CENTRALIZED DATABASE SYSTEM FOR AUTOMATED HEARING SCREENING

MOHD SAAD BIN HAMID

MASTER PROJECT REPORT SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ENGINEERING

FACULTY OF ENGINEERING AND BUILT ENVIRONMENT
UNIVERSITI KEBANGSAAN MALAYSIA
BANGI

2014

SISTEM PENGKALAN DATA BERPUSAT UNTUK SARINGAN
PENDENGARAN AUTOMATIK

MOHD SAAD BIN HAMID

LAPORAN PROJEK SARJANA YANG DIKEMUKAKAN UNTUK MEMENUHI
SEBAHAGIAN DARIPADA SYARAT MEMPEROLEH
IJAZAH SARJANA KEJURUTERAAN

FAKULTI KEJURUTERAAN DAN ALAM BINA
UNIVERSITI KEBANGSAAN MALAYSIA
BANGI

2014

DECLARATION

I hereby declare that the work in this thesis is my own except for quotations and summaries which have been duly acknowledged.

18th January 2014   MOHD SAAD HAMID

P65385

ACKNOWLEDGEMENTS

ABSTRACT

Centralized database software enables simultaneous updates from multiple queries containing data from different machines. It enables the data to be stored safely in a centralized location. Based on the related research work on PC based Automated Hearing Screening kit which store the test results in local storage, it may have problem to perform the data collection from different location of test centers. The main objective of this project is to extend the capabilities of the kit into centralized database architecture for easy access of the test results from multiple test centers locations. The proposed methods used in this project involved database structure design through user requirements analysis, defining data type for each entities and modeling the database using entity relationship diagram (ERD). Web application interface has been proposed in this project as a method to extract the test results from the database. The PC based Automated Hearing Screening kit also has been redesigned in term of user interface and the functionality to submit the data to the centralized database has been added. As part of the project results, the SQL injection test has been performed on web application interface and the web application successfully passed the test by sanitizing user input in the web application. An End User test has been conducted at Audiology & Speech Sciences Clinic, Universiti Kebangsaan Malaysia with 9 participants involved. All participants passed the hearing screening test for the left and right ear except for 1 participant unable to conduct right ear test due to the detected ambient noise level in the test room beyond the permissible level. As the conclusion the objective of this project has been fulfilled and the capabilities of original PC based Automated Hearing Screening kit has been successfully extended based on the results presented in this report.

ABSTRAK

Perisian pengkalan data berpusat membolehkan pelbagai penambahan daripada pelbagai arahan yang mengandungi maklumat daripada berbilang mesin. Ia membolehkan data tersebut disimpan dengan selamat di dalam satu tempat. Berdasarkan hasil kerja penyelidikan berkaitan kit saringan pendengaran automatik yang berasaskan PC yang menyimpan keputusan ujian di dalam simpanan dalaman, ia akan mempunyai masalah untuk melaksanakan pengumpulan data tersebut daripada pelbagai lokasi pusat-pusat ujian yang berbeza. Tujuan utama projek adalah untuk megembangkan lagi keupayaan kit tersebut kepada satu senibina pengkalan data berpusat untuk memudahkan capaian kepada keputusan-keputusan ujian tersebut daripada pelbagai lokasi pusat-pusat ujian. Kaedah yang dicadangkan untuk digunakan dalam projek ini meliputi reka bentuk struktur pengkalan data melalui analisa kehendak penguna, menentukan jenis data untuk setiap entiti dan pemodelan pengkalan data menggunakan rajah hubungan entiti. Antaramuka applikasi web telah dicadangkan sebagai satu kaedah untuk mengeluarkan keputusan-keputusan ujian daripada pengkalan data tersebut. Kit saringan pendengaran automatik yang berasaskan PC tersebut juga telah melalui perubahan dari segi antaramuka pengguna dan fungsi untuk menghantar data tersebut ke pengkalan data berpusat juga telah ditambah. Sebagai sebahagian dariapada hasil daripada projek, ujian suntikan SQL telah dilaksanakan pada antaramuka applikasi web dan applikasi web tersebut berjaya lulus dalam ujian tersebut dengan menapis input pengguna yang digunakan dalam applikasi web tersebut. Satu ujian pengguna akhir telah dijalankan di Klinik Audiologi & Sains Pertuturan, Universiti Kebangsaan Malaysia dengan 9 peserta terlibat dalam ujian tersebut. Semua peserta lulus ujian saringan pendengaran untuk telinga kiri dan kanan kecuali 1 peserta tidak dapat menjalankan ujian telinga kanan disebabkan tahap hingar yang dikesan di dalam bilik ujian di luar tahap yang dibenarkan. Sebagai kesimpulan, objektif utama projek ini telah dipenuhi dan keupayaan asal kit saringna pendengaran automatik yang berasaskan PC tersebut telah berjaya dikembangkan berdasarkan keputusan yang dibentangkan dalam laporan ini.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANSI   American National Standards Institute

ASHA   American Speech-Language-Hearing Association

CLI   Command Line Interface

CSS   Cascading Style Sheet

ERD   Entity Relationship Diagram

GUIDE   Graphical User Interface Development Environment

GUI   Graphical User Interface

HTML   Hypertext Markup Language

HER   Electronics Health Record

HCI   Human Computer Interaction

SQL   Structured Query Language

ODBC   Open Database Connectivity

PC   Personal Computer

UKM   Universiti Kebangsaan Malaysia

WWW   World Wide Web

XSS   Cross Site Scripting Attack

OWASP   The Open Web Application Security Project

LIST OF APPENDICES

CHAPTER I

INTRODUCTION

1.1   OVERVIEW

Nowadays, we live in the era where most of the products, ranging from consumer, industrial, medical and up to military products require human interaction. This interaction requires some interface for human to operate the product. So, manufacturer of the products need to implement good interface which exhibit the usability of the products. Basically, there are two types of interface to help users to interact with the product. The first type of interface is the Command Line Interface (CLI). In the early era of computer most of the applications were written as CLI based, where the user friendly usability factor was not the main focus of the person who develops the interface. For the CLI, the interaction between user and the application mostly by using keyboard. The user will enter the predefined text as command through keyboard to the application to perform the specific action requested by user. However, nowadays most of the developed end user applications focus on the second interface, the Graphical User Interface (GUI) as mentioned by in related paper (Rauf et al. 2010). The application that implements the GUI handles the interaction between user and application mostly by using keyboard and mouse. User will move the cursor on the screen to the desired graphical user interface elements to execute the required action.

As the computer and communication technology evolves rapidly in the recent decade, the development and implementation of software also evolved. From standalone software running on single PC, the software nowadays also can be run from a server or even the cloud computing environment. Even though the atmosphere of the computing platform rapidly changed, the basic need for software still remains unchanged. The basic needs of software are to perform data manipulation and storage.

There are various types of data that can be stored and processed, from day-to-day application data, scientific engineering data and even as Electronics Health Record (EHR).

Centralized database software is a type of software that enables simultaneous updates from multiple queries containing data from multiple users. It can be used to store, update and also delete the stored data. Centralized database software enables users from various computer platforms to store their data in one place. Regardless what happen to the users' computer hardware, the data stored in the centralized database will remain safe.

1.2   PROBLEM STATEMENT

As mentioned in a paper (Gunter 2005), the EHR is an evolving concept defined as a longitudinal collection of electronic health information about individual patients and populations. EHRs may contain various types of data ranging from demographic, medical history information, personal information and even laboratory test result. This valuable information will provide more data for health practitioner to identify any possible problem or to predict the health pattern from the stored EHRs. So it is crucial to develop centralized database software for storing electronic health records.

The current version of Autokit software (Gan et al. 2012) has been developed as a standalone PC application which based on MATLAB scripting environment. The authors successfully conducted clinical test for the hearing screening by using the Autokit software. The test results also have been stored in single PC. The results mentioned in the paper proved that the software can be installed for multiple test centers which extend the coverage of hearing screening. In order to extend the capability of the software to serve for multiple test centers, it has to overcome the problem of collecting test results. Because it will involve different locations and the results have to be extracted from every PC installed with the Autokit software. To overcome this problem, a centralized database for storing the results has to be developed. By using centralized database it is easier to access the results without having to go to the every test center to collect the results. It can be accessed anywhere

by using internet connection. Security authentication procedure has to be implemented to protect the privacy of the data.

The original Autokit software also has to be redesign in term of GUI in order to submit the results into the centralized database. The mentioned clinical test has been performed to the subjects aged between 19 and 79 years old (Gan et al. 2012). So the GUI also has to be redesign to help the end users which are normally not an IT savvy user. It should minimize technical configuration and also should focus on broader range of users, the school children, teenagers and adult users. This will help the user to use the software effectively.

1.3   RESEARCH OBJECTIVE

The objectives for this research as follows:

- To implement and configure centralized database for storing test results.
- To redesign the graphical user interface (GUI) for Hearing Screening Test standalone application.
- To develop web application interface for database management.

1.4   SCOPE OF WORKS

The scope of work for this research project is to redesign the interface for End-User software. The interface design will be done using MATLAB scripting and toolboxes as well as MATLAB GUI development environment. The MATLAB version R2011b.has been used for the purposes. The results from this standalone application will be stored into a centralized database server. For the database server application, MYSQL server application will be used. The process of configuring and implementing database elements is part of the project scope. In order to gain access for the database and easy manipulation of the data from the screening test, a web interface will be design using PHP, CSS, Javascript and HTML programming. Both

MATLAB based software and Web based software will be configured to interact with the centralized MYSQL server.

1.5   ORGANIZATION OF DISSERTATION

This report contains five chapters. The arrangement of the chapter as follows:

Chapter I provides introduction on current style of software. It also discuss about the need of the proposed software as well as the problem statements that motivate the development of the software.

Chapter II mainly includes the literature review and fundamental concept on designing the software as well as the related security issues with web application.

Chapter III explains the methodologies implemented throughout the project. The problem formulations are first identified and the countermeasure proposed to tackle the issues also presented in this chapter.

Chapter IV discusses about the results obtain in this project where it contains figure and tables related to the project. Software assessment result also will be presented in this chapter.

The final Chapter V discusses about the conclusions and future work recommendations based on the results presented and analyzed in Chapter IV.

1.6   SUMMARY

This chapter briefly described about the overview of the human machine interaction and software. Next, the problem statement base on problem in related research work has been explained. The chapter briefly listed the research objective as motivation for this project. The scope of works on this project also has been defined in this chapter. The final part of this chapter explained about the organization of the dissertation.

CHAPTER II

LITERATURE REVIEW

2.1 INTRODUCTION

In order to have a better understanding of this project, it is necessary to discuss the overall background related to basic technology used to complete the software. This chapter highlights the needs for hearing screening together with the related work about the earlier prototype. Then the fundamental explanation pertaining to the process of developing or designing Graphical User Interface, database system as well as the system configuration will be discussed in this chapter. Finally, the security issue related to web application security will be discussed in greater details.

2.2 BACKGROUND

As mentioned in American Speech-Language-Hearing Association (ASHA) websites (ASHA 2013), hearing screening is a quick and cost-effective way to separate people into two groups: a pass group and a fail group. The persons in the pass group is believed to have no hearing impairment while those who belongs to the failed group have the risk of having hearing loss and may need to perform a detailed test conducted by certified audiologist. Based on facts from World Health Organization (WHO) websites (WHO 2013), more than 5% of the world's population or 360 million people – has disabling hearing loss (328 million adults and 32 million children). The loss refers to hearing loss greater than 40 dB in the better hearing ear in adults and a hearing loss greater than 30 dB in the better hearing ear in children. Most of these people come from low and middle-income countries.

There many variation of hearing screening software published. As mentioned in paper (Gan et al. 2012), the authors successfully developed a computer based hearing screening device. The software was developed using MATLAB environment. It is a standalone computer based program with the capability of real time noise level monitoring. It is to confirm that the screening is performed in an environment that follows to the standard (American National Standards Institute S3.6-2004 standard). Figure 2.1 illustrates the interface of the program (Gan et al. 2012). The software stores the result in the Microsoft Access file through Open Database Connectivity (ODBC) connection. However, the test results of the hearing screening only available on the installed computer. The next section of this report will explain the relationship between the paper and the propose work to be implemented to extend the capability of the software.
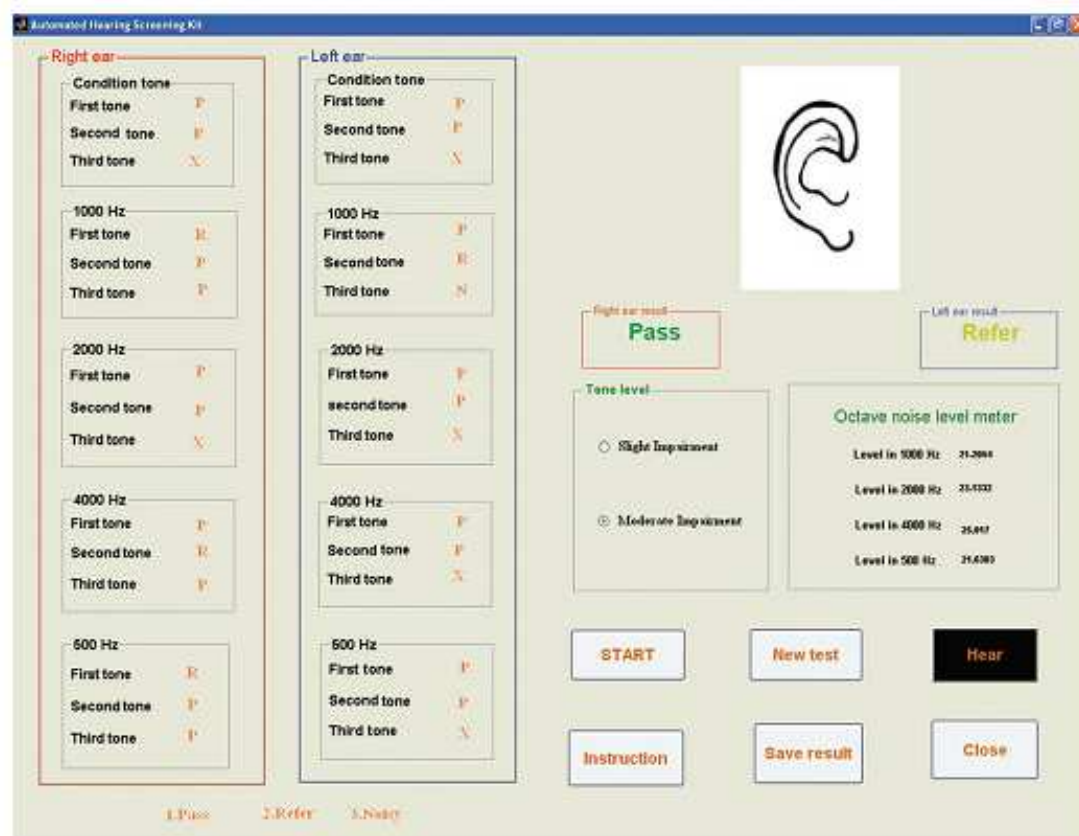

Figure 2.1 Graphical User Interface

Source: Gan et al. 2012

2.3   USER INTERFACE DESIGN RULES

As mentioned in the earlier section, the purpose of having GUI is to enable user friendly factor in the software interface. This is part of Human Computer Interaction (HCI) field. It involves interaction between users (human) and computer (machine). Typical PC application developed with GUI obtains input from user by using keyboard and mouse. User interface design creates effective communication medium between human and a computer (Pressman 2010). So, a software designer should try follows the guiding principle to design the user interface. Theo Mandel in his book (Mandel 1997) highlights three golden rules as user interface design principle.

The first rule is to place user in control. Typically software developer will tend to limit or impose restriction on the interface to simplify the implementation which will caused frustration in users' side. The interface should provide flexible interaction. With the advancement in display and touch sensor technology, software developer has to extend their GUI design to receive another possible input from users such as the touch screen. This will help to improve user experience when they are using the software. The technical internals such as automated procedure should be hiding from casual user. The interface should help the user to control the application and not the other way around. Second rule is to reduce user's memory load. The probability of user to perform error increases when they have to remember too many information. For example try to mimic the real world metaphor. Use appropriate icon in a button to perform certain action. For example, use printer icon to represent print action. If a user has to perform certain amount of tasks, let the users know their current position. For example, display current progress from overall tasks. The third rule mentioned by Mandel is to make the interface constant. It includes the visual style for the interface (color, fonts and background). It should be consistent for multiple interface software. For example, the navigation from one interface to another interface should follow the same style. As mentioned by Aziz (2012), a good GUI design removes the impediment of communication with the computer system and allows the user to work directly on the problem at hand. It shows that the good GUI supposed to help user by automating certain task and focus the real function of the program.

2.4   DATABASE DESIGN

Almost all of the modern databases are relational databases. Its purpose is to provide facilities for fast storage and able to retrieve large quantity of data. One of the characteristic of relational database is to have unique key to identify each rows in a table. It will help to perform sorting of the rows in the table based on their keys. In relational model, every data will be stored once in the same location. The purpose is to avoid data redundancy in relational database. In relational database, we can specify type of data to be stored in the table by constraining the field type into specific type (dates, numbers, small text and other type) and field length into specific length. By following the relational database standards, we should be able to move our database into different platform freely without having to worry that our database will not working. For the electronic health record, as mentioned by Taylor, the schema of any proposed relational database needed to store such clinical data is not known until the information is collected and ready for inclusion into the database (Taylor et al. 2004). So the designed must be ready for future growth.

In order to achieve the level, it is advisable to perform database normalization. A good database design practice also encourage database to be normalized. Database normalization has some significant improvement in terms of the efficiency and maintainability of the database. It enables us to use simple Structured Query Language (SQL) queries to perform complex data selection. It helps to maintain integrity of the data for reliable data storage. As mentioned by Stephens and Plew (2001) the normalization of database will help to reduce redundancies of data in the database. Normalization also provides the crucial aspect of database system which is the scalability of the database as well as the system. Scalability helps to deal with future growth of the system. The database will be able to maintain the performance as the user s' data grows.

The fundamental question when designing or planning a database is to know the purpose of the database. Then identify what are the inputs and output actions that related to the database. After this step we should be able to determine what the tables to be created in the database and what is type of data field to be included for each of the tables. Before creating the tables and fields in the database, we should model the

database by using data modeling tool such as MySQL Workbench. The MySQL Workbench will produce a graphical database diagram or known as the Entity Relationship Diagram (ERD). As mentioned in paper (Sani et al. 2012), ERD is a conceptual database design which is a graphical representation used to depict efficient database design. The ERD will illustrate the relationship of the available entities in the database and its attributes. It is also shows the relationship between entities in the database. By modeling the database earlier, it is easier to analyze any logical error before the creation of the database. The Figure 2.2 illustrates the ERD implemented for Smart Medical Clinical System.
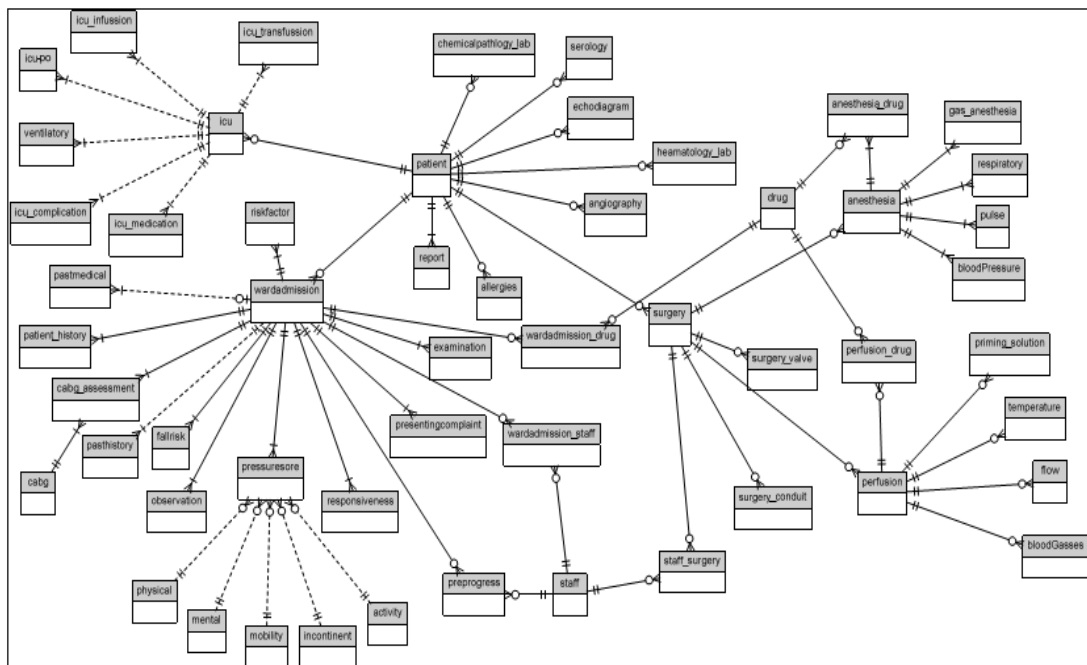

Figure 2.2 Entity Relationship Diagram for Smart Medical Clinical System

Source: Sani et al. 2012

2.5   DATABASE SYSTEM CONFIGURATION

In order to implement the centralized database system for Hearing Screening, the basic database system configuration has to be created. As illustrated in Figure 2.3, we can see that the Internet will be the medium for data to be transmitted from multiple input devices into the database server for data storage. At the other side several services can be provide base on the data analysis performed on the data stored in the

database. For the configuration of the hearing screening database, the database will receive multiple results from multiple test centers such as schools. Then the Certified Audiologists will be able to view and analyses the test results from the services provided from the hearing screening web application. The web application will provide more information for execution of detailed test as well as planning for hearing loss prevention based on demographic and geographical information from the database. This will be based on the system implemented by Nakajima et al. (2012) except that the types of services and equipment might be different.
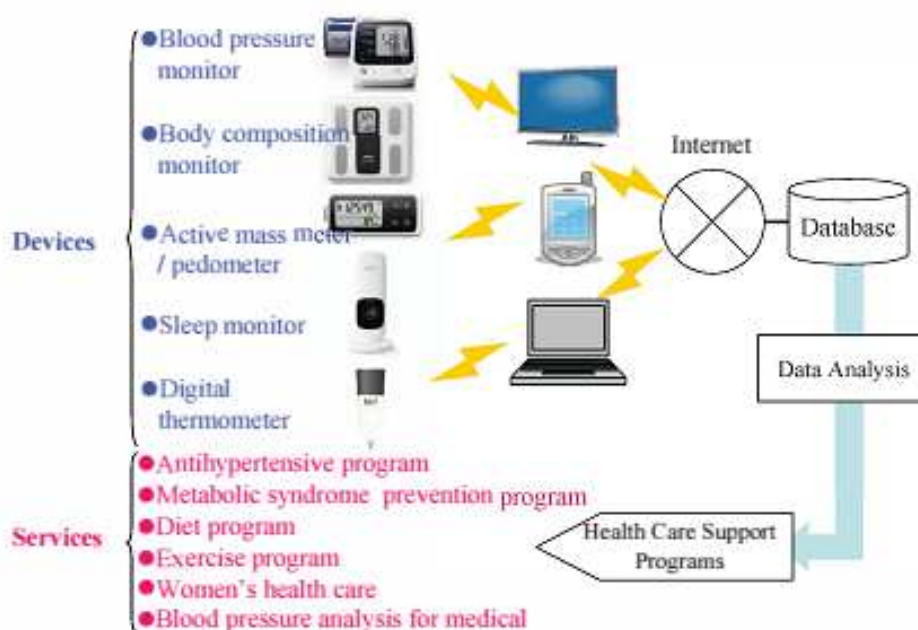

Figure 2.3 Overview of System Configuration

Source: Nakajima et al. 2012

2.6   WEB APPLICATION SECURITY

With the introduction of Web 2.0, web application, blogs, social networking sites that provides information sharing becomes common nowadays. It enables users to interact and collaborate to create user generated content on the web. The way web content being published nowadays becomes more flexible as compared to the era of static web content. In the early era of World Wide Web (WWW) the content are mostly static where user only receive and not able to change the published content. Nowadays, the