

A FRAMEWORK FOR INTELLIGENT VOICE-ENABLED E-EDUCATION SYSTEMS

Azeta A. A.

Ayo C. K.

Ikhu-Omoregbe N. A.

Atayero A. A.

**College of Science and Technology
Covenant University, Ota, NIGERIA**

ABSTRACT

Although the Internet has received significant attention in recent years, voice is still the most convenient and natural way of communicating between human to human or human to computer. In voice applications, users may have different needs which will require the ability of the system to reason, make decisions, be flexible and adapt to requests during interaction. These needs have placed new requirements in voice application development such as use of advanced models, techniques and methodologies which take into account the needs of different users and environments. The ability of a system to behave close to human reasoning is often mentioned as one of the major requirements for the development of voice applications.

In this paper, we present a framework for an intelligent voice-enabled e-Education application and an adaptation of the framework for the development of a prototype Course Registration and Examination (CourseRegExamOnline) module. This study is a preliminary report of an ongoing e-Education project containing the following modules: enrollment, course registration and examination, enquiries/information, messaging/collaboration, e-Learning and library.

The CourseRegExamOnline module was developed using VoiceXML for the voice user interface(VUI), PHP for the web user interface (WUI), Apache as the middle-ware and MySQL database as back-end. The system would offer dual access modes using the VUI and WUI.

The framework would serve as a reference model for developing voice-based e-Education applications. The e-Education system when fully developed would meet the needs of students who are normal users and those with certain forms of disabilities such as visual impairment, repetitive strain injury (RSI), etc, that make reading and writing difficult.

Keywords: e-Education, Intelligent, Voice, VoiceXML, VUI and WUI

INTRODUCTION

The rapid advances in information and communication technology (ICT), including the Internet, have had significant impact on various aspects of the daily lives of mankind and the society. One of the areas that has been highly affected by ICT is education. Technological advances and wide availability of personal computers, Compaq Disks (CDs), the web, broadband access to the Internet, etc, have been used as supporting tools in electronic Education (e-Education) (Kim, 2007), which is often defined as the use of ICT for supporting learning processes.

VoiceXML (also known as VXML) is one of the tools for developing voice-enabled e-Education applications; it is a web-based markup language for representing human-computer dialogs, just like HyperText Markup Language (HTML). But while HTML assumes a graphical web browser, with display, keyboard and mouse, VoiceXML assumes a voice browser with audio output (computer-synthesized and/or recorded), and audio input (voice and/or keypad tones) (Gallivan et al 2002). VoiceXML technology allows a user to interact with the Internet through voice-recognition technology by using a voice browser and/or the telephone. The major goal of VoiceXML is to bring the advantage of web-based development and content delivery to Interactive Voice Response (IVR) system (voiceportalwhitepaper, 2001). With VoiceXML, the product and services available on a corporate website can be accessed via the telephone on a self-service basis.

Human-to-human (H2H) interaction is one of the conventional approaches for communicating and receiving online services using a telephone. With this technique, the speed and success of getting a desired service and completing a task may depend on the human third party. Moreover, majority of present day e-Learning applications only have support for the web user interface (WUI) through the use of a personal computer (PC) and Wireless Application Protocol (WAP) using the mobile phone, with little or no support for voice. This has motivated a lot of research in the provision of voice user interface (VUI) support in the education domain through an IVR system.

Voice-enabled systems are applicable in several areas, such as information providers, financial institutions, e-health, education, etc. For example, in the financial institutions, an IVR system allows customers to call and request information; like calling your bank to confirm your balance (Azeta et al 2008). In educational institutions, a voice response application can provide information about class schedules, availability and course content.

Students can register their courses using the telephone and the application that handles the registration process can also update the database containing enrollment information. A voice response application can call students to inform them of schedule changes or openings in a class for which enrollment has been closed (Mult & Reusch, 2004). Some people have difficulty in typing on the computer keyboard due to physical limitations such as repetitive strain injuries (RSI) and many others. For example, people with hearing difficulty could use a system connected to their telephone to convert the caller's speech to text (Cook, 2002).

Self-service voice-enabled e-Education systems deliver basic learning management services by simply providing a platform for learners to share learning contents and collaborate with one another through the use of a telephone. A typical telephone web e-Education application provides e-Learning materials that can be accessed via the web as well as via the telephone. Voice driven interfaces will also be of great benefit to people who are unable to leave their home due to disability, providing them with a portal to the community simply using a telephone handset (VoiceXMLforum, 2004).

The objective of this paper is to provide a framework for an intelligent voice-enabled e-Education application. The framework was adapted to develop a prototype Course Registration and Examination (*CourseRegExamOnline*) module, as part of an ongoing e-Education development research containing: enrollment, course registration and examination, enquiries/information, messaging/collaboration, e-Learning and library module. The remaining part of the paper is organized as follows: Section 2 presents a review of related literature. In section 3, the statement of the problem is enumerated. Section 4 contains the proposed framework and implementation of *CourseRegExamOnline* module. Section 5 describes the results and discussions, while section 6 contains the conclusion of the paper.

REVIEW OF RELATED LITERATURE

This section presents an overview of voice-enabled e-Education systems. Voice-enabled systems allow users to access information on the Internet or intranet through a telephone interface. It uses technologies such as speech recognition and text to speech (TTS) conversion to create a user interface that enable users to navigate through a dialogue system using telephone and voice commands (Gallivan et al 2002).

In addition to the provision of alternative platform for normal users, voice-enabled systems can be helpful for people with physical access difficulties (e.g. repetitive strain injury, arthritis, high spinal injury) that make writing difficult (Donegan, 2000). It can also be effective for students with reading, writing or spelling difficulties (e.g. dyslexia) and for those with visual impairment (Nisbet & Wilson, 2002). One very interesting application area of speech technology is education. As part of a larger project, a couple of tools for computer assisted language learning environment were developed (Kirschning, 2001). These tools are a first prototype for pronunciation verification and a bilingual dictionary, both for Spanish language students, whose native tongue is American English. Another very interesting aspect of the project is the use of speech technology based systems to support language acquisition for deaf children. Development of voice applications using VoiceXML for higher institutions of learning have remain an open area of research all over the world. For instance, Gallivan et al (2005) presented a VoiceXML absentee system that enables students to report their absence in class through telephone calls. The Absentee System application was developed basically for Pace University students to report class absences and stored in the University database. The VoiceXML absentee system has been designed to include record keeping of absentee calls from students, faculty and university staff. Furthermore, Schindler (2005) developed a chat application for communicating between the deaf and blind.

The goal of the project was to incorporate current speech recognition (speech-to-text) and speech synthesis (text-to-speech) technology into a chat room that is both free and does not require any additional equipment besides a desktop computer. The system was developed using C++ and run on client/server technology with a graphical user interface (GUI) for the client. The system may also be used in educational settings, regardless of students' or teachers' disabilities, as a teaching aid. Chin et al (2006) recommended that one can actually make use of VoiceXML technology to build speech applications that can serve educational purposes or in other words, to build an online learning system that provide better accessibility to users. One of the e-Education applications that can be provided using speech technology are those that deliver basic teaching by simply listening. For example, students can check their scores or other information by simply calling a particular number and get the information they want. The authors went further to develop a prototype based on VoiceXML concepts. However, the prototype was not deployed on the real telephone environment. Instead, a microphone and a speaker were used to simulate the telephone environment.

STATEMENT OF THE PROBLEM

The development of speech applications has been an active area of research. Despite the numerous research prototypes and commercial services, the development of speech applications is yet to reach its prime, and the potential of speech-based interaction has not been fully utilized. In order to facilitate the development of advanced speech applications, we need advanced techniques, models, framework, methodologies and tools to take voice technology beyond its present position (Turunen, 2004). A framework suitable for practical speech applications must provide components for a variety of application requirements, including dialog management, speech recognition and natural language understanding (Turunen et al 2005). In the last couple of years, various advancements in speech technology have been reported. However, voice-enabled applications suffer from some problems when compared to its counterpart graphical/visual applications, such as limited conversation since VoiceXML applications can only handle sentences and commands they are programmed for. Thus, the limitation of the existing e-Education systems is inability to take certain decisions beyond the spoken words or queries of users at a particular time as a result of the insufficient level of intelligence exhibited by these applications. Thus, the elimination of these limitations of dialogue systems will improve the level of user's adaptability, flexibility and intelligence of voice-enabled e-Education applications.

THE FRAMEWORK AND IMPLEMENTATION

The proposed framework, deployment architecture and implementation of *CourseRegExamOnline* module are presented in this section. The information acquired through requirement elicitation was used to develop the framework for intelligent voice-enabled e-Education systems. The prototype *CourseRegExamOnline* module was developed using VoiceXML for the voice user interface (VUI), PHP for web user interface (WUI), Apache as middle-ware, MySQL database as backend. The choice for these tools is because of their advantage as free and open source software (Siemens, 2003).

The framework

Figure: 1 gives the software view of the proposed framework for intelligent voice-enabled e-Education systems. The architecture shows the locations of each of the component services in the system. It consists of the presentation tier, business logic tier (made up of system, intelligent and application services), and database tier.

The presentation tier

The presentation tier provides client access to the system. The client do not store or process any form of data. They only provide an interface for the system through the VUI and WUI. In the case of VUI, data or files or voice browsers are not stored on the client due to resource constraints associated with telephone and hand-held devices. The VUI allows voice browsers (running in the voice gateway) to be used as interface. The information from the database is presented in a compatible form to the VUI client through text to speech translation. The voice browser simply receives any call into the application and submits to the voice gateway for further processing. The WUI provide access to the application through an Internet browser.

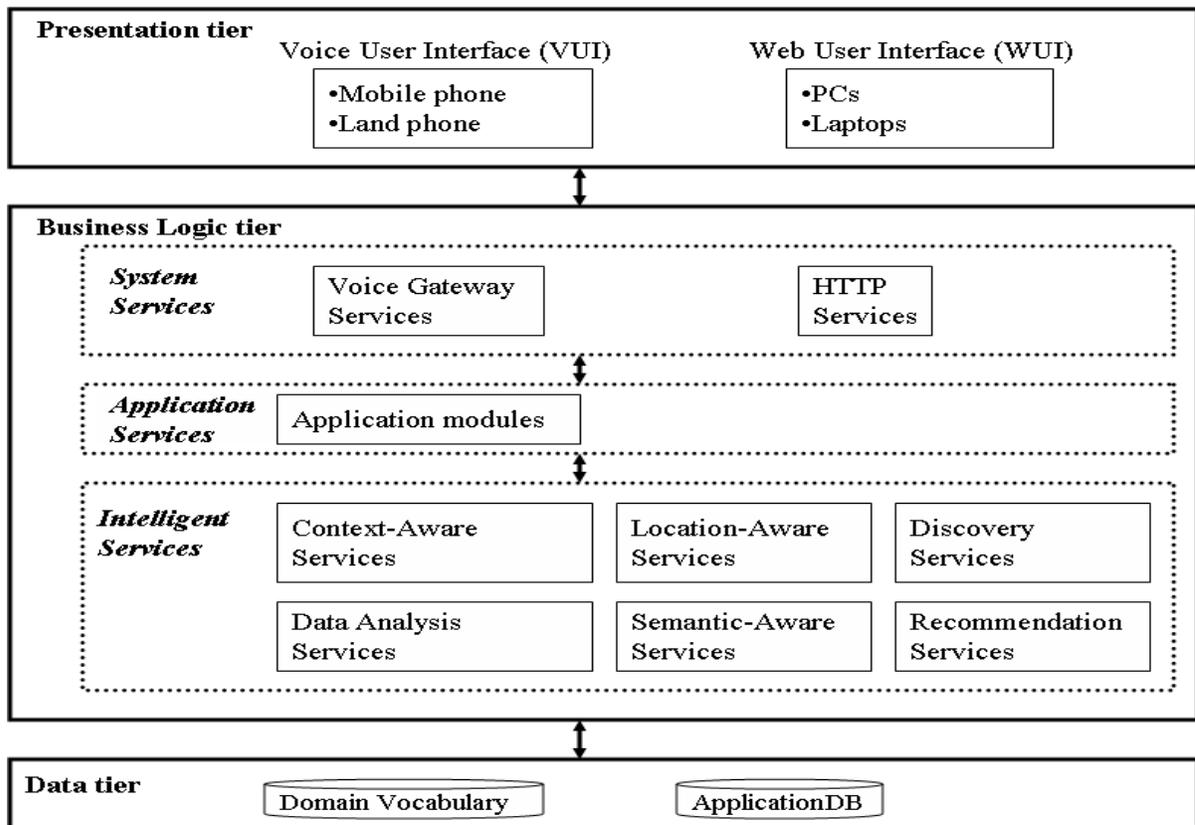


Figure: 1
A framework for intelligent voice-enabled e-Education systems

The business logic tier

The business logic tier comprises *system, intelligent and application services*. The presentation tier communicates with the business logic tier through the *voice gateway services*. The *system services* sub tier contains the *voice gateway and Hypertext Transmission Protocol (HTTP) services*.

A telephone user access the application through the VUI from various mobile and land line phone in real-time. Once a telephone user has been authenticated, the user's query is translated by the *voice gateway services* and passed to the *application services* for further processing. A user can only access the module for which he or she is authorized. The client application engages the VUI and WUI to connect with the business logic tier using the *voice gateway* and the *HTTP services* respectively. The *application services* contain all the application modules for the system.

Data tier

The data tier provides data services and data base management system functions. The data tier is responsible for changing, adding, or deleting information in the database within the system. Any relational database such as MySQL, MS SQL Server, MS Access, etc, may be used for the implementation of the data tier. Technologies exist for enabling each of the services in the business logic tier. For instance, intelligent information retriever system (IIRS) (Belkin 2005) and case-based reasoning (CBR) (Spasic et al 2005) are the technologies for enabling *semantic-aware and recommendation services*.

The Deployment Architecture

The intelligent voice-enabled e-Education application runs on a three layer deployment architecture: client device, server and database (Figure: 2)

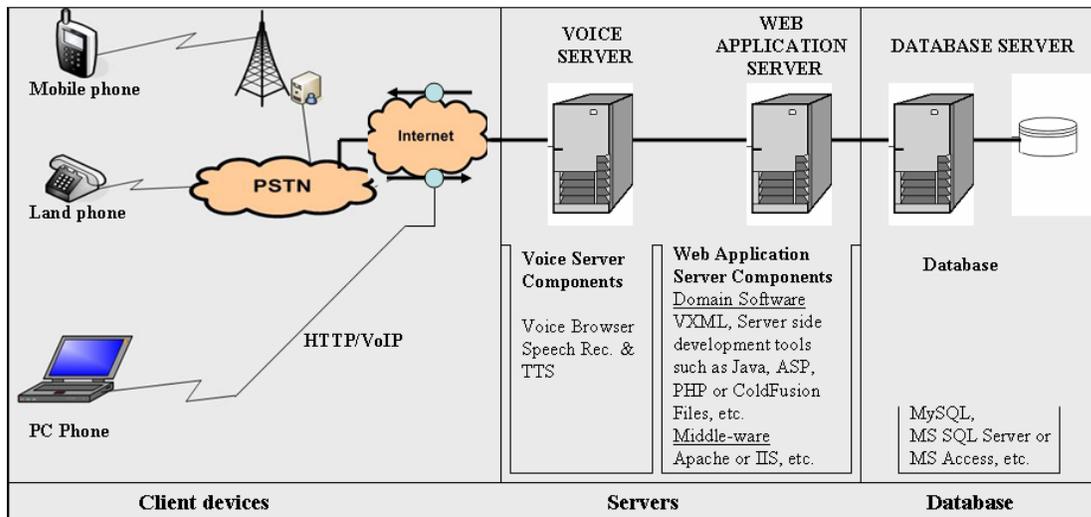


Figure: 2

A three-tier deployment architecture for voice-enabled e-Education systems

The client device includes web and hand-held devices such as mobile phones and PDAs and land telephones. The server includes the voice server and application server. The database contains any relational database.

Implementation of *CourseRegExamOnline* module

The proposed framework was adapted for developing *CourseRegExamOnline* module of e-Education system as a case study.

The cooperation between the different actors in the *Course Registration and Examination* module is modelled with UML collaboration diagram and class diagram in Figure: 3 and 4 respectively.

The collaboration diagram reveals the internal details of *Course Registration and Examination*.

The UML is a visual language that provides a means to visualize, construct and document the artefacts of software systems (Simeon et al 2005).

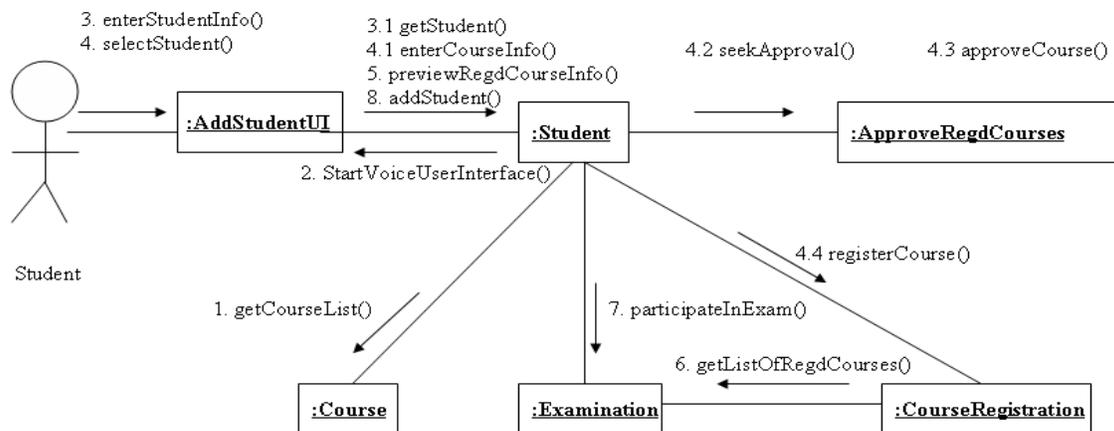


Figure 3:
Collaboration diagram for *CourseRegExamOnline* module

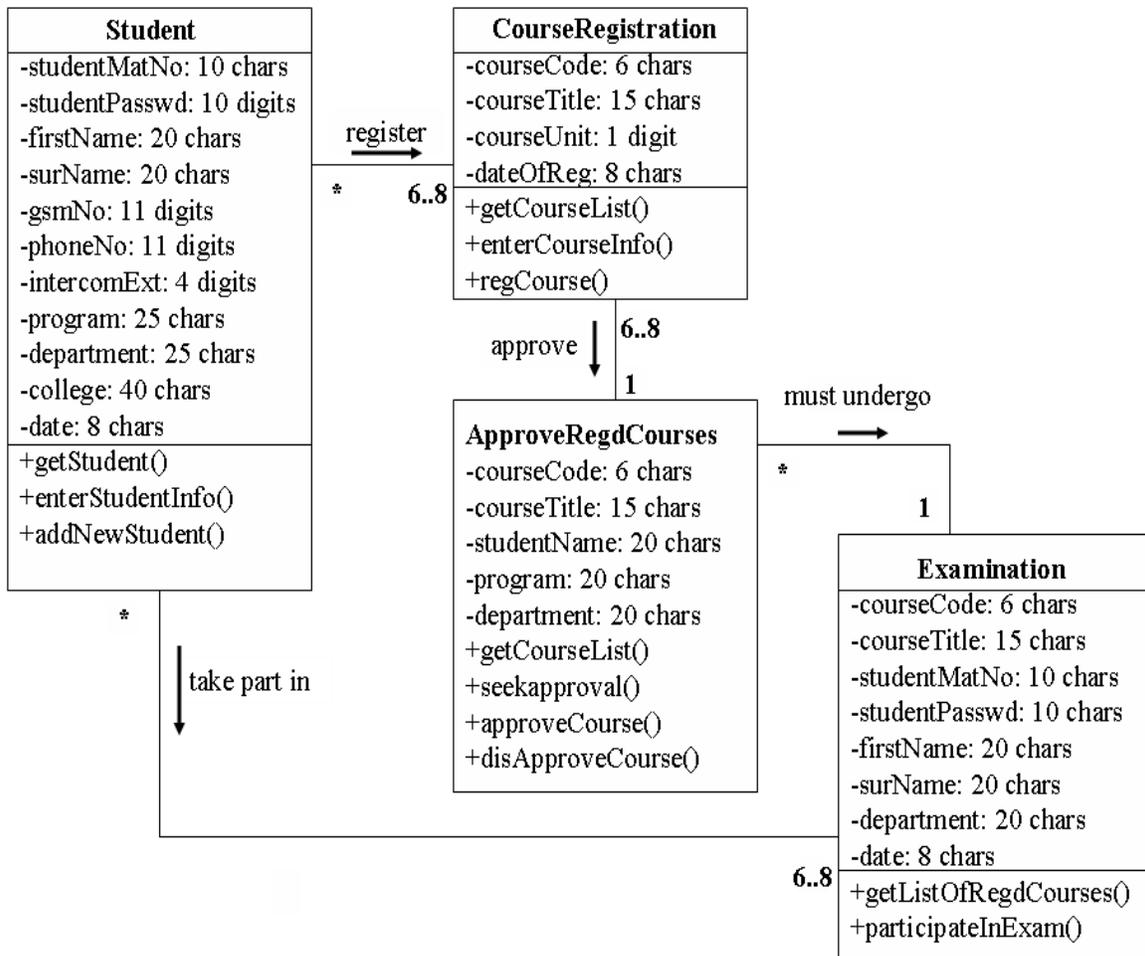


Figure: 4
Class diagram for *CourseRegExamOnline*

Two interfaces would be created for the application, first, the WUI, that would provide enrolment to these services and access to information, and second, a VUI where users call in using mobile or land phones to get services or record their requests. The system will store student records and examination records. Figure 5 shows software architecture of *CourseRegExamOnline* module based on the proposed framework.

A voice browser provides an interface between the caller and the different components of the voice server (automated speech recognition (ASR), text-to-speech (TTS), etc). The user's query is translated by the ASR to text and passed to the database server for execution. The TTS does the reverse of translating text to speech.

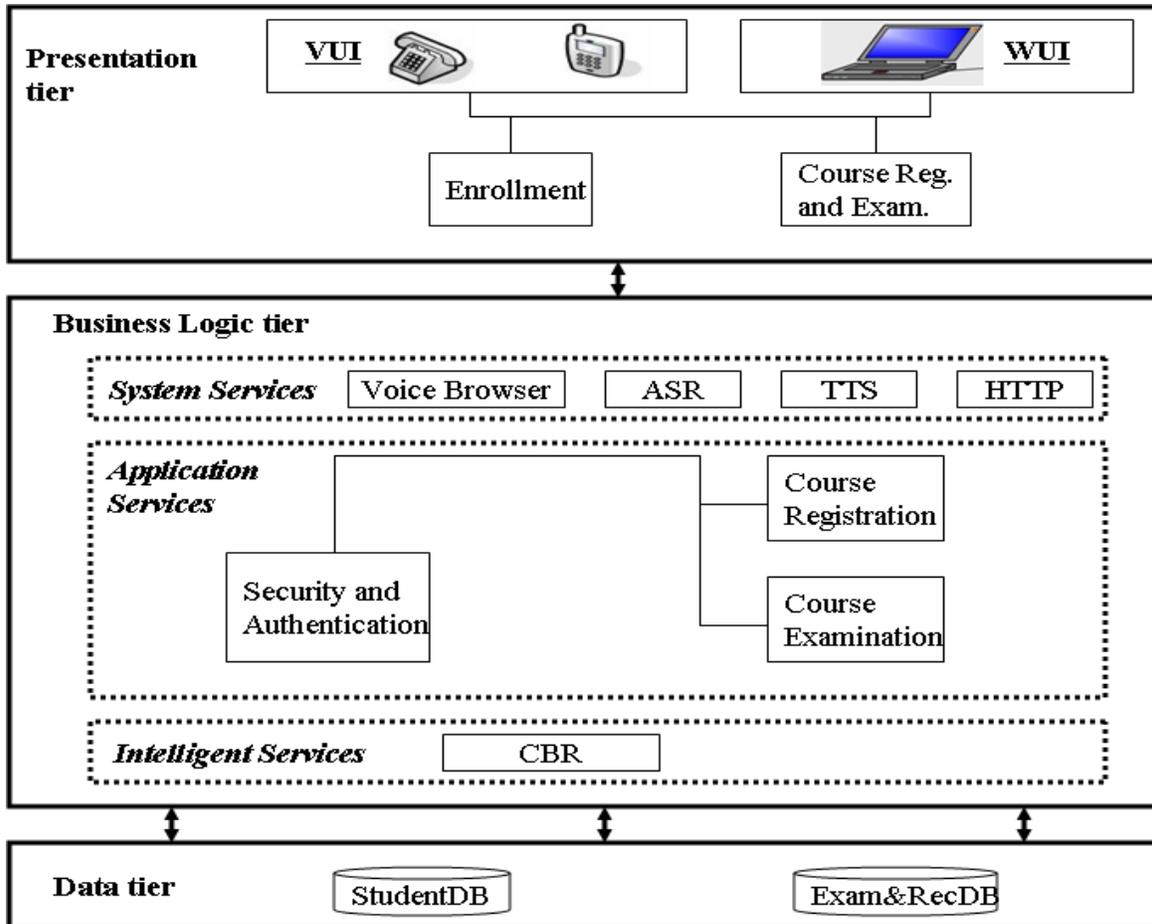


Figure: 5
A software architecture of *CourseRegExamOnline* module

RESULTS AND DISCUSSIONS

The VoiceXML and PHP code were deployed to voice server (<http://community.voxeo.com>) and a free application server (<http://byethost.com/index.php/free-hosting>) respectively. The application is launched using a mobile or land phone by dialling the recommended phone number using the format: <source country code><destination country code><destination area code><generated voice network 7-digit number>.

Dialling the telephone number: 009-1-312-2390285 will connect a caller to the application. The default username is "admin" and password is "admin". Once connected, the system will be prompted with a welcome message and go ahead to authenticate the user's name and password before any transaction can take place.

The system will ask for the services demanded by a student and go ahead to process the request, either course registration or e-Examination as shown in Figure 6.

SYSTEM:	Welcome to Covenant University e-Education Portal.
SYSTEM:	What is your user name?
USER:	admin
SYSTEM:	What is your password?
USER:	admin
SYSTEM:	You have successfully login. Select any of the following options. Say one for course registration. Say two for examination. Say three to exit.
USER:	One
SYSTEM:	Welcome to course registration session. Your login information are as follows. Full Name: Adelabu Uyi Emmanuel. Matriculation number: 04CJ01277. Year of study: 3. Session registered: 2007/2008. Program of study: B Sc Computer Science. Semester: Alpha. Carry over courses: None. Compulsory courses: CSC311, Discrete Mathematics, 3 units. CSC312, Data structure and algorithm, 3 units. CSC313, computer programming 4, 2 units. CSC314, System administration, 2 units. For the following Elective courses, listen and say "select" or "deselect" at the end of each line. CSC316 logic programming, 3 units, say "select". CSM216 Information system analysis and design, 2 units, say "deselect". . . Good bye.

Figure: 6
Sample conversation between the system and a user

CONCLUSION

In this paper, we have presented a framework for an intelligent voice-enabled e-Education application and a prototype *CourseRegExamOnline* module, as part of preliminary work of an ongoing e-Education project containing the following modules: enrollment, course registration and examination, enquiries/information, messaging/collaboration, e-Learning and library. The e-Education application when fully developed will have the capability to improve learning and administrative processes using telephone and web-based technologies. The proposed framework will serve as reference for intelligent voice-based e-Learning applications, which can be adapted in the Education domain. The e-Education application will reasonably reduce human-to-human (H2H) contact (such as teacher to student, etc) by replacing it with human-to-system (H2S) interactivity. The VoiceXML-based e-Education application, when fully developed, will not only serve as an alternative platform for non-physically challenged learners, but be used successfully by students with visual and certain forms of disabilities. The voice-based e-Education system will be helpful for people with physical access difficulties (e.g. repetitive strain injury, high spinal injury) that make reading and writing difficult. In a situation where users are not allowed to carry telephone or where cost is an issue, an alternative for them is to use the PC phone through voice over Internet protocol (VoIP).

BIODATA and CONTACT ADDRESSES of AUTHOURS



Azeta, A. A. is a Ph.D student in the Department of Computer and Information Sciences, Covenant University, Ota, Nigeria. He holds B.Sc. and M.Sc. in Computer Science from University of Benin and Lagos respectively. His current research interests are in the following areas: Software Engineering, Algorithm Design and Mobile Computing. He currently lectures at Covenant University. He is a member of the Nigerian Computer Society (NCS) and Computer Professional Registration Council of Nigeria (CPN).



Charles K. Ayo holds a B.Sc., M.Sc. and Ph.D in Computer Science. His research interests include: mobile computing, Internet programming, e-business and government, and object oriented design and development. He is a member of the Nigerian Computer Society (NCS), and Computer Professional Registration Council of Nigeria (CPN). He is currently the Head of Computer and Information Sciences Department of Covenant University, Ota, Ogun state, Nigeria, Africa. Dr. Ayo is a member of a number of international research bodies such as the Centre for Business Information, Organization and Process Management (BIOPoM), University of Westminster. <http://www.wmin.ac.uk/wbs/page-744>; the Review Committee of the European Conference on E-Government, <http://www.academic-conferences.org/eceg/>; and the Editorial Board, JICTHD.



Ikhu-Omoregbe, Nicholas has a B.Sc degree in Computer science from the University of Benin, Benin city, an M.Sc degree in Computer Sciences from the University of Lagos, and a PhD degree in Computer Science from Covenant University, Ota, Nigeria. His research interests include: Software Engineering, Mobile Computing, Mobile Healthcare and Telemedicine Systems, and Soft Computing. He currently lectures at Covenant University. He is a member of the Institution of Electrical and Electronics Engineers



Dr. Atayero, Aderemi Aaron Anthony holds a B.Sc. degree (summa cum laude) in Radio Engineering and M.Sc. in Satellite Communication Systems (1992, 1994 respectively) both from the Moscow Institute of Technology, and a Ph.D. in Technical Sciences (Speech Processing / Satcom, 2000) from the Moscow State Technical University of Civil Aviation, Moscow, Russia. His current area of research interest is as follows: Speech Processing. FPGA Design, Architecture, and Applications: FPGA–Implementation of Digital Speech Processors. FPGA–Implementation of Digital Filters. FPGA System–on–Programmable Chip Design. He is a member of a number of academic and professional organisations including the Institute of Electrical and Electronic Engineers (IEEE), Nigerian Association of Inventors (NAI). Dr. Atayero is a Senior Lecturer in the department of Electrical and Information Engineering at Covenant University

REFERENCES

- Azeta A. A., Ikhu-Omoregbe N. A., Ayo C. K., & Atayero A. A. (2008). "Development and Deployment of VoiceXML-Based Banking Applications", *Journal of Nigeria Computer Society*, Vol. 15 No.1 June '08 Edition, pg 59-72.
- Belkin N. J. (2005). "Intelligent Information Retrieval: Whose Intelligence", Understanding and Supporting Multiple Information Seeking Strategies, a TIPSTER Phase III Research Project, School of Communication, Information and Library Studies, Rutgers University
- Chin C. C., Hock G. T. & Veerappan C. M. (2006). "VoiceXML as Solution for Improving Web accessibility and Manipulation for e-Education", *School of Computing and IT, INTI College Malasia*. Available online at: <http://intisi.edu.my/INTISJ/InfoFor/StaffResearch/10.pdf> accessed July 2008.
- Cook S. (2002). Speech Recognition HOWTO, available at: <http://tldp.org/HOWTO/Speech-Recognition-HOWTO/software.html> accessed July 2008
- Donegan, M. (2000) *BECTA Voice Recognition Project Report. BECTA*, Available online at : <http://www.becta.org.uk/teachers/teachers.cfm?section=2&id=2142> accessed July 2008.
- Gallivan P., Hong Q., Jordan L., Li E., Mathew G., Mulyani Y., Visokey P. & Tappert C., (2002), "VoiceXML Absentee System, *Proceedings of MASPLAS'02*. The Mid-Atlantic Student Workshop on Programming Languages and Systems Pace University.
- Kirschning I. (2001) "Research and Development of Speech Technology & Applications for Mexican Spanish at the Tlatoa Group". *Proceedings of ACM CHI 2001 Conference on Human Factors in Computing Systems 2001*, pg 49-50. Available online at: http://ict.udlap.mx/people/ingrid/ingrid/CHI_Tlatoa.pdf accessed July 2008.
- Kim W. (2007), "Towards a Definition and Methodology for Blended Learning", *Blended Learning*, pp. 1-8, Pearson, 2007. Workshop on Blended Learning '07, U. K.
- Mult H. C. & Reusch J. A. (2004), "VoiceXML – Applications for E-Commerce and E-Learning. Available online at, accessed July 2008. <http://www.fh-dortmund.de/de/ftransfer/medien/reusch1.pdf>,
- Nisbet, P.D. & Wilson, A. (2002). *Introducing Speech Recognition in Schools: using Dragon Natural Speaking*. Pub. CALL Centre, University of Edinburgh. ISBN 1 898042 22
- Schindler L. (2005), "Web-Based Education: A Speech Recognition And Synthesis Tool", Department of Mathematics and Computer Science, in the College of Arts and Science at Stetson University, DeLand, Florida. Available online at and accessed May, 2008. <http://www.stetson.edu/~helaarag/Laura-proposal.pdf>

Siemens G. (2003). "Open source content in education: Part 2 - Developing, sharing, expanding resources". Available online at and accessed April, 2008.

http://www.elearnspace.org/Articles/open_source_part_2.htm

Simeon B., S. John, L. Ken, *Schaum's Outlines UML*, 2nd Edition, McGraw-Hill International, UK, 2005

Spasic I., Ananiadou S. & Tsujii J. (2005). "MaSTerClass: a case-based reasoning system for the classification of biomedical terms". Available online at:

<http://bioinformatics.oxfordjournals.org/cgi/content/full/21/11/2748>, accessed 5th July 2005.

Turunen M. (2004). 'Jaspis – An Adaptive Speech Application Architecture', *Speech based and Pervasive Interaction Group. IFHO'04*. Available online at and accessed July 2008.

<http://www.cs.uta.fi/hci/spi/reports/Jaspis-IFHOH'04.pdf>

Turunen M., Hakulinen J., Raiha K., Salonen E., Kainulainen A. & Prusi P. (2005). "An

Architecture and applications for speech-based accessibility systems", *IBM Systems Journal*, Volume 44 No 3. pg 485 – 584.

Voiceportalwhitepaper(2001). Available online at and accessed March 2008.

<http://www.medialab.sonera.fi/workspace/VoicePortals.pdf>

VoiceXMLforum(2004). " The VoiceXML forum, a program of IEEE Industry Standard and Technology Organisation (IEEE-ISTO). Available online at: www.voicexml.org

and <http://cnm.open.ac.uk/projects/phonetheweb/access.html> , accessed 25th February 2008.