

Design module for speech recognition graphical user interface browser to supports the web speech applications

Fadya A. Habeeb¹, Suaad M. Saber², Shaymaa Mohammed Abdulameer³, Hassan Muwafaq Gheni⁴,
Ahmed Dheyaa Radhi⁵

¹Department of Mathematics, College of Education for Women, Tikrit University, Tikrit, Iraq

²Department of Computer Science, College of Education, Mustansiriyah University, Baghdad, Iraq

³College of Information Engineering, Al-Nahrain University, Baghdad, Iraq

⁴Department of Computer Techniques Engineering, Al-Mustaqbal University College, Hillah, Iraq

⁵College of Pharmacy, University of Al-Ameed, Karbala, Iraq

Article Info

Article history:

Received Jul 3, 2022

Revised Aug 23, 2022

Accepted Sept 4, 2022

Keywords:

Audio

Input

Speech

SRWB

Visually impaired

ABSTRACT

The web speech API has made it possible to integrate audio data into web applications and make it a unique experience for all customers and users of modern applications. The website can only be accessed through devices equipped with a which stands for graphical user interface (GUI) and screen. For this to be done, there must be a physical attraction with such devices. This paper presents speech recognition using a web browser (SRWB) which permits browsing or surfing the internet with the use of a standard voice-only and vocal user interface (VUL) development. The SRWB system input from the users in form of vocal commands and covers these voice commands to HTTP requests. The SRWB system will send the voice commands to the web server for processing purposes and when the processing is done, the converted or translated HTTP response is outputted to the end-users in a voice format made audible with the attached loudspeakers. SAPI, developed by Microsoft, allows the use of SRWB in Windows applications. The algorithm is implemented by the system to achieve its goal for web content, classifying, analyzing, and sending important parts of web pages back to the end-user.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Fadya A. Habeeb

Department of Mathematics, College of education for women, Tikrit University

Tikrit, Iraq

E-mail: Fadya.habeeb@tu.edu.iq

1. INTRODUCTION

The web speech API has made it possible to integrate audio data into web applications and make it a unique experience for all customers and users of modern applications. Broadly speaking, the web speech API consists mainly of two components: speech synthesis (text-to-speech) and speech recognition (asynchronous speech recognition). This feature has provided meeting the needs of all users of these applications to suit the way they interact and use those services, applications, and available electronic content. Figure 1 show some different applications of speech recognition that use currently in our live. Speech recognition is used for controlling web browsers and applets, enhancing the internet, and filling out forms in numerous ways. Through speech synthesis, web pages are brought back to life, which reduces audio that is sent over the internet. SRWB has many applications such as assisting blind people to make use of the internet via speech and hearing. It will help both the young children and people with disability who can't operate a keyboard to use the internet by just speaking. It grants enablement to people to interface the website contents through oral or audio commands instead of using a keyboard. For this system, Microsoft Access and SQL were used as databases while Visual

Basic.NET was the programming environment used for developing and designing. SRWB is very easy to use, visually independent, and can be used by anyone that is not computer literate [1].

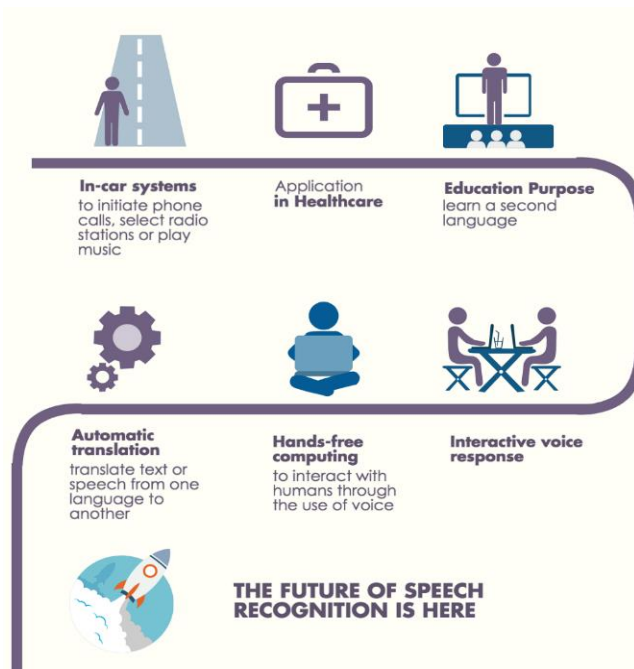


Figure 1. Some different applications of speech recognition that use currently in our live

According to Kazuyuki voice browsing refers to using speech to navigate an application. Speech interface framework parts are used in writing these applications. Web applications and web browsers are written in HTML while voice XML is used for writing speech applications but rendered through a voice browser [2]. There are different voice-driven website applications such as automated telephone receptionists, banking transactions, information about airline arrival and departure and other people use but don't know they are communicating with a web service. The W3C's Voice XML trendy is used employing extra than 85% of cell and telephony interactive voice response (IVR) apps. According to Kazuyuki, there are ten instances greater cellular telephones in the globe than linked PCs, and phone telephones will quickly emerge as the essential gateway to the internet. Speech recognition has no close relationship with 'visual web. However, this is expected to change for different reasons such as the prevalence of cell phones in the area with low literacy rates, shrinking of devices et al. Receiving audio feedback via speech synthesis show how hands-free practical applications can help people using mobile. People with disabilities like those who can't read and those with vision limitations will benefit a lot from voice or speech applications. Voice browsing has many possible applications as illustrated [3].

- Business information accessing such as ATM ordering services, booking cinema, and theatre services, home banking services, support desk, airline information (arrival and departure time), and others.
- Public information accessing such as events, weather reports, school opening and closures, stock market, location, and foreign news and others.
- Private information accessing like shopping lists, calendars, calories counters, telephone lists, and addresses.
- Helping the end-users to interact with people through email and voicemail sending and receiving of messages.

The unique contributions aim to provide the solutions for the problems of proposed web speech applications for the speech recognition in contrast to the current system of handling.

- In this paper we present a prototype speech recognition using a web browser (SRWB) that "converses" with the user by generating novel forms of design and module for the graphical user interface (GUI).
- Our unique contributions include i) a conversational SRWB prototype, ii) an understanding of how existing UI frameworks and toolkits can encourage or inhibit the integration of a SRWB, and iii) a set of guidelines for developers looking to incorporate SRWB into future web-based applications.

- c. Another novelty is the comparison between the voice user interfaces and GUIs for the speech recognition in the applications and in the web browsings.
- d. A computational insight on the software and hardware requirement are briefed in this study, moreover the analysis has been performed for the existing system and proposed system for the speech recognition designs.
- e. The system proposed for the SRWB in this study will be designed to maximize graphical user experience. Moreover, this proposed system differs from existing systems in many ways as discussed in the study.
- f. The unique contributions also include application of SRWB which identifies problems people who are visually impaired, so that they can use this speech recognition system for the user interface.
- g. This project proposes to determine a clear correlation between the speech recognition for the web browsers and applications and the comparing in controlling the multiple intelligent voice personal assistants.
- h. The process of developing a user interface in the browser revealed many opportunities and challenges for working with speech-based systems. We detail several that we encountered for designers and developers to address in their own projects, enumerate various alternatives that were considered for our own work, and conclude with lessons learned.

Uemura *et al.* [4] gave the dynamic interface and vocabulary grammar idea when they created speech user agent. The end-users of this system can communicate with the system in two ways: speak-able hotlist and commands. The end-users of this system can apply the speak-able hotlist. A grammar is associated with a uniform resource locator (URL). Smart pages were used to implement the system which led to grammar definition and many substitute grammars can be launched for a given link. Ma [5] got the main problem with voice browsers which can now read texts on the screen but can't convey the semantics and logical structure of the web content. Figure 2 shows the method to provide speech input in a web browser, current version of each browser is aligned in the same row [6].



Figure 2. The method to provide speech input in a web browser, current version of each browser is aligned in the same row

Instead of changing the entire already existing website, a dynamic conversion approach is required according to our opinion and as such calls for requires a universal solution using existing or new browsers that can convert the old website to a format that can be used. In [7], [8] described the capability of a web of a voice browser in rendering pages that are in audio format or one that supports the interpretation of speech input for navigation. With the support of technology, a good web browser can communicate in a two-way method with the end-user listening to the screen and uttering commands. Figure 3 shows the usage relative method to provide speech input in a web browser [9].



Figure 3. The usage relative method to provide speech input in a web browser

2. METHOD

2.1. Comparing existing system and proposed system

Speech recognition includes accepting discourse through a device's receiver, which is at that point checked by a discourse acknowledgment benefit against a list of language structure (essentially, the lexicon you need to have recognized in a specific app.). When a word or express is effectively recognized, it is returned as a result (or list of comes about) as a content string, and advance activities can be started as a result. The web discourse API encompasses a primary controller interface for this—`SpeechRecognition`—furthermore several closely-related interfacing for speaking to linguistic use, and comes about. By and large, the default discourse acknowledgment framework accessible on the gadget will be utilized for the discourse acknowledgment—most advanced OSe have a discourse acknowledgment framework for issuing voice commands. Think almost correspondence on macOS, Siri on iOS, Cortana on Windows 10, and Android Discourse, below is the explain of analysis of Kalliope Web portal in detail.

2.2. Analysis of kalliope web portal

The Kalliope web portal is a Slovenian web server specialized in voice awareness for those who are visually handicapped or blind. The Kalliope web portal accesses the association of the blind and visually impaired of slovenia's (ZDSSS) electronic information system (EIS) [10]-[12]. Each of Kalliope's net pages complies with the web access initiative's core standards. The machine is made undemanding the use of XML tagging and the introduced communicate module. The portal consists of hyperlinks to different beneficial websites in Slovenia for the blind and visually impaired. It's supposed to be used with the Homer Web browser it truly is included. Only ZDSSS individuals must get admission to the portal [13]-[15]. This is due to the reality that many of the texts in the EIS database are blanketed by using copyright. The textual content documents in the EIS database are saved in an untagged undeniable format. To convert these texts to HTML/XML structure, a special and unique HTML/XML tagger is required [16]-[21].

2.3. Limitations of existing system

The limitation of existing system can list as points show the advantage and disadvantage as:

- The existing system has small screens, and very small for viewing in web phones. It is barely better when you want to view it in palmtops.
- One big setback of existing systems is the accessing speed. Users of this system find it hard to access all kinds of devices. Every device experience slow access.
- The price of existing system is much, making it difficult for many to use.
- It experiences awkward or poor inputting of data. Even to enter short email with Qwerty keyboard, touchtone or palm's Graffiti is awkward.

- e. It has connectivity limitation, and this made it difficult for people in certain location to access and use.
- f. It is not user-friendly [22]-[24].

2.4. Analysis of proposed system

The innovation of voice browsing is quickly advancing these days. Tuning in and talking are the characteristic modes of communication and data gathering. As a result, all are presently heading towards a more voicebased approach of browsing instead of working on literary mode. Speech acknowledgment is gotten to by means of the SpeechRecognition interface, which gives the capacity to recognize voice setting from a sound input (ordinarily through the device's default discourse acknowledgment benefit) and react fittingly. For the most part you'll use the interface's constructor to make a modern SpeechRecognition protest, which features several occasion handlers accessible for identifying when discourse is input through the device's mouthpiece. The SpeechGrammar interface speaks to a holder for a specific set of language structure that your app ought to recognize. Language structure is characterized utilizing JSpeech Linguistic use Forma. These limitations of the existing system were thoroughly investigated to offer better solutions and services with SRWB. The proposed system SRWB will enable the end user to use voice or audio input to enter data/information into the system. This entered data is for browsing or the internet. SRWB will have an attached microphone used for the purpose of collecting audio input from the system users. This solved the input difficult experienced in the old system. The collected audio or voice input will be converted into text for searching with the use of customized web browser. When the outcome is returned, the displayed webpage content will be voiced out or read to the end-users through the loudspeaker attached to the speech recognition web browser system. SAPI was used by the system to gain access to both speech recognition and TTS (text-to-speech). The collected audio or voice input will be converted into text for searching with the use of a customized web browser. When the outcome is returned, the displayed webpage content will be voiced out or read to the end-users through the loudspeaker attached to the speech recognition web browser system. Figure 4 shows the speech recognition interface frame and gives a clearer picture and understanding of how the system works.

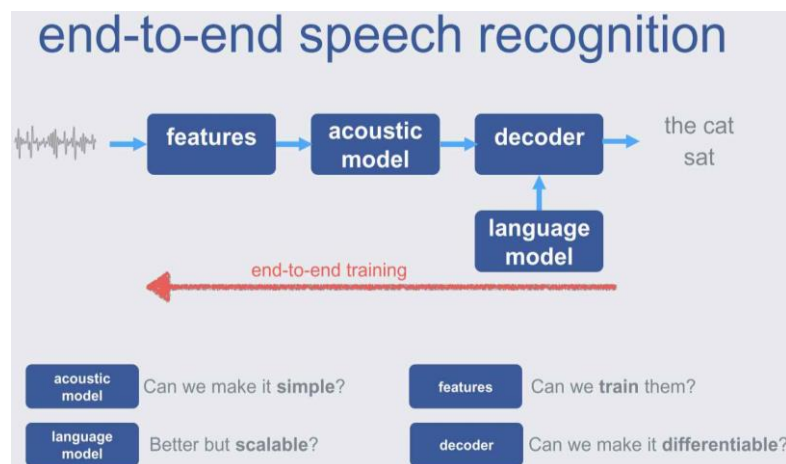


Figure 4. Speech recognition interface frame

The limitations of this system and other speech recognition systems for the blind and visually impaired were solved by SRWB. This existing system has restrictions on people in Slovenia which makes it not accessible to all. The design and user interface of SRWB are better than the old system. Not only is it easier to use, but it also supports multi-users from different countries and at any time, any day, and anywhere. SRWB is also easier to maintain. The user of SRWB will have to first give the command in a voice format by making use of a microphone. The SRWB system will accept the end-users command that comes in form of audio. This entered data or command is for browsing or the internet.

2.5. Speech recognition for user interfaces

2.5.1. Current mechanism of speech recognition

Low fidelity in speech recognition technology has resulted in the development of voice user interfaces (VUIs) that prioritize mechanics at the expense of affordance and feedback in web applications design. The VUI is incorporated into both an alarm and a calendar tool to ground our approach in a common application space in the web browsing. Our system addresses errors through conversation as well, ultimately improving

affordances and feedback while minimizing the loss in mechanics. Recent advances in software capabilities have spurred a growth in "smart" GUIs for hardware, such as phones and computers, integrated with an intelligent personal assistant [10]. These software agents take in user commands or questions and perform various tasks and services. One of the most popular forms of assistant, used by over 50% of US adults as of April 2020 [14], is the voice assistant and graphical assistance built directly into a hardware's operating system. Several major technology companies have commercialized a version of the intelligent voice personal assistant, including Apple (Siri), Microsoft (Cortana), Amazon (Alexa) and Google (Google Assistant). Open-source voice software such as Mycroft AI have also been created to allow anyone to develop voice assistant technology for their own projects.

As such, the principles of GUI software underlying the concept of a voice user interface (VUI), including affordances, feedback, and mechanics according to [25], [26], are an important tool for understanding and designing VUIs. VUIs have several advantages when it comes to GUI design: i) accessibility—they only require speech to control them, and they can be incorporated into ubiquitous devices, such as smart phones, ii) familiarity—speech is a very well known and utilized form of communication across the world; and iii) mechanics—they promote very fast information retrieval and require minimal physical effort.

At the same time, VUIs have several disadvantages as a UI design: i) error rate—they are prone to frustrating errors, both in the recognition of the user's voice as input, and in the synthesis of human-sounding speech when the system responds to the user's input, ii) affordances—they have extremely poor affordances, especially for users that are not tech savvy and may not know what they can do with the graphical interface; and iii) feedback—depending on the available speech recognition and synthesis software, VUIs can often have mixed feedback that is not guaranteed to help the user understand what they've done and whether it was successful.

2.5.2. Proposed mechanism of speech recognition

A proposed system includes a basic GUI to help users debug the web applications. The GUI presents a few graphical input modalities to help the user begin using the tool. Towards the top of the application are voice settings, which control the accent, rate, and pitch of the synthesized voice that responds to the user's input. As stated earlier, it is critical that speech synthesizers prioritize naturalness and intelligibility in their output. We used the SpeechSynthesis interface of the web speech API [26], [27], which offers a multitude of regional accents to help users interpret the synthesized responses from the system. At any time, the user can press the Test voice button to hear a message read aloud by the synthesizer with their current settings. Below this, the Start recording button is used as a switch to allow the browser to begin listening to sounds from the user's microphone and interpret them as human speech. This attempts to mitigate insertion errors through a more affordance-first approach for the speech recognition. The alternative (i.e., a system controlled solely through voice) cannot resolve such errors. Once pressed, recording will commence immediately and will not stop until the system recognizes that the user has spoken something and then stopped speaking. This is the main input modality for telling the system to take in user input, and the only way for the system to initiate recognition and subsequently synthesize a response as proposed in [28]-[32]. Finally, the show/hide output button can be pressed to show or hide the area below the button which automatically records the conversation being had as visual text data. This was provided to show that the system can function with and without visual feedback.

3. METHOD ADOPTED FOR SRWB

The methodology used for the development of SRWB is the waterfall model. It was used because the model gives many benefits for the application system developers. Also, the model's development cycle supports initial discipline which simply means that each phase in the methodology has a beginning and ending point. It is very easy to identify progress. The waterfall model requires more focus on SRWB design and set requirements before the commencement of programming. This ensures little effort and saving of time while no phase was overlooked. Easy usage and management of this method make it the best choice for selection in this research paper or project. The steps passed through include: i) research/analysis definition, ii) basic design, iii) technical design/detailed design, iv) building, v) testing, vi) integration, and vii) management and maintenance.

Algorithm

The voice command processing Algorithm shows how SRWB processes the voice commands:

1. Start
2. Get voice command
3. If SRWB recognizes voice command, then search for command related with command
4. If SRWB don't recognize voice command, then output error message.
5. Stop

The user operation Algorithm how users will use SRWB:

1. Start

2. Login to system
3. If username & password match move, go to step 5.
4. If username & password is not found database output error
5. Enter voice or input commands
6. Operate Voice Browser
7. Stop or End

The voice browser operation Algorithm shows the operation of SRWB:

1. Start
2. Listen for Command
3. If command is to browse prompt user for URL
4. If command is to stop then terminate browser
5. If command is to read results, then read contents of browser
6. If command is to Exit browser, then exit browser application
7. End

3.1. Components of the methodology adopted

The main components of the methodology can divide in four parts start with analysis phase, design phase, coding phase and end with testing phase as below:

- a. Analysis phase: the goal of this component is ensuring definition of data, processes, and boundaries.
- b. Design phase: the goal of this to solve the problem. Focus was shifted to physical other than logical. Data elements were grouped for the formation of physical data files, structures, databases, screens, and reports.
- c. Coding phase: SRWB was created in this stage. There was coding, debugging, and testing. It was at this stage that documents for users was prepared, and files and databases initialized. Procedures were also tested after being written.
- d. Testing phase: SRWB was tested, and remaining problems was attended to. Implementation of the system and its release took place here.

3.2. Hardware requirement of search recognition web browser

SRWB works with many hardware such as:

- a. A personal or color computer system (monitor) that runs on a higher processor
- b. Loudspeaker or PC speaker
- c. Microphone
- d. Random access memory
- e. Gigabyte hard disk drive
- f. Keyboard
- g. Mouse

3.3. Software requirement of search recognition web browser

SRWB has software requirements like:

- a. MS office access
- b. SQL database
- c. Windows operating system
- d. Visual Basic.NET
- e. Net 3.0 Framework
- f. Microsoft voice recognition SDK
- g. Window 7
- h. Anti-Virus protection software
- i. Web based application software (HTML and XML)

3.4. Application of search recognition web browser

SRWB has many applications such as identifying problems people who are visually impaired. It will address the problems visually challenged people are facing. With SRWB people can order ATM services, book other services like cinema, home banking, support desk and theatre services. SRWB can also be used to access telephone and shopping lists, calendars, address, weather report, international and local stock market, event services and others. SRWB will assist people to communicate messages. Helping the end users to interact with people through voicemail sending and receiving of messages without using a keyboard to enter a command.

3.5. System security

To prevent unauthorized users from using the system, an authentication module was incorporated into SRWB in form of system login module, Figure 5 shows the login module interface.

3.6. Interface of search recognition web browser

A vocal interface is not the same with a visual interface. They are not consistent, and this makes voice interfaces to be highly dialog oriented based on response and oral presentation. While a visual interface can give the end users additional information at a time a voice interface listens to the end users data in few amount at once. This made us to develop and design an efficient and reliable that vocal interface pattern. Once the voice of users is accepted, it is translated and outputted to the user immediately.

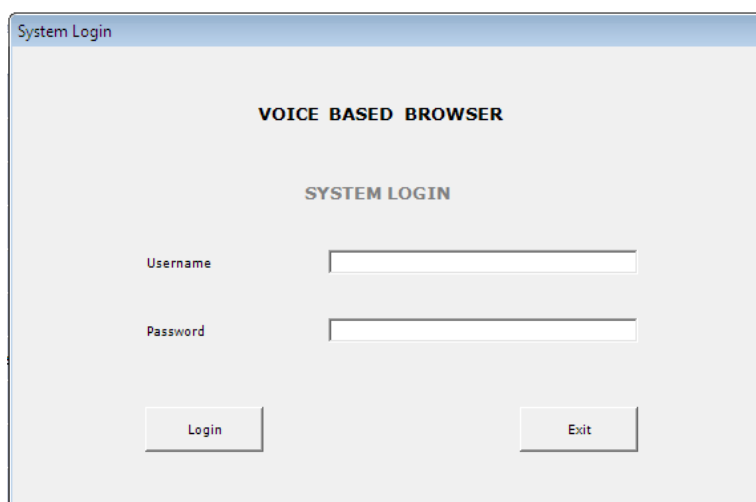


Figure 5. Login module interface

4. RESULTS AND DISCUSSION

The double authentication of the system proves to be effective as only one user can access SRWB with a single password. The SRWB system was able to accept commands from different users which are in audio or voice format. When the web browser had loaded the web page, SRWB was able to read the accepted commands in audio and read the contents of the command of the web browser. It translated and gave out the result to the users with the help of loudspeakers or personal computer speakers. This, therefore, proved that every requirement of SRWB was met. The SRWB system cannot always display the users' commands on the computer screen totally accurately. Also, programs could not understand the language context perfectly as human does. This led to errors because of misinterpretation. The coaching statistics measurement certainly illustrates that the increased the education data, the higher the awareness accuracy. This coaching fact can include these factors in a range of ways, such as speaking in special accents, the usage of the equal phrases stated through male and woman speakers, and the usage of one-of-a-kind phrases delivered below one-of-a-kind situations, such as when the speaker has a sore throat. When the person makes use of a low-quality audio enter device, the device may additionally no longer supply the satisfactory feasible output. The design and development of SRWB is a good addition speech recognition and assisting visually impaired patients and people with low-level of computer knowledge to access the internet and get the help they want. Almost everything is done via internet and an application like this is a boost and an aid for both the blind and the disabled. SRWB is not only a solution to the many problems of visually impaired people who are often victim of fraud from people when they need assistance of the internet. It is also a better system compared to the existing systems. Inputting of patient's information was made easier, and the price is very affordable. Moreover, there is no restriction of usage with SRWB if there is available network connectivity. Some speech recognition systems haven't seen the many benefits of SRWB choose to design their system by copying the design and methodology of SRWB. SRWB has proven to its efficiencies to help the young, blind and people without computer knowledge to access services like airline arrival and departure time, calendars, events, local and international stock market, and many other services. Some speech recognition system has issues in security and breach of unauthorized users, but this is not the case with SRWB because it includes a two-way authentication. SRWB is also efficient and effective as different users with different password can access it at

a particular time. With the easy-to-use interface, the problems of visually impaired patients have drastically reduced.

5. CONCLUSION

Speech is one of the natural and at the same time oldest means of communication among humans. Humans interact and communicate with each other in human-human interface. Many machines that will have the capability of understanding and producing speech just as humans have been tried to be developed. SRWB was one of the systems created to see this desire achieved. SRWB system was designed to assist visually challenged users who desire to use the internet for many beneficial reasons. To develop SRWB to meet these demands a voice-controlled browser was used together with a speech recognizer engine. The resulting voice controlled browser that follows this work accompanying this report can be applied by any researcher who desire to participate in language processing research. To see all the benefits of this system are achieved, the following recommendations have been made available to assist future researchers on this research paper or topics that are related to it: future software application developments should be created to run on different operating system platforms. Future work or research on this specific paper or related ones should also consider integration or joining of voice in voice out speech driven applications. This will send result to visually impaired or challenged user. Provision should be made available for a database backup against loss or damage of current database. Finally, the recognition phase speed is reasonable and more efficient and suitable for limited number of 15 persons. However, there is more than 15 voices of persons in the database the speed of the recognition decreases. SRWB will give the best result if the user uses a high-quality audio input equipment such as microphone.

ACKNOWLEDGEMENTS

The authors are grateful to the Iraqi Ministry of Higher Education and Scientific Research (MOHESR) for technically supporting the current research.




REFERENCE

- [1] J. C. Long and R. J. Toal, "Modeling patterns for javascript browser-based games," *Internet and Multimedia Systems and Applications/ 747: Human-Computer Interaction*, 2011, doi: 10.2316/p.2011.746-018.
- [2] Y. Astuti and I. R. Wulandari, "An arrangement of the number of K-grams in the performance of Rabin Karp algorithm in text adjustment," *Indonesian Journal of Electrical Engineering and Computer Science (IJECS)*, vol. 26, no. 3, pp. 1388-1394, 2022, doi: 10.11591/ijeecs.v26.i3.pp1388-1394.
- [3] H. Zahr, H. Al Haj Hassan and J. Haydar, "Web-Based framework for assisting users using speech recognition," *2018 International Arab Conference on Information Technology (ACIT)*, 2018, pp. 1-6, doi: 10.1109/ACIT.2018.8672581.
- [4] T. Uemura, S. Takagi, Y. Kosuda, and R. Sasaki, "A proposed web access control system request policy framework for cooperation of DNS and a web browser," *2011 Sixth International Conference on Availability, Reliability and Security*, pp. 549-554, Aug. 2011, doi: 10.1109/ARES.2011.84.
- [5] C. Y. Ma, "Multilingual speech recognition and its application in a multilingual voice browser," M.Phil, Hong Kong University of Science and Technology, Hong Kong, 2001, doi: 10.14711/thesis-b712897.
- [6] A. S. Hasan, M. A. Akraa, and S. J. Abbas, "Tackling Graphical Natural Language Processing'S Problems With Recurrent Neural Networks," *Journal of Southwest Jiaotong University*, vol. 54, no. 5, pp. 1-9, 2019, doi: 10.35741/issn.0258-2724.54.6.35.
- [7] R. K. Jain, V. Sharma, R. Kardam, and M. Rani, "Artificial Intelligence Based A Communicative Virtual Voice Assistant Using Python & Visual Code Technology," *World Journal of Research and Review (WJRR)*, vol. 13, no. 5, pp. 23-26, 2021.
- [8] N. J. Qasim and I. Barazanchi, "Unconstrained joint face detection and recognition in video surveillance system," *Journal of Advanced Research in Dynamical and Control Systems-JARDCS*, vol. 11, no. 1, pp. 1855-1862, 2019.
- [9] I. A. Barazanchi and H. R. Abdulshaheed, "Adaptive illumination normalization framework based on decrease light effect for face recognition," *Journal of Advanced Research in Dynamical and Control Systems-JARDCS*, vol. 11, no. 01, pp. 1741-1747, 2019.
- [10] J. A. Larson, "VoiceXML 2.0 and the W3C speech interface framework," *IEEE Workshop on Automatic Speech Recognition and Understanding, 2001. ASRU '01*, 2001, pp. 5-8, doi: 10.1109/ASRU.2001.1034576.
- [11] K. Henriksen and J. Indulska, "Adapting the web interface: an adaptive web browser," *Proceedings Second Australasian User Interface Conference*, pp. 21-28, doi: 10.1109/AUIC.2001.906272, 2001.
- [12] M. T. Ahmed, M. Rahman, S. Nur, A. Z. M. T. Islam, and D. Das, "Natural language processing and machine learning based cyberbullying detection for Bangla and Romanized Bangla texts," *TELKOMNIKA Telecommunication Computing Electronics Control*, vol. 20, no. 1, pp. 89-97, 2022, doi: 10.12928/TELKOMNIKA.v20i1.18630.
- [13] E. H. Muktafin, P. P. Pramono, and K. Kusriani, "Sentiments analysis of customer satisfaction in public services using K-nearest neighbors algorithm and natural language processing approach," *TELKOMNIKA Telecommunication Computing Electronics Control*, vol. 19, no. 1, pp. 146-154, 2021, doi: 10.12928/TELKOMNIKA.V19I1.17417.
- [14] P. R. Dixon and S. Furui, "Exploring web-browser based runtimes engines for creating ubiquitous speech interfaces," *Interspeech*, pp. 630-632, Sep. 2010, doi: 10.21437/Interspeech.2010-240.
- [15] A. Yayla, H. Korkmaz, A. Buldu, and A. Sarikas, "Development of a remote laboratory for an electronic circuit design and analysis course with increased accessibility by using speech recognition technology," *Computer Applications in Engineering Education*, vol. 29, no. 4, pp. 897-910, 2021, doi: 10.1002/cae.22340.




- [16] M. Boukabous and M. Azizi, "Crime prediction using a hybrid sentiment analysis approach based on the bidirectional encoder representations from transformers," *Indonesian Journal of Electrical Engineering and Computer Science (IJECS)*, vol. 25, no. 2, pp. 1131–1139, 2022, doi: 10.11591/ijeecs.v25.i2.pp1131-1139.
- [17] A. Abdulkareem, T. E. Somefun, O. K. Chinedum, and F. Agbetuyi, "Design and implementation of speech recognition system integrated with internet of things," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 2, pp. 1796–1803, 2021, doi: 10.11591/ijece.v11i2.pp1796-1803.
- [18] H. Kamau, D. O. McOyowo, and D. O. Okoyo, "An enhanced browser reference model," *International Journal of Computer Applications Technology and Research*, vol. 8, no. 8, pp. 299–302, Jul. 2019, doi: 10.7753/IJCATR0808.1001.
- [19] Z. Weiqiang and F. Lina, "Design and development of interactive web-based voice broadcasting system," *2009 International Forum on Computer Science-Technology and Applications*, 2009, pp. 124–126, doi: 10.1109/IFCSTA.2009.151.
- [20] C. Muangmee, S. Kot, N. Meekaewkunchorn, N. Kassakorn, S. Tiranawatananun, and B. Khalid, "Students' use behavior towards e-learning tools during COVID-19 pandemics: Case study of higher educational institutions of Thailand," *International Journal of Evaluation and Research in Education (IJERE)*, vol. 10, no. 4, pp. 1166–1175, 2021, doi: 10.11591/IJERE.V10I4.21821.
- [21] J. B. Liu, "Teach network class over Internet using browser-based technologies," *IECON'99. Conference Proceedings. 25th Annual Conference of the IEEE Industrial Electronics Society (Cat. No.99CH37029)*, 1999, pp. 106–110 vol.1, doi: 10.1109/IECON.1999.822180.
- [22] Ö. Salor, B. L. Pellom, T. Ciloglu, and M. Demirekler, "Turkish speech corpora and recognition tools developed by porting SONIC: Towards multilingual speech recognition," *Computer Speech & Language*, vol. 21, no. 4, pp. 580–593, Oct. 2007, doi: 10.1016/j.csl.2007.01.001.
- [23] R. Zhang and G. Kikui, "Integration of speech recognition and machine translation: Speech recognition word lattice translation," *Speech Communication*, vol. 48, no. 3–4, pp. 321–334, Mar. 2006, doi: 10.1016/j.specom.2005.06.007.
- [24] K. Sugita and M. Yokota, "Implementation of voice guidance function for sightseeing contents using web speech API," *2016 10th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS)*, 2016, pp. 133–138, doi: 10.1109/IMIS.2016.106.
- [25] T. Hannagan, J. S. Magnuson, and J. Grainger, "Spoken word recognition without a TRACE," *Frontiers in Psychology*, vol. 4, p. 563, 2013, doi: 10.3389/fpsyg.2013.00563.
- [26] T. C. Kietzmann, P. McClure, and N. Kriegeskorte, "Deep neural networks in computational neuroscience," In S. Murray Sherman (Ed.), *Oxford research encyclopedia of neuroscience*, Oxford University Press, 25 January 2019, doi: 10.1093/acrefore/9780190264086.013.46.
- [27] V. Dongre, S. Ghosh, and A. Katkar, "Speech Recognition Based Web Browsing," *International Journal of Engineering Research & Technology (IJERT)*, vol. 2, no. 11, pp. 2489–2495, 2013.
- [28] H. You, and J. S. Magnuson, "TISK 1.0: An easy-to-use Python implementation of the time-invariant string kernel model of spoken word recognition," *Behavior Research Methods*, vol. 50, no. 3, pp. 871–889, 2018, doi: 10.3758/s13428-017-1012-5.
- [29] V. Bhardwaj et al., "Automatic speech recognition (ASR) systems for children: a systematic literature review," *Applied Sciences*, vol. 12, no. 9, 2022, doi: 10.3390/app12094419.
- [30] L. Chen, J. Ren, X. Mao, and Q. Zhao, "Electroglottograph-based speech emotion recognition via cross-modal distillation," *Applied Sciences*, vol. 12, no. 4338, pp. 1–18, 2022, doi: 10.3390/app12094338.
- [31] N. Washani and S. Sharma, "Speech recognition system: a review," *International Journal of Computer Applications*, vol. 115, no. 18, pp. 7–10, 2015, doi: 10.5120/20249-2617.
- [32] M. Zuchowski and A. Göller, "Speech recognition for medical documentation: an analysis of time, cost efficiency and acceptance in a clinical setting," *British Journal of Healthcare Management*, vol. 28, no. 1, pp. 30–36, 2022, doi: 10.12968/bjhc.2021.0074.

BIOGRAPHIES OF AUTHORS






Fadya A. Habeeb    received the Bachelor's degree in Computer Science from the College of Science, University of Tikrit, Iraq, 2006. Received a Master's degree from the College of Computers and Information, Mansoura University, Egypt, in 2016. Now working in the College of Education for Girls, University of Tikrit, Iraq. Her research interests are in AI and information security, IoT, and networking. She can be contacted at email: fadya.habeeb@tu.edu.iq.






Suaad M. Saber    received the Bachelor's degree in Computer Science from the College of Science, Al-Mustansiriya University, Iraq in 2011. Received Master's degree from the College of Science, Al-Mustansiriya University, Iraq, 2017. Now working in the College of Education, Al-Mustansiriya University, Iraq. Her research interests are in AI and information security, IoT, and networking. She can be contacted at email: suaad.m.saber@uomustansiriyah.edu.iq.






Shaymaa Mohammed Abdulameer    is an Asst. Lecturer at the Department of Information and Communications Engineering, College of Information Engineering, Al-Nahrain University, Iraq, where she has been a faculty member since 2021. Shaymaa graduated B.Sc. Computer Information Systems from Al-Rafidain University College, Iraq, in 2001, and an M.Sc. in Computer Science from Iraqi Commission for Computer and Informatics/Informatics Institute for Post Graduate Studies, Iraq in 2020. Her research interests are primarily in the area of data mining as well as wireless communications and networks, IoT, intrusion detection system. She can be contacted at email: shaymaa.m@nahrainuniv.edu.iq.



Hassan Muwafaq Ghani    Received his Bachelor (B.Sc) of Electrical & Electronic Engineering from Department of Electrical Engineering, Babylon university-Iraq-hilla in June 2016. In February 2018, he entered the master's program at the Faculty of Electrical and Electronic Engineer, Universiti Tun Hussein Malaysia. He is a lecturer at Al-Mustaqbal University College/Department of Computer Techniques Engineering. His research interest is optical communication, IoT, wireless sensor network, communications, V2V system and artificial intelligent. He can be contacted at email: hasan.muwafaq@mustaqbal-college.edu.iq and hassanmuwafaqghenih@gmail.com.



Ahmed Dheyaa Radhi    obtained a Bachelor's degree in Information Technology from the Software Department at the College of Information Technology at Babylon University, Iraq in 2013. He received a master's degree in information technology from the same college in 2018. Currently, work as a lecturer at the Faculty of Pharmacy, Al-Ameed University in Iraq, as well as in charge of the Systems and Software Division at the university. His current interests are programming, artificial intelligence, website design, databases, and server management. He can be contacted at email: ahmosawi@alameed.edu.iq.