewa Luger. Chatbots for social good. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018.
- Forbes.com. Electronic, 2019. URL <https://www.forbes.com/sites/gilpress/2019/10/02/ai-stats-news-86-of-consumers-prefer-to-interact-with-a-human-agent-rather-than-a-cl#2d2cb9432d3b>.
- GrandViewResearch. Electronic, 2017. URL <https://www.grandviewresearch.com/industry-analysis/chatbot-market>.
- IFCN. Ifcn fact checking organizations on whatsapp. online, 2020. URL <https://faq.whatsapp.com/general/ifc-n-fact-checking-organizations-on-whatsapp>.
- interaction design.org. The 7 factors that influence user experience. online, 2020. URL <https://www.interaction-design.org/literature/article/the-7-factors-that-influence-user-experience>.
- Andrew Kohut, Carroll Doherty, Michael Dimock, and Scott Keeter. Americans spending more time following the news. *Pew Research Center*, 2010.
- Zhaojiang Lin, Peng Xu, Genta Indra Winata, Farhad Bin Siddique, Zihan Liu, Jamin Shin, and Pascale Fung. Caire: An end-to-end empathetic chatbot. In *AAAI*, pages 13622–13623, 2020.
- Luis Enrico Lopez, Diane Kathryn Cruz, Jan Christian Blaise Cruz, and Charibeth Cheng. Transformer-based end-to-end question generation, 2020.
- Rada Mihalcea and Paul Tarau. TextRank: Bringing order into text. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, pages 404–411, Barcelona, Spain, July 2004. Association for Computational Linguistics. URL <https://www.aclweb.org/anthology/W04-3252>.

- Jeff Z Pan, Siyana Pavlova, Chenxi Li, Ningxi Li, Yangmei Li, and Jinshuo Liu. Content based fake news detection using knowledge graphs. In *International semantic web conference*, pages 669–683. Springer, 2018.
- Nicole M Radziwill and Morgan C Benton. Evaluating quality of chatbots and intelligent conversational agents. *arXiv preprint arXiv:1704.04579*, 2017.
- Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. Exploring the limits of transfer learning with a unified text-to-text transformer, 2019.
- Pranav Rajpurkar, Jian Zhang, Konstantin Lopyrev, and Percy Liang. SQuAD: 100,000+ questions for machine comprehension of text. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 2383–2392, Austin, Texas, November 2016. Association for Computational Linguistics. doi: 10.18653/v1/D16-1264. URL <https://www.aclweb.org/anthology/D16-1264>.
- Stephen Robertson and Hugo Zaragoza. *The probabilistic relevance framework: BM25 and beyond*. Now Publishers Inc, 2009.
- Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter*, 19(1): 22–36, 2017.
- Heung-Yeung Shum, Xiao-dong He, and Di Li. From eliza to xiaoice: challenges and opportunities with social chatbots. *Frontiers of Information Technology & Electronic Engineering*, 19(1):10–26, 2018.
- Emily K Vraga and Melissa Tully. News literacy, social media behaviors, and skepticism toward information on social media. *Information, Communication & Society*, pages 1–17, 2019.
- Xinyi Zhou and Reza Zafarani. A survey of fake news: Fundamental theories, detection methods, and opportunities. *arXiv preprint arXiv:1812.00315*, 2018.
- Xinyi Zhou and Reza Zafarani. Network-based fake news detection: A pattern-driven approach. *ACM SIGKDD Explorations Newsletter*, 21(2):48–60, 2019.
- Xinyi Zhou, Reza Zafarani, Kai Shu, and Huan Liu. Fake news: Fundamental theories, detection strategies and challenges. In *Proceedings of the twelfth ACM international conference on web search and data mining*, pages 836–837, 2019.