

## Development and testing of an FPT.AI-based voicebot

Duc Chung Tran<sup>1</sup>, Duc Long Nguyen<sup>2</sup>, Mohd Fadzil Hassan<sup>3</sup>

<sup>1,2</sup>Computing Fundamental Department, FPT University, Hoa Lac Hi-Tech Park, Vietnam

<sup>3</sup>Department of Computer and Information Sciences, CERDAS, Universiti Teknologi PETRONAS, Malaysia

### Article Info

#### Article history:

Received Mar 24, 2020

Revised May 30, 2020

Accepted Jun 15, 2020

#### Keywords:

Development

FPT.AI

Speech-to-text (STT)

Text-to-speech (TTS)

Vietnamese

Voicebot

### ABSTRACT

In recent years, voicebot has become a popular communication tool between humans and machines. In this paper, we will introduce our voicebot integrating text-to-speech (TTS) and speech-to-text (STT) modules provided by FPT.AI. This voicebot can be considered as a critical improvement of a typical chatbot because it can respond to human's queries by both text and speech. FPT Open Speech, LibriSpeech datasets, and music files were used to test the accuracy and performance of the STT module. For the TTS module, it was tested by using text on news pages in both Vietnamese and English. To test the voicebot, Homestay Service topic questions and off-topic messages were input to the system. The TTS module achieved 100% accuracy in the Vietnamese text test and 72.66% accuracy in the English text test. In the STT module test, the accuracy for FPT open speech dataset (Vietnamese) is 90.51% and for LibriSpeech Dataset (English) is 0% while the accuracy in music files test is 0% for both. The voicebot achieved 100% accuracy in its test. Since the FPT.AI STT and TTS modules were developed to support only Vietnamese for dominating the Vietnam market, it is reasonable that the test with LibriSpeech Dataset resulted in 0% accuracy.

*This is an open access article under the [CC BY-SA](#) license.*



### Corresponding Author:

Duc Chung Tran,  
Computing Fundamental Department,  
FPT University,  
Hoa Lac Hi-Tech Park, Hanoi, 155300, Vietnam.  
Email: chung.tranduc89@gmail.com

## 1. INTRODUCTION

The 21<sup>st</sup> century has witnessed the spectacular progress of artificial intelligence (AI) [1-4]. Instant messaging systems have been fully implemented as one of the primary communication technologies [5, 6]. Nowadays, chatbots, an application of AI, are replacing humans in interacting with users of the systems [7-11]. They have demonstrated their value of improving the overall user experience in different situations and contexts such as customer services [12], education [13] and online business [14]. They are opening up a new way for companies to connect with the world and, most critically, with their clients by increasing the reach of messaging applications. As far as we know, people cannot handle their work, for which in this case, is replying messages, for 24 hours a day and 7 days a week due to the limit of human strength and agility. Therefore, chatbots emerge as a suitable replacement for humans [15, 16]. They have been playing a crucial role in many systems that require high stability and fast response [17], as most companies understand the importance of customers' experience that just a small problem could make the customers losing their interest in the companies' services. An unpleasant experience could lead to losing a customer to the company's competitors [18-21].

In Figure 1, the message response times of two different management pages having the same kind of service, Homestay, is presented (some information has been blurred due to privacy reasons). The first one

typically replies instantly, while the second one usually replies within a few hours. Assuming that these two services have the same quality, price, fame, open and close times, and are locating near each other. Most customers' preferred customer services' value is saving their time. They will also use the services that answers their questions quickly and accurately, often putting price of services and products the second [22, 23]. Therefore, most customers would choose the first service.

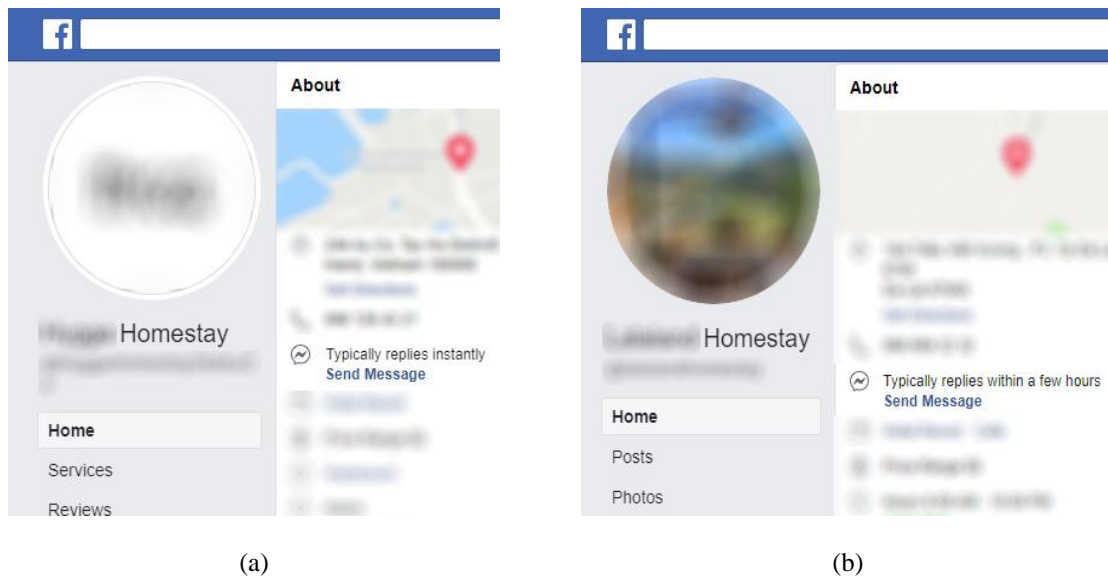


Figure 1. Message response speeds between two homestay service's fanpages, (a) The page replying instantly, (b) The page replying within a few hours

However, in communication, using voice is the most natural and effective way [24]. Therefore, voicebots are introduced. Typically, a voicebot is a combination of a chatbot, a TTS module and an STT module. After receiving the customer's voice message, the bot will pass it to the STT module to analyze and convert it into text. Then, the bot generates a response message then converts it into an audio file by the TTS module and plays it in the user interface.

Our objective is to create a professional voicebot that can run as a web application on users' Internet browsers. It will be convenient for users as it does not require any installation on users' computers. The user interface should be as friendly and simple as possible to minimize loading time and improve user experience. The remaining of this work is organized as follows: In section 2, we present the research methodology. In section 3, results and analysis will be discussed. Finally, section 4 draws our research conclusion about the developed FPT.AI-based voicebot.

## 2. PROPOSED METHOD

### 2.1. Development

#### 2.1.1. Front-end and back-end development

In this work, the front-end languages used to create a graphical user interface (GUI) are hypertext markup language (HTML), cascading style sheets (CSS), and JavaScript (JS). Bootstrap framework and JQuery library were also utilized because of their customization, speed of development and ease of use. For back-end development, PHP is used. In addition, WebAudioRecorder JS library [25] is in place to record voice commands and encode them to MP3 files before performing any analyzing tasks.

#### 2.1.2. FPT.AI's utilized modules

FPT.AI's modules [26] are chosen because it is free and supports PHP. Besides, contacting for support in Vietnam (where the population is approaching 100 million) is easier and faster than through other commercialized tools supported by Google or Microsoft since their headquarters are in overseas countries. In addition, the provided tool [26] has advantages in supporting local language, Vietnamese, over other similar products offered by Google or Microsoft.

## 2.2. Dataset

### 2.2.1. Module testing

For TTS module testing, random texts from CNN and VnExpress news pages were used. Ten (10) text blocks which are small-length text (words or phrases), medium-length text (sentences) and a very-long-length text (paragraphs) were chosen to be tested. For STT module testing, 15-second-long files in FPT open speech dataset (FOS) [27, 28] and LibriSpeech Datasets (LS) [29] and two music files were chosen. The audio files available in FOS are already in MP3 format while audio sources from LS must be converted to MP3 before testing as the modules do not support the processing of non-MP3 audio files.

### 2.2.2. Voicebot testing

Messages about homestay service were sent to check whether the voicebot understands or not. Also, it should be tested with off-topic messages to see how it replies.

## 2.3. Voice recorder's configuration

Figure 2 presents the configuration of the voice recorder integrated into the developed voicebot. Here, *encodingType* defines the type of output file after encoding the recorded data. *workerDir* presents the directory containing the script WebAudioRecorder.JS. *numChannels* stands for Number of Channels of the encoded file, i.e., 2, because MP3 files only support 2 channels. The time limit for recording is set to 15 seconds, as voice commands need to be short and easy to understand. *encodeAfterRecord* represents the encoding method. We set it to *true* because encoding after recording is safer than encoding while recording. *bitrates*, which is the number of bits per second encoded in the MP3 file, is set to *160kbps* keeping the record not only light-weight but also acceptable for the module to process.

```
recorder = new
WebAudioRecorder(input, {
  workerDir: "js/",
  encoding: "mp3",
  numChannels: 2
});
recorder.setOptions({
  timeLimit: 15,
  encodeAfterRecord: true,
  mp3: {bitRate: 160}
});
```

Figure 2. Voice recorder's configuration

## 2.4. Testing process

### 2.4.1. Module testing

For modules testing, a PHP file is created to receive text or audio sources and make requests to the server by its application programming interface (API). Log files are then written. In this work, the TTS module testing process is presented as a flowchart in Figure 3(a). Firstly, the tester inputs testing text into a text field and click on Submit button. StartTime will be written. The application must check whether it can connect to the server or not. If not, it displays an error message and ends the testing process. Otherwise, it will send the input text to the server and wait for the server's TTS conversion stage. After that, it receives the processed audio file and plays in the user interface. Then the resultant EndTime and Processing Time will be written in the log file. Finally, the test process ends.

STT module testing process is presented as a flowchart in Figure 3(b). Firstly, audio source information will be sent to the application. StartTime will be written. The application must check whether it can connect to the server. If not, it displays an error message and ends the testing process. Otherwise, it will send the audio source to the server and wait for the server's processing. After that, it receives a text converted from the audio source and shows in the user interface. Then the results, EndTime and Processing Time, will be written in the log file. Finally, the test process ends.

In Table 1, testing data for each module are presented. Five blocks of Vietnamese text quoted from VnExpress [30, 31] and five blocks of English text quoted from CNN [32, 33] were requested to generate an audio file corresponding to each block. In addition, ten audio files containing Vietnamese voices and ten audio files containing English voices were used to test the STT module. To challenge the module more, 2 music songs performed by both Vietnamese and American artists were tested. The English music file to be chosen in this test, which is a 15-second cut of the song called *Boulevard of Broken Dreams*-performed by

*Green Day*, is named *English\_Music.mp3*. The Vietnamese one which is also a 15-second cut is named *Vietnamese\_Music.mp3*. It is the cut version of the song called *Dù có cách xa* - performed by *Đinh Mạnh Ninh*. All of the testing audio files have durations of 15 seconds to match with the voice recorders configure which is shown earlier.

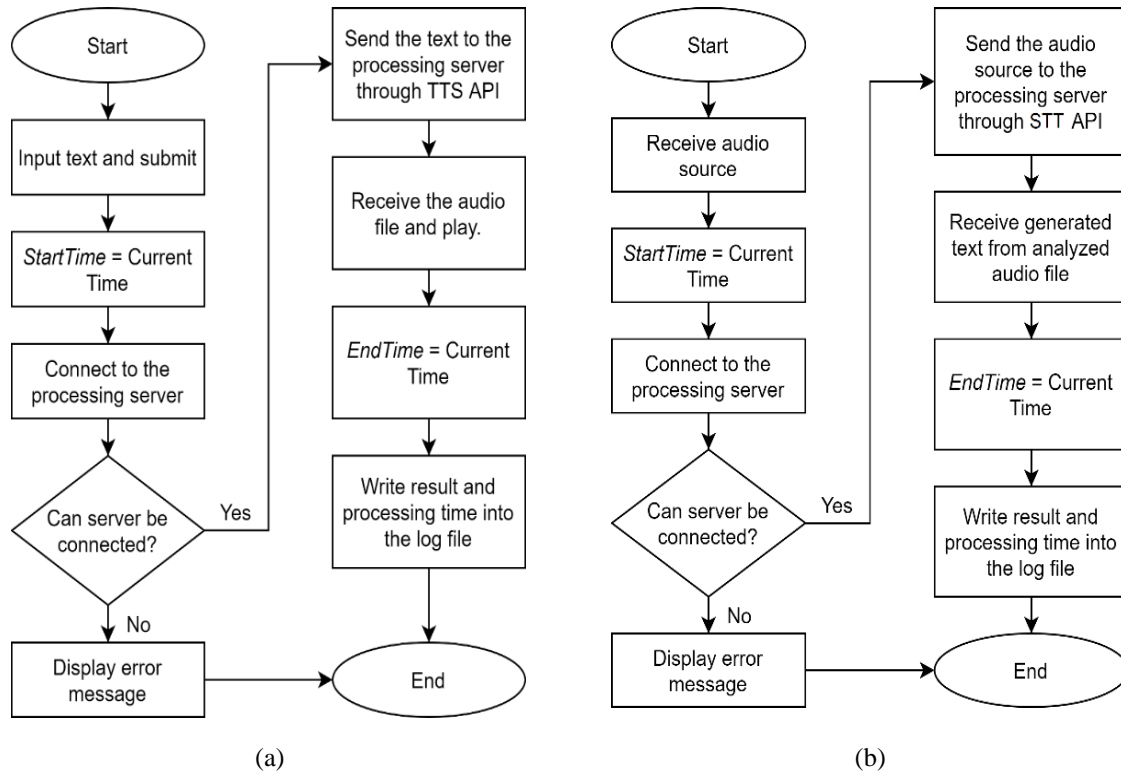


Figure 3. Modules testing process flowchart, (a) TTS module, (b) STT module

Table 1. Module testing data

Module	Data
TTS	<p>Bệnh nhân. (2 words)</p> <p>Cùng với giảm giá. (4 words)</p> <p>Phong tỏa bệnh viện Lao Phổi Quảng Ninh. (8 words)</p> <p>Sở Y tế Quảng Ninh sáng 15/3 cho biết đã cách ly toàn bộ nhân viên y tế và 113 bệnh nhân nội trú Bệnh viện Lao và Phổi, do bố mẹ "bệnh nhân 52" làm việc tại đây. (39 words)</p> <p>Phương án điều hành giá lần này, theo nhà chức trách, đảm bảo giá bán xăng dầu trong nước phản ánh xu hướng giá thành phẩm thế giới, phù hợp với tình hình kinh tế xã hội. Mặt khác, mức trích lập Quỹ Bình ổn giá xăng dầu ở mức hợp lý để có dư địa điều hành giá xăng dầu trong những kỳ tiếp theo trong bối cảnh giá xăng dầu đang có diễn biến khó lường, giảm giá ở tất cả mặt hàng. (85 words)</p> <p>Infected people. (2 words)</p> <p>New studies in several countries. (5 words)</p> <p>US Secretary of Health and Human Services. (7 words)</p> <p>These officials have emphasized that the virus is spread mainly by people who are already showing symptoms, such as fever, cough or difficulty breathing. If that's true, it's good news, since people who are obviously ill can be identified and isolated. (41 words)</p> <p>Albert II and Charlene have two young children -- twins Princess Gabriella and Prince Jacques. Gabriella was born two minutes earlier than Jacques, but he will one day take over the monarchy. As a man, he will take precedence over his older sister. Albert II has two older children, who were born out of wedlock and as such are not considered royal. The royal house received \$54.4 million in 2020, according to the principality's official budget. Of that, \$14.6 million goes to the Prince. (84 words)</p>

Table 1. Module testing data (*continue*)

Module	Data
STT	FPTOpenSpeechData_Set001_V0.1_002331.mp3 FPTOpenSpeechData_Set001_V0.1_006070.mp3 FPTOpenSpeechData_Set001_V0.1_008636.mp3 FPTOpenSpeechData_Set001_V0.1_011371.mp3 FPTOpenSpeechData_Set001_V0.1_012565.mp3 FPTOpenSpeechData_Set001_V0.1_013642.mp3 FPTOpenSpeechData_Set002_V0.1_000201.mp3 FPTOpenSpeechData_Set002_V0.1_004204.mp3 FPTOpenSpeechData_Set002_V0.1_004461.mp3 FPTOpenSpeechData_Set002_V0.1_007016.mp3 672-122797-0001.mp3 672-122797-0042.mp3 672-122797-0064.mp3 1089-134686-0018.mp3 1089-134691-0012.mp3 1188-133604-0011.mp3 1188-133604-0024.mp3 1221-135766-0010.mp3 1221-135767-0016.mp3 1284-1180-0019.mp3 English_Music.mp3 Vietnamese_Music.mp3

#### 2.4.2. Voicebot testing

Table 2 shows message templates to test whether the voicebot answers correctly or not. Besides, input random text or random voice commands will be given to see the bot's responses. Each question has its keywords to help the bot determine which question the user is asking. The voicebot will reply depending on the identified keywords.

For example, the keyword of this question “*Homestay còn phòng không?*” (Is this homestay available for rental?) is “*còn phòng*” (available/available for rental); or for this question “*Sau khi thuê có thể đổi phòng hoặc hủy đăng kí không?*” (After renting, could I change to another place or cancel renting?) are “*đổi phòng*” (change to another place) and “*hủy đăng kí*” (cancel renting).

Table 2. Voicebot testing data

Questions
Homestay còn phòng không? (Is this homestay available for rental?)
Homestay bao gồm những gì? (What does this homestay consist of)
Địa chỉ homestay ở đâu? (What is the address of this homestay)
Từ homestay mình có thể đi chơi những đâu? (Where can I go for a trip from this Homestay?)
Giá thuê homestay như thế nào? (What is the rental price?)
Làm thế nào để đặt chỗ? (What I need to do to rent this homestay?)
Hãy cung cấp thông tin chuyển khoản. (Please provide banking information.)
Tôi cần cung cấp thông tin cá nhân tại đâu? (Where do I need to provide my personal information?)
Giờ check-in và check-out thế nào? (What is check-in and check-out time?)
Sau khi thuê có thể đổi phòng hoặc hủy đăng kí không? (After renting, could I change to another place or cancel renting?)

### 3. RESULTS AND DISCUSSION

#### 3.1. Module testing

Test results of TTS and STT modules are presented in Table 3. In the TTS test, the module had mean trial processing time (MTPT) of 6.720 seconds for Vietnamese. In total, it processed 138 Vietnamese words and achieved 100% accuracy. In addition, it had MTPT of 8.508 seconds for English. In total, it processed 139 English words but the module only achieved approximately 72.66% accuracy. Meanwhile, the STT module achieved 90.51% accuracy in the Vietnamese audio test. It can be seen that the accuracies after testing with LS dataset and music files is 0%.

Figure 4. presents an example of a wrong case in English audio conversion. The expected output was “*The place he had was a very good one, the sun shone on him, as to fresh air there was enough of that, and round him grew many large-sized comrades, pines as well as firs*”. It can be seen that FPT.AI's modules only support Vietnamese well rather than other languages, i.e., English. The reason is that the modules were mainly developed for targeting the Vietnamese market. Therefore, English and other languages text and audio cannot be processed well, and the outputs cannot be expected to be as highly accurate as

Vietnamese ones. It should be noted that the conversion results were also depending on other factors, such as audio noise [30], audio quality [31], audio volume [32], tone and accent of the speaking person [33].

Table 3. Modules testing result

		Quantity	Correct	Mean Trial Proc. Time
TTS	Vietnamese	138 words	138 words	6.720 seconds
	English	139 words	101 words.	8.508 seconds
	FOS	432 words	391 words	11.422 seconds
STT	LS	395 words	0 words	23.417 seconds
	Music	2 files	0 files	21.058 seconds

```
File name: 672-122797-0001.mp3;
Length: 00:00:15;
Analyzed content: Quây xì gà Cuba
Real sân Anfield cà rau. Thank you
may Need More sai hôm ra khỏi sự chú
ý ở sở;
Processing time: 00:00:23.40.
```

Figure 4. Detailed results of a wrong case of English audio file processing

From the results, it is observed that the STT module cannot handle music audio sources, despite which language of their content is. Figure 5 presents the detailed result after processing the two music files done by STT module. The expected output for the English file is *"I'm walking down the line that divides me somewhere in my mind on the border line of the edge and where I walk alone"*, and for the Vietnamese file is *"Ngày hôm qua em chợt mang nắng tới, chợt làm ngẩn ngơ đắm say anh chiếc hôn bối rối, chợt làm muốn câu ca, chợt gọi mùa xuân qua, đừng rời tay trong tay anh nghe mùa xuân ôi đắm say, và..."*. Thus, the system's actual results are different from the expected outputs. Hence, the developed modules cannot be used to process music audio sources.

```
File name: Vietnamese_Music.mp3;
Length: 00:00:15;
Analyzed content: Ngày hôm nay. Trời
oi, nhớ nam say you do. Gọi mùa xuân
qua đường dây đeo trên tay anh.
Đình.;
Processing time: 00:00:20.698.
```

(a)

```
File name: English_Music.mp3;
Length: 00:00:15;
Analyzed content: Sao. Đi. Qua.;
Processing time: 00:00:21.418.
```

(b)

Figure 5. Detailed results after processing the two music files, (a) Vietnamese music file, (b) English music file

### 3.2. Voicebot testing

In this test, the voicebot achieved 100% accuracy for the prepared questions. The processing time of the voicebot started from when the user clicked on the *Stop recording* button to when the user received the corresponding response message. The voicebot took 24.738 seconds in total to convert ten questions of the user from speech into text to process; 60.174 seconds in total to find ten answers and convert them from text into audio files to play in the user interface. On average, the trial processing time is 4.246 seconds for 20 questions and answers. However, if the user asks in different ways, such as asking the shortened questions, or asking questions containing different keywords, the bot will not understand the intention of the user, and it will reply to off-topic messages. For off-topic messages, the voicebot randomly gets one of three messages: *"Xin lỗi! Tôi chưa hiểu ý bạn là gì."* (*I am sorry! I didn't understand what you meant*), *"Bạn vui lòng nêu rõ câu hỏi nhé!"* (*Please specify your question!*) and *"Xin lỗi bạn! Tôi chưa hiểu bạn nói gì"* (*I am sorry! I didn't understand what you said*) and replies to the user.

#### 4. CONCLUSION

In this research, FPT.AI-based STT and TTS modules were utilized to develop and test a voicebot supporting Homestay Service's queries. As seen from the obtained results, one can conclude that the modules are only suitable for Vietnamese text and speech processing due to their imprecision in processing another language (i.e., with English, the accuracies are 72.66% and 0% for TTS and STT respectively). For Vietnamese conversation content, the accuracies of TTS module and STT module are 100% and 90.51% correspondingly. The voicebot works well with the prepared questions and statements. However, the bot's detection accuracy still needs to be improved in the future to be able to recognize more complex questions and statements.

#### ACKNOWLEDGEMENTS

The authors would thank FPT University and FPT Corporation for supporting this work by providing modules, APIs and testing data to develop the voicebot.

#### REFERENCES

- [1] Y. Ong and A. Gupta, "AIR5: Five Pillars of Artificial Intelligence Research," in *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 3, no. 5, pp. 411-415, Oct. 2019.
- [2] J. Liu *et al.*, "Artificial Intelligence in the 21st Century," in *IEEE Access*, vol. 6, pp. 34403-34421, 2018.
- [3] X. Ren and Y. Chen, "How Can Artificial Intelligence Help With Space Missions - A Case Study: Computational Intelligence-Assisted Design of Space Tether for Payload Orbital Transfer Under Uncertainties," in *IEEE Access*, vol. 7, pp. 161449-161458, 2019.
- [4] W. Chen, Y. Lin, F. Ng, C. Liu and Y. Lin, "RiceTalk: Rice Blast Detection Using Internet of Things and Artificial Intelligence Technologies," in *IEEE Internet of Things Journal*, vol. 7, no. 2, pp. 1001-1010, Feb. 2020.
- [5] E. Sahafizadeh and B. Tork Ladani, "A Model for Social Communication Network in Mobile Instant Messaging Systems," in *IEEE Transactions on Computational Social Systems*, vol. 7, no. 1, pp. 68-83, Feb. 2020.
- [6] A. Irene, L. Chang-Tsun, and C. Roberto, "Social Network Identification Through Image Classification With CNN," in *IEEE Access*, vol. 7, pp. 35264-35273, 2019.
- [7] H. Yeung Shum, X. Dong He, and D. Li, "From Eliza to XiaoIce: challenges and opportunities with social chatbots," *Frontiers of Information Technology & Electronic Engineering*, vol. 19, no. 1, pp. 10-26, 2018.
- [8] A. N. Vaidyam, H. Wisniewski, and J. D. Halamka, "Chatbots and Conversational Agents in Mental Health: A Review of the Psychiatric Landscape," *Canadian Journal of Psychiatry*, vol. 64, no. 7, pp. 456-464, 2019.
- [9] K. Chung and R. C. Park, "Chatbot-based healthcare service with a knowledge base for cloud computing," *Cluster Computing*, vol. 22, pp. 1925-1937, 2019.
- [10] L. Ciechanowski, A. Przegalinaska, M. Magnuski, and P. Gloor, "In the shades of the uncanny valley: An experimental study of human-chatbot interaction," *Future Generation Computer Systems*, vol. 92, pp. 539-548, 2019.
- [11] L. Zhou, J. Gao, D. Li, H. Y. Shum, "The Design and Implementation of XiaoIce, an Empathetic Social Chatbot," *Comput. Linguist.*, vol. 46, no. 1, pp. 53-93, 2020.
- [12] M. Chung, E. Ko, H. Joung, and S. J. Kim, "Chatbot e-service and customer satisfaction regarding luxury brands," *Journal of Business Research*, 2018.
- [13] P. Marsh, "The use of a chatbot in radiology education," *J. Med. Imaging Radiat. Oncol.*, vol. 62, pp. 97, 2018.
- [14] G. Daniel, J. Cabot, L. Deruelle, and M. Derras, "Xatkit: A Multimodal Low-Code Chatbot Development Framework," in *IEEE Access*, vol. 8, pp. 15332-15346, 2020.
- [15] Y. Wang, W. Rong, Y. Ouyang, and Z. Xiong, "Augmenting Dialogue Response Generation with Unstructured Textual Knowledge," in *IEEE Access*, vol. 7, pp. 34954-34963, 2019.
- [16] M. Nuruzzaman and O. K. Hussain, "A Survey on Chatbot Implementation in Customer Service Industry through Deep Neural Networks," *2018 IEEE 15th International Conference on e-Business Engineering (ICEBE)*, Xi'an, pp. 54-61, 2018.
- [17] Z. Liu, C. Long, X. Lu, Z. Hu, J. Zhang, and Y. Wang, "Which Channel to Ask My Question?: Personalized Customer Service Request Stream Routing Using Deep Reinforcement Learning," in *IEEE Access*, vol. 7, pp. 107744-107756, 2019.
- [18] Tran Duc Chung, Hahong Son profile imageHa Hong Son, Alexandra Khalyasmaa profile imageAlexandra Khalyasmaa, "A Question Detection Algorithm for Text Analysis," in *2020 5th International Conference on Intelligent Information Technology (ICIIT)*, pp. 61-65, February 2020.
- [19] L. Witell *et al.*, "Characterizing customer experience management in business markets," *Journal of Business Research*, Volume 116, pp. 420-430, August 2020.
- [20] Ashley S. Otto, David M. Szymanski & Rajan Varadarajan, "Customer satisfaction and firm performance: insights from over a quarter century of empirical research," *Journal of the Academy of Marketing Science*, vol. 48, pp. 543-564, 2020.
- [21] Nadine Schirmer, Christian M. Ringle, Siegfried P. Gudergan, Matthias S. G. Feistel, "The link between customer satisfaction and loyalty: the moderating role of customer characteristics," *Journal of Strategic Marketing*, vol. 26, no. 4, pp. 298-317, 2018.



- [22] Johanudin Lahap, Rabiatal Iylia Azlan, Khairil Anuar Bahri, Noraslinda Mohd Said, Dahlan Abdullah, Razlan Adli Zain, "The effect of perceived waiting time on customer's satisfaction: A focus on fast food restaurant," *International Journal of Supply Chain Management*, vol. 7, no. 5, pp. 259-266, October 2018.
- [23] J. De Vries, D. Roy, and R. De Koster, "Worth the wait? How restaurant waiting time influences customer behavior and revenue," *Journal of Operations Management*, vol. 63, pp. 59-78, November 2018.
- [24] Manjul Tiwari and Maneesha Tiwari, "Voice-How humans communicate?," *Journal of natural science, biology, and medicine*, vol. 3, no. 1, pp. 3-11, 2012.
- [25] Higuma, "WebAudioRecorder.js," 2016, [Online], Available at: <https://github.com/higuma/web-audio-recorder-js>
- [26] T. D. Chung, M. Driberg, M. F. Bin Hassan and A. Khalyasmaa, "End-to-end Conversion Speed Analysis of an FPT.AI-based Text-to-Speech Application," *2020 IEEE 2nd Global Conference on Life Sciences and Technologies (LifeTech)*, Kyoto, Japan, pp. 136-139, 2020.
- [27] FPT Corporation, "30 years-FPT shares 30 hours of recorded voice data," *FPT Corporation*, 2018, [Online], Available at: <https://techinsight.com.vn/language/en/30-years-fpt-shares-30-hours-of-recorded-voice-data/> (accessed Mar. 20, 2020).
- [28] L. C. Mai, "Vietnam Country Report 2019," *2019 22nd Conference of the Oriental COCOSA International Committee for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques (O-COCOSA)*, Cebu, Philippines, 2019.
- [29] V. Panayotov, G. Chen, D. Povey and S. Khudanpur, "Librispeech: An ASR corpus based on public domain audio books," *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brisbane, QLD, pp. 5206-5210, 2015.
- [30] Q. Thủy, "Bệnh viện Lao và Phổi Quảng Ninh ngừng tiếp nhận bệnh nhân (Quang Ninh Tuberculosis and Lung Hospital has stopped accepting patients)," *VnExpress*. 2020. [Online], Available at: <https://tinnhanhchungkhoan.vn/thoi-su/benh-vien-lao-va-phoi-quang-ninh-ngung-tiep-nhan-benh-nhan-318318.html>.
- [31] Anh Minh, "Giá xăng, dầu đồng loạt giảm sâu (Gasoline and Oil prices have plummeted)," *VnExpress*. 2020. [Online], Available at: <https://thuvienphapluat.vn/tintuc/vn/thoi-su-phap-luat/tai-chinh/27782/gia-xang-dau-dong-loat-giam-sau>.
- [32] I. Kottasová, "Here's how much Europe's royal families really cost," *CNN*. 2020. [Online], Available at: <https://edition.cnn.com/2020/03/14/europe/european-royal-families-intl/index.html>.
- [33] E. Cohen, "Infected people without symptoms might be driving the spread of coronavirus more than we realized," *CNN*. 2020. [Online], Available at: <https://edition.cnn.com/2020/03/14/health/coronavirus-asymptomatic-spread/index.html>.

## BIOGRAPHIES OF AUTHORS



**Duc Chung Tran** completed bachelor and Ph.D degrees, both in Electrical & Electronic Engineering, at PETRONAS University of Technology, Malaysia in 2014 and 2018 respectively. During the studies, his majors were Instrumentation & Control and WirelessHART networked control system correspondingly. He had worked for several companies namely Intel Technology Sdn. Bhd. Malaysia, R&D Centre - Sony EMCS Sdn. Bhd. Malaysia, Dasan Zhong Solutions Vietnam, FPT Technology Research Institute Vietnam. He has involved in authoring more than 30 indexed publications. Currently, he is with FPT University, Hanoi, Vietnam, researching in emerging technologies and topics including Natural Language Processing and Generation (NLP, NLG), and Applied Artificial Intelligence.



**Duc Long Nguyen** joined FPT University in 2018. He is a sophomore studying Software Engineering specialization. With the help and cooperation of his research team, he is currently working on research topics related to Artificial Intelligence and Natural Language Processing.



**Mohd Fadzil Hassan** is an alumnus of the Malay College Kuala Kangsar (MCKK) and graduated in 1999 from the Colorado State University, USA with a BSc (cum-laude) in Computer Information Systems. He obtained his MSc in Artificial Intelligence in 2001 and PhD in Informatics in 2007 from the University of Edinburgh, UK. He was the former Dean, Centre for Graduate Studies, Universiti Teknologi PETRONAS (UTP) and currently the Director, Institute of Autonomous Systems, UTP. His research interests are in the area of Artificial Intelligence, Multi-Agent Systems and Service Oriented Architecture (SOA). He is actively involved in research works focusing on these areas and has secured numerous research grants as a principal investigator namely FRGS, ERGS, PRGS, and Technofund from the Malaysian government. He is also actively involved with international collaborative research works particularly with universities from the Middle East, Korea, and ASEAN region. He has involved in authoring more than 100 indexed publications. He is a senior member of IEEE and was an executive committee for the IEEE Computer Society Malaysia in 2018.