

Received: 7 Nov, 2017; Accept: 23 Dec, 2017; Published: 29, Dec 2017

SamBot - Intelligent Conversational Bot for Interactive Marketing with Consumer-centric Approach

Aditya Pradana¹, Goh Ong Sing² and Yogan Jaya Kumar³

¹Universitas Dian Nuswantoro, Jl. Nakula I No. 1-5, 50131,
Semarang, Indonesia,
zastix@yahoo.com

^{2,3}Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka,
76100 Durian Tunggal, Melaka, Malaysia.,
²goh@utem.edu.my, ³yogan@utem.edu.my

Abstract: Interactivity of corporate website is important as marketing products through it become popular. Moreover, the usage of Artificial Intelligence application is growing in many sectors. In order to improve the interactivity and effectiveness of corporate website in providing information, a conversational bot called SamBot is developed. As a part of Artificial Intelligence application, it can respond to users' question and prolong the conversation with its intelligence. It is integrated to Samsung IoT Showcase website as the corporate website. SamBot contains the knowledge of Samsung marketing domain which is useful to deliver appropriate answer to the users' questions. Its knowledge base includes Samsung promotion, Samsung product Frequently Asked Question, and general knowledge. In default, a conversational bot will generate random answers if matching knowledge cannot be found. In this research, the authors would like to overcome this problem by improving its knowledge maintenance capability either performed by botmaster or users. A function is added to support the knowledge maintenance process by botmaster, namely conversation log-based knowledge maintenance module. As to support consumer-centric knowledge approach, several features are developed such as random answer enhancement and consumer-centric learning capability. This will allow user to directly teach the bots a new knowledge. In addition, enhancement to reuse recently asked question as questions recommendation or autocomplete questions is developed in order to solve a problem where users need to be updated to current issues. With this enhancement, SamBot can give recommendation of questions for users to ask. All of the enhancements are measured using a world-wide Turing Test or Loebner Prize chatbot competition scoring system. The result shows that SamBot is capable of handling Samsung related questions very well with 26.4 points scored.

Keywords: Artificial Intelligence, conversational bots, deep learning, AIML, marketing.

I. Introduction

Nowadays, companies tend to use their corporate websites to promote their products and sell directly from there. This method is not only used by companies but also salespersons.

Marketing products through website is more efficient due to the growth of e-commerce websites like e-bay, Alibaba, Amazon, etc. Companies do not want to lose their grasp on online marketing. Therefore, they put more effort on their corporate websites to compete with e-commerce websites. Marketing becoming a dominant force in competitive business environment. It requires attention to customer's needs, desires, creative ideas and planning to find out what works best. As a result, customers tend to expect a certain level of interaction on a company's corporate website, regardless of the nature of the company and its services. The rapid growth of social media and crowd sourcing techniques could lead corporate websites toward extinction if companies do not create the tools, technologies, and applications needed to deliver a customer-centric website [1].

In Web 2.0 era, interactivity is a requirement that all websites should fulfill. It changed the way users interact to a web page [2]. However, users' demand of interactivity is still increasing. Therefore, some methods were proposed to give more interactive feel of the website such as live chat, FAQ and chatbot. The latter method is what this research will look into. A chatbot or conversational bots is an implementation of Artificial Intelligence (AI) in a form of software or application which users can interact by having conversations [3]. It can understand natural language and receive text or voice input. It is even intelligent enough to remember user's name and can prolong a conversation. In marketing sector, the chatbot can act as a salesperson to help companies advertise their products. As the bots can talk and influence people, it can help to attract more people, thus increase the advertisement power of the company [4]. In addition, it also has low maintenance cost and serves 24/7.

A corporate website usually contains rich information which spread across the links. In order to find a particular information, visitors need to explore the links by opening them one-by-one. It is inefficient, less interactive and time wasting.

In order to overcome this problem, a prototype of conversational bot called SamBot is introduced in this research. The name of SamBot is derived from the words Samsung Bots which will assist the Samsung IoT Academy Showcase. SamBot will have all information regarding the company, including the products. Therefore, visitors can simply ask to the bot when they need such information through text or speech. By this way, the information needed can be retrieved efficiently in an interactive approach.

Information of the website is stored in chatbot's knowledge base. The knowledge can be a domain-specific or a general knowledge. Knowledge of a domain-specific chatbot is usually built by a botmaster who gathered it from external resources by various methods. Huge amount of knowledge can be gathered and stored to the knowledge base. However, new information of that particular domain always emerges. Since there is no knowledge related to the new information in the knowledge base, a chatbot will improperly respond to such input and generate random answer. The improper response might annoy users who interact with the bots which can lead to unsuccessful improvement of interactivity aspect. Therefore, this research proposed a way to maintain the knowledge of a chatbot to keep up with new information.

Botmaster will be provided a list of unanswered questions from conversation logs to be evaluated. From the logs, botmaster can provide proper answers to the questions that previously are answered using random answer. Thus, after the knowledge is updated, the bots will be able to answer those particular questions. This kind of method will be called as Knowledge creation. Another method is Learning Capability where the conversational bots will be improved by learning new knowledge from online users. In conversational bots studies, there is a concern where usually conversational bots can only provide information to users, not the other way around. As mentioned earlier, the information stored in the bots is provided by a botmaster. Therefore, the information is limited to the botmaster's knowledge only. This study will address this concern as conjunction with Knowledge creation by considering users' perspective. This concern is important because users' knowledge may be useful to other users. This function will let users to teach the bots when it generated random answers or they were not satisfied with the bot's answer. In addition, visitors of a website usually need fresh information regarding what other visitors concern. They want to keep updated with the trend of such information. In a conversational bots there is a database which stores all conversation from the users. It is called as conversation log. It might be useful to consider the usage of conversation log in order to let users know what other users are concerning about. It will also helpful to help users to decide the questions to be posed to the conversational bots.

II. Literature Review

This research will integrate conversational bots into a corporate website. Conversational bots also known as Conversational Agent (CA) is a computer software which interacts and adapts with conversation environment. It is a part of AI's intelligent agent which understands natural language. Some examples of earlier conversational bots are ELIZA and

ALICE. Within the last 10 years, the development of conversational bots is promising. Most of them are domain-specific conversational bots which are implemented in various domain such as education, social, politics, networking, entertainment, business, health, tourism, and marketing. This section will discuss some conversational bots applied in the domains as well as different types of conversational bots models.

As other conversational bots, the conversational bot in this research will be developed based on AIML (Artificial Intelligence Markup Language). AIML contains the knowledge of conversational bots which is maintained by botmaster. Botmaster is a person who develops the bots and is responsible for maintaining bot's knowledge. This research proposed a method to evaluate conversational bot's knowledge by filtering unanswered question and update its knowledge base. This section will also review several important AIML elements.

A. Corporate Website

There are two main characteristics of a corporate website such as usability and interactivity [1]. The more usable a corporate website is the easier for users to browse. Interactivity aspect concerns with psychology aspect of users to interact with. Thus, interactivity is an important requirement of a corporate website. Interactivity is defined as the ability of a website to maintain vivid interaction and communication with the users. Several research found that interactivity gives significant effects on user performance. Moreover, a website's IT-enabled interactivity influences customer's trust in online services provided as well as marketing practices. As marketing is usually the main objective of a company, it is important to improve it along with the current technology. This research will integrate conversational bots into corporate website in order to improve the interactivity. The bots can interact with users either by text or speech. It is loaded with information related to the company and the products. Thus, it can support users to find information more effectively.

Some examples of corporate website are reviewed to show each company's method on supporting their marketing aspect of the website. As many as 10 corporate websites are reviewed such as Apple, Hewlett Packard (HP), Asus, Honda, General Motors (GM), Toyota, Walmart, AirAsia, Australia Taxation Office (ATO), and Nike website. Most of them implement conventional and less interactive Frequently Asked Question (FAQ) section where customers have to choose the topic before choosing the exact questions. Apple, GM and Toyota applied this method while HP, Honda and Walmart improved the section with a search engine. The search engine functioned as quicker way to find a question in question list. Meanwhile, Asus applied a slightly different FAQ section where customers have to choose a product first before going to FAQ section. AirAsia and Nike are simply used search engine in their FAQ section. The only website which implements chatbot is ATO. Some of them also provide live chat with customer service but this method can only be done in specific time. By using FAQ section, customers are presented a less interactive way to find information. Integrating a chatbot is more interactive, effective and dynamic method to support the marketing. The website will produce a livelier interaction with chatbot and

capability to serve online 24/7.

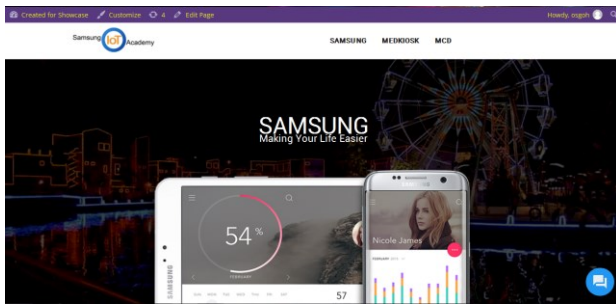


Figure 1. Samsung IoT Academy Showcase

In this research, Samsung IoT showcase website is enhanced with the integration of conversational bot called SamBot. The homepage of the website is shown in Figure 1. SamBot gives satisfying responses to customers who ask product features, product specifications and other questions related to Samsung electronic products. The interaction between SamBot and customers will raise effectiveness marketing as well as the interactivity of this portal.

B. Conversational Bots

Conversational bots also known as Conversational Agent (CA), chatbot or chatter bots is software robot which is placed in a conversation environment to interact between computer and human. Chatbots can operate according to a set of predefined rules or, using machine learning, searching algorithm [5], work based on various degrees of artificial intelligence [6]. They can be accessed through chat tools such as Facebook Messenger or other types of messaging apps [7]. Basically, a chatbot interacts with the user in the same way between human-human conversations. Theoretically, the bots gets smarter as it interacts with more people and learns new things, which opens up some fascinating instructional avenues to explore. It also understands natural language thus can respond naturally [8].

One of the earliest conversational bots was developed by Weizenbaum (1966) called as ELIZA. ELIZA played a role as psychotherapist to open up patients' personal opinion which may prolong the conversation. Pattern recognition with decomposition rules is implemented in ELIZA thus it can rephrase the patient's statement in to a question. Beside ELIZA, another remarkable conversational bots called ALICE was developed by Dr. Richard Wallace in 2000. It won the Loebner Prize in 2000, 2001 and 2004. Loebner Prize is a world-wide chatbot competition held annually. ALICE is using AIML created by Dr. Richard Wallace and Alicebot community between 1995 and 2000 [9]. Since then, AIML has been used widely to build conversational bots and the number of conversational bots development is increasing.

Conversational bots has been implemented widely and become popular. Besides ELIZA and ALICE, recent conversational bots also contribute in various domains such as education, social, politics, networking, entertainment, business, health, tourism, and marketing. In education domain, O. S. Goh & Lim (2016) [10]; Mikic, Burguillo, Llamas, Rodríguez, & Rodríguez (2009) [11]; and Orlando & Giovanni (2008) [12] developed conversational bots to support education process in e-learning environment. They

built bots to assist education process by monitoring the learning process or by replacing the teaching team in Massive Online Open Courses (MOOC) websites. The conversational bots are very assistive in e-learning environment since people from various country with different time zones are lined up in different time. This issue is difficult to handle for human who have uncertain and limited online time. For bots which is a software, online time is not an issue. They are literally "online" continuously once integrated to the website. They can answer same questions asked by different students tirelessly and they are not affected by users' emotion. Another example from education domain is built by Ghose & Barua (2013) [13] who sees from different view. They developed a bot as an undergraduate advisor which can assist human undergraduate advisor to answer same questions many times. Conversational bots are also applied in social and politics such as in e-government systems. Mahapatra, Sharma, Trivedi, & Aman (2012) [14] integrated conversational bots in e-government systems to improve the interface interactivity while Rodrigo & Abraham (2012) [15] developed a bot representing Twitter user to control social opinions regarding particular issue.

Health is also a domain used as an object of domain-specific conversational bots. One of them is a bot which is developed by O. S. Goh, Fung, Wong, & Depickere (2007) [16]. They showed that conversational bot's knowledge can be maintained to current issue which in this case is Bird Flu pandemic crisis. Another popular domain which is as popular as education domain is marketing. In marketing domain, as Chakrabarti & Luger (2015) [17] and Weerawarna et al. (2011) [18] built, their conversational bots were used as customer service who can answer Frequently Asked Question tirelessly. It also adds more interactivity to the website when customers are looking for information. In marketing domain, similar developments have been done continuously within four years such as Chakrabarti & Luger (2015) [17]; Moradi, Aghaie, & Hosseini (2013) [19]; and Weerawarna et al., (2011) [18].

On the evaluation of the chatbot, some researches evaluated the performance of conversational bots instead of building the domain knowledge of the bots. Torrey, Powers, Marge, Fussell, & Kiesler (2006) [20] evaluated the quality of human-robot conversation whether it can communicate using effective natural language. Similarly, Goh et al. (2007) [21] evaluate the quality of conversational bots. Unlike Torrey et al. (2006) [20] who evaluated the natural language usage, they evaluated the response quality of domain-specific conversational bots using black-box approach. Some other researches emphasized the knowledge building of conversational bots and tried to enhance it. In 2007, G Pilato, Augello, Vassallo, & Gaglio (2007) [22] combined rational reasoning and associative reasoning brain in order to enhance the intuitiveness of conversational bots. They used ontological approach to enhance the knowledge of conversational bots. In 2008, O. S. Goh & Fung (2008) [23] conducted a study whereby trust based knowledge acquisition was done to enhance the knowledge of conversational bot. Another approach using dynamic knowledge representation was done by Giovanni Pilato, Augello, & Gaglio (2011) [24] to enhance the adaptiveness of conversational bots. Previous researches showed that knowledge enhancement is important in conversational bot development. Thus, in this research a new

approach on knowledge creation will be introduced to enhance the knowledge of SamBot in order to maintain the knowledge updated along with current issues. This research will implement the conversational bot into marketing domain. With the knowledge creation, the conversational bots can be more adaptive, intuitive and dynamic in responding users' input.

C. Artificial Intelligence Markup Language (AIML)

Most of conversational bots existing nowadays are based on AIML. Conversational bots built in this research will also be implemented based on AIML. AIML was designed by Dr. Richard Wallace and Alicebot free software community to create stimulus-response chatbots. It is a derivative of XML with a specific structure. AIML is an XML-based language for conversational bots knowledge building. ALICE AI Foundation (n.d.) [9] stated that 80% of conversational bots development are based on AIML. This statement also proved by the use of AIML in recent conversational bots development discussed in previous sub-chapter.

In AIML, there are three main elements such as “<category>”, “<pattern>”, and “<template>”. The “<category>” tag represents a single knowledge which means, the more “<category>” an AIML file has, the more knowledge it contains. Inside of “<category>” tag, “<pattern>” and “<template>” contains input statement and respond statement respectively. The “<pattern>” tag may contain words, spaces, and the wildcard symbols ‘_’ and ‘*’. Besides the main elements, there are several optional elements such as “<that>”, “<topic>”, “<system>”, and “<javascript>”. The “<that>” tag is a referral tag to other knowledge while “<topic>” tag is used to organize knowledge with similar topic together. AIML supports the interaction with other programs and language which is represented with “<system>” tag and “<javascript>” tag. The “<system>” tag is used to execute other program while “<javascript>” is used to execute JavaScript code inside the templates. All of the elements are contained in “<aiml>” tag which opens and ends the AIML file.

D. Retrieval-Based Model

Conversational bots is usually developed on top of deep learning model. There are two types of deep learning models such as retrieval-based model and generative model. Retrieval-based model decides a matched pair of question and response from a repository in which the pair is already pre-defined either by a simple rule based technique or a more complex machine learning classifiers. Thus, it just picks a proper response from a fixed knowledge repository without generating new responses itself. In contrast, generative model can respond without any pre-defined knowledge. It is usually based on Machine Translation technique with a modification to make it translates input to an output or response. Both models have their own advantages and limitations. As stated in WildML (2016) [25], retrieval-based model is easier to implement and more accurate in term of finding responses. However, it is not as smart as generative model since it cannot respond properly to an undefined input or grammatical mistakes whereas generative model can respond to all kind of inputs. Unfortunately, the implementation of generative model is almost impossible since the outputs are usually have

grammatical errors or incomplete sentences which lead users to a confusion. By considering the drawbacks of generative model and the advantages of retrieval-based model, this research will develop a conversational bot based on retrieval-based model. As mentioned earlier, this model needs a knowledge repository or knowledge base. This knowledge base contains either the knowledge created by botmaster, taught by user, or updated knowledge as a result of knowledge maintenance. A botmaster will create knowledge based on the domain where the conversational bot is placed. As much as it is created, there is always new information raised which is not included in its knowledge base. This kind of information can be considered as new knowledge as addition to the earlier knowledge and it is a responsibility of botmaster to maintain the novelty of a conversational bots. In addition, to complement the maintenance done by botmaster, knowledge base may also contain certain knowledge created by considering user's perspective.

Some researches tried to minimize the intervention of human in knowledge creation. OntBot built by Al-Zubaide & Issa (2011) [26] gained its knowledge dependent to ontology-to-relational-database mapping techniques and regular knowledge base. Despite the use of ontological approach, it still needs high interference of botmaster to update the knowledge by defining new rules and to update its knowledge base. Similarly, Wu et al. (2008) [27] tried to automate the knowledge acquisition of conversational bots. They combined classification model based on rough set from online forum and ensemble learning. However, the intervention of human expert is still heavily needed to identify the relevance of thread replies in a discussion. Simply said, they only automate the classification process of domain selection rather than the knowledge building itself. Thus, trying to minimize human intervention in knowledge building is not recommended if knowledge-able bots is the objective. Using human intervention leads to better accuracy and performance (Goh, 2008) [28]. It also expands the flexibility of domain change to adapt with the current issue. Unlike other research works done, the knowledge creation in this research is supported with enhancement on the system. A database will be utilized to store random answers from conversation logs. The random answers can be filtered from overall conversation by creating new function inside the bots and let the botmaster to evaluate the questions. The knowledge creation will also be done to support learning process directly from users.

III. SamBot Architecture

The system architecture in this research is focused on the website components as well as the conversational bot's knowledge creation. The knowledge creation and knowledge enhancements will improve the intelligence of the conversational bots and prolong the conversation. The architecture contains several databases, programs, files and entities representing important elements in the system. For the conversational bots itself, the knowledge is based on AIML language which will be processed in an AIML interpreter. The AIML interpreter runs on PHP as programming language, Apache as web server and MySQL as database. Thus it can be concluded that the AIML interpreter is the core of SamBot.

Then SamBot can be integrated with the website and accessed by any platform such as PC, laptop and mobile devices. In order to develop the bots knowledge, botmaster needs to write some AIML files which contain of Samsung products promotion, product FAQ, and AAA (Annotated ALICE AIML) knowledges. Samsung promotion and product FAQ are domain-specific knowledge concerning Samsung’s products. AAA knowledge are collection of general responses, emotions, and any other common knowledge which might be asked by the users. All of the knowledge are stored in a database called AIML Knowledge Base. The AIML interpreter will refer to this database when a question is posed by a user. The user can access the bots through Livebase (shown in Figure 1) which is a chat box plugin and can be integrated with any website using Wordpress platform. This website and AIML interpreter are secured with Apache as the web server.

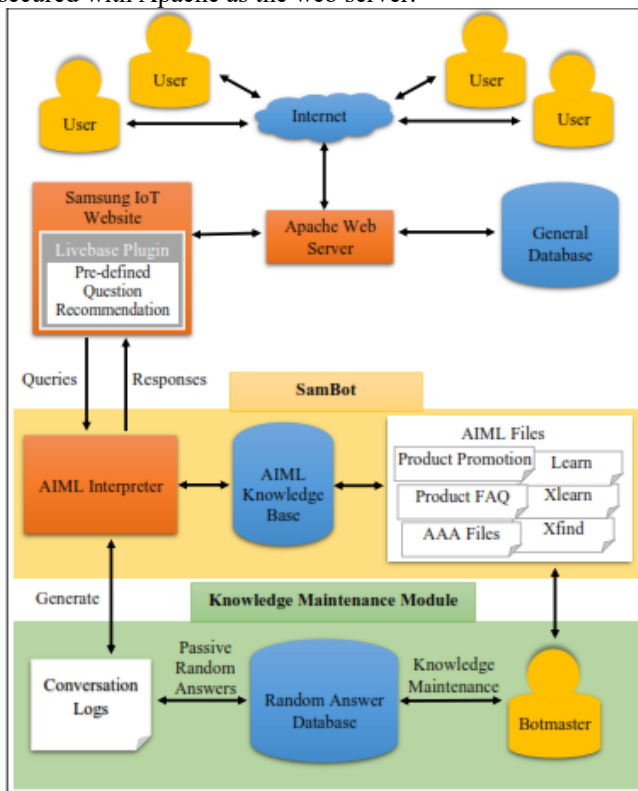


Figure 2. SamBot’s System Architecture

Beside knowledge creation, the system will have four additional features. First, the ability to filter random answers from conversation logs and store them into a database. This improvement will ease the knowledge maintenance process later. Second, enhancement on the random answer generation to produce two types of random answers which are active random answer and passive random answer. Third feature is the ability to learn new knowledge from users. With this feature, the bots can learn new knowledge from users and become more intelligent. The fourth feature is to add new function on Livebase interface plugin which will recommend the user a pre-defined questions came from users’ previous inputs. The function of the pre-defined questions was similar with Autocomplete proposed by Google [29]. Saving a few keystrokes is great during the questions input from the users, but the real benefit appears to be helping users use keyword or phrase that are hard to spell. Figure 2 illustrates the system architecture.

A. Website and Livebase Plugin

The prototype of the Samsung IoT Academy Showcase website contains latest features of Samsung products as well as their latest products. The website has attractive transition among page section. There are ten sections in this page and each section has its own content and function such as products section, gaming section, and design section. It also has unique shortcut to identify the section. This shortcut is useful in integration with the bot’s knowledge where the bots can change the section according to user input. The website built using PHP language and works on AJAX (Asynchronous Javascript and XML) method to lively update the chat without reloading the page. It also runs Web Speech API which makes it able to receive voice input as well as reading the text response to user.

B. Speech Input

In the development, a speech input is included to help visually impaired users and produce better interactivity. Web Speech API is used for this speech input. Web Speech API is a JavaScript and embedded with Text-to-speech (TTS) which allows a web application to have the ability of speech recognition and speech synthesis. With this function, SamBot can receive users’ voice and recognize the words spoken thus it understands the question. SamBot can also respond by reading the text respond if the device’s speaker is on. However, this feature is currently available only on Google Chrome and Apple Safari [10].

C. Knowledge Base

The knowledge of SamBot consists of Samsung promotion, product FAQ, and general knowledge in AAA knowledge base. The Samsung promotion knowledge contains the common features of the Samsung as a company. Product FAQ knowledge contains the specifications of several latest specification of Samsung products including smartphones, tablets, and accessories. AAA knowledge contain general knowledge which might be asked by users. AAA files are basically ALICE bot’s which was awarded as most human conversational bots. The AAA knowledge base are very important to prolong the conversation with users. All of these knowledge bases are written in AIML files which can be opened and edited in text editor.

```

1 <aiml ver="1.0">
2
3 <category>
4 <pattern>* SAMSUNG GALAXY S6 EDGE
5 SPECIFICATION</pattern>
6 <template>Here you go! <script
7 language="JavaScript"> var winObj
8 =window.open('http://www.samsung.com/my/con
9 sumer/mobile-devices/smartphones/galaxy-
10 s/SM-G925FZDAXME', 'winObj', 'width=800,
11 height=640'); </script>
12 </template>
13 </category/>
14
15 </aiml/>

```

Example responses

User : Hello

Bot : Hello there.

User : Do you have Samsung Galaxy S6 Edge specification?

Bot : Here you go! (Redirect user to Samsung Galaxy S6 Edge page)

Figure 3. Product FAQ Scenario

1) Samsung Promotion

Samsung promotion knowledge makes SamBot can persuade users with latest promotion available and show the latest features of Samsung devices. It has the ability of AIML to run JavaScript to go to a particular section of the page. This particular JavaScript will load a specific page or section of the website as requested which will improve the visibility and interactivity of the website.

2) Product FAQ

Users tend to ask about product specifications on a corporate website. Thus, this knowledge is included to enrich SamBot. Product specification is important since users need to know more about the product and may influence their decision on buying a product. The conversation of product FAQ is illustrated in Figure 3. The `<script>` tag shows the ability of the bots to execute other language. The JavaScript loads a URL in a new window to show the specification of Samsung Galaxy S6 Edge specification. The products included in this knowledge will be smartphones, tablets, wearables, and accessories.

3) AAA Knowledge Bases

Another important knowledge is contained in the AAA (Annotated ALICE AIML) knowledge bases. The AAA knowledge is collection of human interest topic. The AAA knowledge is a free copy of ALICE bots brain which is a three times winner of world-wide chatbot competition with some revision by ALICE Foundation to accommodate current issues. With this knowledge, it will prolong the conversation where most of time, users will like to ask common sense questions. The knowledge includes astrology, emotion, sports, geography, and etc. This knowledge itself contains more than 60,000 categories.

D. SamBot

SamBot is built based on AIML interpreter which is very important in this architecture. AIML interpreter is core of SamBot interprets AIML files, accepts text and voice inputs, responds queries, performs pattern matching, and produces conversation through text or TTS output. The pattern matching algorithm performs depth-first search and returns the first matching response to answer a question. SamBot used

MySQL database to store the knowledge and information including conversation logs. SamBot is design to be able to filter random answers and store it to a database. In addition, SamBot also have capability to learn from users in order to improve its knowledge base.

E. Random Answer

Random answer is basically a response of question which the bots cannot understand. It is generated in order to keep the conversation alive or prolong the interaction. The questions usually contain either new issues or unrelated topic which are not available in the knowledge base. In a fresh installation of the bots, it will generate random answers when the knowledge of questions posed are not found from database and normally it called undefined questions. These random answers taken from randomized single sentence such as "Let me think about it", "Have you tried a web search?", and "Interesting question". This research will enhance the way random answers are generated by adding the answer so that it will persuade users to teach the bots indirectly. Firstly, two groups of answers are made such as passive random answers which are stored in Xfind.aiml and active random answers which are stored in Xlearn.aiml. Passive random answers are the answers which only answer without asking the user to teach the bots, whereas active random answers are the opposite. Active random answers will trigger the usage of consumer-centric learning while passive random answers will trigger the usage of conversation log-based knowledge. These two groups are inserted in `<random>` tag of all undefined question templates. Therefore, every time a user posed an undefined question, the bots will give random response between active and passive random answers. It makes the respond more natural as different users might get different responses.

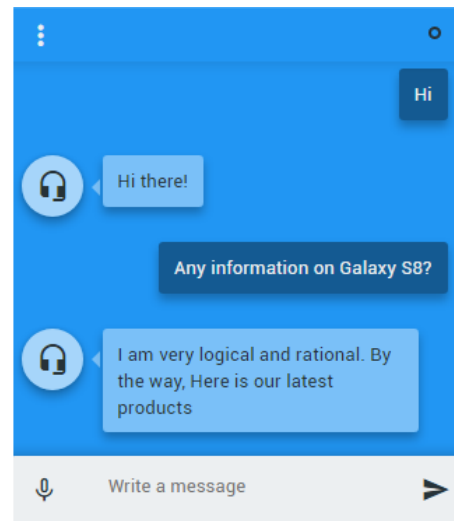


Figure 4(a). Passive Random Answer Scenario

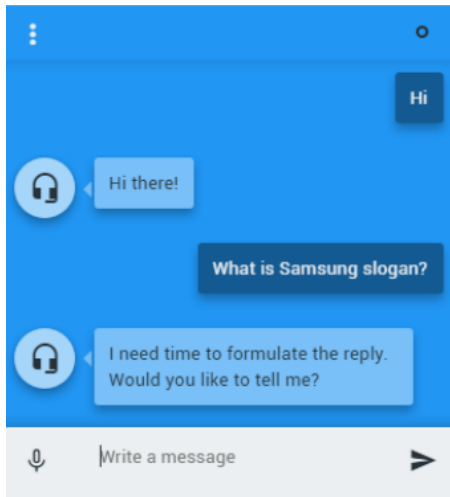


Figure 4(b). Active Random Answer Scenario

This research also enhanced the passive random answers to be more Samsung-oriented by adding Samsung promotion knowledge at second sentence. This enhancement provides better answers by letting users engaged to the offers and promotions rather than just getting boring replies/answers. The enhancement utilizes the <random> tags between the first and the second sentence so there is a random combination between these two. The first sentence contains the original random answer and the second sentence contains the Samsung promotion knowledge. By using the <random> tags the answer can be more fluid and flexible. Figure 4(a) shows a scenario of passive random answer and Figure 4(b) is active random answer after this enhancement.

F. Conversation Log-Based Knowledge Module

This module tries to improve SamBot knowledge base by utilizing its conversation logs. The conversation logs contain users’ dialogue with the bots and by default is sorted by users’ name. By utilizing the logs, botmaster will be able to know what the user expect from the bots. Thus, botmaster can add more knowledge to the bots based on the new issue. This research improved SamBot with new feature to display random answers by creating “Unanswered questions” table. Basically, this function filters the conversation logs to display only passive random answer responses.

Through this table, botmaster can easily choose a question to be answered and update the random answer to be the correct answer. This method is efficient since botmaster can choose which questions are relevant to Samsung domain and only answer the relevant questions. With this feature, botmaster doesn’t need to look into all the logs and read the questions one by one which is time consuming. As the knowledge are updated, SamBot is ready to answer accurately.

G. Consumer-centric Learning Approach

Another important feature in this research is to make the bots learn new knowledge from users. This function is made to complement the active random answers generated by the bots. It will help the development of bots knowledge bases due to a different users might have different knowledge to teach. SamBot able to learn by creating a new AIML file called Learn.aiml which contains possible patterns to let user teach the bots such as <pattern>* IS *</pattern>. The asterisks will

contain phrases which will be saved in the database. Then when one of the asterisk is called in a question, the other asterisk will be produced as a response. As an example, the user input is “The latest Samsung tablet is Galaxy Tab S3” then when the bots is asked “What is the latest Samsung tablet?”, it will answer “Galaxy Tab S3”. In this example, the phrase “The latest Samsung tablet” is on the position of the first asterisk while “Galaxy Tab S3” is on the other one. Therefore, when the phrase on the first asterisk is used as question then the phrase on the second asterisk is used as the answer. It also works in either way in which the user input is “Galaxy Tab S3 is the latest Samsung tablet” and the question is “What is Galaxy Tab S3?” or “What is the latest Samsung tablet?”.

H. Pre-defined Questions Recommendation

Pre-defined Questions Recommendation feature worked similar with Autocomplete proposed by Google. This function is developed to show the questions asked previously by other users. This function generates a list of recent questions which match to the input of the user. This function works on AJAX method so that it can update the list according to the last character typed by the user and matched it to the database. When the list pops up, the user can choose a question from the recommendation list or continue typing to pose different question. With this recommendation, users will easily pose a question without thinking too much. This function is using a modified JQuery autocomplete function. JQuery framework is a feature-rich JavaScript library which can manipulate almost all elements on HTML. The autocomplete function runs in AJAX method and works by matching character-by-character inside a textfield to the existing array of strings. The parameter of matching characters can be set manually and in this research the parameter is set to nine characters. Thus, before nine characters are input, the autocomplete function will not return any recommendation. A modification is to make the array is filled by queries retrieved from the chatbot database, in conversation_log table.

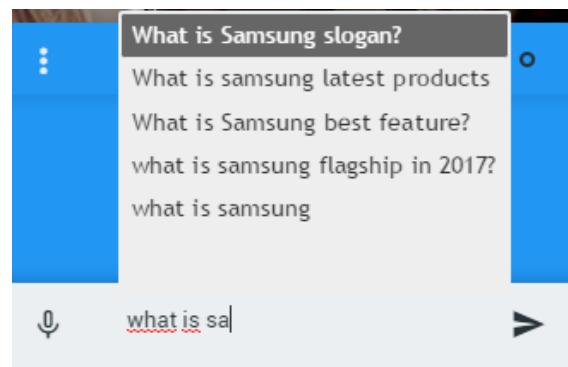


Figure 5. Pre-defined Question Recommendation Scenario

As an example, the questions on previous examples are displayed in the pre-defined questions recommendation list as shown in Figure 5. As mentioned earlier these questions are retrieved from conversation log table which contains all conversation logs from all users. Thus, these recommendation questions reflect the current trend that users concern about.

IV. Evaluation and Results

Our review of the evaluation literature indicates that natural language processing (NLP) systems have largely been evaluated using a black-box, functional, approach. This kind of approach is commonly used in qualitative research of conversational bots [4], [30]. It will test the conversational bots by evaluate the functionality of the system. This test consists of two different evaluations.

```
"936","Hello, my name is Andrew. what is your name?","My name is SamBot.", "2017-10-06 07:27:22"
"937","Why don t you tell me a little more about yourself?","How much time do you have?","2017-10-06 07:27:45"
"938","I am a researcher in Artificial Intelligence at Goldsmiths University, do you know what that is?","Maybe you can give my Botmaster a job.", "2017-10-06 07:28:28"
"939","Will you tell me about your dreams?","I dream that I will become a better undefined.", "2017-10-06 07:28:47"
"940","Can you rephrase that?","I think I can, don't you? That is a very original thought. By the way, <script language=""JavaScript""> window.location.href = \ 'https://samsungiotacademy.com/product/#galaxy7edge\ ' ; </script> Samsung have bold and elegant design.", "2017-10-06 07:29:06"
```

Figure 6(a). Sample of datasets on judges’ questions with SamBot

```
"457","what is samsung?","Samsung is a South Korean multinational conglomerate company headquartered in Samsung Town, Seoul. It comprises numerous subsidiaries and affiliated businesses, most of them united under the Samsung brand, and is the largest South Korean chaebol.", "2017-05-10 13:38:46"
"458","I want to buy samsung J7","Thank you for considering Samsung as your choice. <script language=""JavaScript""> var winObj =window.open(\ 'http://www.lazada.com.my/samsung-galaxy-j7-2016-j710-16gb-black-14314561.html\ ', \ 'winObj\ ', \ 'width=800, height=640\ '); </script>", "2017-05-10 13:39:10"
"459","where is nearest service center in my city","Here you go! <script language=""JavaScript""> var winObj =window.open(\ 'http://www.samsung.com/my/support-t/service/location/\ ', \ 'winObj\ ', \ 'width=800, height=640\ '); </script>", "2017-05-10 13:39:39"
"460","where can i buy samsung galaxy s8","Give me your name and phone number and I'll have someone call you.", "2017-05-10 13:39:59"
"461","what is samsung s8 specification","Here you go! <script language=""JavaScript""> var winObj =window.open(\ 'http://www.samsung.com/my/smartphones/\ ', \ 'winObj\ ', \ 'width=800, height=640\ '); </script>", "2017-05-10 13:40:18"
```

Figure 6(b). Sample of Online Conversation Between Users and SamBot

The first evaluation is based on the Loebner Prize Competition environment where the bots will be asked 20 questions by judges. Therefore, a set of judges’ questions within five years from 2013 until 2017 is collected. These questions are available online on www.aisb.org.uk. This test is aimed to measure the ability of SamBot compared to the others chatbots participated in the competition. The second evaluation is an online user interaction. This test is performed by online users since the system can be accessed online at <https://samsungiotacademy.com>. A total of 100 questions are developed. A number of users were invited to do the test by having conversation with the bots through several rules. The

rules are made to drive the questions related to Samsung and to fulfil the expectation of 100 questions. Users are expected to ask several questions related to Samsung domain and then the collection of responses is scored.

A total of 200 questions and answers from Loebner Prize Competition judges’ questions and from online users have been collected. Figure 6 shows samples of the datasets. These datasets are exported directly from database to CSV files.

Both of the test is scored using the latest Loebner Prize scoring system which can be accessed in <http://www.aisb.org.uk/events/loebner-prize>. This scoring system is chosen because it is used in a world-wide conversational bots competition. The response to each question is evaluated in three criteria such as:

1. **Relevance:** the response should be relevant to the question or in the context of the question. The relevancy ignores correctness.
2. **Correctness:** the response is correctly answered the question either factually or in subjective concern.
3. **Plausibility and Clarity of Expression or Grammar:** The response is grammatically understandable and shows understanding of the question. This criterion is usually partially met when the response cites back the question but has meaning misconception.

Depending on the extent to which the response meets the criterion, 0, 1, or 2 points will be assigned. 0 point is assigned if the criterion is not met at all by the response. 1 point is assigned if the criterion is partially met by the response. 2 points are assigned if the criterion is fully met by the response.

A. Evaluation 1: Black-box Loebner Prize Questions

This evaluation is set similarly to Loebner Prize Competition. Judges question sets from 2013 until 2017 are collected in which each year has 20 questions. Each of the questions are input to SamBot and each pair of the question and answer are recorded in the conversation logs. After all of the questions are answered, the answers are scored. SamBot’s final score in each year is then compared to the score of Loebner Prize competition chatbots. The comparison of each year includes the score of the winner chatbot (highest score), the average score of all chatbots, the lowest score, and SamBot’s score. The comparison of the scores is shown in Figure 7.

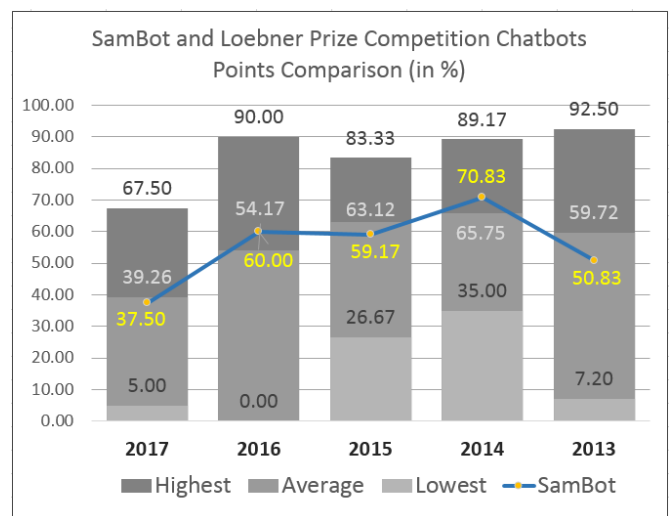


Figure 7. Evaluation 1: Scoring Result

Based on the chart, we concluded that SamBot stands close with the average scores in every year. Even in 2016 and 2014 SamBot collects more points than the average scores with 60.00% over 54.17% and 70.83% over 65.75% respectively. It also clearly shows that SamBot is better than the other chatbots with the lowest scores in each year. Even SamBot is close enough with the winner of each year in which the closest gap is 18.34% in 2014. This result is positive for a SamBot that is not intentionally made for the competition. In conclusion, SamBot is capable in handling common knowledge questions since the questions of Loebner Prize Competition are considered as common knowledge.

B. Evaluation 2: Online Users Interaction

The evaluation result is retrieved from conversation log database. Questions which are in form of greeting, compliment, or gratitude are excluded. Then a total of 100 questions and answers are retrieved. Each answer is scored based on the scoring system stated earlier.

From the result shown in Figure 8, can be seen that most of the responses meet at least one of the criterion with total of 62 questions. This means that SamBot is capable of answering wide variety of Samsung related questions including Samsung stores, profile, and products. Furthermore, it shows that the knowledge creation of Samsung promotion and Samsung product FAQ has successfully enhanced the bots to be a domain specific bots. Only eight responses which partially met the criterion which makes SamBot collects 132 points and average of 26.4. Although there are 30 questions which are not answered properly, it might show the usage of the enhanced random answers. When passive random answers appear, the users are engaged with Samsung promotion so even though the bots cannot answer, it still responds in a good way by trying to persuade users with Samsung promotion. On the other hand, when active random answers appear, the users may teach SamBot new knowledge.

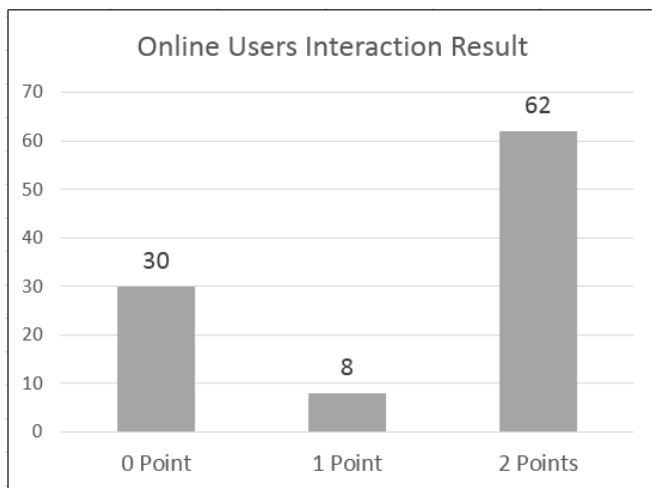


Figure 8. Evaluation 2: Scoring Result

C. Comparison of Evaluation 1 and Evaluation 2

Since SamBot is purposely made to be a domain specific chatbot, it is necessary to have the ability to answer both common knowledge and domain specific knowledge. Therefore, a comparison between the result of Evaluation 1 and Evaluation 2 is made. This comparison shown that SamBot is capable interact with Loebner Prize datasets with

100 questions from judges and from the 100 questions from online users.

In Figure 9, it is clearly shown that the result of Evaluation 2 is better that Evaluation 1 result. It has less questions improperly answered with 41 over 30 and more questions properly answered with 38 over 62. In total, SamBot collected 97 points in Evaluation 1 compared to 132 points in Evaluation 2. This comparison shows that SamBot is more capable of handling Samsung related questions than common knowledge questions. Therefore, it can be concluded that the enhancements being made have shown positive results towards domain specific chats/queries.

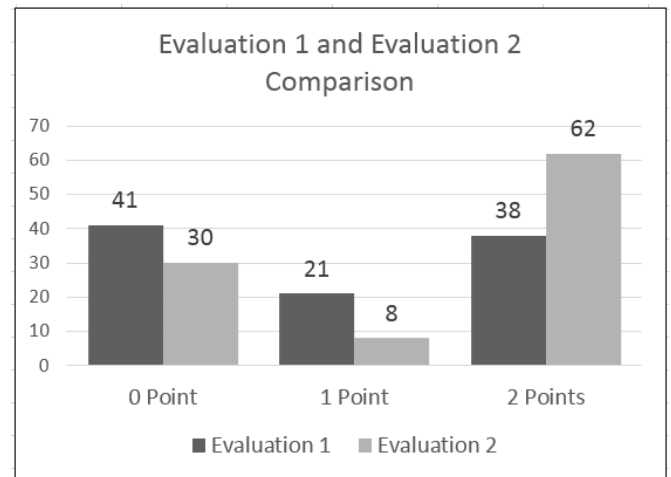


Figure 9. Comparison of the Evaluation 1 and Evaluation 2

V. Discussion

From the evaluation conducted, it can be seen that the results reflect the versatility of SamBot where it is able to response both general knowledge and domain specific knowledge. This ability shows that the design of the system architecture target has been achieved. The knowledge base enhanced with AAA knowledge has made SamBot stood more than average among other chatbots while the Samsung knowledge has helped SamBot in performing well in Evaluation 2. The result in Evaluation 2 also depicted the performance of other enhancements such as Random Answer enhancement, Consumer-centric Learning Approach, and Pre-defined Questions Recommendation. Random Answer enhancement has made the 30 questions of 0 point become more interactive. Both Consumer-centric Learning Approach and Pre-defined Questions Recommendation helped 62 questions of 2 points more likely to ask the same question of previous user and keep the questions related to Samsung domain.

VI. Conclusion and Future Work

This research proposed a new way to improve marketing aspect of a corporate website as well as its interactivity. Corporate website is chosen as an object in marketing domain since previous research still struggles to get the attention or engagement from online users. In order to improve it, a conversational bot called SamBot is developed. SamBot is design to be deployed at the Samsung IoT Academy Showcase website as the corporate website. From the blox-box evaluation, the consumer-centric learning approach also comes in play when the user is not satisfied with the answer of SamBot and enriches its knowledge bases. The pre-defined

question recommendation is proven to be useful in deciding the questions to be posed. These successes were proven from the evaluation carried out using international competition Loebner Prize standard and real online user interaction with Sambot. The results reflect that SamBot is capable of handling common knowledge questions which is proven by the scoring collected in Evaluation 1 as well as highly capable of handling Samsung domain questions which is proven by the scoring collected in Evaluation 2.

However, based on this research, there are several possibilities to improve and further continue this research. One biggest concern is to improve the knowledge creation to be automated using generative models of deep learning on big conversation log data or any other big data resources. It will make the knowledge creation are more objective, automated and have larger knowledge bases. Another possibility is to implement conversational bots in other domain such as politics, hobbies, history, and many other untouched domains other than marketing domain.

Acknowledgment

This research work is supported by Samsung IoT Academy, Universiti Teknikal Malaysia Melaka and grant from Ministry of Education, Malaysia (PJP/2015/FTMK(5B)/S01438), and scholarship from Universitas Dian Nuswantoro, Indonesia.

References

- [1] R. Zollet, "Interactivity of Corporate Websites: An Integrative Review of the Literature," *IEEE Trans. Prof. Commun.*, vol. 57, no. 1, pp. 2–16, 2014.
- [2] P. Anderson, M. Hepworth, B. Kelly, and R. Metcalfe, "What is Web 2.0? Ideas, Technologies and Implications for Education," *JISC Technol. Stand. Watch*, vol. 60, no. 1, p. 64, 2007.
- [3] S. Russell and P. Norvig, *Artificial intelligence—a modern approach*, Prentice Hall. Series in Artificial Intelligence, Englewood Cliffs, NJ., 3rd ed. Englewood Cliffs, New Jersey: Prentice Hall, 2010.
- [4] O. S. Goh and C. C. Fung, "An Intelligent Personal Robot Assistant," no. November 2014, 2016.
- [5] K. W. Wong, L. T. Koczy, T. D. Gedeon, A. Chong, and D. Tikik, "Improvement of the cluster searching algorithm in Sugeno and Yasukawa's qualitative modeling approach," *Comput. Intell. Theory Appl. Proc.*, vol. 2206, pp. 536–549, 2001.
- [6] A. Khan, N. Salim, and Y. Jaya Kumar, "A framework for multi-document abstractive summarization based on semantic role labelling," *Appl. Soft Comput. J.*, vol. 30, pp. 737–747, 2015.
- [7] C. L. Officer, "Chatbots for Learning," 2016. [Online]. Available: <http://www.clomedia.com/2016/09/12/chatbots-for-learning/>. [Accessed: 20-May-2017].
- [8] A. S. Lokman and J. M. Zain, "Extension and prerequisite: An algorithm to enable relations between responses in chatbot technology," *J. Comput. Sci.*, vol. 6, no. 10, pp. 1212–1218, 2010.
- [9] ALICE AI Foundation, "AIML Implementations." [Online]. Available: <http://www.alicebot.org/downloads/programs.html>.
- [10] O. S. Goh and S. L. Lim, "Intelligent Conversational Bot for Massive Online Open Courses (MOOCs)," 2016.
- [11] F. A. Mikic, J. C. Burguillo, M. Llamas, D. A. Rodríguez, and E. Rodríguez, "CHARLIE: An AIML-based Chatterbot Which Works as an Interface Among INES and Humans," in *20th EAAEIE Annual Conference, EAAEIE 2009 - Formal Proceedings*, 2009.
- [12] D. Pietro Orlando and F. Giovanni, "An Integrated System, with Natural Language Management, for the Monitoring Activities in E-learning Environments," in *Proceedings - CISIS 2008: 2nd International Conference on Complex, Intelligent and Software Intensive Systems*, 2008, pp. 965–970.
- [13] S. Ghose and J. J. Barua, "Toward the Implementation of a Topic Specific Dialogue Based Natural Language Chatbot as an Undergraduate Advisor," in *2013 International Conference on Informatics, Electronics and Vision, ICIEV 2013*, 2013, pp. 0–4.
- [14] R. P. Mahapatra, N. Sharma, A. Trivedi, and C. Aman, "Adding Interactive Interface to E-Government Systems Using AIML Based Chatterbots," in *2012 CSI 6th International Conference on Software Engineering, CONSEG 2012*, 2012.
- [15] S. M. Rodrigo and J. G. F. Abraham, "Development and Implementation of a Chat Bot in a Social Network," in *Proceedings of the 9th International Conference on Information Technology, ITNG 2012*, 2012, pp. 751–755.
- [16] O. S. Goh, C. C. Fung, K. W. Wong, and A. Depickere, "Embodied Conversational Agents for H5N1 Pandemic Crisis," *J. Adv. Comput. Intell. Intell. Informatics*, vol. 11, no. 3, pp. 1–7, 2007.
- [17] C. Chakrabarti and G. F. Luger, "Artificial Conversations for Customer Service Chatter Bots: Architecture, Algorithms, and Evaluation Metrics," *Expert Syst. Appl.*, vol. 42, no. 20, pp. 6878–6897, 2015.
- [18] N. T. Weerawarna, H. M. H. R. B. Haththella, A. R. G. K. B. R. Ambadeniya, L. H. S. S. Chandrasiri, M. S. L. Bandara, and S. S. Thelijjagoda, "CyberMate : Artificial Intelligent Business Help Desk Assistant with Instance Messaging Services," in *2011 6th International Conference on Industrial and Information Systems, ICIIIS 2011 - Conference Proceedings*, 2011, pp. 420–424.
- [19] M. Moradi, A. Aghaie, and M. Hosseini, "Knowledge-collector Agents: Applying Intelligent Agents in Marketing Decisions with Knowledge Management Approach," *Knowledge-Based Syst.*, vol. 52, pp. 181–193, 2013.
- [20] C. Torrey, A. Powers, M. Marge, S. R. Fussell, and S. Kiesler, "Effects of Adaptive Robot Dialogue on Information Exchange and Social Relations," *1st ACM SIGCHI/SIGART Conf. Human-Robot Interact. (HRI '06)*, pp. 126–133, 2006.
- [21] O. S. Goh, C. Ardil, W. Wong, and C. C. Fung, "A Black-box Approach for Response Quality Evaluation of Conversational Agent Systems," *Int. J. Comput. Intell.*, vol. 4, no. 6, pp. 26–34, 2007.

- [22] G. Pilato, A. Augello, G. Vassallo, and S. Gaglio, "Sub-symbolic Semantic Layer in Cyc for Intuitive Chat-bots," in *ICSC 2007 International Conference on Semantic Computing*, 2007, pp. 121–128.
- [23] O. S. Goh and C. C. Fung, "Trust Based Knowledge Acquisition for Conversation Agents," *IAENG Int. J. Comput. Sci.*, vol. 33, no. 2, pp. 1–9, 2008.
- [24] G. Pilato, A. Augello, and S. Gaglio, "A Modular Architecture for Adaptive ChatBots," *2011 IEEE Fifth Int. Conf. Semant. Comput.*, pp. 177–180, 2011.
- [25] WildML, "Deep Learning for Chatbots, Part 1 - Introduction," 2016. [Online]. Available: <http://www.wildml.com/2016/04/deep-learning-for-chatbots-part-1-introduction/>. [Accessed: 20-May-2017].
- [26] H. Al-Zubaide and A. A. Issa, "OntBot: Ontology Based ChatBot," *2011 4th Int. Symp. Innov. Inf. Commun. Technol. ISIICT'2011*, pp. 7–12, 2011.
- [27] Y. Wu, G. Wang, W. Li, and Z. Li, "Automatic Chatbot Knowledge Acquisition from Online Forum via Rough Set and Ensemble Learning," *Proc. - 2008 IFIP Int. Conf. Netw. Parallel Comput. NPC 2008*, pp. 242–246, 2008.
- [28] O. S. Goh, "A Framework and Evaluation of Conversation Agents," 2008.
- [29] "Search Using Autocomplete." [Online]. Available: <https://support.google.com/websearch/answer/106230?hl=en>. [Accessed: 20-Oct-2017].
- [30] S. J. du Preez, M. Lall, and S. Sinha, "An Intelligent Web-based Voice Chat Bot," in *IEEE EUROCON 2009*, 2009, pp. 386–391.

both his bachelor degree and master degree from Universiti Sains Malaysia (USM), in the field of Mathematical Science in year 2003 and 2005. He completed his PhD studies at Universiti Teknologi Malaysia, in 2014 in the field of Computer Science. Currently, his research involves in the field of Text Mining, Information Extraction and AI applications.

Author Biographies



Aditya Pradana received his Bachelor of Computer Science (B.Sc) from Universitas Dian Nuswantoro (UDINUS), Indonesia in 2015. In 2017 he received his Master degree in Software Engineering and Intelligence from Faculty of Information Technology and Communication, Universiti Teknikal Malaysia Melaka (UTeM), Malaysia. His research interests are in intelligent agent and data mining. Currently he is working at Samsung IoT Academy UTeM as a web application developer.



Prof. Dr Goh Ong Sing started his career as an academician since 1990 at Universiti Sains Malaysia (USM). In August 2002, he joined Universiti Teknikal Malaysia Melaka (UTeM) and in 2015 he was appointed as a Director of Industry Liaison Centre and Lead Samsung IoT Academy. Now he is working as Assistant Vice Chancellor at Office of Industry and Community Network. His main research interest is in the development of intelligent agent, natural language processing and speech technology to facilitate graceful human-computer interactions, conversational robot and mobile services. He is the author of more than 100 peer-reviewed scientific journal, books, conference papers, and book chapters. He has led and worked on research grants funded by Malaysian Government's Intensified Research, Malaysia Technology Development Corporation, Big Data Technology Pty Ltd, Murdoch University, and Australian Research Centre for Medical Engineering at University of Western Australia.



Dr Yogan Jaya Kumar is a Senior Lecturer at the Department of Intelligent Computing and Analytic in the Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM). He earned