Journal of Advances in Mathematics and Computer Science

29(6): 1-15, 2018; Article no.JAMCS.44344

ISSN: 2456-9968

(Past name: British Journal of Mathematics & Computer Science, Past ISSN: 2231-0851)

Development of an Improved Electronic Patient Health Record Management System with Speech Recognition

Justice O. Emuoyibofarhe[1,2,3], Keji K. Adewuyi[1] and Elizabeth A. Amusan[1*]

[1]Department of Computer Science and Engineering, Ladoke Akintola University of Technology, Ogbomoso, Nigeria.
[2]Center for Excellence in Mobile e – Services, University of Zululand, South Africa.
[3]Hasso Plattner Institute, University of Potsdam, Germany.

Authors' contributions

This work was carried out in collaboration between all authors. Author KKA designed and implemented the system and wrote the first draft of the manuscript. Author JOE managed the analyses of the study. Author EAA managed the literature searches. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JAMCS/2018/44344
Editor(s):
(1) Dr. Kai-Long Hsiao, Associate Professor, Taiwan Shoufu University, Taiwan.
Reviewers:
(1) Ali Sever, Pfeiffer University, USA.
(2) R. Nedunchelian, Anna University, India.
(3) Lanhua Zhang, Taishan Medical University, China.
Complete Peer review History: http://www.sciencedomain.org/review-history/27527

Original Research Article

Received: 26 July 2018
Accepted: 04 October 2018
Published: 01 December 2018

Abstract

Electronic Patient Health Record (EPHR) is an individual official patient health document shared among multiple facilities and agencies. However, most existing EPHR systems are dependent on keyboard and mouse only and do not support human speech interaction. This study, therefore, developed an improved EPHR management system, characterized by human speech interaction. Oral interview was conducted, the information acquired was used to design the improved system whose components are speech recognition architecture, speech recognition algorithm and voice command algorithm. The improved EPHR system database was developed using Microsoft SQL server 2016, Legacy ActiveX Data Objects (ADO.net) with Object Relational Mapping. Microsoft speech library was used for the speech recognition module. The improved EPHR system was implemented using C# programming language (.NET 4.5) and Visual Studio 2017. The performance of the improved EPHR system with speech recognition was evaluated with 50, 100, 150 and 200 words using correctness, accuracy and Word Error Rate (WER). The performance of the improved EPHR system yielded correctness, accuracy and WER values of (96, 96 and 4.0%), (96, 95 and 4.0%), (95, 95 and 5.0%) and (93, 94 and 6.0%) for 50, 100, 150 and 200 words

**Corresponding author: E-mail: eaadewusi@lautech.edu.ng;*

respectively. This study developed an improved EPHR management system with speech recognition and voice command which can improve user interactivity and help in disabilities or hands-free environment.

Keywords: Electronic health record; health; patient; speech recognition.

1 Introduction

Electronic Patient Health Record (EPHR) is an individual official patient health document shared among multiple facilities and agencies [1]. The accessibility of complete medical information when needed brought the novelty of storing the patient's information by electronic means. Electronic health record systems are designed to store data correctly and to catch the condition of a patient across time. Electronic Health Records (EHRs) may include laboratory test results, a variety of records, vital signs, medication and allergies, medical history, immunisation status, individual statistics like age and weight, radiology images, and billing data [2]. It removes the need to trail down a patient's former paper medical records and helps in certifying data is accurate and readable. It can decrease the risk of data duplication, as there is only one adjustable file, which means the file is more expected to be up to date and drops the risk of lost bookkeeping. Speech recognition and speech synthesis were added to EPHR to improve user interactivity with the system. Speech Recognition (is also known as Automatic Speech Recognition (ASR) or computer speech recognition) is the process of converting a speech signal to a sequence of words, by means of an algorithm implemented as a computer program [3]. Speech recognition is the manner by which a computer identifies spoken words [4]. Basically, it means "talking" to a computer and having it correctly recognize what was said [5]. It is also identified as "Automatic Speech Recognition" (ASR), "computer speech recognition", or just "Speech to Text" (STT) [6]. Some SR systems use "training" (also called "enrollment") where a single speaker reads the manuscript or isolated vocabulary into the system. The system examines the person's exact voice and uses it to adjust the recognition of that individual's speech, give rise to increased accuracy. Systems that do not use training are called "speaker independent" systems while the systems that use training are called "speaker dependent"[7]. The complement of SR that involves the computer or other devices talking back is the speech synthesis or simply Text-to-Speech (TTS). Speech Synthesis is the artificial production of human speech [8]. It is the computer-generated imitation of human speech [9]. This is used to convert written information to aural information where it is more suitable, particularly for mobile applications such as voice-enabled e-mail and unified combined. It is similarly used to help the vision-impaired so that, for instance, the contents of a display screen can be repeatedly read aloud to a sightless user [10]. Speech synthesis is the complement of speech or voice recognition. Most existing EPHR systems are dependent on keyboard and mouse only and do not support human speech interaction, hence, this research developed an improved EPHR management system with speech recognition.

2 Research Methodology

The software development methodology used is a waterfall design method, which includes oral interview, requirement gathering and analysis, system framework analysis and design, prototype development, system development and performance evaluation (testing). The methodology approach used to develop the system is shown in Figure 1.

The interview was conducted with some staff, medical personnel and some potential staff to find out what difficulties they encountered with the current system and share their feelings and experiences about the current system. Through this, raw data of the existing system were gathered and analyzed. The functional and non-functional requirement of the system were gathered, some of the functional requirements include:

i) Users should be able to search the database for the patient (names like omowonuola, chuckuemeka, Patrick etc.) using speech recognition technology and the system should readout the data or information of the patient if found using speech synthesis.
ii) Voice command feature should be added to enable the user to navigate around the system.

iii) For security reason, the system should check the login attempt and close the application after three unsuccessful logins.
iv) Users should be able to log out and switch user.
v) The system should be able to auto-generate a patient number or patient ID and should only allow an image of 128 by 128 resolution.
vi) The system should be able to populate the surname and first name textbox with the corresponding value of inputted patient ID.
vii) The system should not allow the user to supply alphabet in number textbox (e.g. phone number).
viii) User should be able to choose or select the desired data column to export to excel or print.
ix) Users should be able to search by patient ID, surname, first name from any angle (beginning, middle or end) and should be able to sort data column in ascending or descending order.
x) The system should have different user access level or right (admin and doctor), i.e. it should have Role Based Access control (RBAC)
xi) The system should use captcha with a combination of numbers and uppercase alphabet and allow users to be able to refresh or select another captcha value for security reason etc.

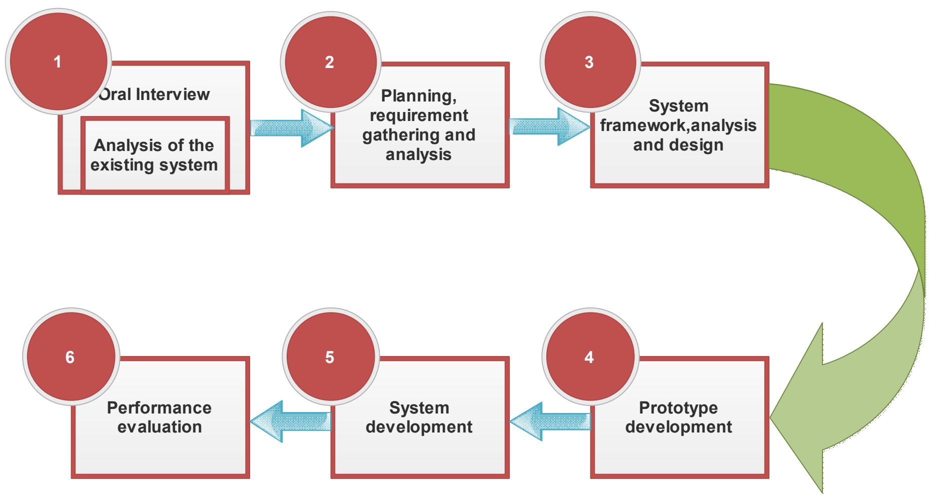

Fig. 1. Methodology approach

Different paper forms were gotten from the record manager, collated and designed in a bid to implement the appropriate EPHR system that would accurately automate the usual manual process. The paper forms were transformed into different views in the system.

The system architecture is shown in Figure 2.

The topmost level of the application is the user interface or presentation layer, which includes the Graphic User Interface (GUI) that allows users to interact with EPHR system through visual indicators and graphical icons. This layer interacts with the process/business layer, which coordinates business processes and provides building blocks for aggregating loosely coupled services as a sequencing process aligned with business goals. The control/data flow from the GUI to this layer and then to Data Access Layer(DAL) or database layer which provides simplified access to the system data stored in persistent storage of an entity-relational database. This layer ensures the communication between the application and database.

Speech recognition (architecture, flowchart and algorithm), speech synthesis (architecture, flowchart and algorithm) and voice recognition algorithm designed to handle the human speech interaction features. Figure 3 shows speech recognition and speech synthesis architecture. When the user launches the application, the application connects to the database to retrieve all patients' name and store in a temporary location (grammar). When the user speaks to the microphone to search for a patient, the system checks in the grammar. If the name is not found, the system returns "patient not found" message else, the system uses the name to search for the patient record from the database, and the speech synthesis reads out the retrieved record. Figures 4-6 shows the system algorithms while Figure 7 shows the speech recognition flowchart.

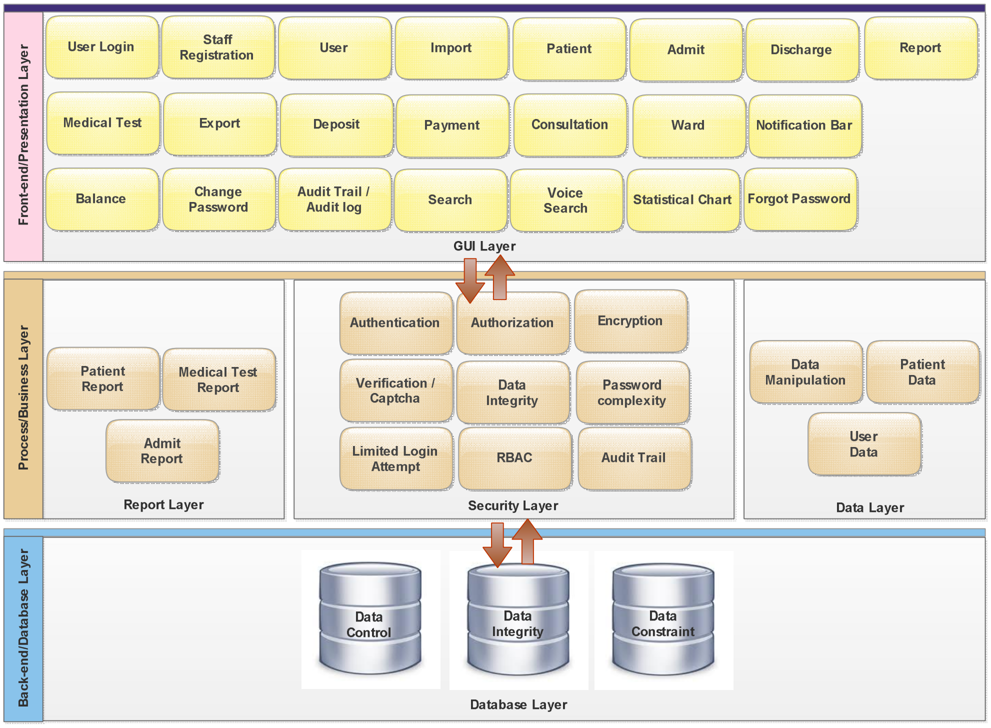

Fig. 2. System architecture

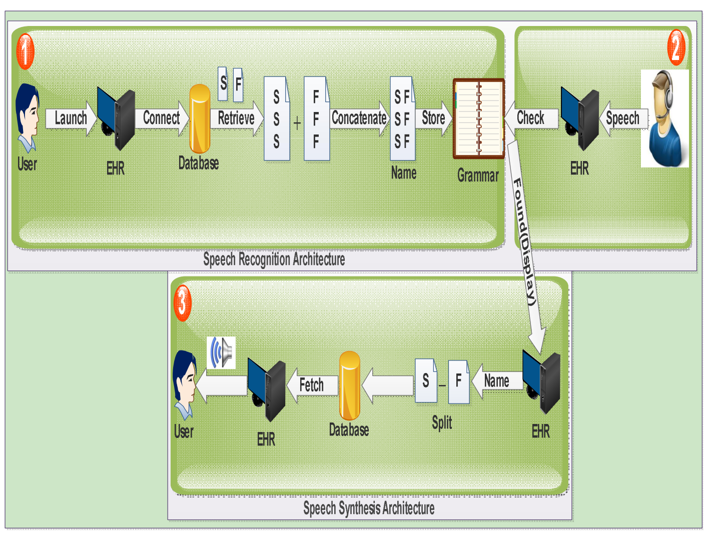

Fig. 3. Speech recognition and speech synthesis architecture

Step 1: Start
Step 2: Connect to database
Step 3: Get all patient surname and last name from database
Step 4: Concatenate the name to full name
Step 5: Store the name in grammar/dictionary
Step 6: Create speech recognition engine using Microsoft speech library
Step 7: Set audio input to default microphone
Step 8: Hook audioLevelUpdated and speechRecognized event
Step 9: Start listening for recognized word
Step 10: If word is recognized
        Step 11
    Else
        Step 9
Step 11: Display name in textbox
Step 12: Split word to surname and last name
Step 13: Search database and retrieve patient record
Step 14: Start speech synthesizer engine using Microsoft speech library
Step 15: Readout each field
Step 16: Stop

Fig. 4. Speech recognition algorithm

Step 1: Start
Step 2: Create speech recognition engine using Microsoft speech library
Step 3: Hook SpeechRecognized event
Step 4: Load grammar/dictionary with words
Step 5: Set audio input to default microphone
Step 6. Check audio level using default microphone
Step 7: Start listening for recognized word
Step 8: If word is recognized
        Step 9
    Else
        Step 7
Step 9: Perform action
Step 10. Stop

Fig. 5. Voice command algorithm

Step 1: Start
Step 2: Initialize Microsoft speech synthesizer
Step 3: Cancel all speech synthesis queue
Step 4: Check record from speech recognition
Step 5: If record was found from speech recognition
        Step 6
    Else
        Step 4
Step 6: Display record found message
Step 7: Readout patient name record found
Step 8: Readout out field name asynchronously
Step 9: Readout corresponding value asynchronously
Step 10: Dispose speech synthesizer object
Step 11: Stop

Fig. 6. Speech synthesis algorithm

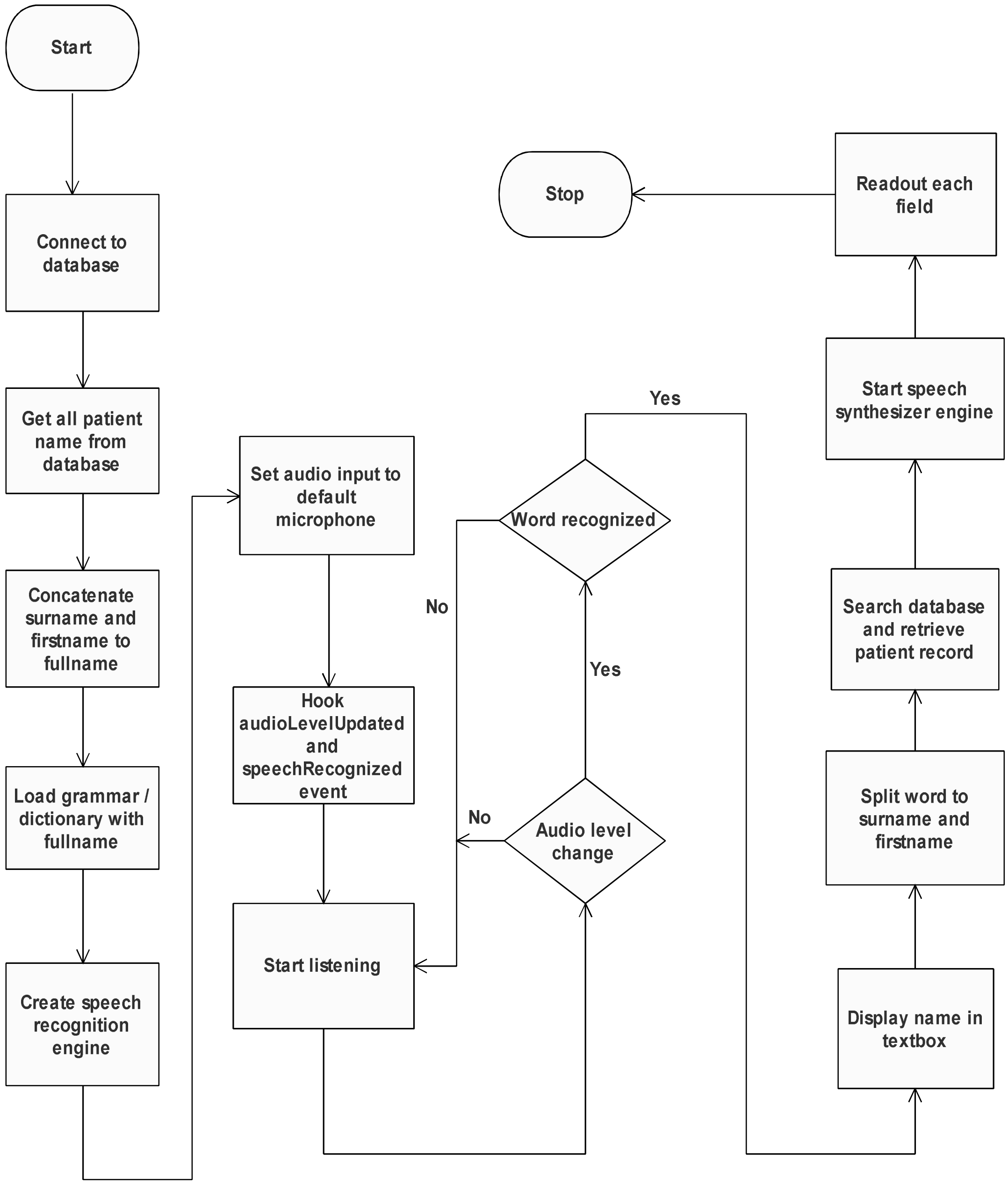

Fig. 7. Speech recognition flowchart

The system was broken down to modules and modules to submodules and the structural chart was devised. Fig. 8 shows the system structure chart that expresses the breakdown of the system to its lowest manageable levels. A structure chart is used to show the hierarchical arrangement of the module in a structured program; each rectangular box represents one module, the name of the module is written inside the box [11].

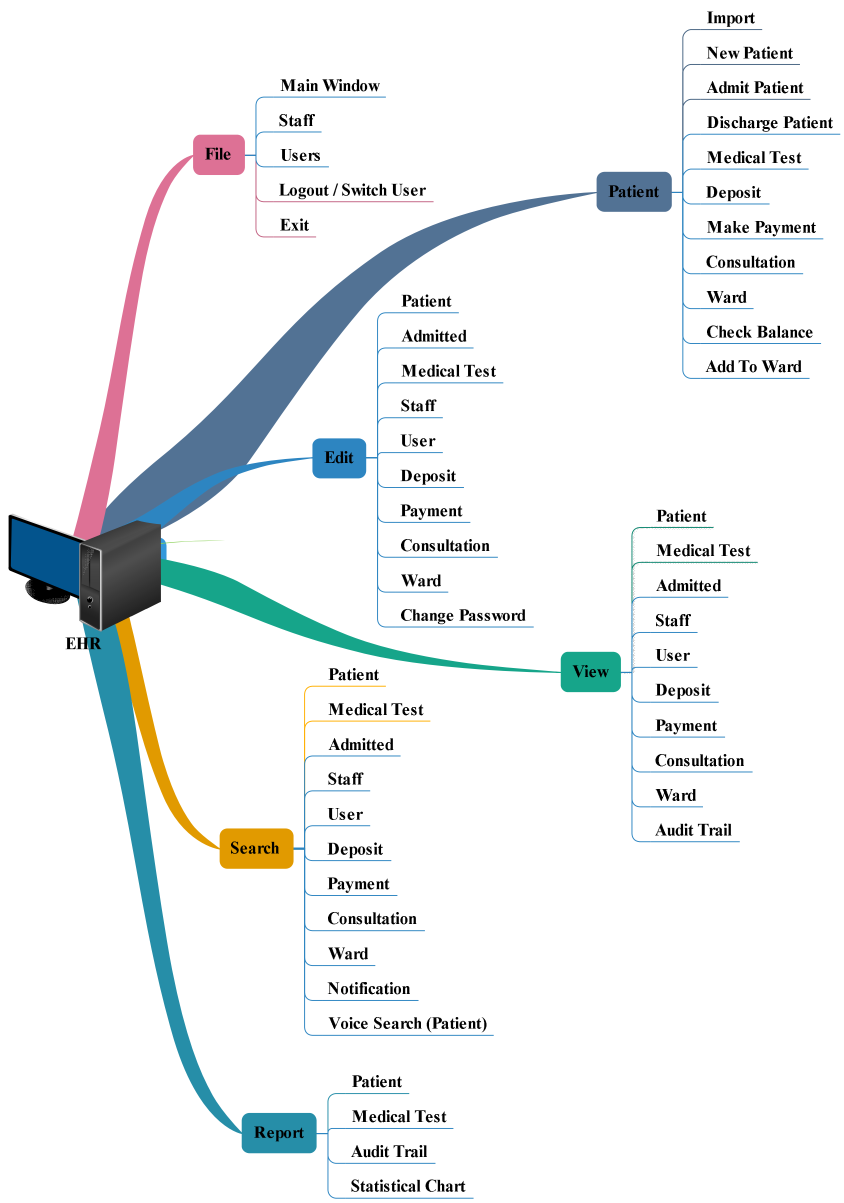

Fig. 8. System structure chart

The development phase of this work was carried out using C# programming language (.NET 4.5). It was implemented using Microsoft Visual Studio 2017 and Blend for Visual Studio 2017. The database was designed using Microsoft SQL server management studio 2016. Fields or columns are extracted from all the paper forms collated, entities where formed and a data model was created. Figure 9 shows the database Entities Relationship (ER) diagram. The software development platform used in this work is Windows Presentation Foundation (WPF), which is a graphical subsystem by Microsoft for rendering user interfaces in Windows-based applications. Model View View Model (MVVM) software architectural pattern was used. Entity framework, legacy ADO.net, Object Relational Mapping (ORM), which maps/converts entity/table to an object used in object-oriented programming languages where used to access the database.

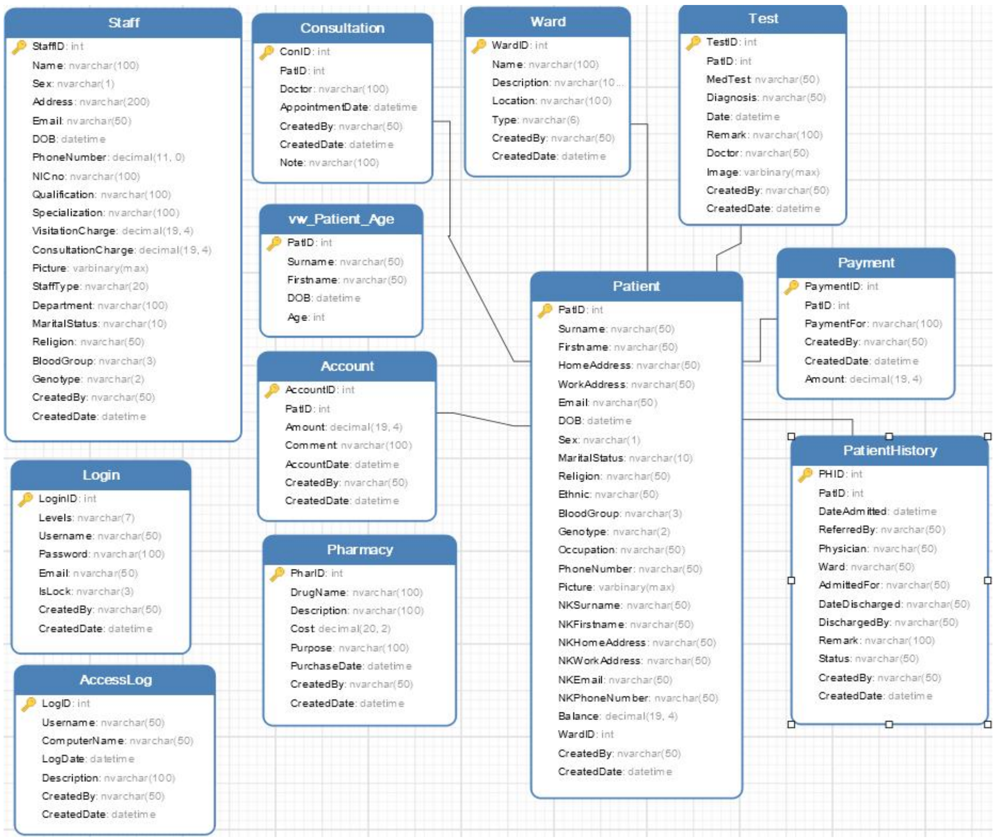

Fig. 9. Database Entity Relationship (ER) diagram

3 Results and Discussion

The system was broken down to modules and then to sub modules, which is further broken into its lowest manageable levels. The modules of the system are file, patient, edit, view, search, report and help. Each module has several sub-modules.

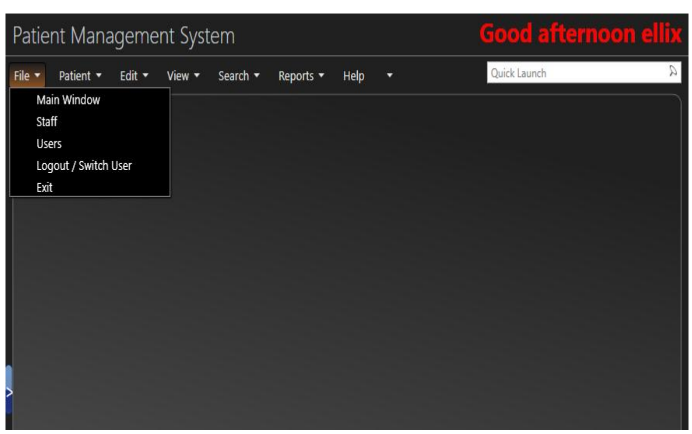

Fig. 10. File module and submodule

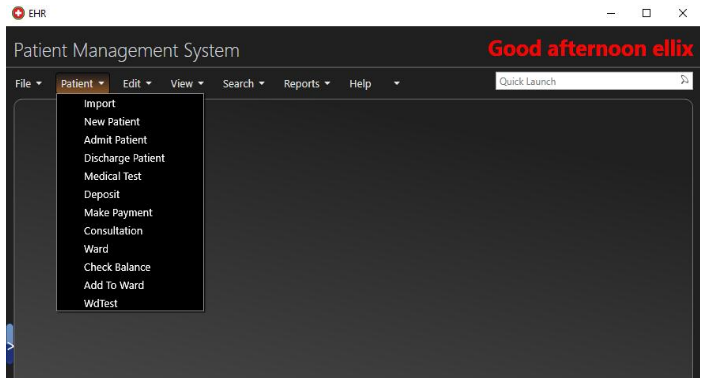

Fig. 11. Patient module and submodule

The system has voice command feature, which allows the user to navigate to some views by calling the view command. This gets rid of using mouse for navigation. Users can easily operate the system with their hands full or while doing other tasks. The system is speaker-independent, so it can respond to multiple voices, regardless of accent or dialectal influences. Commands like main window, edit staff, edit user, change password, open discharge, open medical test, open statistical chart and more can be called for action. Fig. 12 shows the staff registration view after "open staff" was called while Fig. 13 shows the main window of the application when "main window" was called.

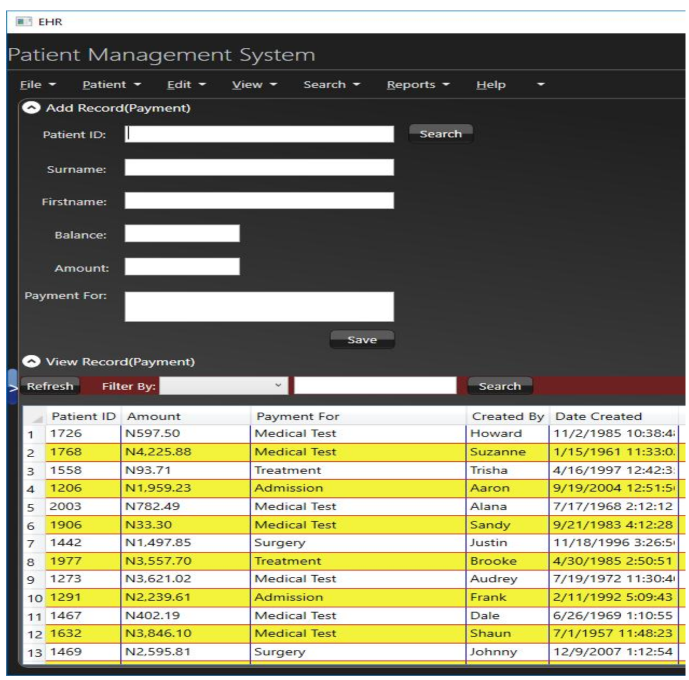

Fig. 12. Staff registration form by voice command

For speech recognition, Microsoft speech library that contains windows desktop speech technology types for implementing speech recognition was used. The speech recognition feature allows the user to search for the patient record using speech. Different names can be searched, names like Omowonuola Adenike, Chuckemeka Chibuzor, Kasali Kazeem, Maxwell John etc. Fig. 14 shows the sound property menu used to open sound property. The sound property was opened to set/adjust the property of the microphone. Fig. 15 shows the checking microphone status, Fig. 16 shows setting the decibels and microphone volume, Figs. 17-18 shows the voice search form when record was found.

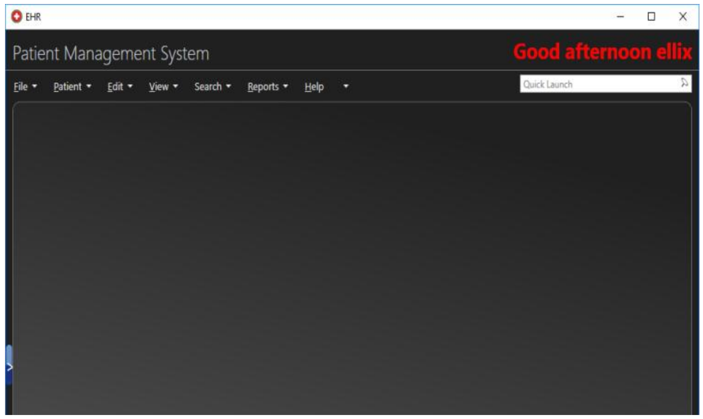

Fig. 13. Main window by voice command

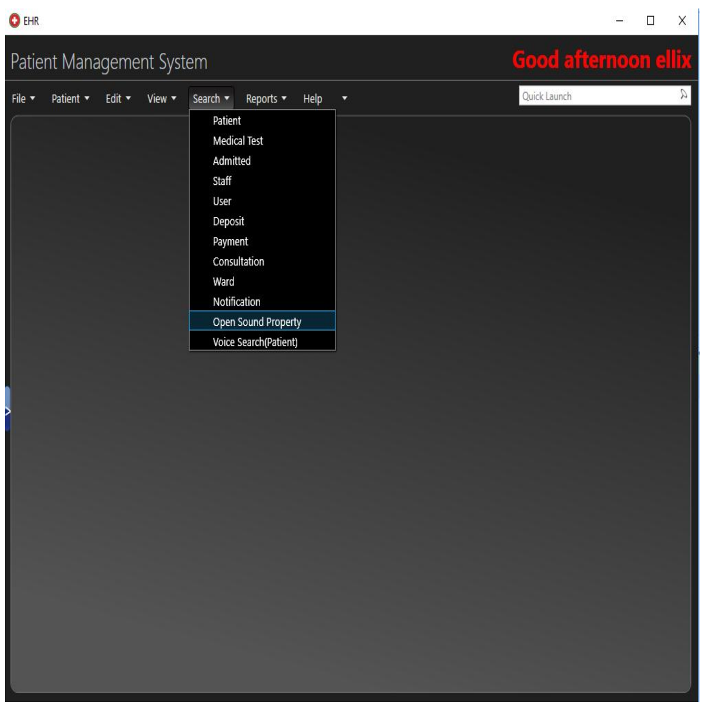

Fig. 14. Sound property menu

When a patient name “UmoruRasak" was searched, the speech recognition engine fires up to check the audio level that increases when the name was called. Since the name is in the grammar, the application searches the database to retrieve patient record using the name found in the grammar. The record of the patient was retrieved and the fields are added to the corresponding control. The speech synthesis then fires up to readout all the record retrieved.

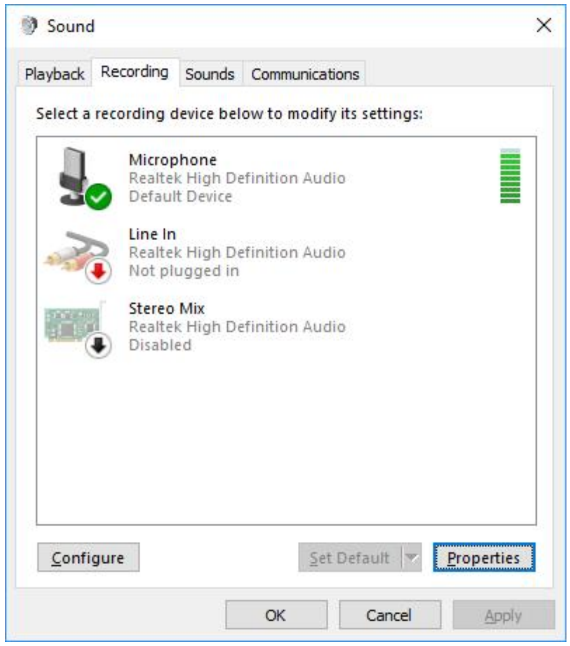

Fig. 15. Checking microphone status

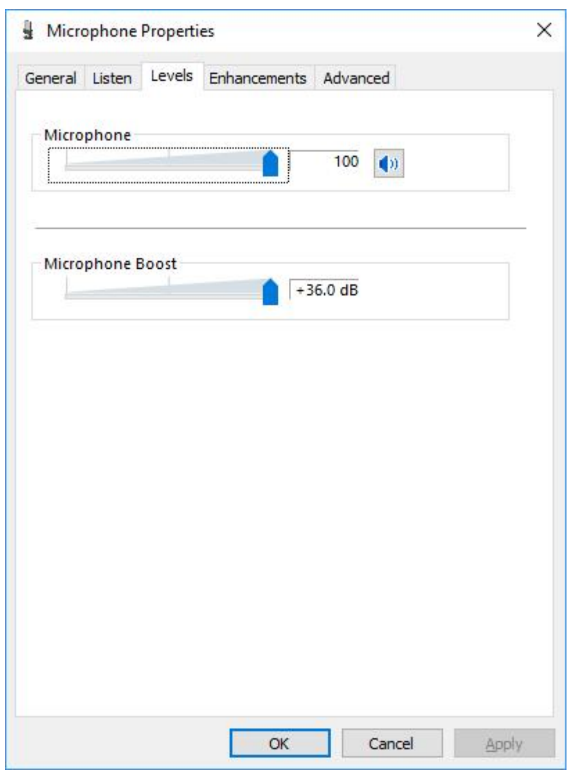

Fig. 16. Setting decibels and microphone volume

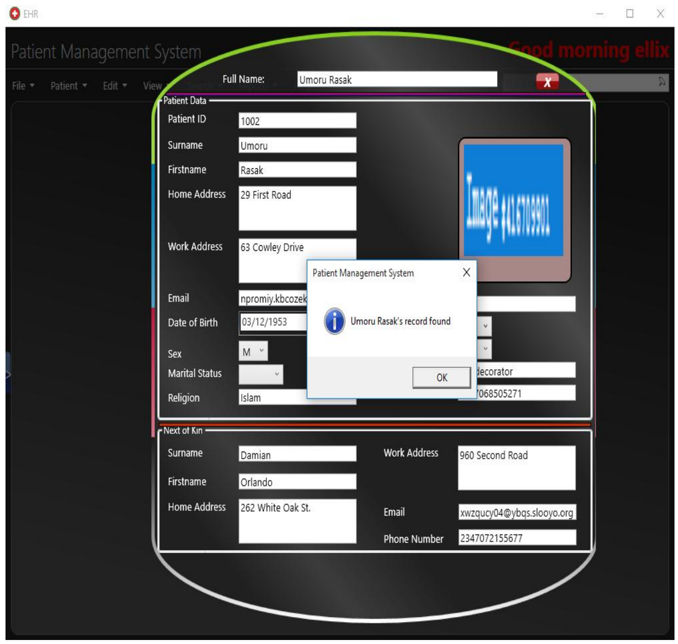

Fig. 17. Voice recognition form for patient Umoru

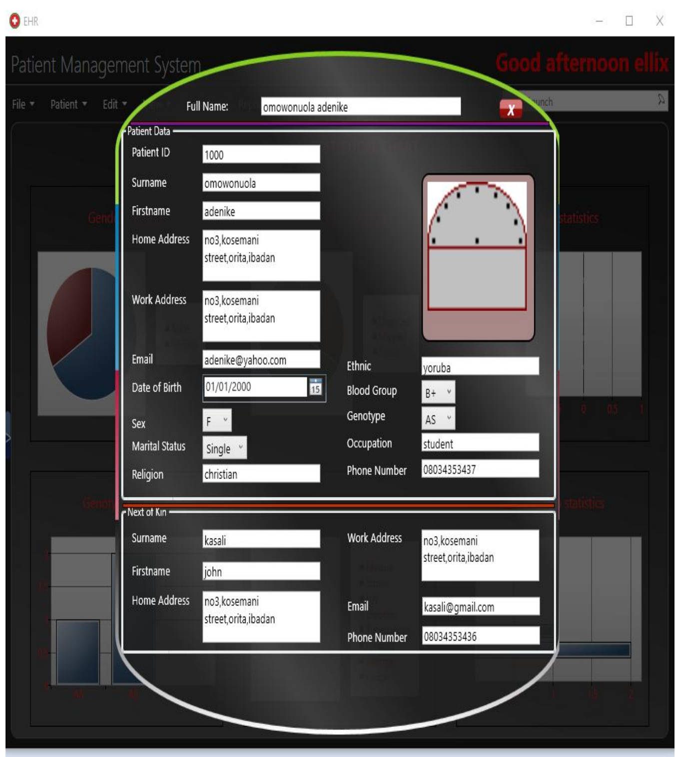

Fig. 18. Voice recognition form for patient Omowonuola

Word Error Rate (WER) was used as a performance evaluation metrics to test for the speech recognition as shown in Fig. 19.

50, 100, 150 and 200 words were tested and the correctness, accuracy and WER value were recorded in Table 1.

$$\text{Correctness (\%)} = \frac{N-D-S}{N} X \frac{100}{1} \quad \begin{aligned} &where\ S = No\ of\ substitutions \\ &D = No\ of\ deletions \\ &I = No\ of\ Insertions \\ &N = No\ of\ words\ in\ the\ reference \end{aligned}$$

$$\text{Correctness (\%)} = \frac{50-0-2}{50} X \frac{100}{1} = 96\%$$

$$\text{Accuracy (\%)} = \frac{N-D-S-I}{N} X \frac{100}{1} \quad \begin{aligned} &where\ S = No\ of\ substitutions \\ &D = No\ of\ deletions \\ &I = No\ of\ Insertions \\ &N = No\ of\ words\ in\ the\ reference \end{aligned}$$

$$\text{Accuracy (\%)} = \frac{50-0-2-0}{50} X \frac{100}{1} = 96\%$$

$$\text{Error(WER) (\%)} = \frac{D+S+I}{N} X \frac{100}{1} \quad \begin{aligned} &where\ S = No\ of\ substitutions \\ &D = No\ of\ deletions \\ &I = No\ of\ Insertions \\ &N = No\ of\ words\ in\ the\ reference \end{aligned}$$

$$\text{Error (\%)} = \frac{0+2+0}{50} X \frac{100}{1} = 4.0\%$$

Fig. 19. Performance evaluation based on word error rates for 50 words tested

Table 1. Showing correction, accuracy and word error rate

No of word test	Correctness (%)	Accuracy (%)	Error rate (%)
50	96 %	96 %	4.0 %
100	96 %	95 %	4.0 %
150	95 %	95 %	5.0 %
200	93 %	94 %	6.0 %

This work followed closely the works of [12] and [13] in which authors developed a streamlined Electronic Medical Record system to computerize the actions of patient medical record system in a typical Nigerian hospital. The system is plagued with the following among others: manual generation of patient ID by the admin, low security levels, old technology design (window forms) and the absence of speech recognition and speech synthesis capabilities. To improve on this, the developed application automatically generates patient ID, developed with recent tools/ technology and has speech recognition for easy user interaction. When compared to a study by [14] that used linear SVM regression model, where 90% accuracy was obtained, Table 1 shows the performance of the speech recognition system has been improved.

4 Conclusion

This study developed an improved EPHR management system with speech recognition that supports human speech interaction. Architectures, flowcharts and algorithms of the system were designed to develop the application. The system was developed with C# programming language (DotNet 4.5) using visual studio 2017. Latest and recent technology/tool like Windows Presentation Foundation (WPF), Entity Framework, Object Relational Mapping (ORM) and more have been used to develop the system. Future research work may be geared toward developing a web-based version of the system.

Consent

As per international standard or university standard written participant consent has been collected and preserved by the authors.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Margaret Rouse . Electronic health record (EHR); 2017.
Available: https://searchhealthit.techtarget.com/definition/electronic-health-record-EHR
(Accessed 13 August 2018)

[2] Top Mobile Trends. Mobile Tech Contributions to Healthcare & Patient Experiences; 2014. Available:http://topmobiletrends.com/mobile-technology-contributions-patient-experience-parmar/. (Accessed 11 May 2016)

[3] Anusuya MA, Katti SK. Speech recognition by machine: A review. International Journal of Computer Science and Information Security. 2009;6(3).
Available: https://arxiv.org/ftp/arxiv/papers/1001/1001.2267.pdf (Accessed 13 August 2018)

[4] Sunanda M, Neelam T, Dipali B. isolated word recognition system for speech to text conversion using ann. Special Issue - Energy, Environment, And Engineering; 2016.
Available: https://www.iioab.org/articles/IIOABJ_7.11_78-91.pdf (Accessed 17 September 2018)

[5] Potale S, Kharpude P, Patil R, Ajay K. Design and development of word recognition for marathi language. International Journal of Computer Science and Information Technologies. 2016;7(5):2337-2340.
Available: http://ijcsit.com/docs/Volume%207/vol7issue5/ijcsit20160705041.pdf.
(Accessed 17 September 2018)

[6] Parwinder S, Bhupinder S. Speech recognition as emerging revolutionary technology. International Journal of Advanced Research in Computer Science and Software Engineering. 2012;2(10).
Available:https://pdfs.semanticscholar.org/5a49/ddd68b765d7ff7c2ce93db85959119ae7eb7.pdf
(Accessed 17 September 2018)

[7] Fifth Generation Computer Corporation. Speaker Independent Connected Speech Recognition. Fifthgen.com. Archived from the original on 11 November 2013.
Available: http://www.fifthgen.com/speaker-independent-connected-s-r.htm
(Accessed 13 August 2018)

[8] Smita SH. A review on speech synthesis an artificial voice production. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering. 2014;3(3).
Available:http://www.rroij.com/open-access/a-review-on-speech-synthesis-an-artificialvoice-production.pdf (Accessed 17 September 2018)

[9] Kaladharan N. An English text to speech conversion system. International Journal of Advanced Research in Computer Science and Software Engineering. 2015;5(10).
Available: http://ijarcsse.com/Before_August_2017/docs/papers/Volume_5/10_October2015/V5I10-0228.pdf (Accessed 17 September 2018)

[10] Shikha K, Ritika A. Auto spell suggestion for high quality speech synthesis in Hindi. International Journal of Computer Applications (0975 – 8887). 2014;87(17).
Available: https://arxiv.org/ftp/arxiv/papers/1402/1402.3648.pdf (Accessed 23 August 2018)

[11] James M, Carma LM. Structured techniques: The Basis for CASE. 1988;56.

[12] Ajala F, Awokola J, Emuoyibofarhe O. Development of an Electronic Medical Record (EMR) system for a typical Nigerian hospital. Journal of Multidisciplinary Engineering Science and Technology (JMEST). 2015;2(6).
Available: http://www.jmest.org/wp-content/uploads/JMESTN42350149.pdf. (Accessed 11 May 2016)

[13] Mary Y, Boniface E, Ikenna I. Automated Patient Information Systems for Federal Government Hospitals in Nigeria; 2012.
Available:http://www.ajol.info/index.php/wajiar/article/viewFile/86645/76454
(Accessed 11 May 2016)

[14] Michael L, Shuangyu C, Bruce B, Nick K. End-To-End speech recognition accuracy metric for voice-search tasks. Speech at Microsoft, Microsoft Corporation; 2012.
Available:https://www.microsoft.com/en-us/research/wp-content/uploads/2012/03/Levit-SearchBasedAccuracy.pdf. (Accessed 17 September 2018)

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
http://www.sciencedomain.org/review-history/27527