International Journal of Engineering & Technology, 7 (3.7) (2018) 274-277

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

The Use Of Voice in Opac Using Google API Voice Recognition & Speech Synthesis and Fullproof Algorithm as Faster Searching Device

Maisyatus Suadaa Irfana

Faculty of Vocational Studies, Universitas Airlangga Surabaya, Indonesia
**Corresponding author Email: maisyah_irfana@vokasi.unair.ac.id*

Abstract

The development of technology provides convenience to all human beings including public librarians. OPAC becomes a solution for library users to search books in a short time with accurate data in which they previously used a manual system. However, there are some users, especially users with certain disabilities, such as users who cannot see, are unable to use OPAC. The idea came up to make OPAC with voice as a substitute for another alternative of text input use. OPAC with voice has been used in several big libraries in Indonesia. One of them is in State Islamic University of Sunan Kalijogo. However, because there are still some shortcomings in terms of voice recognition in which one of them is the limited number of recognizable languages, so in this study, it will be discussed the use of Google API for voice recognition and voice synthesis which is then integrated with OPAC, and the algorithm used in the search is Fullproof algorithm. The methods used in this research were literature study, system design, prototype application creation, and application testing. From the test results, the application is able to search the data book in OPAC using voice input from various languages with a relatively fast searching time. Applications are also able to voice the book found in the form of sound.

Keywords: *Opac, Google API, Speech Recognition, Speech Sintetic, Fullproof.*

1. Introduction

The speed offered by information and communication technology has caused many people to become more dependent and cannot be separated from technology. It also cannot be denied by the library officials who really need the existence of information and communication technology in the library. The current library users who are referred to Net Generation demand that the library should meet the needs of information with a fast time, so information and communication technology are inevitably present in the library to help meeting the demands of users. As Tapscott [1] notes that one of the characteristics of Net Generation is speed, that is, they do their work quickly and do not like "waiting" activities.

OPAC or Online Public Access Catalog is an automation technology in the library that seems to bring fresh air to adjust the library with the characteristics of library users, so the library becomes a user oriented library. OPAC helps library users as a tool of "information retrieval" to find the information they need faster and more accurately when compared to the library card catalog.

On the other hand, libraries also have different community patrons from other community patrons. They are those who have "special needs" or people with disabilities. According to the results of the National Socio-Economic Survey (Susenas) conducted by the Central Bureau of Statistics (BPS) in 2012, the number of persons with disabilities in Indonesia is 6,008,661 people. Out of these numbers, 1,780,200 persons are persons with visual disability, 472,855 persons with speech disability, 402,817 persons with intellectual disability, 616,387 persons with physical disability, 170,120 persons have disability which is difficult to take care of themselves, and about 2,401,592 people experience double disability [2].

Although they are different, they still have the same rights as other library community patrons. It is the right to access information quickly and appropriately which is legalized by law. According to the Law of Indonesia Number 4, Article 1 Paragraph 1[3] and Government Regulation Number 43 Year of 1998, Article 1 Paragraph 1 states that, as other citizens, persons with disabilities have the right to have similarities, positions, rights and duties, in having total role and integration according to their ability in all aspects of life and livelihood.

This should be a consideration for the library to make the library which is user oriented, which is able to provide information quickly and accurately to all community patrons of library. Until now, there are not many facilities that support the people with visual disability. Provision of access to people with disabilities is only on facilities and infrastructure. As for the provision of non-physical, it is only a sign (signage) that supports physical services [5].

Currently, most OPACs that are used as fast and precise "information retrieval" tools have limitations. It is OPAC-based typing that can only be used by people who are able to type on the keyboard. It cannot be used independently by users with disabilities especially person with visual disability. Although in some libraries in Indonesia, there are already libraries that implement OPAC with the sound system, but it has not been maximized by voice recognition and speed in searching the data.

OPAC with voice of UIN Sunan Kalijaga Yogyakarta was launched on 1 March 2012 by the rector of UIN Sunan Kalijaga with the following sophistication [6].

1. Library users can interact with computer using voice input. The users can do a search through utterance in Indonesian language and OPAC will reply with search results in Indonesian language too by mentioning the position of the bookshelf.

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

2. OPAC can also display based-ranking book just like the standard search engines, so the library users can rely on the natural perception that the precision of proximity is based on the most accurate findings in top positions which are sorted from the top to the bottom.

From here, the authors conclude that the language used in this application is only Indonesian language Indonesia. Furthermore, an appropriate search of algorithm is needed because the more data will affect the searching time.

Looking at some of the problems above, it is necessary to conduct a research to make OPAC with multi-language voice support using the appropriate technology. In this study, we will discuss the possible technological options to solve this problem, how to integrate it with OPAC with the selection of appropriate search algorithms, and how the results of application testing.

2. Methodology

The methodology used in this study consists of four steps, i.e. literature review, system design, making application, and application testing. Details of these methods are as follows.

Literature review

In this stage, the process was to find documents on the current condition of library users. After that, the study of the literature on technology that was suitable to fix the problem was conducted. Results from this stage were in the form of a literature review about problem to be solved in the library along with alternative technology to be used for making applications' prototype. Thus, the decision of what technology used in this research were decided in this step.

System design

At the stage of system design, the step was to create database design. After that, we created application flow and access rights of each user. The results of this phase were the database design, application design flow and interface design.

Making Applications

After the creation process, we continued the system design by making the database based on the design of systems that had been made previously. Once the database was completed, the next step was to make the application based on design of the application from previous step. The next stage was to make an application in accordance with the permissions that had been created on the system design process. The result of this process was the application that was ready to be tested.

Testing Applications

Applications that had been made was then tested thoroughly in accordance with the system design. The result of this phase was the application that has been created and was ready to use by the user.

3. Background

Comparison

Since there are a lot of voice recognition and voice synthetic technologies, it is important to choose the appropriate technology. Below are some technologies for voice recognition and speech synthetic.

Table 1. Alternative Technology for Voice Recognition and Speech Synthetic

Technology	Google [6]	Nexmo [7]	Speechnotes [8]
Number of language supported	80	NA	72
Integration API	Yes	Yes	Yes
Accurate in Noisy Environments	Yes	NA	NA
Works With Apps Across Any Device	Yes	NA	NA
Word Hints	Yes	NA	Na
Real-time Audio Support	Yes	Real-time	Real-time
Pre-recorded Audio Support	Yes	No	Yes
Support speech synthesis	Yes	No	No

Table 1 describes some of the most popular speech recognition technologies available today. Judging from the number of supported languages, Google has the advantage of recognizing 80 languages compared to others. Judging from its ability to be integrated with other apps, Google and Nexmo have this capability, whereas speechnotes.co does not. This means that Google and Nexmo are able to collaborate with OPAC applications with API. Related to the technology to reduce the noise or distortion in voice input, Google is able to filter the noise so that the conversion of voice to text is more accurate. Google also has other features: the ability to work on different hardware. This feature is not possessed by other products yet. In this study, it will be tested on computer and android mobile phone. A Word hint is a feature to select one of the predefined lists of options. With the existence of word hints, the selection process will be more accurate with the predefined output result choice. All three vendors are capable of supporting voice recognition in real-time, but only Google and Speechnotes is capable of supporting previously recorded sounds. Google also has the technology to make sounds from text. This capability is not possessed by other products. Of the many advantages of voice recognition technology offered by Google, so, in this research, the voice recognition technology used is Google.

Because Google has role as a server and is able to work in many different types of tools, such as computers, and mobile phones, it needs a programming language that can support many devices as well. In this study, PHP [9] is used to meet this need. The output of PHP that is sent by the server is a combination of HTML and JavaScript. HTML organizes the in-app display such as font, font size, image size, and more [10], while JavaScript directs the interactivity of the application with the user [11].

Fulproof Algorithm

The use of Google's voice recognition technology and the Google's client server technology that require a stable Internet connection urges the writers to rethink how the system works to improve data search speed on OPAC server.

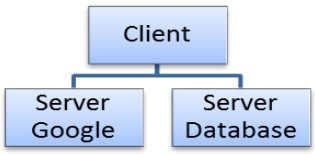


Figure 1. How the System Works

The figure above shows how the application works in general. Google uses AJAX (Asynchronous JavaScript and XML) to transmit voice and make sounds. Client needs communication with Google server continuously to translate voice from user. Furthermore, the client must also communicate with the database server to search for the book sought. Next, the client must communicate to Google server to utter the voice. Communication runs continuously on the client side so that the delay occurs at the time of its use. An alternative technology is needed to handle this. The most likely is to reduce communication with the database server by copying the database to the client computer. Furthermore, the searching process does not use the communication to the database server. It simply uses the data in the client computer. One of the most technologically available alternatives is Fullproof [12].

Figure 2 is the stages of FullProff algorithm. Below are the descriptions of each stage.

1. The server will download all the data of Books in the database server which then the data will be converted into JSON format and sent to the Client.
2. Parsing

In this process, the Client changed the shape of JSON into ARRAY.

3. Indexing
   This stage creates 2 lists of contents or indexes. Firstly, the purpose of making these two indexes is to speed up the search. The first index uses Light Normalization to convert all uppercase letters to lowercase letters, and also to omit double letters contained within a word. For example, Marrios become Mario. The second index uses aggressive normalization in which this index has function to make abbreviations in a word so that it can shorten the search time. For example: Mariois becomes MR and Rosalina become RSLN.
4. Analyze
   After the stage of making the Index, then the process of Analyze is conducted which aims to change all the text into a simple form in accordance with the indexes that have been created before.
5. Initial Process

In this process, Array in the step number 2 is inserted into the Index in the step number 3, then the result is ready to be used for searching process.

6. Callback

This process aims to call the result in the second index if the first index fails to outlaw the result.

7. Sound Input

This stage is the stage where the users input a sound to find the required book, then, by system, the input is converted to text by using Google API.

8. Search in Index

The system try to find the book collection in Index made before.

9. Results search

Searching results are displayed in the form of sound.

Below is the process stage of Foolproof Algorithm:

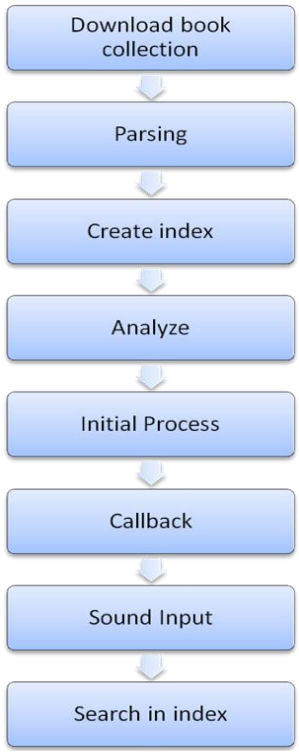


Figure 2. The Stages of Fullproof Algorithm Process

4. Experimental Result

System Design

After doing the literature review, the system is designed with the steps as follows.

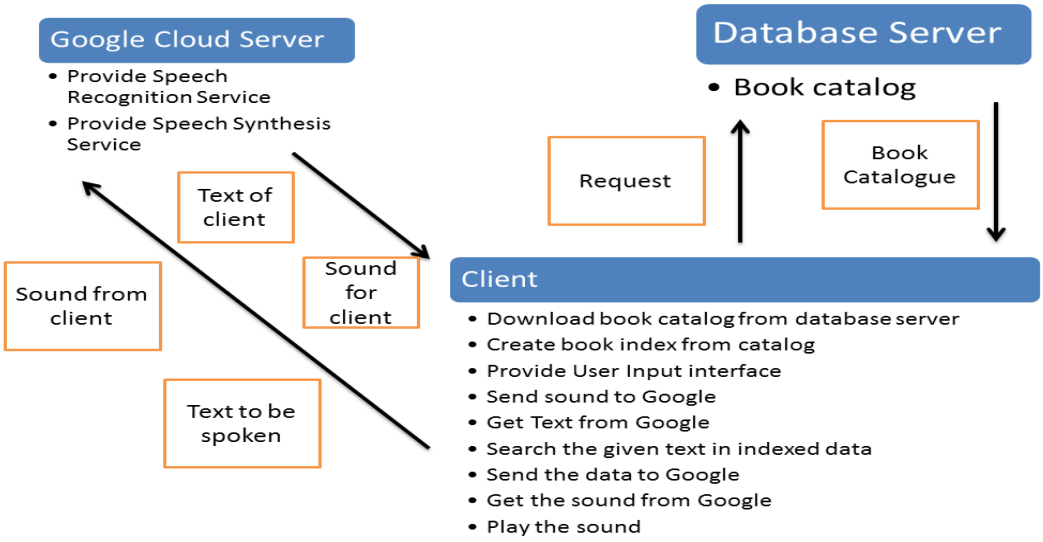


Figure 3. System Design

Figure 3 shows the order in which the system works. System works with the following steps.

1. System will download all catalog books from the database server to the client.
2. Next, system creates an index of catalog book.
3. System will invite users to say the title of books they look for.
4. Users then voice the title of the book they look for.
5. System does conversion sound to in-text by using Google API.
6. The text is then searched in the index that is previously created, and it will result the result lists of the searched book.
7. Next, the results of the book search are voiced with the synthetic sound using Google API.
8. Next, it is back to the process 3.

B. *System Testing*

Applications were then created using a combination of programming languages of PHP, JavaScript and HTML. PHP acts as an application on the server side, while HTML and JavaScript play a role in the interaction with the user. The application was then tested using the Book-Crossing Dataset [13]. Book-Crossing Dataset contains 3 (three) tables, they are:

- BX-Users

This table contains about the data of users with the number of 278,858 data.

- BX-Books

This table contains the book data with a very large number which is 271,379 data. There are several columns including ISBN, Book Title, Author, Publication Date and others.

- BX-Book-Ratings

This table contains about the rating of the book. The rating scale is between 1 until 10. The higher score means to describe a high appreciation. If there is no appreciation, it has 0 score.

Furthermore, the system was tested using a laptop and mobile phone with the following brands.

1. TOSHIBA SATELLITE L740-1219U LAPTOP
2. SMARTFREN E2 + MOBILE PHONE.

Below is the image of the finished app.

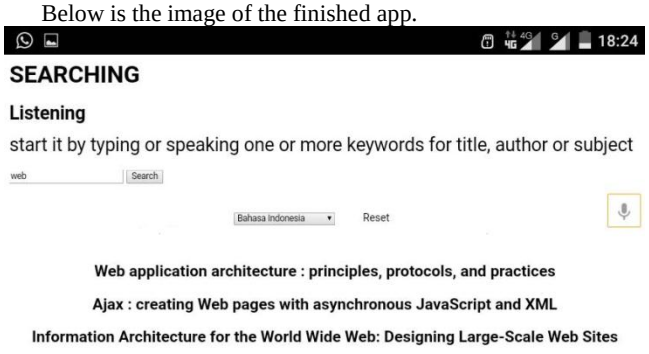


Figure 4. Book Search Results

Figure 4 shows the search results in the form of a list of book titles found by the system. The system requires a microphone and speakers to run. If there is no microphone, then the following display will appear.



Figure 5. ***Error display when there is no microphone***

Figure 5 illustrates if the user does not allow applications to use microphones or no microphone is available that can be used as voice input. This does not happen in the mobile phone because the permission of the use of microphone and speaker is automatic and in general the phone is equipped with microphone and speakers. Here are the results of testing applications.

Table 2. System test results

No	Feature	A	B
1	Listen to sound	✓	✓
2	Searching the book from collection	✓	✓
3	Recite the search result	✓	✓
4	Restart application automatically	✓	✓

From the test results in the table 2, it can be concluded that the prototype of the application already functions properly so that it can support the implementation in daily life in library.

5. Conclusion and Future Work

This study aims to implement OPAC by using input and output in the form of sound. Previous research was only able to use 1 language as input and its output. This research has successfully used multi-language as input and its output by using Fullproof search algorithm to speed up the process of searching data by reducing the communication to server. Because of the use of Google API, this system is very dependent on the internet. Subsequent research can focus on how to develop a multi-language voice recognition system offline without having to communicate with Google servers so that the system does not depend on the Internet network. This research has not considered the security side of library data books that are downloaded to the client computer. Subsequent research can also focus on how to design data storing process so that it cannot be opened by parties with certain interests.

References

[1] Tapscott, D. (2009). Growing Up Digital: The Rise of The Net Generation. New York: Mc-Graw Hill.

[2] Kemensos-RI. (2015, September 16). Pelayanan Penyandang Disabilitas Dalam Menggunakan Berbagai Sarana Aksebilitas. Retrieved May 23, 2017, from Kementrian Sosial Republik Indonesia: http://www.kemsos.go.id/modules.php?name=News&file=article&sid=18765

[3] Indonesia. (1997). Undan-Undang tentang Penyandang Cacat No. 04 Tahun 1977. Indonesia.

[4] Mutia, F. (2012, 12 19). Kecacatan dan aksesibilitas dalam memperoleh informasi melalui media komputer. Retrieved 5 24, 2017, from Web Dosen: http://fitri-m-a-fisip.web.unair.ac.id/artikel_detail-69475-Artikel-Kecacatan%20dan%20Aksesabilitas%20Dalam%20Memperoleh%20Informasi%20Melalui%20Media%20Komputer.html

[5] Speechnotes. (2017, 05 03). Speech to Text Online Notepad. Retrieved 05 23, 2017, from Speechnotes: https://speechnotes.co/

[6] Google. (2017, 05 2). Speech API - Speech Recognition. Retrieved 05 24, 2017, from Google Cloud Platform: https://cloud.google.com/speech/

[7] Nexmo. (2017, 05 02). Nexmo - APIsfor SMS, VOice and Phone Verifications. Retrieved 05 24, 2017, from Nexmo: https://www.nexmo.com/

[8] Speechnotes. (2017, 05 03). Speech to Text Online Notepad. Retrieved 05 23, 2017, from Speechnotes: https://speechnotes.co/

[9] PHP. PHP:Hypertext Preprocessor.Retrieved 05 23, 2017, from http://php.net

[10] HTML Language Explained. (2012, 08 21). HTML Source : HTML Tutorials. Retrieved 05 23, 2017, from HTML Language Explained: http://www.yourhtmlsource.com/starthere/whatishtml.html

[11] All About Programming Language. (2017, February 21). What is Javascript. Retrieved 05 24, 2017, from ThoughtCo: https://www.thoughtco.com/what-is-javascript-2037921

[12] Reyes, R. (2013, March 12). Fullproof - a javascript search engine library for the browser. Retrieved 05 23, 2017, from Fullproof: https://reyesr.github.io/fullproof/

[13] Book-Crossing Dataset. (2). Retrieved from Institut For Informatik Freiburg: Book-Crossing Datase