A Case-Based Reasoning Approach for Speech-Enabled e-Learning System

A. A. Azeta, C. K. Ayo, A. A. Atayero and N. A. Ikhu-Omoregbe
College of Science and Technology
Covenant University, Ota, Nigeria
azeta_ambrose@yahoo.com, ckayome@yahoo.com, atayero@ieee.org, omoregbe@ieee.org

Abstract* - *E-Learning plays an important role in our society today; hence, higher institutions now offer courses through distance learning. Several studies and methodologies towards improving e-Learning have been proposed and provided. However, not too many works are dedicated to the design and implementation of e-Learning for the visually impaired learners. Sight challenge is a serious form of disability, yet, the existing e-Learning platform (web, mobile, etc) have not devoted enough attention to the plight of the visually impaired particularly in the area of usability. The objective of this paper is to present an intelligent speech-based e-Learning system with dual interface – Voice User Interface (VUI) and Web User Interface (WUI). Case-Based Reasoning (CBR) was engaged to provide intelligent services. Voice Extensible Markup Language (VoiceXML) was used to develop the VUI, Hypertext Preprocessor (PHP) for the WUI and Apache as the middle ware. The VUI and WUI are accessed through mobile phone by dialing a telephone number and the WUI using the Internet respectively. The e-Learning system will especially be useful for students who are visually impaired and those with dyslexia ailment that make reading, writing and spelling difficult. The application will complement the existing e-Learning systems such as web-based learning, m-Learning and others.

Index Terms - **CBR, e-Learning, VoiceXML, VUI and WUI.**

I. INTRODUCTION

The rapid advances in Information and Communication Technologies (ICTs), including the Internet, have had significant impact on various aspects of the daily lives of mankind and the society. In spite of the rise of Internet technology, e-Government and e-Business, telephone still remains the dominant means by which people communicate [1]. In most societies, students in tertiary institutions of learning are among the highest category of mobile telephone users. In many developing countries, majority of institutions of higher learning particularly the government owned (public) schools have a low ratio rate of teachers to learners. Demand for access to education remains much higher than the system's physical capacity to accommodate students in public schools [2]. Consequently, so many students are left unattended to by the few teachers available. The process of conducting course registration, lectures, examination and result is cumbersome. More so, some students are neglected when they are not getting the desired attention from their teachers. Currently, several universities and college campuses are suffering from over-enrollment of students in courses; students often stand in corridors and around open spaces in classrooms during lectures [3]. Telephone-based learning using land or mobile phone provides one of the e-Learning platforms capable of eliminating this common problem by having excess students remain in their dormitories or residencies and access the same course contents on-demand basis by dialing a telephone number.

Within the context of the aforementioned technology, not only are institutions able to offer more courses for the non-physically handicapped learners, and the physically challenged alike particularly the visually impaired learners, but also increase enrolment without the dilemma and huge cost of constructing additional classrooms and laboratories. The ability to combine the functionalities provided by a mobile phone for making calls for the purpose of conversation and for connecting speech-enabled or voice-enabled e-Learning application opens new opportunities for institutions of higher learning equipped with telecommunication infrastructure.

Voice Extensible Markup Language (VoiceXML) technology allows a user to interact with the Internet through speech recognition technology by using a speech browser and/or the telephone [4]. VoiceXML is the new standard for developing speech and touch-tone applications. The technology brings Interactive Voice Response (IVR) into the Internet age, leveraging the same content and infrastructure used by web applications, and capitalising on open standards to reduce cost and increase efficiency [5]. A VoiceXML platform is the foundation for developing and operating voice automation application [6]. During the human computer interaction, it executes the commands and logic specified by applications written in VoiceXML. It also provides the speech processing capabilities (speech recognition, speech synthesis, voice authentication, etc). VoiceXML platform architecture is based on HyperText Transfer Protocol (HTTP) protocol, and uses both phone and Internet. The web server is often connected to a database which the user can query and update.

Speech-enabled systems are applicable in several areas including information providers such as stock,

978-1-4244-3523-4/09/$25.00©2009 IEEE

flight, financial institutions [7], and e-health. Speech applications are also used for providing services to customers. For instance, Envox Worldwide provides a full range of IVR software, customers' services and solutions [8]. The Cisco Unified Customer Voice Portal (CVP) enables customers to efficiently and enjoyably retrieve the information they need from the contact center [9]. One very interesting application field for voice technology is in education [10].

A VoiceXML absentee system that enables students to report of their class absence through a telephone call was developed in [11]. The Absentee System application was provided basically for Pace University to store students' absentee request and report class absences. A couple of tools for computer assisted language learning environment were developed in [10]. These tools are used for pronunciation verification and a bilingual dictionary, 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. An architecture was presented in [12] to enable multi-channel interaction (web, SMS/MMS and voice) with services that support collaboration on a blog server.

Similarly, a chat application was developed in [13] 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 it run on client/server technology with a graphical user interface (GUI) for the client. A web-based learning tool called LINC (Learn IN Context) was developed in [14] and delivered in the form of a collaborative and voice-enabled presentation.

Artificial intelligence (AI) is one of the techniques used for providing recommendation services. AI is a system that mimics the behavior of humans. Case-Based Reasoning (CBR) is a general AI paradigm for reasoning from Experience [15]. The CBR framework includes case acquisition; case representation; case Indexing; case search; case retrieval; and case reasoning. In searching and retrieval of previous cases, terms in the domain knowledge are usually identified by their word stems. Stemming reduces the storage requirements by decreasing the number of words maintained. Porter's stemming algorithm is probably the most widely used stemmer in information retrieval (IR) research. The Porter stemming algorithm removes the commoner and morphological endings from English words. It does not usually matter whether the stems generated are genuine words or not – thus, "computation" might be stemmed to "comput" [16]. Examples of products using stemming algorithms are search engines such as Lycos and Google, and also thesauruses and other products using Natural Language Processing (NLP) for the purpose of IR [17]. The suffix stripping process of porter's stemming algorithm reduces the total number of terms in the IR system, and hence reduce the size and complexity of the data in the system, which is always advantageous [17]. A smaller dictionary size results in savings in storage space and processing time.

Thus, the problem with existing speech-enabled e-Learning systems is a lack of support for intelligence such as inability to take decisions that will adapt to learners' request based on requirement. The key issues in contemporary speech applications are adaptive and flexible interaction methods [18]. A suitable system will be an application that will not only provide solution but recommend a viable solution even in the absence of enough teachers to attend to learner's problems.

II. E-LEARNING

One of the areas that has been highly affected by ICT is education. Technological advances and wide availability of personal computers, Compact Disks (CDs), the web, broadband access to the Internet, etc, have been used as supporting tools in e-Learning [19]. E-Learning is the use of ICT for supporting learning processes. The various forms of learning environment can be classified into different categories. The traditional form of education is the first category. Here, teachers and students have to be physically present in a classroom setting. Face to face interaction is the most prevalent means of communication between teachers and students.

The second category is by distance education, with the development of communication and transportation technologies in the industrial age, teachers and students were separated by space and time. Communication between teachers and students is achieved by using traditional mail, phone, radio and television. Electronic revolution and the advert of the Internet led to e-Learning which is the third category. Since then different types of media like text, audio, video, hypertext, simulations and two-way communication have been used commonly in education. The m-Learning constitutes the fourth category. It involve the use of mobile communication tools like mobile phone, pocket PC, etc. m-Learning is a learning paradigm that takes place anytime, anywhere with the help of a mobile device. Telephone based learning when used with only a mobile device, is seen as a variant of m-Learning. However, when it is used with land and mobile device, it is seen as voice learning. Voice learning is the use of land and/or mobile phones to access learning contents in the Internet or Intranet on real-time by dialing a telephone number.

III. SYSTEMS DESIGN

Although VoiceXML is easy to learn, building a successful VoiceXML application requires not only software development skills, but also other skills like understanding human factors for the telephone interface, linguistics, speech recognition and audio production. The VoiceXML application development life cycle [20] contains the steps for building speech applications. The steps are similar to web application development stages but include voice user interface (VUI) design and speech recognition system. The development cycle consists of five phases; they include problem definition; systems design; systems development; systems testing; and pilot and deployment.

A. VUI Design

Every authenticated user of the application will undergo some questions and answers session, which will be matched against the content of the database, and the result received by the user through voice response. The sample call flow for the VUI speech-based e-Learning system is presented in Fig. 1.

IVR: Welcome to an Intelligent voice enabled e-Learning system. To login into the system, you must enroll and register, to obtain user name and password.
IVR: What is your password.
Student: Say [password]
IVR: Say "1" For course registration.
IVR: Say "2" For voice learning
IVR: Say "3" For Tutorial
.

Fig. 1: Sample call flow of intelligent speech-based e-Learning application.

B. Psuedocode and Flowchart

The partial psuedocode for the intelligent e-Learning application and CBR search/retrieval module are contained in Fig. 2. The query in the psuedocode represent tutorial question asked by a student requesting for an answer through voice response. The Porter Stemming algorithm has been integrated with CBR to realize an intelligent processing agent for the speech-enabled e-Learning application. Fig. 3 describes a program flowchart of the e-Learning application. A student submits a query and the query is match against case knowledge base, if a match is found then it is reported to the student through IVR and the system will terminate, otherwise the term is tokenized and stemmed, and the process starts all over again until a match is found, otherwise it is treated as unresolved question. Unresolved questions are handled by the teachers and feedback sent to the student through email.

Step 1: The user submits a query and the query is compared with the case knowledge base terms for a match. If the query is found, the corresponding solution is retrieved, and the system terminate, otherwise the query is sent to ***step 2***.
Step 2: The query is tokenized.
Step 3: Generated tokens of the query are compared with the generated index keywords in the domain knowledge for a match. If there is a match then token is replaced with index keyword, goto ***step 1***. If there is no match then the token is sent to stemming algorithm and the stemmed tokens is compared with index keywords and goto ***step 1***.

Fig. 2: Partial psuedocode for intelligent e-Learning application using CBR and stemming approach.

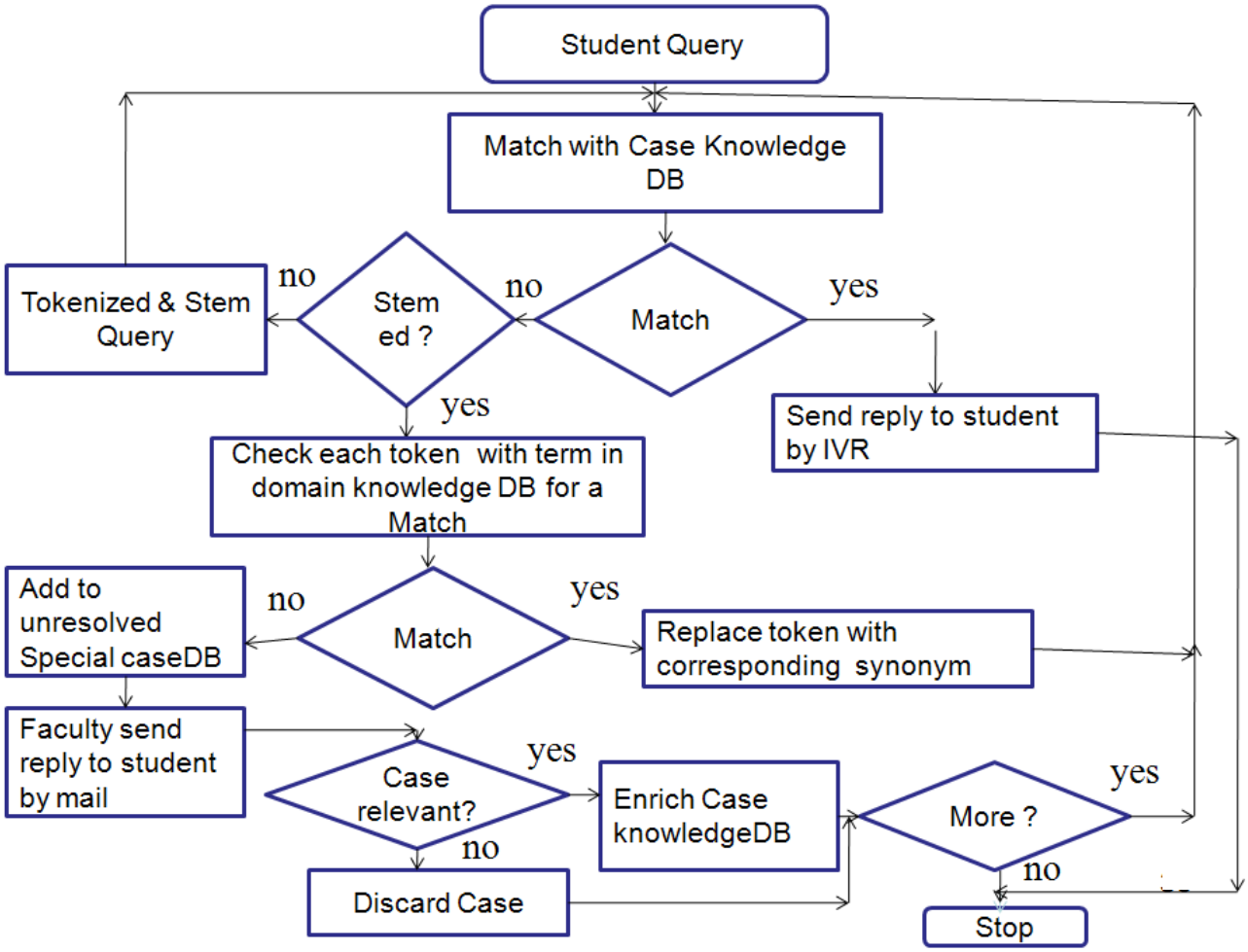

Fig. 3: Flowchart for intelligent e-Learning application using CBR and stemming approach

IV. DEPLOYMENT ARCHITECTURE

Fig. 4 presents a Unified Modeling Language (UML) deployment diagram for speech-enabled e-Learning system. It contains client devices, servers and database.

A. Client Devices

The client systems include student's telephone such as land phones and mobile phones (Personal Digital Assistant (PDA), cell phone and smart phone and other handheld communication devices). In institutions where students are not allowed to use mobile phones or where cost is an issue, then the alternative device is a PC phone such as Skype on Voice over Internet protocol (VoIP) platform.

B. Servers

The server comprises of the voice and web server. The voice server contains the automatic speech recognition (ASR), text-to-speech (TTS) and voice browser. The web server stores the e-Learning application and is used to maintain Internet connectivity with the voice gateway. The robust web server provides real-time access to the e-Learning database.

C. Databases

The database used is MySQL. It stores the student's profile and other information related to the learning content. MySQL is a server application (for database) able to carry out a great number of SQL commands [21].

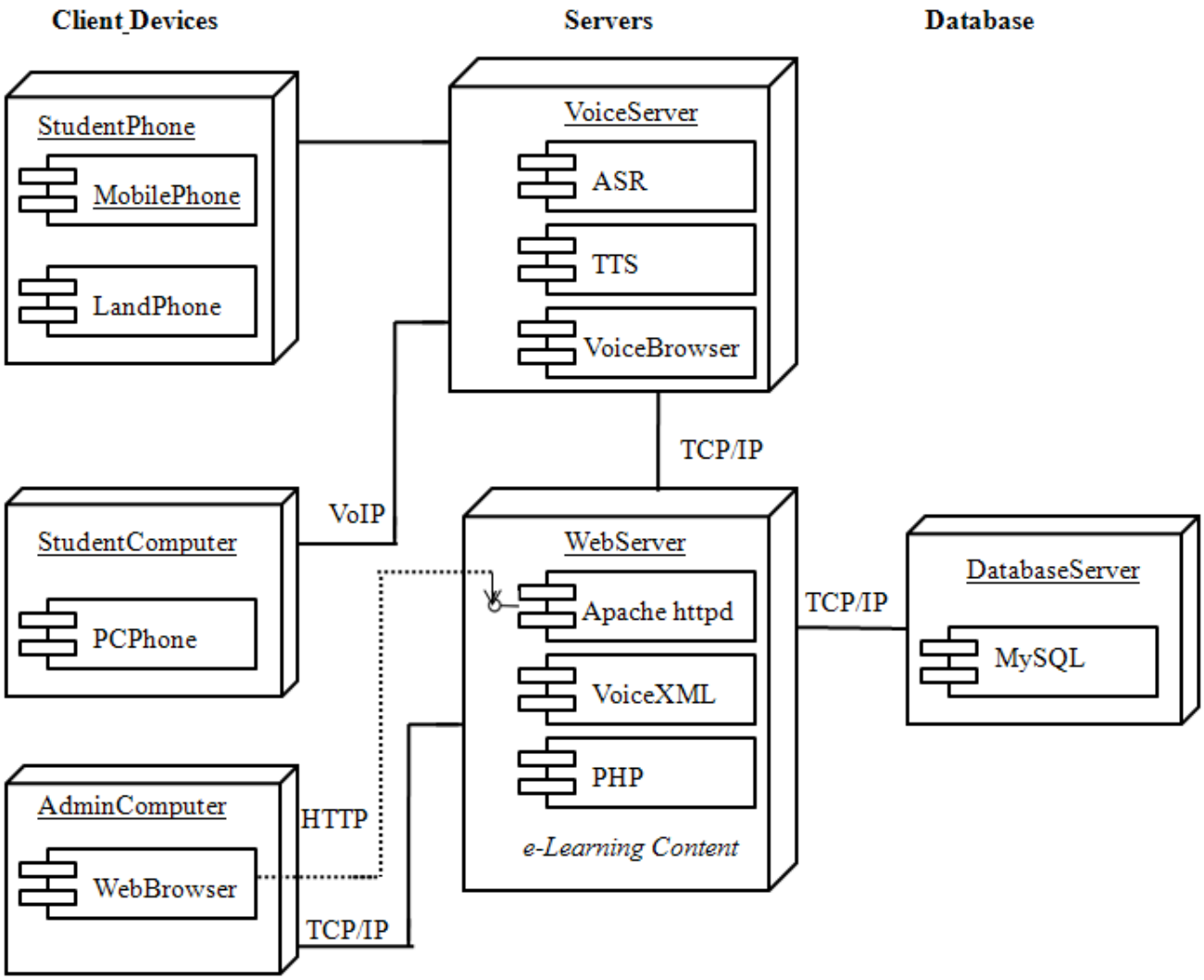

Fig. 4: A UML deployment diagram for speech-based e-Learning system

V. RESULTS AND DISCUSSION

The application provides access to learners through VUI. The VUI was developed using VoiceXML, Hypertext Preprocessor (PHP) was used for the web user interface (WUI), Apache was used as the middleware and MySQL as the database. A Voxeo Prophecy phone emulator was downloaded [22] and installed on a local computer and used to develop the VUI application. A headset was connected to the local computer for the caller to get voice response and also be able to supply voice input. Clicking the Dial button from the Voxeo Prophecy SIP Softphone keypad (see Fig. 5) allowed connection to the application for access to the various services provided. Once connected, the application prompts with a welcome message and goes ahead to authenticate the user's name and password before any transaction can take place. The application will ask for the services demanded by a student and then process the request.

The web user interface (WUI) is used by the system administrator for user management, account management and configuration management. All the students access the application through the VUI while the administrator accesses the application through the WUI.

Fig. 5: Voxeo Prophecy SIP Softphone keypad Source: [22]

VI. CONCLUSION

In this paper, an intelligent speech-enabled e-Learning application has been presented. The Intelligent component of the application was realized using CBR and porter's stemming algorithm.

The UML deployment diagram will provide a reference model and platform for developers to create intelligent speech-based e-Learning applications. It also makes a contribution in the area of ubiquitous learning. The e-Learning application will especially be useful for students who are physically challenged such as the visually impaired. It will also be useful for people who have medical history of reacting to repetitive strain injury (RSI) as a result of seating too long in one place using the keyboard. Therefore this system gives every citizen equal right and access to quality education irrespective of his or her physical disability.

The e-Learning system offers an alternative platform of learning without any physical disability. It complements the existing e-Learning systems such as web-based learning, m-Learning, etc. The e-Learning system offers the possibility to learn anytime and anywhere there is telephone access, regardless of the availability of Internet services.

REFERENCES

[1] J. Markowitz, 'Voice Authentication in the RealWorld", *Keynote Speech*, National Centre for Biometric StudiesConference on Voice Authentication for Identity Management, 2006. Retrieved 14th September 2009 from http://www.voicexml.org/insights/JMarkowitz Consulting/jmc_voice_auth_2006.pdf

[2] NUC. The state of Nigerian universities" Nigerian university system. Newsletter, (2), (1), 2004.

[3] P. O. Bobbie, I. Fordjour. D. O. Aboagye, C. Dzidonu, O. Darkwa and K. Nyantakyi, "Technology Enablers for collaborative Virtual Education", 3rd International Conference on ICT for Development, Education and Training. AICC, Accra, Ghana. May 28 – 30, 2008. pp. 191-193.

[4] Palo, Wireless, VoiceXML resource center, 2007. Retrieved 19th March 2009 from http://www.palowireless.com/voicexml/tutorials.asp

[5] G. Meyer, "Voxpilot a eureka software company", benefits of VoiceXML, 2008

[6] J. Rouillard, 'Web services and speech-based applications around VoiceXML" Journal of networks, vol. 2, no. 1, February 2007, pp.27-35.

[7] A. A. Azeta A. A., N. A. Ikhu-Omoregbe, C. K. Ayo & A.A. Atayero, "Development and Deployment of VoiceXML-Based Banking Applications", Journal of Computer Science & its Application, An International Journal of the Nigeria Computer Society (NCS), Vol. 15 No.1 June 2008, Edition, pp. 59-72

[8] Envox, "Envox Worldwide", 2009. Retrieved March 2009 from http://www.ivr.envox.com/

[9] Cisco, "Intelligent Voice and VideoSelf-Service", Voice portal, 2009. Retrieved January 2009 from http://www.cisco.com/en/US/products/sw/custcosw/ps1006/index.html

[10]I. Kirschning, "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, pp 49-50. Retrieved August 2009 from ict.udlap.mx/people/ingrid/ingrid/CHI_Tlatoa.pdf

[11]P. Gallivan, Q. Hong, L. Jordan, E. Li,G. Mathew, Y. Mulyani, P. Visokey and C. Tappert, VoiceXML Absentee System, Proceedings of MASPLAS, 2002. The Mid-Atlantic Student Workshop on Programming Languages and Systems Pace University.

[12]Peter Y., Vantroys T. and Lepretre E., "Enabling Mobile Collaborative Learning through Multichannel Interactions". An article published in the 4th International Conference on Interactive Mobile and Computer Aided Learning, IMCL 2008, pp. 1-4. Retrieved May 2009 from http://telearn.noe-kaleidoscope.org/warehouse/Peter-Yvan-2008_(001784v1).pdf

[13]L. Schindler, "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, 2005. Retrieved July 2009 from http://www.stetson.edu/~helaarag/Laura-proposal.pdf

[14]S. Selouani, T. Le, C. Moghrabi, B. Lanteigne and J. Roy, 'Online Collaborative Learning System Using Speech Technology', Proceedings of World Academy of Science, Engineering and Technology Volume 15 October 2006 ISSN 1307-6884. Retrieved August 2009 from http://www.waset.org/pwaset/v15/v15-11.pdf

[15]M. S. Abdel-Badeeh and A. Rania, "A Case Based Expert System For Supporting Diagnosis of Heart Diseases", AIML Journal, Volume (5), Issue (1), March, 2005[14] C. A. Kumar, A. B. Gupta and S. Trehan, "Latent Semantic Indexing-Based Intelligent Information Retrieval System for Digital Libraries", Journal of Computing and Information Technology - CIT 14, 2006, 3, pp. 191–196

[16]Selouani S., Le T., Moghrabi C., Lanteigne B., & Roy J. (2006), 'Online Collaborative Learning System Using Speech Technology', *Proceedings of World Academy of Science, Engineering and Technology* Volume 15 October 2006 ISSN 1307-6884. Retrived July 2009 from http://www.waset.org/pwaset/v15/v15-11.pdf.

[17]B. S. Singh "Search Algorithms", DRTC Workshop On Digital Libraries: Theory and Practice, March 2003, DRTC, Bangalore. pp. 1-9. Retrieved May 2009 from https://drtc.isibang.ac.in/bitstream/handle/1849/32/E_Searchalgo_brijesh.pdf?sequence=2

[18] M. Turunen, "Jaspis - an Adaptive Speech Application Architecture." In Proceedings of 7th IFHOH World Congress: 2004, pp. 51-57. Retrieved March 2009 from http://www.cs.uta.fi/hci/spi/reports/Jaspis-IFHOH'04.pdf

[19]W. Kim, "Towards a Definition and Methodology for Blended Learning", Blended Learning, pp. 1-8, Pearson, 2007. Workshop on Blended Learning '07, U. K.

[20]VoiceXML, "VoiceXML application development life cycle", Hewlett Packard Development Company, 2007

[21]E. Edelhauser, "Client Server Versus Distributed Network Applications in Human Resource management", Proceedings of the International Conference on Theory and Applications of Mathematics and Informatics –ICTAMI 2004, Thessaloniki, Greece. pp. 129 –136. Retrieved May 2009 from http://www.emis.de/journals/AUA/acta8/Edelhauser1.pdf.

[22]Voxeo,"The Voxeo Prophecy Platform – Free Download", 2008. Retrieved February 2009 from http://www.voxeo.com/prophecy 2008.