ECAI 2018 - International Conference – 10th Edition
Electronics, Computers and Artificial Intelligence
28 June -30 June, 2018, Iasi, ROMÂNIA

Chatbot Application using Search Engines and Teaching Methods

Petre Anghelescu, Stefan Vladimir Nicolaescu
Department of Electronics, Communications and Computers
University of Pitesti
Pitesti, Romania
petre.anghelescu@upit.ro, nicolaescu.vladimir@gmail.com

Abstract – This paper is focused on the design and the implementation of a Chabot system using search engines and teaching techniques. Chatbots are imitating human natural conversations interacting with diverse people, learning things, trying to make us believe that it is a human person behind the screen. Nowadays artificial intelligence (AI) methods are a lot more advanced than few years ago, things that help us in our day to day activities with many high level techs that make our lives easier. To perfectly simulate a human dialogue, the bot must analyze the input given by the user correctly and formulate a response that should be relevant and appropriate. The Chabot was realized as a website application based on search engines, custom learning, custom cascading style sheets and JavaScript for making conversations as real as possible.

Keywords - chatbot, knowledge base, teaching methods, AI search

I. INTRODUCTION

"A computer would deserve to be called intelligent if it could deceive a human into believing that it was human." – Alan Turing [1]

In today's connected world, things evolve very rapidly and the "people" that interact online are not all human beings. The chabots are everywhere and we have to recognize that in the near future almost all the techs things will have a bot that can interact with people in many kinds. Nowadays speech interaction between bots and humans has received a lot more interest from the biggest companies in the wide world such as Google, Android and IOS. In [2] are presented various chatbots implemented in time. A chatbot application will be used in many fields in the future and this is almost predictable because of the very high speed of improving technological methods for easier and better lives of the human being. Even industrial robots are being prepared to start interacting with people that use them. Our complicated lives will become more comfortable because the bots will do a lot of things for us and the things we could do in a lot of time will now be done by a bot instead of us. It is time to learn more about this great feature which already exists in our phones, PCs, lives. This paper covers the techniques used to make our assistant, a chatbot, which is becoming more intelligent with every question that people ask and with every person that teach it the answer to these questions, when it does not know them. The paper is organized as follows. The following section presents basic theoretical foundations of this research work. We present some basics of AI chatbots applications. Section III presents the implementation of the proposed chatbot and the technology used. Section IV contains testing and results obtained and section V concludes the paper and presents future research directions goals.

II. BASICS OF CHATBOTS APPLICATIONS

A. The AI speech interaction

First of all, the artificial intelligence chatbot means more than some databases and creativity, it means more than one person's thoughts. A good overview of chatbot's design techniques is presented in [3].

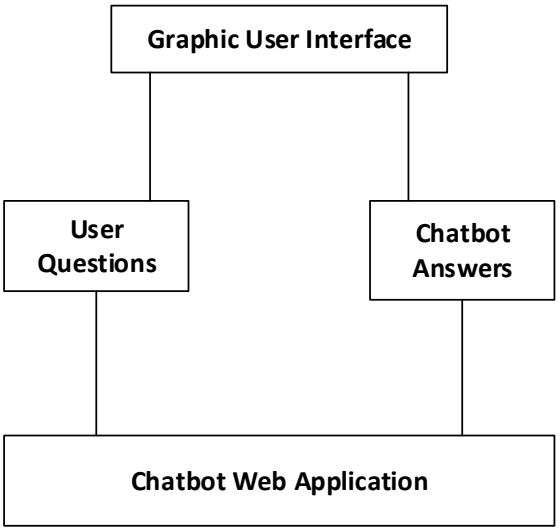

Figure 1. The diagram of user and chatbot interaction

To start with, as you can see in Fig. 1, the interaction between the user and the chatbot is about asking and answering. An artificial intelligence chatbot is based on different levels of questions and answers. The questions could be personal and, in this case, the intelligence of the machine should appear, in order to answer as a real human. There are also questions that the chatbot can easily answer by searching on the World Wide Web using search engines like Google and Wikipedia. It returns the essential information found on these search engines to

978-1-5386-4901-5/18/$31.00 ©2018 IEEE

create a very good and short answer. The search engines, are very useful when it comes to the normal questions. But you have to teach it how to answer personal questions by giving him the answers. Or you can make it act as it does not know all the answers. Our chatbot, when it does not know what to answer, it has the possibility to return the following message: "I do not know!" or "I do not know the answer to this question, please teach me!". This is the moment when we, in the position of administrators, give the good answers to the questions it does not know or approve the answers received from others. It is like you grow up a baby and you have to teach it a lot of things. Moreover, it comes a time when it should study by searching on the search engines. It is clear that it is not so easy to make your own cyber "child", to teach it things, to make it search on the web by itself and other smaller things than these.

B. When it becomes personal

There are some moments in the conversations when the questions are personal. These are the keys to simulate a human being. Personal questions reflect the personality of a human and this should trick people, making us think that the chatbot can really be a human. Things are very complicated to be in the best position, but the words used by the chatbot and the sarcasm or the irony, fool the people. That makes it intelligent even if you teach it those answers. When you have the opportunity to learn a machine something, it is a good time to be creative, to think more about what other people will think when they will interact with the chatbot. The personal questions do not just help the chatbot to act like humans but they are needed to improve artificial intelligence of a machine too. That's why, making a panel for administrators helps you control what happens with the database itself and change answers and even questions as you consider suitable. To end with, go for the most personal questions you can ask, to find out if it is a chatbot or a human. This will help both the machine's creator and yourself.

C. The normal questions

An artificial intelligence based machines should always be connected to the internet, more precisely, to the IoT (Internet of Things). The databases and more other internet things, should always be up-to-date, it will help the chatbot a lot. When it comes to normal questions, meaning the "simple" questions, the machine does not need you too much to answer them. And this is the moment when search engines help a lot. The chatbot makes some searches and returns the best answer it can get from searching engines. Google and Wikipedia are the biggest search engines where you can find the most daily searched things. These are good tools to search on the normal questions and return the best answers. You can base on these search engines a lot because they have the biggest databases in the world and the connectivity between websites and the functionality to search whatever you need to find. These are day-to-day questions that the chatbot can answer easily in every conversation that you will have with it. Of course, you can alternate every questions you want to ask the chatbot and it will return the answers it gets from the databases. As the personal ones, you have the possibility in the administrator panel to change any answer you think it is not returned as you wish, or even the question. This is a good feature thinking about some people that just want to have fun and ask bad things. Having the badwords table filter and this, you can manage easily all the chatbot questions and answers.

D. Get into the databases

The more questions asked, the more answers you get.

Figure 2. The answers and questions relation

Both questions and answers are saved in the databases, as you can notice in Fig. 2. Bigger databases, more questions with at least an answer known. Searching on the web is a good fact for a chatbot because you will search in other databases, such as Google, which will help the improvement of performance and the answering speed. That means you will not save all the questions and the answer in your database because you will make the application slower. Above this fact, searching on the web will always bring you the newest things and the updated ones. For example, the president of a country is changed once in some years and the local database could not be updated. This kind of errors will definitely lower the usage of your application because people will not trust your product anymore. The conclusion of this is stay live and updated all the time for a better application and for happy customers.

Figure 3. Keywords and badwords tables

To continue with the database, we create two more tables for keywords and badwords, as you can see in the Fig. 3. These are used as filters. The keywords table is used for searching on the internet. There is a word by word search in proposition after the keywords. If any of them are found in the proposition, the question is asked on the search engines. Keywords such as what, where, who and weather, are used to search on the web the "normal" questions. The first three of them are going to Google and Wikipedia search engines, and the last one is going to an API called Open Weather Map [9], which

returns the answers to what is the weather like in somewhere questions.

The badwords table is for the people who are having fun, trying to learn chatbot bad words. This is a manual bad language filter which helps you keep the databases without any vulgarity. This is searching word by word in propositions and where it founds a word from the table it returns the message: "You used bad language!". Then it requests you to ask something else without any badwords.

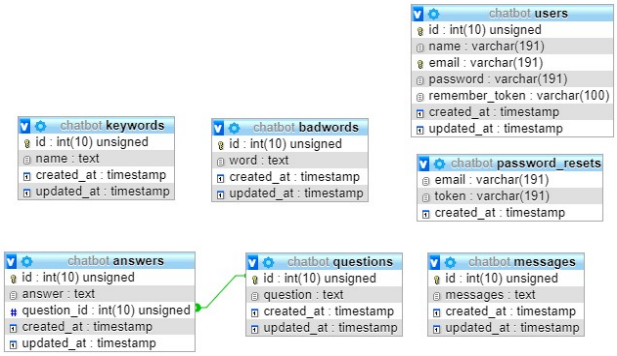

Figure 4. The database tables

As you can see below in Fig. 4, in addition to answers, questions, keywords and badwords, there are other tables used for registration of new people coming to the website. You have to register to use the chatbot entirely. And the password_resets table is used by the users that lost their passwords. Having users registered is a good feature for a chatbot website that needs the databases to grow.

The messages table is used to save message by message, user-chatbot conversations. This can be used to find the preferences of a user and you can give him the possibility to see the last conversations as on messenger.

III. IMPLEMENTATION PROCESS

The chatbot is realized as a website application based on search engines, custom learning, custom cascading style sheets and JavaScript for making conversations more real.

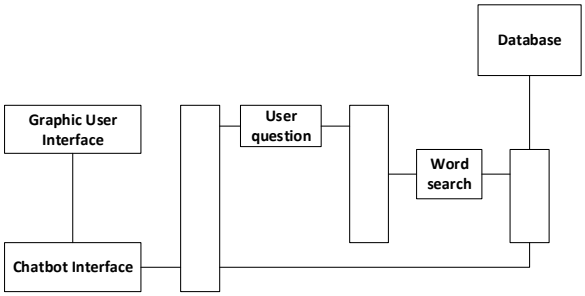

Figure 5. Diagram representing design of the Chatbot

In this part of the article is explained how to make a chatbot, what technologies are used and how everything is combined together for a smart application that shows us the level of tech nowadays.

A. Laravel, database and other technologies used

Firstly, Laravel PHP framework is good to be used because it is simple and topical. The coding languages used here are: PHP, HTML, CSS, JS, MySQL, Bootstrap, Ajax and JSON. These are the elements that mixed together can get you to the next level of technology on making websites. They are continually growing, leveling up, updating to newer versions, so you can make an application without being afraid that these coding languages will be deprecated. You can find documentations freely on the internet with any costs and about everything you need to do.

1) Laravel – the framework, itself

The PHP Framework for web artisans, as it is presented in [4]. This is the motto they are using on the first page of their website. We all love beautiful code, they do to, so let's do it. It has beautiful syntax and it is expressive. This framework is good even if you are coding solo or for a team with 20 people. They can all sync together. An amazing ORM, painless routing, powerful queue library, and simple authentication give you the tools you need for modern, maintainable PHP.

2) Hypertext Preprocessor

PHP is a widely-used open source for general purpose scripting language that is especially suited for web development and can be embedded into HTML [5]. Instead of many commands to output HTML, PHP pages has HTML with embedded code that can do "something". What is different to PHP from websites using client-side JavaScript is that the code is executed directly on the server, generating the HTML code which is sent to the client. He would get the final results after running that script, but he will not know what the underlying code was. You can process all your HTML files just by configuring the web server to do that and then there is no chance that users can tell what you have up your sleeve. PHP's good things are that it is very simple even for a newbie and offers a lot of advantages for a senior programmer.

3) HyperText Markup Language

HTML is the language used to create simple web pages that can be shown in a browser. The scope of HTML is more to present the information: paragraphs, fonts, tables and other kind of text. The specs of HTML are according to the World Wide Web Consortium. HTML, CSS and JavaScript form a triad of cornerstone technologies for the World Wide Web. Browsers are getting the HTML documents from a web server or from a local storage and render the documents into multimedia web pages.

4) Cascading style sheets

CSS is a style sheet language which is used to describe the presentation of a document written in a markup language like HTML. The styles can be attached to the elements of HTML by some extern

files or by the document itself. You can recognize the element of CSS coding start by the tag <style> or the attribute style. CSS can be used even for formatting the elements of XHTML, XML and SVGL. CSS3 represents an upgrade and it brings some new features that develop new concepts in web design.

5) JavaScript

JS is a coding language that is object oriented, prototype-based. It is used more for some websites functionalities. JavaScript coding is rendered by the browser. It is known that JS is used for the construction of websites, but it also has access to embedded objects from other applications. It was firstly developed by Brendan Eich from Netscape Communications Corporation who firstly called it Mocha, then LiveScript, and then finally JavaScript.

6) MySQL databases

MySQL is one of the most popular SGBD open-source at the moment, being an important feature for the LAMP (Linux, Apache, MySQL and PHP). Even if it is used frequently with coding language PHP, MySQL can be used to build a variety of other applications with different coding languages. MySQL is an integrated component for platforms like LAMP and WAMP. In many specialty books it's mentioned that MySQL is easier to learn and use than many other SGBD.

7) Bootstrap

This part is for web design mostly, as it is presented in Fig. 6. Bootstrap is an open source toolkit for developing with HTML, CSS and JS. If you want to build a responsive application, to have mobile responsiveness on any screen size, this is world's most popular front-end component library. Quickly prototype your ideas or build your entire applications with their Sass variables and mixtures, responsive grid system, extensive prebuild components and powerful plugins built on jQuery [6].

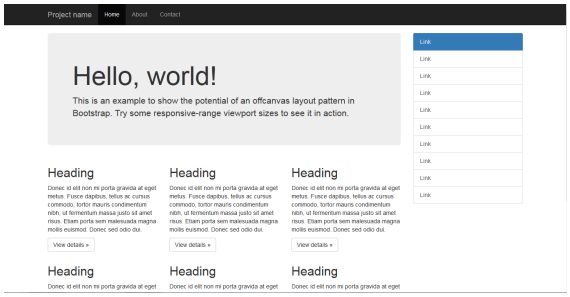

Figure 6. Web design application

8) Ajax and JSON

Ajax is the abbreviation for Asynchronous JavaScript and XML. It is a good technique to code with for the building of interactive websites applications. The intention of it is that the websites become faster and more acceptable by changing in background some small information with the server. This means that the client should not refresh the web page every time he did something. Ajax has the goal to grow the interactivity, to make them faster and to use easily the websites [7]. Ajax is not a technology by itself. The term is used to define the websites that use a set of technologies. Ajax is used for the transfer of data between the server and the database to the client, as it is presented in Fig. 7.

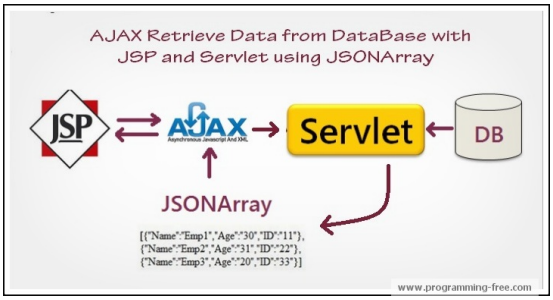

Figure 7. The relations between Ajax, JSON and database

JSON is the abbreviation for JavaScript Object Notation that it is a format for representing and interchanging the data between informatics applications. In Fig. 8 you can see that it has a clear for people text format, used to represent objects and other structures of data and it is used mostly to send structured data through the network, and the process is called serialization. JSON is the simple and easier alternative of XML coding language [8]. The elegance format of JSON comes from the fact that it is a subset of the JavaScript, being used along to this coding language.

```
{
  "firstName": "John",
  "lastName": "Smith",
  "age": 25,
  "address": {
    "streetAddress": "21 2nd Street",
    "city": "New York",
    "state": "NY",
    "postalCode": "10021"
  },
```

Figure 8. JSON code design

B. Creating the first three free questions feature

When it comes to this amazing project, you can think about both commerce and chatbot in a single application. Firstly, the chatbot welcomes you to the website, then it asks you "How are you today?". This is the point where you can start asking yourself: "is it human or not?". After this, the bot will show you a notification message where it says that you can ask it three questions, as it is shown in Fig. 9. After these three questions you have to login or register if you are not already a user. Once you are logged in, there is a panel where you have three possibilities to take: three questions again because it is other design, chat with the bot like on text messages or teach him something that it does not know. As it was written a couple of rows below, both commerce and the chatbot were used in this application. You have a free trial that means 30 days of free use for everything on the website. Then you have to pay some money to extend the trial period.

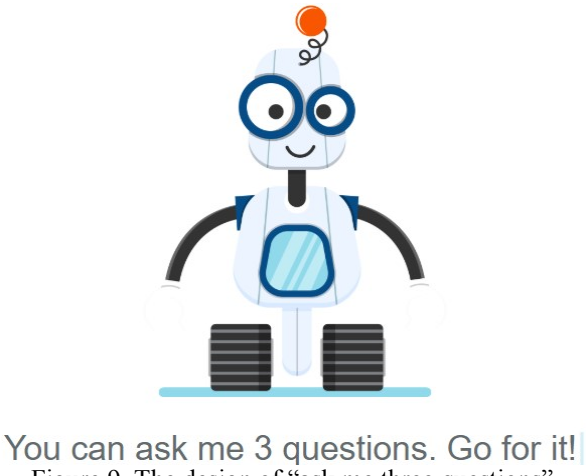

Figure 9. The design of "ask me three questions"

C. *The chat itself*

The chat developed is almost like Messenger from Facebook and you can talk with the chatbot whenever and whatever you want, as Fig. 10 shows.

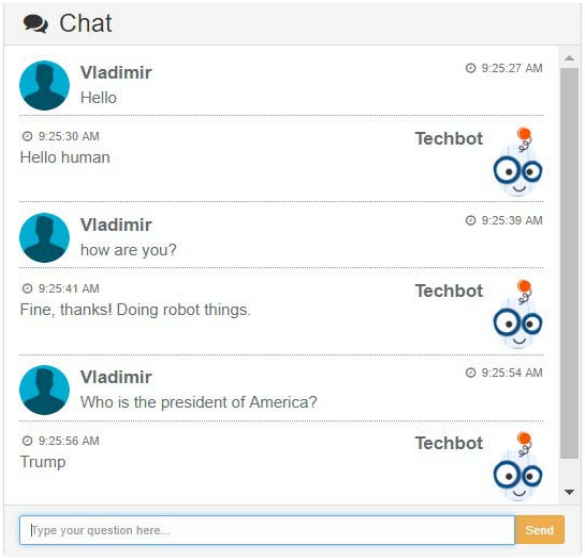

Figure 10. The design of the chat

The answers of the chatbot have a random delay because it seems to reproduce a chat with a human indeed. This is one of the things that is making a realistic chatbot. Another thing is that when the bot "thinks" what is the answer to your question, the three dots of writing appears and makes you think that someone is writing behind the screen, as you can see in Fig. 11.

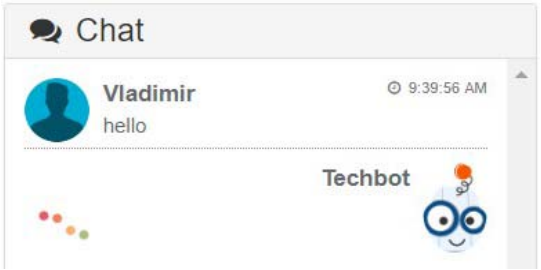

Figure 11. The design of the three dots delay

You can ask the chatbot every question you want to know even if they are personal, historical, musical, weather, or everything you want to find out. It will answer to you even if it knows or does not know the answer because the chatbot is your loyal assistant.

D. *The teaching method*

Of course being a machine is hard to know all the answers in the world but the hardest thing is to make him respond to personal questions. And this is the moment where we can step in and teach the "child" the best answers we can give to him. This is hard too because the answers have to be a little bit tricky to fool people, making them think it is a real person. We can make it our as-we-wish loyal chat partner. When you ask it a question that it does not know the answer to, it brings you an input to write the correct answer. After this step, it thanks you for your support, and indeed it helps the application a lot because you grow the database and it will know what to answer the next time you ask it the same question. You can see all this stuff below, in Fig. 12.

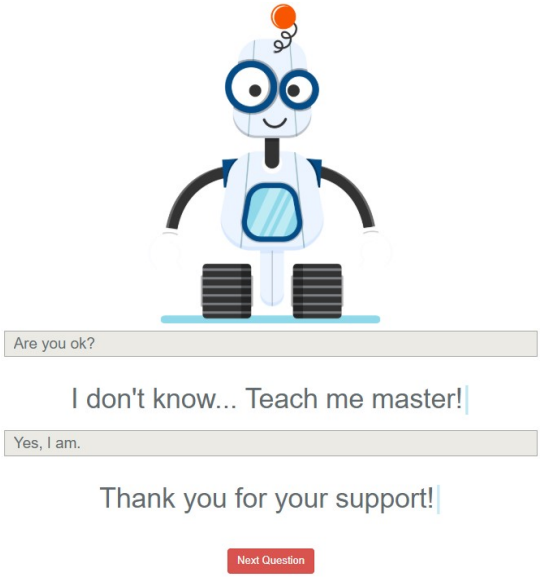

Figure 12. The design of the teach feature

E. *Some other features integrated*

One of them, when you use the teaching method, it is to build a table list with swearing words for people that want to learn the chatbot bad language, as it is explained above. The transitions and animations between the questions and answers are made with CSS and JS for a more realistic application that tricks a little bit people's minds. All the images used are edited in Photoshop and Illustrator and the user register and login are coming with Laravel software.

For the weather questions it is used an API from Open Weather Map which returns the exact degrees for any places in the world [9].

IV. TESTING AND RESULTS

Once you start to create the chatbot, the tests are inevitably every time you finish a feature because it has to be tested to see if it works or not. For this

reason, it is necessary to make a lot of tests before getting the results you want. Ask the chatbot different questions and try to get the best answers from Google or Wikipedia for normal questions, from Open Weather Map for the weather all over the world and from yourself for the personal questions.

Secondly, the results are great for the work that was done, of course it is not very intelligent without a big database of answers but the time will resolve all the problems. More users, more answers and in this way bigger databases.

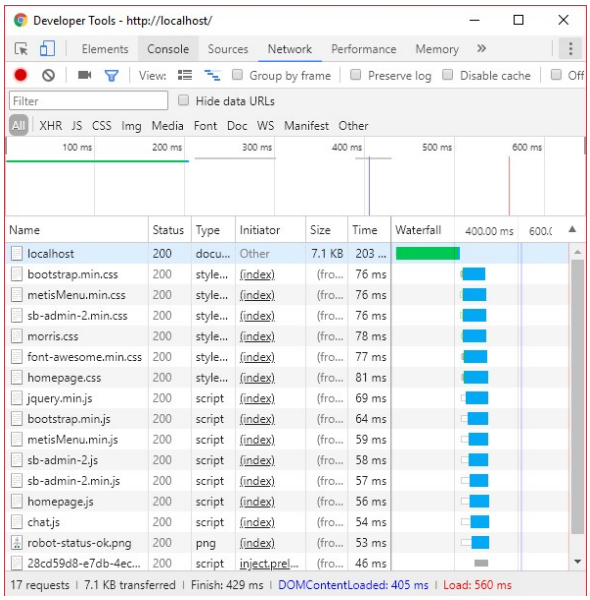

Figure 13. The website analyzed results

We have to mention that the tests presented in Fig. 13 were made on the actual progress of the chatbot application. For the actual source code, the results were encouraging and with no errors.

V. CONCLUSION AND FUTURE RESEARCH

We have proposed in this paper a Chatbot system that imitating human natural conversations, interacting with people, learning things, and trying to make us believe that it is a human person behind the screen. This work will grow up with every user that will use the application, and this is why we have to make something user friendly in order to attract a lot of people that will help tables from the chatbot database to be full of questions and answers.

The development in this field has started, the base artificial intelligence is already implemented and is growing every day and this is the thing that should make us confident that it will be greater in the near future.

REFERENCES

[1] https://www.brainyquote.com/quotes/alan_turing_269231

[2] Aditya Deshpande, Alisha Shahane, Darshana Gadre, Mrunmayi Deshpande, Prof. Dr. Prachi M. Joshi, "A survey of various chatbot implementation techniques", International Journal of Coumputer Engineering and Applications, Vol. 11, ISSN 2321-3469, 2017.

[3] Sameera A. Abdul-Kader and Dr. John Woods, "Survey on Chatbot Design Techniques in Speech Conversation Systems", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 6, No. 7, 2015.

[4] Vishal V. Parkar, Prashant P. Shinde, Sanket C. Gadade, Prasad M. Shinde, "Utilization of Laravel Framework for Development of Web Based Recruitment Tool", IOSR Journal of Computer Engineering (IOSR-JCE), e-ISSN: 2278-0661, p-ISSN: 2278-8727, pp. 36-41, www.iosrjournals.org, 2018.

[5] Official website information: http://php.net/manual/en/preface.php

[6] Official website: http://www.getbootstrap.com

[7] Jang, Sang Dong "Performance Analyses of Web Page Model Using Ajax Technologies", Advanced Science Letters, Vol. 22, No. 11, PP. 3437-3440(4), 2016.

[8] Zhang Yu, "Research of Conversion Method of Entity Object and JSON Data", The 2nd International Conference on Computer Application and System Modeling, published by Atlantis Press, France, 2012.

[9] Official website: http://www.openweathermap.org