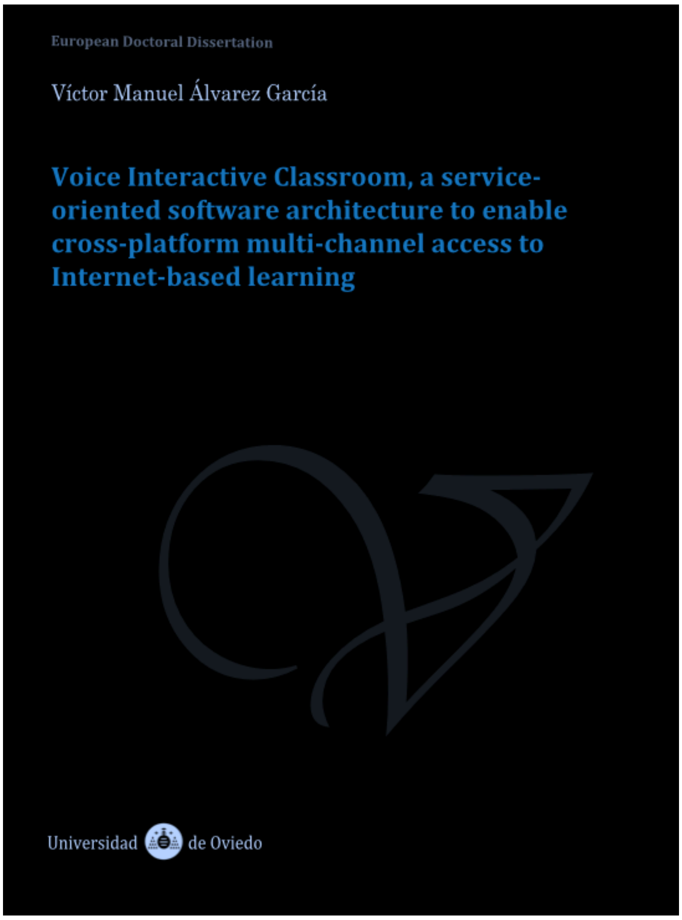
Víctor Manuel Álvarez García (2011) "Voice Interactive Classroom, a service-oriented software architecture to enable cross-platform multi-channel access to Internet-based learning", Ph.D. Thesis Dissertation, University of Oviedo, Spain



European Ph.D. Dissertation
Defence: 17th February 2011

Supervisors:
Dr. María del Puerto Paule Ruiz, *University of Oviedo, Spain*
Dr. Marcus Specht, *Open University of the Netherlands, The Netherlands*

Reviewers:
Dr. Marie Joubert, *University of Bristol, England*
Dr. Lorella Giannandrea, *University of Macerata*, *Italy*

Committee:
Prof. Dr. Manuel Ortega Cantero, *University of Castilla – La Mancha, Spain*
Prof. Dr. Rob Koper, *Open University of the Netherlands, The Netherlands*
Prof. Dr. Juan Manuel Cueva Lovelle, *University of Oviedo, Spain*
Prof. Dr. Manuel Pérez Cota, *University of Vigo, Spain*
Dr. Juan Ramón Pérez Pérez, *University of Oviedo, Spain*

Voice Interactive Classroom, a service-oriented software architecture to enable cross-platform multi-channel access to Internet-based learning

Víctor Manuel Álvarez García

UNIVERSITY OF OVIEDO
DEPARTMENT OF COMPUTER SCIENCE

In Partial Fulfilment of the Requirements for the Degree of
EUROPEAN DOCTOR OF PHILOSOPHY

Aula de Voz Interactiva, una arquitectura software orientada a servicios para habilitar un acceso multiplataforma y multicanal al aprendizaje basado en Internet

Voice Interactive Classroom, a service-oriented software architecture to enable cross-platform multi-channel access to Internet-based learning

Víctor Manuel Álvarez García

UNIVERSIDAD DE OVIEDO
DEPARTAMENTO DE INFORMÁTICA

Tesis Doctoral depositada el día 5 de Noviembre de 2010 para la obtención del título de Doctor Europeo por la Universidad de Oviedo

Dirigida por:

- Dra. María del Puerto Paule Ruiz (Universidad de Oviedo)
- Dr. Marcus Matthäus Specht (Centre for Learning Sciences and Technologies, Open Universiteit Nederland)

Abstract

Software technology is creating a ubiquitous context for human living and learning in which new modes of interaction are gradually being incorporated. At the same time, interaction with Internet-based learning systems has evolved from using the traditional access from a personal computer or laptop's web browser to more flexible access from mobile devices. However, in both cases, e-learning systems have created a context in which interaction with the user is carried out mainly by using visual perception. This situation restraints the number of scenarios in which students can make use of a learning management system.

Audio-based software brings a new way of interacting with the Internet. Visual access to Internet resources can be complemented with audio interfaces to adapt e-learning systems to different educational settings and learning styles, but auditory access to web resources needs to be more broadly supported by software architectures.

This dissertation proposes the use of a service-oriented approach, supported by the e-learning specifications IMS Abstract Framework and The Open Knowledge Initiative, to provide a complementary audio communication channel and enable the interoperability of audio features among the various e-learning platforms and components. In addition, the use of the W3C recommendations of VoiceXML, along with the methods and techniques of adaptive educational hypermedia, allows overcoming the limitations of using unusual or non-standard languages to describe voice dialogues and better adapt the on-line learning process to innovative e-learning scenarios.

The components of this model have been mapped to a software architecture named "Voice Interactive Classroom" that serves as a starting point for the design of research case studies which aim to demonstrate the benefits of the proposed solution. The validation of this thesis has been achieved by setting up and following an active strategy for scientific outreach. Activities include planning, coordinating and supervising final year projects, master projects and master dissertations, presenting articles and communications in scientific meetings, and the publication of this research study in a prestigious scientific journal.

Resumen

La tecnología software está creando un contexto ubicuo para la vida y el aprendizaje humano en el que gradualmente se incorporan nuevos modos de interacción. Además, la interacción con los sistemas de aprendizaje basados en Internet ha evolucionado desde el acceso tradicional a través del navegador web de un ordenador personal o portátil a accesos más flexibles desde dispositivos móviles. Sin embargo, en ambos casos, los sistemas de e-learning han creado un contexto en el cual la interacción con el usuario se realiza principalmente utilizando la percepción visual. Esta situación limita el número de escenarios en el que los estudiantes pueden hacer uso de los sistemas de gestión del aprendizaje.

El software basado en audio proporciona una nueva manera de interaccionar con Internet. El acceso visual a los recursos de Internet puede ser complementado con interfaces de audio para adaptar los sistemas de e-learning existentes a diferentes entornos educativos y estilos de aprendizaje, pero el acceso auditivo a los recursos web necesita encontrar mayor soporte en las arquitecturas software.

Esta tesis propone el uso de un enfoque orientado a servicios, apoyado en las especificaciones de e-learning IMS Abstract Framework y The Open Knowledge Initiative, para proporcionar un canal de comunicación complementario y auditivo, y permitir la interoperabilidad de las aplicaciones de audio con las plataformas y componentes de e-learning. Adicionalmente, el uso de las recomendaciones VoiceXML del W3C, junto con los métodos y técnicas de hipermedia adaptativa para sistemas educativos, permiten superar las limitaciones de usar lenguajes inusuales o no estándar para describir los diálogos de voz y adaptar mejor el proceso de aprendizaje en línea a escenarios de e-learning innovadores.

Los componentes de este modelo han sido mapeados a una arquitectura software denominada "Aula de Voz Interactiva" que sirve como punto de partida en el diseño de casos de estudio que persiguen demostrar los beneficios de la solución propuesta. La validación de esta tesis se ha conseguido mediante el planteamiento y seguimiento de una estrategia de alcance científico. Las actividades incluyen la planificación, coordinación y dirección de proyectos fin de carrera y trabajos fin de máster, la presentación de artículos y comunicaciones en reuniones científicas y la publicación de esta investigación en una prestigiosa revista científica.

Acknowledgements

This Ph.D dissertation has been possible only because of the help and support of a large number of professors, colleagues, students, relatives and friends. Thanks for being part of my work and my life, this achievement is also yours.

I am heartily thankful to **Dr. Andrés Sampedro Nuño** and **Dr. María del Puerto Paule Ruíz**. Without their guidance, support and encouragement over the years this thesis would not exist.

My deep thanks to **Dr. Juan Manuel Cueva Lovelle**, **Dr. Víctor Guillermo García, Dr. Jose Ramiro Varela** and **Dr. María Camino Rodríguez Vela**, their support and confidence in my work made it possible to carry out this research at the **University of Oviedo**.

My most sincere gratitude to **Dr. Rob Koper**, **Dr. Marcus Specht** and the **Centre for Learning Sciences and Technologies**, for hosting me in the Netherlands, allowing me to work as part of their team, co-supervise and positively influenced this thesis and my research.

Thanks to **Steinn E. Sigurðarson**, **Sebastian Kelle** and **Simone Carletti** for making my days in the Netherlands so enjoyable.

Thanks also to **Dr. Marie Joubert** and the members of the **STELLAR Network of Excellence on Technology Enhanced Learning** for supporting this dissertation.

I am in debt with **Isaac Rubio de Gabriel**, **Moisés Riestra González** and **Manuel Pérez González**, the students who developed the prototypes of this thesis. My deep gratitude and appreciation to **Dr. Juan Ramón Pérez** and **Remko van Dort,** for their invaluable support and contributions during the whole research and publication processes.

Thanks to all my **colleagues** at the **Department of Computer Science** and my fellows from the **Laboratory of Object Oriented Technologies**, our research team at the University of Oviedo. It is a great privilege to work and learn everyday from you.

Special thanks to **Dr. Ángel Martínez Nístal**, for giving me my first professional opportunity, and to my colleagues at the **Image Processing Service**, for teaching me that work can be a most enjoyable experience.

Also thanks to my **professors** and **students** at the University of Oviedo for encouraging me over the years with their work and enthusiasm.

I would like to express my gratitude to **Dr. Raúl Izquierdo**, **Dr. Francisco Ortín** and **Dr. José Manuel Redondo** for their support and encourage. Also to **Dr. Cesar Acebal** and **Dr. Daniel Lanvin** for their help and reference concerning this dissertation.

Thanks to all the **editors**, **symposium chairs**, **members of technical programs** and **anonymous reviewers** for making it possible to communicate and publish our research and for providing constructive comments that helped to improve the quality of this dissertation.

Thanks to **David Knell** and **Alejandro Pérez** for giving me the chance to join their development team in England, and to my colleagues in London. Our time together has deeply influenced my life and this thesis.

My gratitude and love to the **Batesons**, my English family, for all their love and support. It was one morning at their home in Ely, Cambridgeshire, that I decided to start my Ph.D studies.

My deepest gratitude to my **sports mates** along the years and specially to my karate instructor, **Jose Luis Prieto** and my **karate** fellows for teaching me discipline, respect and willingness to improve, something that has been with me all my life.

Thanks to my **travel mates**, for making every adventure unique.

It is finally to my **family**, **friends** and **beloved ones** that I owe the least articulable but most profound thanks. I will never been grateful enough for all your care, love, understanding, encourage and unconditional support throughout the years. Thank you for teaching me the greatest thing I will ever learn.

This thesis is dedicated to **nature**.

Contents

Illustration Index

Table Index

I

Introduction

This initial chapter sets the hypothesis of this thesis:

Visual access to Internet resources can be complemented with audio interfaces to adapt e-learning systems to different educational contexts and learning styles. This chapter aims to answer the question of what this thesis is about and sets the foundation upon which the rest of the dissertation is based.

Because of its introductory nature, it will inevitably make assumptions and refer to concepts that will be either demonstrated or thoroughly discussed in the body of the dissertation. Finally, it outlines the structure of the whole dissertation, with a brief summary of every chapter.

Origin of this Thesis

This thesis reflects a decade of research and professional experience in the computer science and learning technologies fields. It is not a single person, idea or project which has inspired the conception of this thesis, but the sum of all experiences I was involved with within this period.

While studying computer science I was privilege to participate in research and professional experiences on e-learning and social networking for web and mobile environments. These first out-of-class projects allowed me to start putting into practice the skills and knowledge acquired during my studies, as well as specialise more my areas of expertise.

At the same time, I maintained my interest in the English language, which I had been learning from young age. Besides the usual reading of technical material, during the last years at the faculty I started using more English in my regular life. By the end of my studies I was accustomed to write software programs and technical documentation in English, which led me to the next step.

At the end of my studies I started my overseas experience in England, where I was lucky to join a telecommunications company and gain knowledge in new areas. Besides learning how to provide audio communications via Internet, I got also involved in free software developments for telecom environments.

In the autumn of 2006 I went back to Asturias to start my Ph.D. studies. At the same time, I obtained a position as a lecturer at the University of Oviedo and continued giving freelance consultancy for several companies in the United Kingdom involved in Voice over IP and Interactive Voice Response development.

It was then that, together with my students, I started to consider new scenarios and study techniques to enable audio access to a number of systems, such as customer services and visual-only web-based environments.

This thesis focuses on the use of software architectures and audio technologies for browsing web-based learning management systems.

In summer 2007, I had meetings with Dr. Andrés Sampedro and Dr. María del Puerto Paule which supported the decision of making use of my previous grounding and focus my Ph.D. research in exploring the use of software architectures and audio technologies for browsing Web-based learning management systems.

For the next three years, we evolved this research in the context of our university and research group, and with the help of colleagues and students. We reviewed the scientific literature on the topic, enabled new e-learning scenarios, designed, created and evaluated technologies and prototypes, and communicated our research in congresses, events and scientific journals. We also arranged a research visit at the Centre for Learning Sciences and Technologies (Celstec) in the Netherlands, where I was privileged to join Dr. Rob Koper, Dr. Marcus Specht and their research team and gave a final impulse to this thesis.

Summary

It has been more than a decade since the release of the first Internet-based learning systems. During this period, educational institutions worldwide have been incorporating on-line educational tools at an increasing rate to complement their classroom teaching (blended learning)(Masie, 2002) and in some instances even replace traditional classroom settings with online educational programs (distance learning)(Meyer, 2002).

E-learning and technology enhanced learning has evolved during this period trying to keep pace with the World Wide Web (WWW), incorporating new technologies as well as adding pedagogical principles to a wide variety of software tools. At the same time, interaction with Internet-based learning systems has evolved from using the traditional access from a personal computer's Web browser to more flexible access from mobile devices, which has stimulated the research in mobile learning (m-learning)(O'Malley et al., 2003) and ubiquitous learning (u-learning)(Ogata & Yano, 2004).

Celstec's Learning Media Programme[1] assumes that human living and learning will be subject to progressing virtualisation, which allows and requires new modes of interaction, communication and knowledge creation accommodated with networked electronic media and devices.

However, e-learning and technology enhanced learning have created a context in which interaction with the user is carried out mainly by using visual perception. The fact that we have so far left behind alternative modes of interaction, such as the use of audio interfaces, imply that there are only a limited number of situations in which students can make use of a learning management system.

This thesis states that there are only a limited number of situations in which students can make use of a learning management system.

Furthermore, the learning process emphasizes the importance of providing aural access to e-learning systems.

[1] http://celstec.org/content/learning-media/

Learning styles, being the different ways in which a person can learn, are related to an understanding of learning as an active process. Looking at the visual/verbal learning style dimension (Felder & Silverman, 1988) we can clearly distinguish between a visual learner, who prefers using pictures, images and spatial understanding, and an aural learner, who is a very good listener and likes to talk. Therefore, if a learning management system offer two types of interaction, both visual and aural, it provides students the possibility to select a learning style best suitable for their needs.

This thesis states that complementing visual interaction with audio interfaces allows us to enable new e-learning scenarios.

Complementing visual interaction with audio interfaces allow us to enable new learning scenarios for students. However, an exhaustive review of the learning management systems and software frameworks for developing electronic learning platforms shows the absence of specific software structures to accommodate voice access to current Internet-based learning.

Our initial attempts to figure out a software architecture suitable for enabling voice access to e-learning end up in restricted ad-hoc solutions. During this period, Isaac Rubio de Gabriel developed Feedo, a new personalised software environment for web and voice-based learning using feed syndication (Rubio, 2008). Initial stages of another development by Moisés Riestra González delivered voice modules which could be used only upon one of the existing e-learning and voice platforms (Riestra, 2009). However, the latest prototypes in Riestra's study showed already a transition into a more general approach.

As part of the preliminary study for this thesis, we conducted a survey focused on the use of web-based learning management systems for higher education at the University[2], which showed the lack of a clearly predominant e-learning system, as well as an evolution in the adoption of different solutions. At this stage of the research, we found ourselves involved in a technological context which included many heterogeneous e-learning systems already in use.

This situation lead us to go one step further towards a more general solution which could be used and ported among the various existing platforms, and a software paradigm which could allow a platform independent plug-and-play approach for the voice modules.

Although not widespread, we found support in the software frameworks for developing e-learning platforms, as they were coincident in the need for a service-oriented approach to enable interoperability among the various

[2] http://www.di.uniovi.es/~victoralvarez/survey/

platforms and their components. As our goal was to add voice modules as components of the existing platforms, we started to use the approach proposed by IMS Abstract Framework[3] and The Open Knowledge Initiative[4]. Furthermore, we also found a platform independent manner to describe the voice dialogues, the W3C recommendations of VoiceXML[5].

By following a service-oriented architecture we designed, prototyped and evaluated a proposal consisting of a middleware approach to enable voice access to e-learning platforms (Álvarez et al., 2010). The proposed software architecture was used by Manuel Pérez Gónzalez to enable voice access to Sakai (Pérez, 2010) and Moisés Riestra González to reproduce the experiment with Moodle (Riestra, 2010). The voice modules, written in VoiceXML, were for the first time capable of being shared among platforms, and the learning resources provided by the platforms were processed following the same OKI schema. Although with some limitations, these software systems successfully achieved our goal of providing cross-platform multi-channel access to Internet-based learning.

Rationale

E-learning and the enabling learning technologies are concerned with making learning experiences in all types of settings more effective, efficient, attractive and accessible for learners (Koper & van Es, 2004). This statement, which we completely agree with, arises two questions, what types of settings can enable an e-learning process? And, who is the learner in e-learning?

Define an educational and technological setting for an e-learning process is not a static problem at all, as it is determined by the evolution of formal and informal education and technology enhanced learning, which is at the same time very influenced by the technological trends e-learning can benefit from. A first global classification can include the following four settings:

- Blended learning
- Distance learning
- Mobile learning
- Ubiquitous learning

[3] http://www.imsglobla.org/af/
[4] http://www.okiproject.org/
[5]http://w3.org/TR/voicexml21/

As for the challenging question “Who is the learner?”, one may initially take for granted that the learner is one of the students belonging to a traditional learning setting and that e-learning occurs during the same period of time and within the same institution. This affirmation, which may be true for many blended learning settings, becomes very limited when we take into account the geographical, cultural, social, age and physical boundaries which are being overcome by the Internet. In the technological era of the Internet, any person, at any time, from any place and using any device can adopt the role of e-learner.

The individual characteristics of the learner are remarked in “The Universal Design for Learning” (UDL)(Rose & Meyer, 2002), which reflects an awareness of the unique nature of each learner and the need to accommodate differences, creating learning experiences that suit the learner and maximise his or her ability to progress (Center for Applied Special Technology)[6]. Moreover, if we transfer the question into higher education, The Universal Declaration of Human Rights (1948)7 states the right to education and equal access to higher education for all on the basis of merit (Art26).

A thorough analysis of the state of art on e-learning reveals that learning management systems and e-learning frameworks have been developed focusing on a web environment that was created for visual communication. Thus, the access to electronic learning tools is mainly supported by the use of a single communication channel, reducing in this manner the number of scenarios in which e-learning can take place and the number of learners who can benefit from the various e-learning processes.

This dissertation focuses on the flexibility provided by the combination of visual and auditory interfaces to enhance learning experiences in all type of settings and for every learner.

This dissertation focuses on the flexibility provided by the combination of visual and auditory interfaces to enhance learning experiences in all type of settings and for every learner. The use of the auditory channel can be applied for a variety of situations in which voice can be used as a means of communication and enable the use of e-learning platforms in multiple scenarios.

The use of visual-only interaction implies that there are only a limited number of situations in which students can make use of a learning management system. Technology enhanced learning allows the emergence of new modes of interaction, such as voice interaction, to best suit students’ personal needs and learning contexts, But adding audio interfaces can't be correctly approached from schemas designed for a single interaction mode and thus, it requires an evolution in the design of the architectural structures of e-learning platforms to better accommodate new access modes.

[6]http://www.cast.org/
[7]http://www.un.org/en/documents/udhr/

This thesis is committed to enhance current learning management systems, enabling the aforementioned scenarios, proposing a new software architectural design to allow auditory access to current web-based learning systems and adapting the learning process to best suit each individual student needs.

Objectives of this Thesis

The study for this thesis was built upon initial premises and expanded to explore the use of audio technologies for browsing web-based learning management systems. In support of finding an effective way of approaching this study we defined a list of objectives expressed in such a way that a reader could determine whether these goals have been achieved or not.

The main objective of this thesis is to propose a reference software architecture to support building cross-platform and multi-channel applications for Internet-based learning systems.

The main objective of this thesis is to propose a reference software architecture to support building cross-platform and multi-channel applications for Internet-based learning systems.

Other specific objectives of this thesis can then be summarised as to:

- Identify e-learning scenarios which can benefit from the usage of voice interaction.
- Propose the adoption of technologies to fill the lack of multi-channel access to Internet-based learning.
- Design and develop speech-enabled software applications for Internet-based learning.
- Enable multi-channel access to existing Internet-based e-learning platforms.
- Propose a solution to enable the creation of voice interactive features compliant with existing e-learning and voice interaction guidelines and specifications.
- Design and develop speech-enabled software applications following the proposed solution.
- Adapt the user interface of learning management systems for the audio communication channel.

Research Methodology

The research for this thesis has a constructive and exploratory orientation.

The research for this thesis has a clear practical orientation throughout, and thus it has followed a primarily constructive research approach (March & Smith, 1995). Research activities related with this methodology include build, evaluate, theorise and justify. Nonetheless, some stages of this research had a more exploratory focus and hence a thoughtful review of the scientific literature on related topics as well as technology reviews and associated proposals have also been included. All the outputs from this research have been communicated and validated through publications in national and international congresses and scientific journals.

Prior to the preliminary state-of-the-art review, and as part of my teaching experience at the University, I set up a laboratory work involving the development of prototypes to explore the combination of a self-developed web-based software solution with telephone-based interfaces. The success of this experience inspired a later publication and partially motivated the subsequent research.

The study of the state-of-the-art was scheduled for my Master of Advance Studies dissertation and included an exhaustive review on e-learning platforms and e-learning frameworks, as well as a survey research on the use of e-learning at universities. This work, which was presented in a national congress, raised the matter of the heterogeneity of the e-learning solutions and the need of adopting a service-oriented approach to enable interoperability among systems, along with other side questions.

From this point on, my research adopted a more constructive research dimension. I established an initial software architecture for the construction of voice interactive learning modules, which lead to the construction of two new prototypes for enabling voice interactive access to e-learning platforms by following a service-oriented schema, which allowed a certain level of interoperability among the two selected platforms. The development of the prototypes was taken on by two undergraduate students and evaluated and refined following an iterative user-centred design process.

In a final European oriented stage, our theoretical based and justification of the proposal was reasserted, and the paper which covers this research was accepted for publication in a prestigious scientific journal. The newly adopted European link also established a foundation for ongoing and future research.

Organisation of the document

After the present introduction, the core of the dissertation has been organised in three major sections.

Before examining the contributions in detail, Chapter II, III, IV, and V provide state-of-the-art information on the various topics and issues addressed by this dissertation. As this thesis falls into several fields, these sections provide a background on web-based e-learning platforms, e-learning frameworks, adaptive educational hypermedia and audio interfaces for Internet applications.

Then, it starts the part of the dissertation that discusses the problem and proposes a solution. Chapter VI presents initial case studies that allowed us to state the solution requirements (Chapter VII). Chapter VIII proposes the solution to the problem of creating a software architecture to enable cross-platform multi-channel access to e-learning. Chapter IX presents a design of case studies based on the proposed solution, which is later demonstrated in Chapter X with the final case studies.

The last chapters describe the scientific outreach (Chapter XI), summarise the major contributions of this thesis and present further research directions (Chapter XII).

A brief description of each chapter is provided below.

Chapter I. Introduction

This introductory chapter aims to provide the reader with a general understanding of what this thesis is about and sets the foundation upon which the rest of the dissertation is based. It also gives a brief inside on the research process.

Chapter II. Web-based e-learning platforms

Chapter II, below, describes the evolution and current level of development of web-based e-learning platforms. It also includes a summary on the main developments at the University of Oviedo and a review on the use of learning management systems in Spanish public universities.

Chapter III. E-learning frameworks

Once the findings for the development of e-learning platforms are presented, Chapter III analyses the evolution and main characteristics of the international frameworks which support the development of such systems.

Chapter IV. Voice-enabled Internet applications

Chapter IV reviews how speech-enabled Internet applications currently work. If firstly summarises the evolution of the research in create machines and mechanisms to simulate human speech communication capability. Then, this section focuses on interactive voice response technologies, and particularly in those which allow the creation of voice applications for the Internet and enable the integration with other Internet-based software.

Chapter V. Adaptive educational hypermedia

Once the evolution, current status and perspectives of e-learning and voice-enabled Internet applications have been described, Chapter V completes the state-of-the-art by analysing how adaptive educational hypermedia can provide methods and techniques to adapt user interaction. This section reviews the evolution of adaptive hypermedia and summarises the main existing approaches for adapting the presentation and navigation of traditional web-based hypermedia systems.

Chapter VI. Case studies

After describing the state-of-the-art in the various fields concerned, Chapter VI provides case studies which are used to gain insight into the best way to design and develop a voice interactive learning application. This section helps to understand research limitations and opportunities for the study, and is employed to properly capture the requirements of the solution.

Chapter VII. Solution requirements

Previous chapters present the studies which helped us to understand the limits and possibilities of this research, and permitted us to state the requirements for the solution, which are presented within Chapter VII.

Chapter VIII. Proposed solution: Voice Interactive Classroom

Chapter VIII describes the Voice Interactive Classroom architecture, the proposed solution to the problem of designing a software architecture to enable cross-platform multi-channel access to Internet-based learning.

This chapter extends an article published in Elsevier's Journal of Network and Computer Applications, of which the author of this thesis is corresponding author, adding further explanations for design decisions and discussing the alternatives in more detail.

Chapter IX. Design of case studies

Chapter IX complements the described solution with an insight into the process for designing, developing, documenting and evaluating case studies based on the Voice Interactive Classroom. It is intended to provide students and practitioners with indications they can use to plan, implement and evaluate case studies based on the proposed software architecture.

Chapter X. Demonstration: Case studies revisited

Chapter XI extends the case studies described in Chapter VI and completes the contributions of this dissertation. This chapter is aimed to demonstrate the benefits of the proposed solution, and describe how it contributes to enable cross-platform multi-channel access to Internet-based learning.

The approach followed for such demonstration consists on describing case studies conducted during the last year of the dissertation to address the experimental aspects derived from it.

Chapter XI. Scientific communication and validation

After providing the solution and completing the contributions of this dissertation, Chapter XI describes the scientific communication and validation derived from it. Scientific outreach includes final year projects, master projects and master dissertations, articles and communications presented in scientific meetings, and the publication of this research study in its entirely in a prestigious scientific journal.

Chapter XII. Conclusions and further research

Chapter XII summarises, in a compressed form, which are, in our opinion, its major contributions. This section also describes briefly the areas of research that we believe will lead to further progress in the near future.

Finally, a last chapter reproduces Chapter XII in Spanish, and an additional final section provides the bibliographical references of the selected articles and the scientific literature reviews which support the various studies described within this dissertation.

II

Web-based E-learning Platforms

This chapter analyses the evolution and current level of development of web-based e-learning platforms. It reviews the main developments and points out future directions.

The information is completed with a description on the development and use of learning management systems at the University of Oviedo, and a research survey which shows the use of e-learning platforms in Spanish universities in the year 2009.

Introduction

The development of web-based e-learning platforms emerges in the 1990s. The term includes software applications such as Learning Management Systems (LMSs), Learning Content Management Systems (LCMSs) and Virtual Learning Environments (VLEs). In general, a Web-based e-learning platform is a software application which relies on the Web to give support to on-line learning processes. Although historically they were mainly used to provide course management functionalities, current e-learning platforms often include a wide variety of on-line software tools, including applications for assembling and delivering multimedia learning materials, design and monitor learning experiences, and provide social interaction and communications.

Current e-learning platforms include a wide variety of on-line software tools.

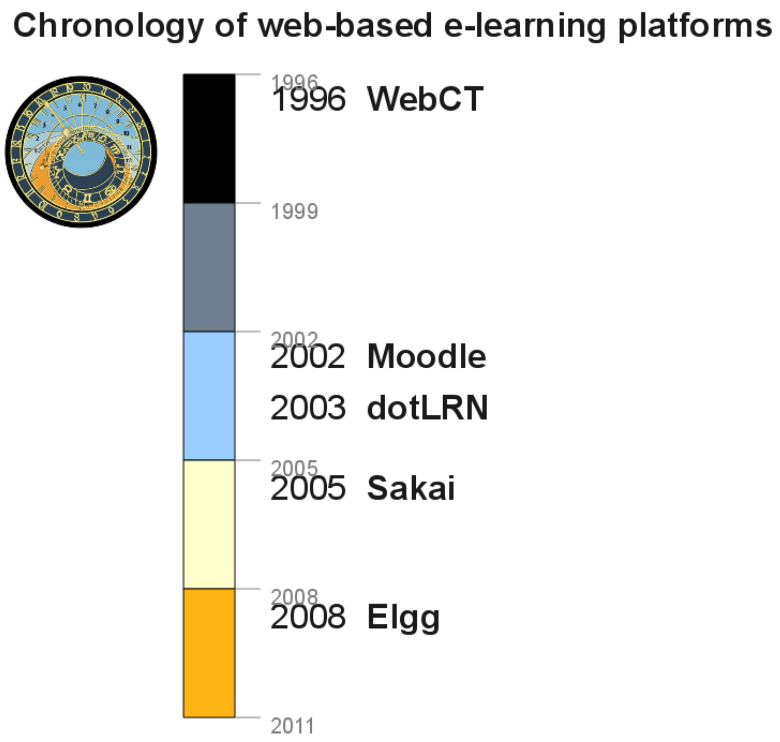


Illustration 1: Chronology of web-based e-learning platforms

The first generation of e-learning platforms (c. 1993 onwards) provided, in essence, black box e-learning solutions, where it was common to find a 1:1 mapping between systems and courses, with very limited user tracking, if any (Dagger et al., 2007). This first approach soon gave way to a more flexible modular design where the various system functionalities were mapped into separated modules which could be recombined to some extent.

Black-box e-learning solutions gave way to a more flexible modular design.

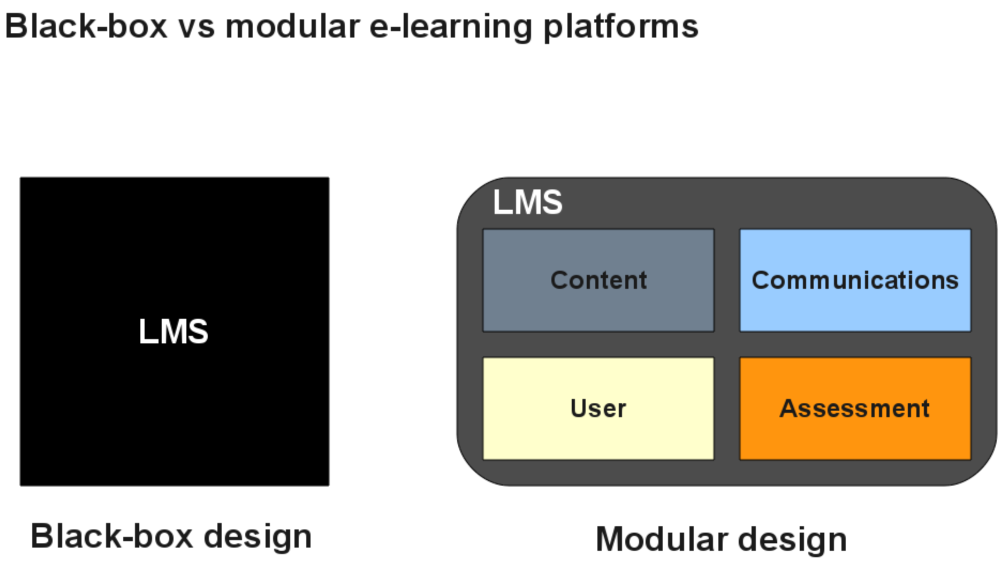


Illustration 2: Black-box vs. modular e-learning platforms

Examples of recognised modular e-learning platforms are WebCT, Moodle and dotLRN.

WebCT

Web-based Course Tools system (WebCT) is one of the first web-based learning environments and it has been commonly situated at the top of the list of most recommended/used learning packages.

WebCT was originated from a research conducted by professor Murray W. Goldberg at the Canadian University of British Columbia (UBC) in 1996 (Goldberg et al., 1996). Goldberg, a researcher in web-based systems for education, aimed to make courses and course material available using the Internet and founded WebCT Educational Technologies Corporation, a spin-off company of the UBC, in 1997. WebCT progressively became popular among students and faculty. Goldberg's company was latterly acquired by its competitor Blackboard Inc[8], a company which licenses software applications and related services with special focus on education, and current owner of WebCT.

WebCT addresses modularised and customised e-learning by providing extension toolkits such as Blackboard Building Blocks[9] and application interfaces such as Blackboard Powerlinks, which makes it possible for external applications to interface with Blackboard e-learning platforms.

However, over the years, WebCT found a strong competitor in free/open source software, which allowed institutions to reduce license costs and gave developers direct access to the source code. GNU/GPL licensed systems like Moodle and dotLRN increasingly gained attention.

Moodle

Illustration 3: Moodle, Modular Object-Oriented Dynamic Learning Environment

Moodle[10], which stands for Modular Object-Oriented Dynamic Learning Environment, is a free and open-source e-learning platform first released in 2001 and created by Martin Dougiamas, a former WebCT administrator at the University of Curtin, Australia. Dougiamas authored several publications in the field of e-learning pedagogy before the defense of his thesis entitled "The use of Open Source software to support a social constructionist epistemology of teaching and learning within Internet-based communities of reflective inquiry". The results obtained from this research conducted the conception of Moodle, whose design and development are based in the adoption of pedagogical principles, particularly social constructivism and social constructionism (Dougiamas & Taylor, 2003).

Since its first release, Moodle has evolved and gained interest among academic and business communities. The version 1.9.9, released in June 2010, has been translated into 82 different languages and the rapidly growing user-base is spread among Moodle web-sites in 209 countries.

[8] http://www.blackboard.com/
[9] http://www.blackboardextensions.com/
[10] http://moodle.org/

dotLRN

dotLRN[11] or .LRN (dot learn) is another open modular e-learning system originated at the Massachusetts Institute of Technology (MIT)[12] and built upon OpenACS[13], a software toolkit for building scalable and community-oriented web applications. .LRN 1.0 was released in April 2003 and it has been adopted by a large number of institutions including higher education, government and non-profit organisations.

Sakai

Notwithstanding the advantages in the adoption of modular designs and pedagogical strategies, service orientation, a new software architectural approach, started to be introduced with the goal of increasing the level of granularity and thus enhance the flexibility and interoperability of the e-learning platform components.

Service orientation enhances the flexibility and interoperability of the components.

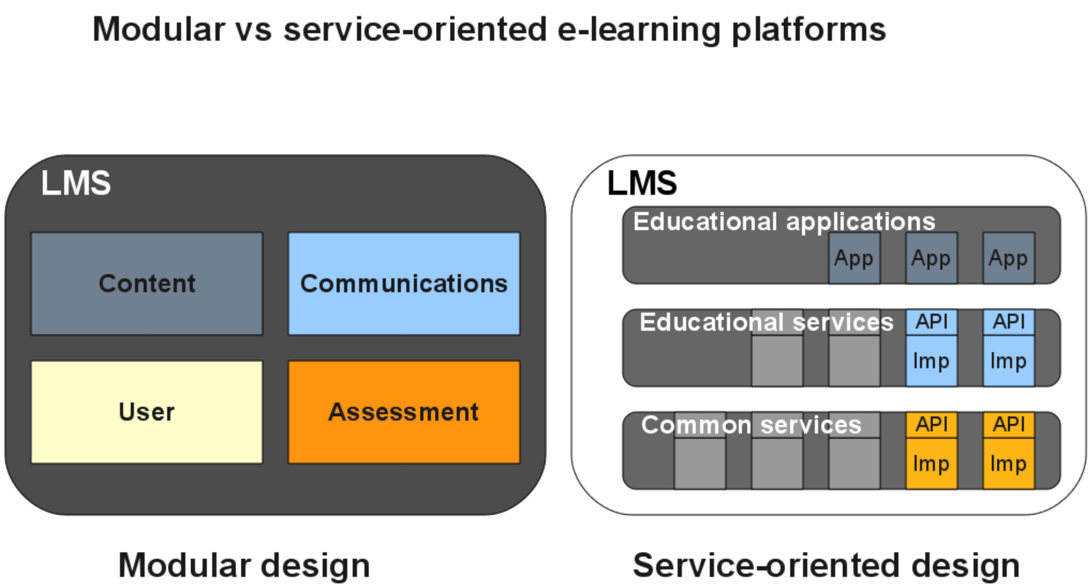


Illustration 4: Modular vs. service-oriented e-learning platforms

[11] http://dotlrn.org/
[12] http://web.mit.edu/
[13] http://openacs.org/

Just about two years after the release of .LRN, a community of academic institutions, from which the Massachusetts Institute of Technology is also part, made available the first service-oriented e-learning platform, Sakai[14].

Sakai is a free and open-source e-learning platform licensed under the Educational Community License, which is mainly intended for use by academic institutions. Sakai Collaboration and Learning Environment (Sakai CLE) was conceived during the period from January 2004 to March 2005, with the support of the Andrew W. Mellow foundation[15] and contributions from a number of universities and academic institutions.

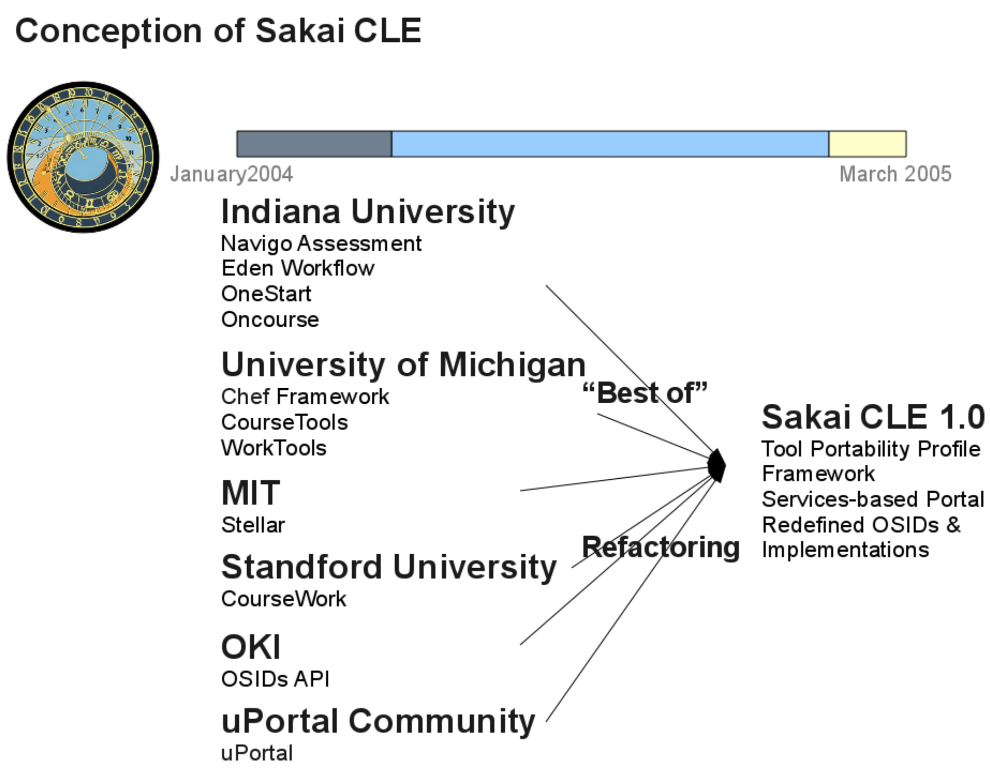


Illustration 5: Conception of Sakai CLE

In addition to course management features, Sakai has been intended as a collaborative tool for research and group projects. The 2008 Sakai Community Map included over 200 educational institutions, most of them located at the United States of America.

[14] http://sakaiproject.org/
[15] http://www.mellon.org/

Alternatives

Many more web-based learning platforms exist and new proposals are continuously appearing. A significant number of academic and educational institutions have focused more on self-developed e-learning solutions in order to better fit their needs. Despite their isolated character, some self-developed solutions have been published and open for use in other institutions. This is the case of Segue[16], a free open source collaborative learning system by the University of Middlebury.

By the middle of the 2000s, institutions started to consider educational opportunities through the use of web-based social networking software (Klamma et al., 2007). Virtual reality (Mikropoulos & Natsis, 2010) became popular in web-based learning with the support of a computer simulated 3-D environment, Second Life[17], which was first released in 2003. One year later, David Tosh, a doctoral candidate in e-learning at the University of Edinburgh, and Ben Werdmuller, a web developer with experience in on-line communities, created Elgg[18], a social networking approach to e-learning. Elgg combines elements of an electronic portfolio (Zubizarreta, 2009), weblog and social networking with the goal of connecting learners and creating communities of learning. The first version of Elgg became available on August 2008, one month after the completion of the state of the art for my Master thesis dissertation.

Illustration 6: Elgg web-site in summer 2008

Web-based e-learning at the University of Oviedo

Our previous experiences were performed at the University of Oviedo, one of the pioneers in Spain in complementing classroom teaching with on-line education. Between 1997 and 1999 the University worked on two self-developed learning solutions, the Welcoming Environment for Lifelong Learning in Pathology (Sampedro et al., 1999) and a more general virtual classroom initially called Aulanet (Perez et al, 1999), along with other e-learning solutions such as the III Virtual Spanish American Congress of Pathological Anatomy (Martinez et al, 1999). I was privileged to be part of the research team for the first and the former.

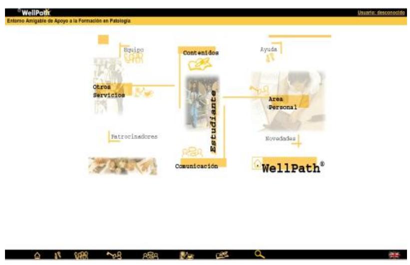


Illustration 7: Wellpath, Welcoming Environment for Lifelong Learning in Pathology

[16] http://segue.middlebury.edu/sites/segue/
[17] http://secondlife.com/
[18] http://elgg.org/

Illustration 8: Virtual Classroom at the University of Oviedo

At a later stage, WebCT was the choice for the University's virtual classroom. But in October 2006, three years after the release of Moodle, the University of Oviedo decided to migrate its virtual classroom from WebCT to Moodle. This migration has also been performed by many public universities in Spain and other universities and organisations all around the world (Jamieson and Verhaart, 2005)(Corich, 2005)(Ahmed, 2005).

Web-based e-learning at Spanish universities

In order to extend this information and be able to establish the level of use of Web-based e-learning platforms at Spanish public universities, we performed an analysis based on the publicly available information or directly contacting the universities when necessary. This tracking started in January 2008 and was continued until October 2009, date when the results hereby presented were taken.

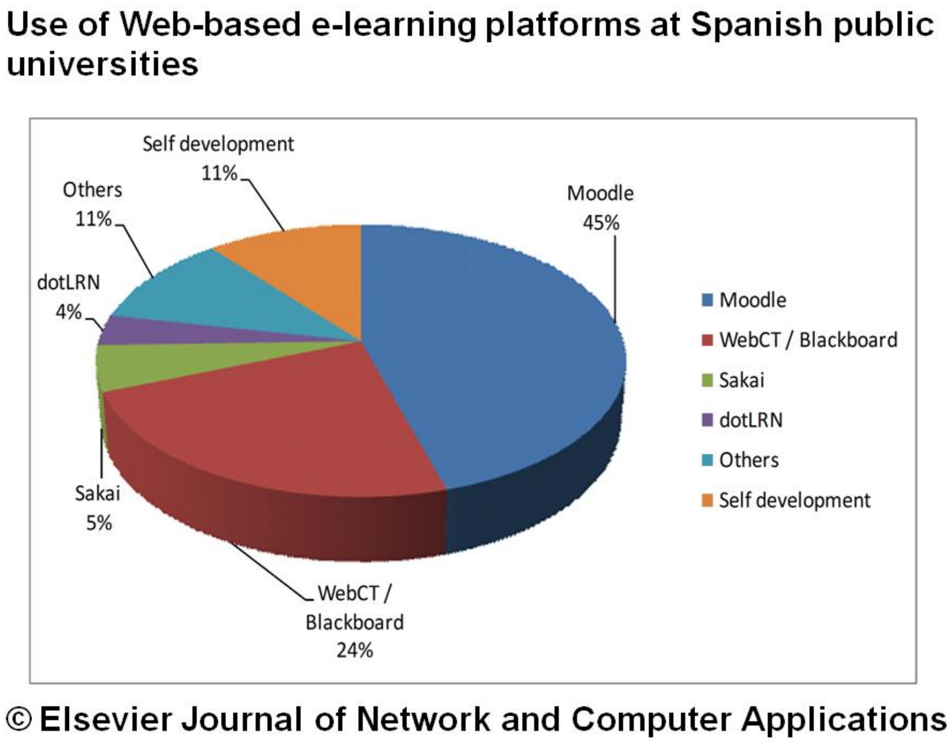


Illustration 9: Use of web-based e-learning platforms at Spanish public universities

In October 2009, Moodle and WebCT were the most used learning management systems at Spanish public universities. But besides Moodle and WebCT it is also possible to find other learning management systems being used at Spanish public universities, such as Sakai and dotLRN, as well as self-developed solutions. Also virtual environments and social networking tools are present, such as Second Life or Elgg.

The information on the use of e-learning platforms in Spain was extended in a final year project by student Adrián Rodríguez, which includes data from private universities and a comparison on e-learning technologies between the years 2005 and 2009 (Rodríguez, 2009).

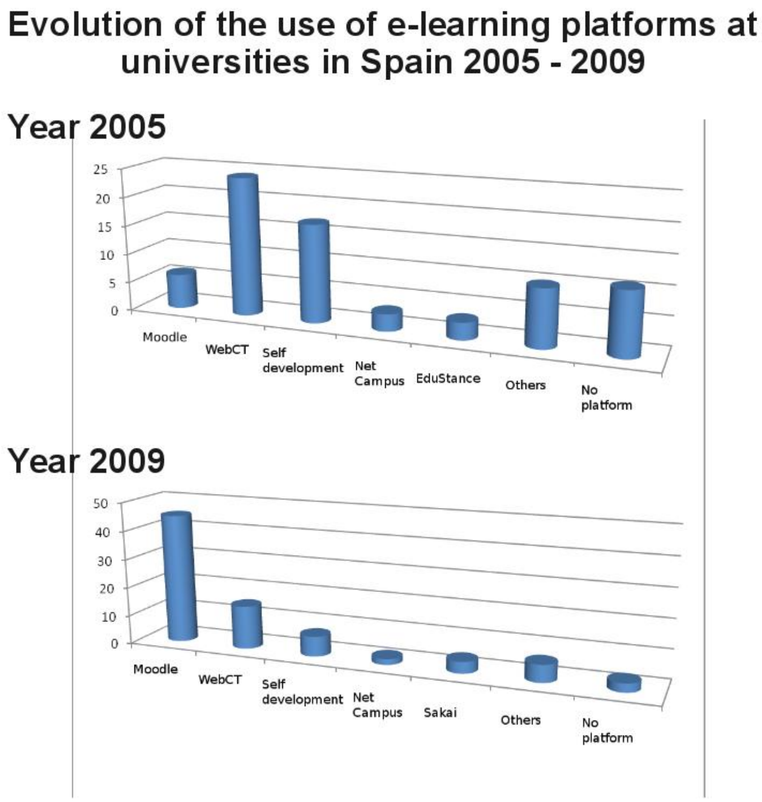


Illustration 10: Evolution of the use of e-learning platforms at universities in Spain 2005-2009

In order to understand the evolution of the use of e-learning platforms at universities in Spain, Adrián Rodríguez suggests looking into the significant increase in the adoption of e-learning platforms, technical challenges, reluctance of teachers and students to switch platforms and importance of economic costs.

Conclusions

Traditional e-learning is being replaced by more flexible e-learning.

Learning management systems have dominated Internet-based education for the last two decades. Traditional passive e-learning is being replaced by more flexible and dynamic e-learning. This implies the demand for more flexible software architectures in order to meet the needs of new learning scenarios. Architectural flexibility is being addressed by providing toolkits which support customisation, making source code available for modification or allowing service composition.

This state-of-the-art shows that there are a number of web-based e-learning platforms in use at academic and educational institutions. Although Moodle and WebCT are predominant at the time of carrying out this study, there is a dispersal focus owing to the continuous evolution of the e-learning platforms field. Hence, it is difficult for organizations and researchers to adopt the right strategy towards the development and adaptation of e-learning management systems.

Some of the factors determining the use of e-learning are technical challenges, reluctance of teachers and students to switch platforms and importance of economic costs. These factors may also explain that none of the e-learning platforms which we reviewed include dialogues as an interactive tool the user can benefit from. Perhaps this is due to the fact that the architecture of such platforms needs to be modified and such changes are often problematic and costly.

III

E-learning Frameworks

Requirements for flexibility and intra-domain/inter-domain interoperability in e-learning platforms are addressed by software architecture design guidelines. This chapter analyses the evolution of software frameworks for developing web-based e-learning platforms. It reviews the main international contributions and summarises their features and main elements.

Introduction

E-learning frameworks provide architecture design guidelines for developing e-learning platforms.

E-learning frameworks provide software architecture design guidelines for the development of web-based e-learning platforms. Since the release of IMS Abstract Framework in 1999, e-learning frameworks have focused the attention on the development of service-oriented architectures for e-learning platforms. Thus, they define and describe general and educational specific services and components, development patterns and guidelines and, exceptionally, implementation examples.

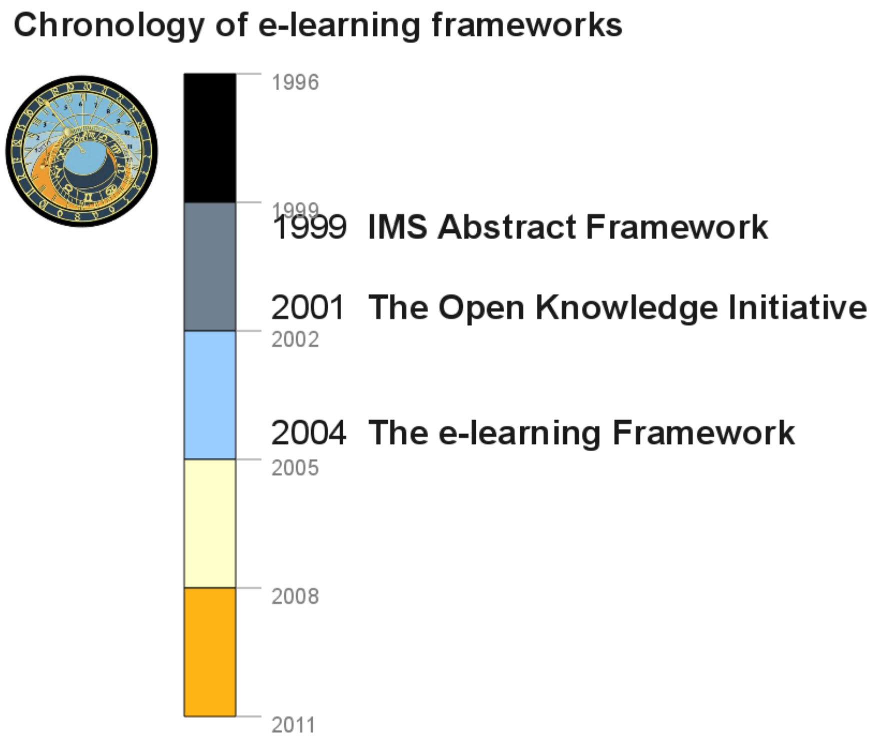


Illustration 11: Chronology of e-learning frameworks

MIT has promoted major academic e-learning initiatives.

The Massachusetts Institute of Technology has promoted major academic e-learning initiatives and has participated in the elaboration of the e-learning frameworks IMS Abstract Framework and The Open Knowledge Initiative.

IMS Abstract Framework

IMS Global Learning Consortium, or IMS Global[19], started its activity in 1997 contextualised within the initiative 'National Learning Infrastructure' by EDUCASE[20], an international non-profit association with base in the United States of America, and whose mission is to advance higher education by promoting the intelligent use of information technology.

IMS Abstract Framework (IAF)[21] is a work by the IMS Global Consortium with the partnership of the Massachusetts Institute of Technology and Blackboard, among others. IAF describes the general architectural assumptions that underlie IMS specifications and promotes interoperability among e-learning systems by adopting a series of software development guidelines. IAF provides an abstract representation for the services and components of an interoperable e-learning platform.

IAF describes general architectural assumptions.

IAF rest on the following described underlying principles:

- Interoperability

  The specifications are focussed on the exchange of information between systems. The specifications make no assumptions on how the data is managed within the communicating systems.

- Service orientation

  The exchange between the systems is to be defined in terms of the services being supplied by the collaboration of the systems.

- Component-based

  The set of services will be supplied as a 'sea of components' that can be mixed and matched to form a particular service. A single component may provide all or a sub-set of a service.

- Layering

  The total set of services required to make an e-learning system will be modelled as a set of layers, with each layer providing a clearly defined set of services.

- Behaviours and data models

  A service will be defined in terms of its behaviours and data model.

[19] http://imsglobal.org/
[20] http://educase.edu/
[21] http:/imsglobal.org/af/

- Multiple bindings

The information model for a specification (behaviour and data) can be supported through multiple bindings, e.g. as Java, XML, web services, etc.

- Adoption of specifications

New specifications will only be created as required.

The IMS Abstract Framework is available through the IMS public website. All organizations interested and active in the domain of e-learning are encouraged to download and review the IAF document set.

The first version of IMS Abstract Framework was conceived between the years 1999 and 2003, and culminated with the publishing of a second version in August 2003. IAF 1.0 is described in three documents, a white-paper[22] which provides general information about the framework, a second document with a list and description of applications, services and components[23], and a glossary of terms[24].

IAF service-oriented abstract model is layered in four subsections.

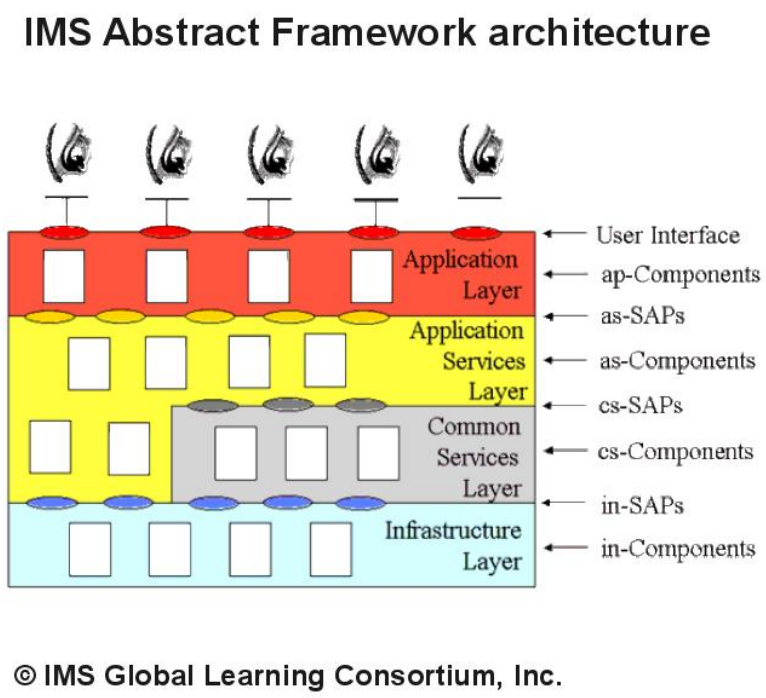


Illustration 12: IMS Abstract Framework architecture

[22]http://www.imsglobal.org/af/afv1p0/imsafwhitepaperv1p0.html
[23]http://www.imsglobal.org/af/afv1p0/imsafascv1p0.html
[24]http://www.imsglobal.org/af/afv1p0/imsafglossaryv1p0.html

- Application layer

  It consists of tools, systems, agents, etc. which interface and present the appropriate application services to the user, hiding when possible the internal system composition.

- Application Services layer

  It is composed from a set of services that provide the required e-learning functionality to the applications. An application service, which is the equivalent of educational service in other proposals, may make use of one or more common services.

- Common Services layer

  It is constituted by a set of services that are available to the application services. Common services may use other common services. Therefore, a common service is available to any other service.

- Infrastructure layer

  This layer provides end-to-end transactions and communication services for the application and common services layers.

The concept of service interface is treated in IMS Abstract Framework as Service Access Point (SAP). Access to a service is made through the appropriate SAP. A component may support one or more SAPS, each service having a single one.

IMS proposals focus on the exchange of interoperable data. Without discarding the possibility of adopting a different strategy, IMS proposes the use of WSDL for the service definition, SOAP for messaging implementation and HTTP for its transportation.

IMS addresses the implementation of the services by using a set of components called 'sea of components'. The components cooperate and are combined to support the specification of the e-learning architecture. A concept latterly referred as service orchestration or service choreography.

Services in IAF are defined in an abstract form and no detailed API with routines, data structures or object classes is given. Nonetheless, there is a close connection between the former and the e-learning framework by The Open Knowledge Initiative, which provides a more specific description of the service interfaces and their implementation, as it is analysed below.

The Open Knowledge Initiative

The Open Knowledge Initiative (OKI)[25] is an organization initially supported by the Andrew W. Mellow foundation and leaded by the Massachusetts Institute of Technology with collaboration from the universities of Stanford, Michigan, Cambridge, Dartmouth, North Carolina, Pennsylvania and Wisconsin-Madison.

OKI describes how the components of a software environment communicate with each other and with other systems.

OKI was started in January 2001 with the goal of offering a software framework for the development of e-learning systems in higher education. OKI is an initiative which develops and promotes specifications that describe how the components of a software environment communicate with each other and with other enterprise systems. O.K.I. specifications enable sustainable interoperability and integration by defining standards for Service Oriented Architectures. Through this work, OKI seeks to open new market opportunities across a wide range of software application domains.

OKI proposes the adoption of a service oriented architecture that decouples service interfaces from the underlying implementation, in such a way, that the learning services built upon this architecture can interoperate across a wider set of Web-based learning management systems.

Following this objective, OKI has defined a set of services which can be developed from technology independent interfaces called Open Service Interface Definitions (OSIDs). OSIDs follow a contract-oriented strategy to allow the compatibility with most other technologies and specifications, as well as the integration of the components in heterogeneous systems and guarantee the interoperability among e-learning platforms.

OKI guidelines specify how the components of a learning technology environment communicate with each other and with other campus systems. By clearly defining points of interoperability, the architecture allows the components of a complex learning environment to be developed and updated independently of each other. This leads to a number of important benefits[26]:

- Learning technologies which are appropriate for a range of teaching and learning requirements can be integrated together into a common environment.

[25]htttp://www.okiproject.org/
[26] http://web.mit.edu/oki/learn/whtpapers/

- The specifications are focussed on the exchange of learning technology, and content can be more easily shared among educational institutions.

- There is a lower long term cost of software ownership because single components can be replaced or upgraded without requiring all other components to be modified.

- Modularity makes learning technology more stable, more reliable, and able to grow with increased usage, and allows components to be updated without destabilizing other parts of the environment.

The architecture offers a standardised basis for learning technology software development. This can reduces development effort and encourages the development of specialised components which can be integrated into larger systems.

OKI announced the first public version of the OSIDs in June 2003. In February 2004, the Andrew W. Mellow foundation offered funds for the Sakai project, a service-oriented web-based e-learning platform written in the Java programming language. OKI became a founder member of the Sakai project, which was partially built following its premises.

Although OSIDs were originally considered for Java, in order to facilitate developers to employ different programming languages for the bindings, the next version of the OSIDs used neutral XML representations of the services named XOSIDs. Services represented as XOSIDs can be complemented by XSL transformations (XSLT) to provide an interface in a particular programming language (i.e. Java).

In February 2005, a month before Sakai 1.0 was released, developers from University of Middlebury's e-learning platform Segue met Sakai developers to discuss the implementation of the OSIDs using the PHP programming language. Middlebury contextualised this development as part of the Harmoni PHP software development framework[27], which includes a set of implementations of the OSIDs that rely on features of Harmoni[28]. OSIDs include service definitions for authentication, authorisation, course management, grading and messaging to mention but a few.

[27]http://harmoni.sourceforge.net/
[28]http://www.harmony-framework.com/

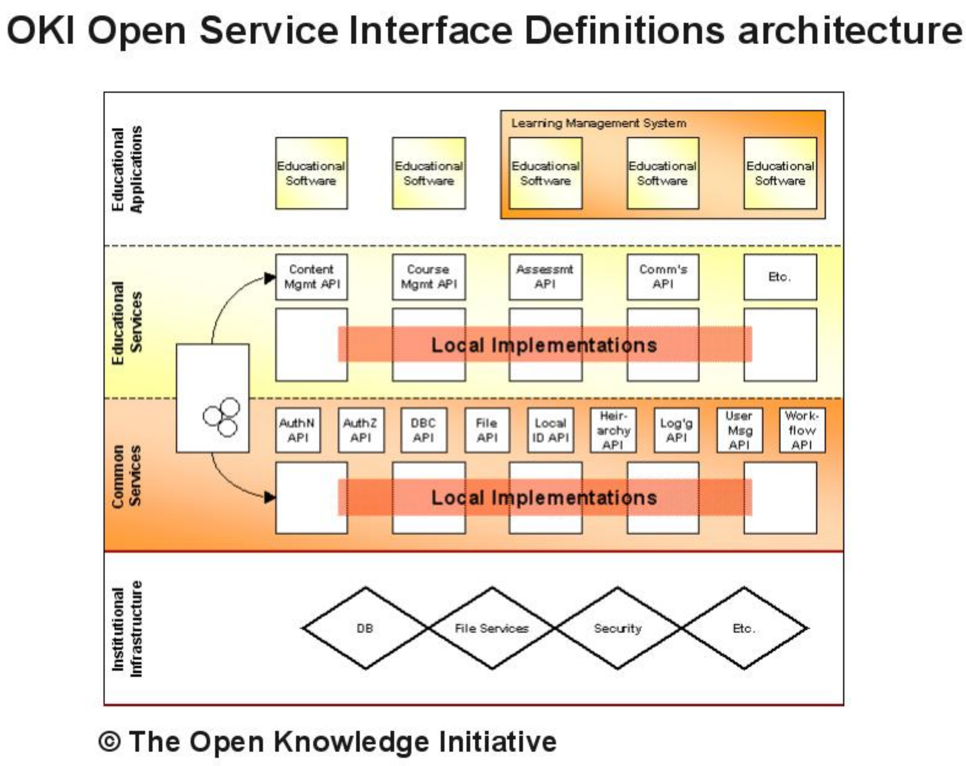


Illustration 13: OKI Open Service Interface Definitions architecture

OKI has continued its evolution, addressing the numerous issues that developers have raised over the years and adding new educational services. The OKI architecture follows the same principles as IMS Abstract Framework, with software components and services vertically organised in four layers. An upper application layer represents the various e-learning tools which can be made from the composition of services. Two services layers distinguish educational from common services, whereas the underlying infrastructure layer provides the communication links as well as "back end" capabilities like file systems, databases, registries and authentication servers.

The use of OKI SOA contract-first design enables e-learning platforms to interoperate. Using Java, PHP or any other programming language, predetermined OSID interfaces can be implemented and integrated into the e-learning platform to provide general software contracts between a particular platform and service consumers.

Following contract-first practices, The Open Knowledge Foundation along with the universities of Middlebury, MIT and the Open University of Catalonia, have set up the OSID bridging project[29]. OSID bridging is an attempt to create an XML-RPC bridge between OSID v3 repository implementations written in both PHP and Java.

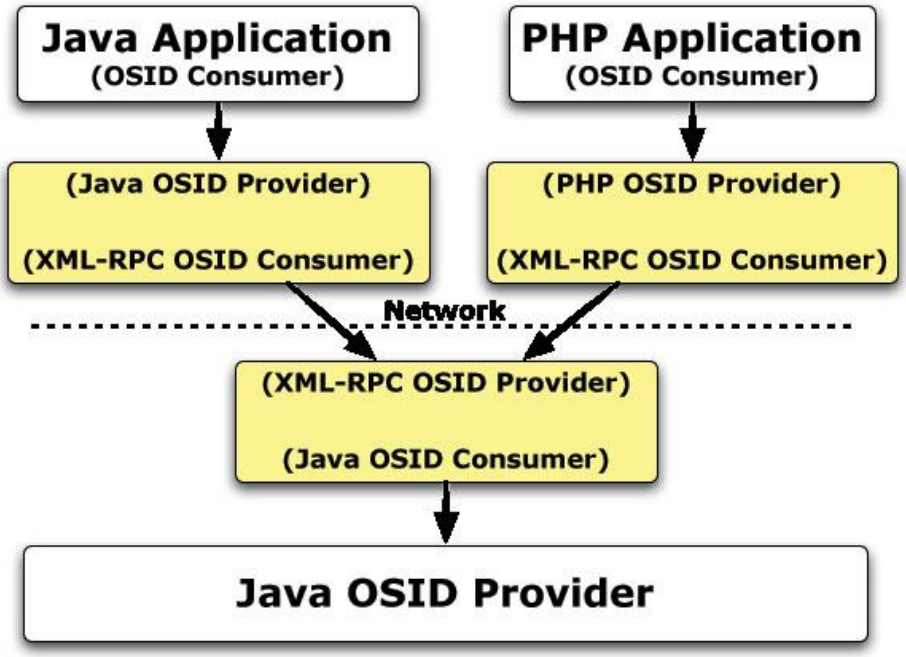


Illustration 14: OKI OSID bridge

OSIDs can also be implemented in a middleware layer to offer a common interface and provide interoperability among heterogeneous learning management systems. Following contract-first design, software developers can create service consumers independently of any particular e-learning platform or version. This approach can be seen in the Campus project[30], a project leaded by the Open University of Catalonia in Spain.

Illustration 15: Campus Project

[29]http://sourceforge.net/apps/mediawiki/harmoni/index.php?title=OSID_Bridging
[30]http://www.campusproject.org/

The e-Framework for Education and Research

e-Framework provides information on investing and using information technology infrastructures.

The e-Framework for Education and Research[31] is an international initiative that provides information to institutions on investing and using information technology infrastructures. The primary goal of this framework is to facilitate technical interoperability within and across education and research through improved strategic planning and implementation processes.

The e-Framework was born in May 2004 under the name The e-learning Framework (The ELF) and promoted by the Joint Information Systems Committee (JISC)[32], a board made up of different institutions from the United Kingdom to lead the innovative use of digital technologies and help to maintain the United Kingdom's position as a global leader in education.

In order to extend and further spread the proposal, JISC, provider of the original JISC e-Learning Framework, developed a partnership with the Department of Education, Employment and Workplace Relations of Australia[33] and Industry Canada, the department of the Government of Canada with responsibility for regional economic development, investment and research.

One of the most significant works at the time of launching The e-learning Framework was conducted by the Technology-Assisted Lifelong Learning Unit at the University of Oxford (TALL)[34]. TALL conducted a study to review a predefined set of existing e-learning developments[35]. The scope of the study consists of the following three levels:

- E-learning implementations, including virtual learning environments (VLEs), e-learning platforms and larger integrated e-learning systems, such as WebCT, Moodle and Sakai.
- E-learning frameworks, standards and specifications, including IMS Abstract Framework, OKI and e-GIF.
- Technology platforms, including Microsoft's .NET[36] and Sun's Java 2 Enterprise Edition[37].

[31] http://e-framework.org/
[32] http://www.jisc.ac.uk/
[33] http://www.deewr.gov.au/
[34] Http://www.tall.ox.ac.uk/
[35] http://www.elframework.org/projects/tall/
[36] http://www.microsoft.com/net/
[37] http://java.sun.com/j2ee/

The results extracted from the study were reassessed and mapped into The e-learning Framework to update and link the former JISC proposal with the newest and most relevant work in the e-learning field at the time the study was made (summer 2004).

During the subsequent period 2004-2006, The e-learning Framework settled down the basis for addressing the level of development and future of information systems for teaching and learning. To that end, the Centre for Educational Technology and Interoperability Standards (JISC-CETIS)[38] organises an annual conference with the collaboration of its partners. This event is held in a different location in the United Kingdom every year.

The e-Learning Framework proposes a software architecture oriented to the use of services and open standards for the development and integration of computer systems in the sphere of learning, research and education administration. This e-learning framework is based on a service-oriented factoring of a set of distributed core services required to support e-learning applications, portals and other user agents. Each service defined by the framework is envisaged as being provided as a networked service within an organisation, typically using either web services or a REST-style HTTP protocol (Fielding, 2000).

Built on earlier initiatives, The e-learning Framework shows a set of sample user agents which relays on two sets of services defined within the framework. The upper set of boxes identifies services specifically within the domain of e-learning; the lower set identifies services that may be common across multiple domains.

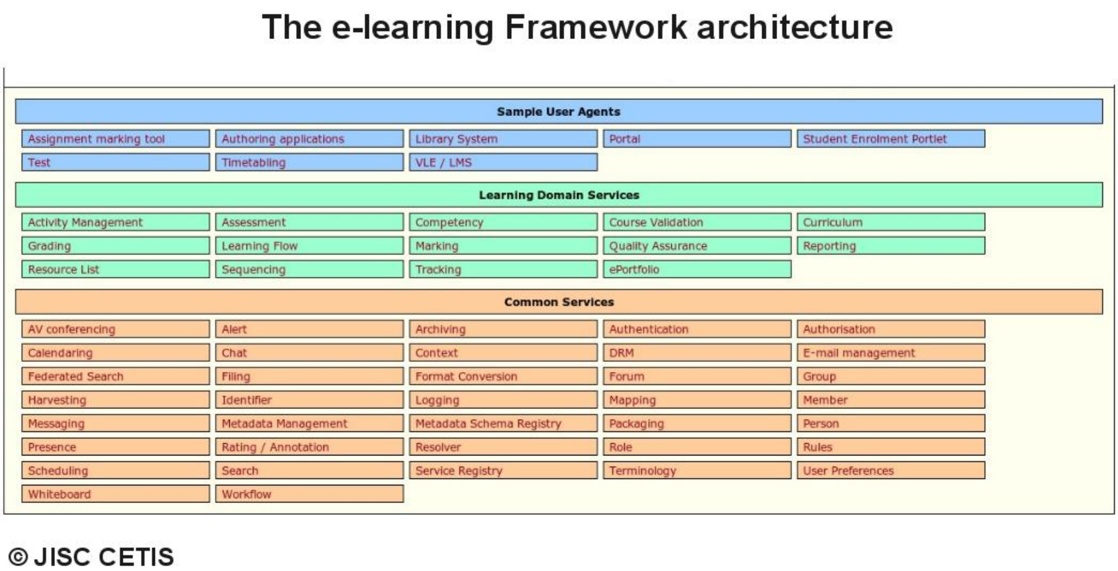


Illustration 16: The e-learning framework architecture

[38] http://jisc.cetis.ac.uk/

The structure of the framework did not suffer any major modification in 2004-2006. During the same period, and as a consequence of the multiple collaborations and meetings, a number of projects and publications arose.

In the year 2007 the proposal got a new boost and welcomed two new partners, the Ministry of Education of New Zealand[39] and the organisation for higher education institutions and research institutes of the Netherlands (SURF)[40]. It was then that the framework was subsumed into the wider initiative The e-Framework.

The transition did not result in any substantial modification from the principles and goals formerly established. The e-Framework advocates service-oriented approaches to facilitate technical interoperability of core infrastructure as well as effective use of available funding. From a technical point of view, The e-Framework attempts to follow an approach with smaller, reusable services in order to produce more efficient and, ultimately also more economical, e-learning systems. To this end, a significant change was made in the element composition of the service-oriented framework, which became categorised in services, service usage models and guides.

- Services

Within the e-Framework, a service is used in the sense of a technical service that promotes interoperability between systems. It relies on service-oriented architectures (SOA) and web services (WS).

- Service Usage Models (SUMs)

SUMs provide a description of the needs, requirements, workflows, management policies and processes within a domain, and the mapping of these to a design of service genres and service expressions, resources, associated standards, specifications, data formats, protocols, bindings, etc., that can be used to implement software applications within the domain. In other words, SUMs model how services meet businesses needs.

- Guides

Guides are used to explain contributors to the framework how to write and submit documentation.

Following the aforementioned model, The e-Framework has defined an evolving service-oriented knowledge base, built on contributions from members of its community. Moreover, The e-Framework preserves the goal of assisting other international education and research communities in planning, prioritising and implementing their IT infrastructure in a better way.

39 http://www.minedu.govt.nz/
40 http://www.surf.nl/

Conclusions

There is a common trend in e-learning platforms towards improving software extensibility, flexibility, pluggability and reusability. E-learning frameworks share the idea of promoting the adoption of service-oriented approaches which help to facilitate technical interoperability to the internal software infrastructure (intra-domain) and among heterogeneous e-learning systems (inter-domain). Although the basis is maintained, every proposal has specific particularities. From a technical point of view, we find The Open Knowledge Initiative particularly appealing, as it directs the service-oriented philosophy to a very practical end, providing developers with service interface definitions oriented to the development of e-learning platforms. These characteristics of the OSID interfaces can be used to integrate visual and voice services into any e-learning platform, making them very convenient for our goals.

E-learning frameworks promote the adoption of service-oriented approaches to facilitate technical interoperability.

IV

Voice-enabled Internet Applications

This chapter firstly summarises the evolution of research to create machines and mechanisms to simulate human speech communication capability. Then, this section focuses on interactive voice response technologies which allow the creation of voice applications for the Internet and enable integration with other Internet-based software.

Introduction

Using machines to mimic the capability of speaking naturally has intrigues engineers and scientists for centuries.

Speech is the primary means of communication for human beings. While not all human languages have a written form, every language has a spoken form. The challenge of using machines to mimic human behaviour, and particularly the capability of speaking naturally, has intrigued engineers and scientists for centuries (Juang & Rabiner, 2005). Initial attempts to develop machines which can simulate human speech communication capability consisted in mechanical artefacts which appear to have started in the second half of the 18th century.

In 1773, the Russian scientist Christian Kratzenstein, a professor of physiology in Copenhagen, succeeded in producing vowel sounds using resonance tubes connected to organ pipes (Kratzenstein, 1782). Later, Wolfgang von Kempelen, a Hungarian writer and inventor, constructed in Vienna an "Acoustic-Mechanical Speech Machine" (1791) (Dudley & Tarnoczy, 1950) and in the mid-1800's Sir Charles Wheatstone built a version of von Kempelen's speaking machine using resonators made of leather (Wheatstone, 1879).

The first electronic devices appeared during the first half of the 20th century. In the 1930s, Homer Dudley, a researcher of telephone transmission for Bell Telephone Laboratories[41], and who was greatly influenced by Harvey Fletcher's research on speech perception, proposed a system model for electronic speech analysis and synthesis which culminated in the development of a synthesizer called the VODER (Voice Operating Demonstrator)(Dudley et al.,1939).

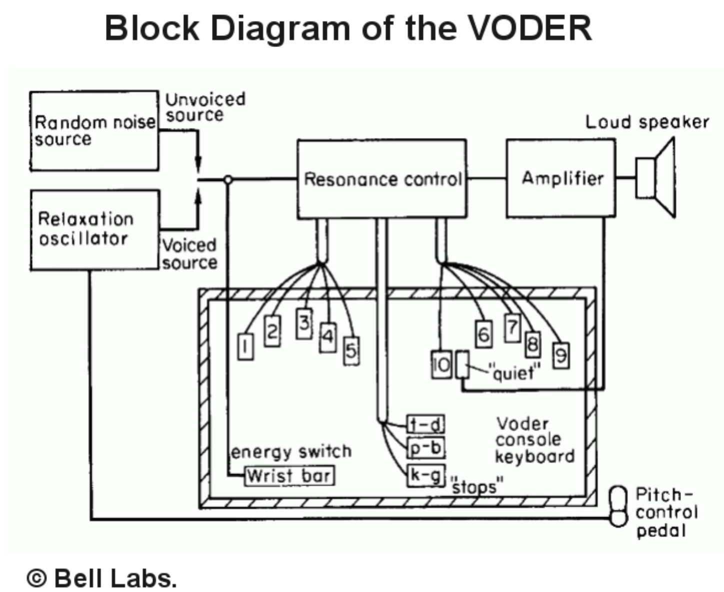


Illustration 17: Block diagram of the Voice Operating Demonstrator

[41] http://www.bell-labs.com/

It was only a few years after Dudley's speech synthesiser that speech recognition was firstly attempted. First designs for automatic speech recognition were mostly guided by the theory of acoustic-phonetics, which describes the phonetic elements of speech (the basic sounds of the language).

In the 1950s, Bell, RCA and MIT laboratories hosted research projects to attempt basic speech recognition. In 1952, Davis, Biddulph, and Balashek, of Bell Laboratories built a system for isolated digit recognition for a single speaker (Davis et al., 1952). Shortly later, Olson and Belar of RCA Laboratories built a system to recognise 10 syllables of a single talker (Olson & Belar, 1956) and at MIT Lincoln Laboratory, Forgie and Forgie built a speaker-independent 10-vowel recogniser (Forgie & Forgie, 1959).

The problem of speech synthesis and automatic speech recognition has been approached progressively, from a simple machine that produces or responds to a small set of sounds to a sophisticated system that generates and responds to fluently spoken natural language and takes into account the varying statistics of the language in which the speech is produced.

In 1995, Philip R. Cohen and Sharon L. Oviat (Cohen & Oviat, 1995), assessed potential benefits of spoken interaction with machines and identified a number of situations in which spoken communication with machines may be advantageous:

There are a number of situations in which spoken communication with machines may be advantageous.

- When the user's hands or eyes are busy.
- When only a limited keyboard and/or screen is available.
- When the user is disabled.
- When pronunciation is the subject matter of computer use.
- When natural language interaction is preferred.

Spoken interaction has progressively found application in areas such as medical devices information, educational software, customer care and leisure. According to a Forrester Research market study in 2006 (Herrell, 2006), the speech recognition market has reached an inflection point. After many years of slow growth, the year 2006 showed increased spending on speech technology with a compound annual growth rate of 52% versus 11% for IVR spending over the previous five years. The percentage of companies who had deployed speech grew fourfold from 10% in 2000 to 40% in 2003%, with an additional 32% planning to deploy speech every year.

Spoken interaction has progressively found application areas.

Speech technology gets more attention than usual these days because it is a good fit for pervasive computing solutions.

Speech technology gets more attention than usual these days because it is a good fit for pervasive computing solutions (Lai, 2000). Pervasive computing devices are often used in hands-busy, eyes-busy settings and usually lack both usable displays and keyboards. Either of these characteristics would be a solid foundation for using speech. Combined, they make the argument that much more compelling.

However, even though there are many fields in which spoken interaction can play a significant role, and even though users show a supportive attitude and strong intentions to use speech interfaces, there are still many technical difficulties in simulating a human conversation using human-machine interfaces.

Despite small vocabulary is becoming a familiar feature for users of telephone-based interactive voice response (IVR) systems, contemporary automatic speech recognition systems are not able to fulfil the requirements demanded by many potential applications, and their performance is still significantly short of the capabilities exhibited by human listeners (Moore, 2003).

By far the most comprehensive comparison between automatic and human speech recognition accuracy was performed by Richard P. Lippmann at MIT's Lincoln Laboratory in 1997 (Lippmann, 1997). Lippmann compiled results from a number of well-known sources and presented comparative word error rates (WER) for a range of tasks and conditions. The results indicated clearly that, in terms of word error rate scores, automatic speech recognition performance lags about an order of magnitude behind human performance.

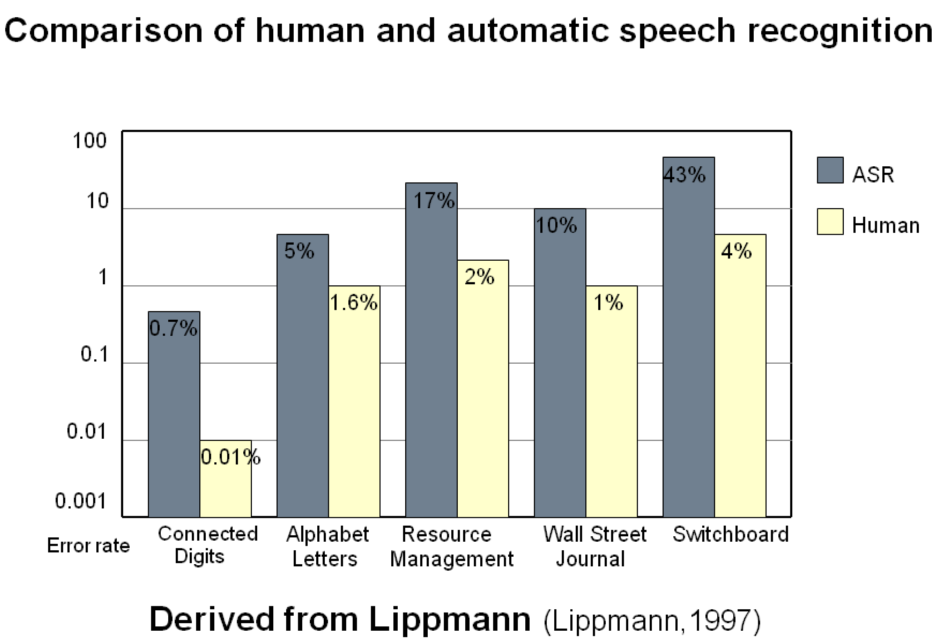


Illustration 18: Comparison of human and automatic speech recognition

Interactive voice response applications

Interactive voice response applications are designed to accept and return audio to the user. They commonly require the use of automatic speech recognition and speech synthesis technologies. Commercial providers and academic researchers have addressed this need by offering a varied list of Automatic Speech Recognition (ASR) and Text to Speech (TTS) modules, which have progressively added resources and support for more languages.

Examples of automatic speech recognition and speech synthesis providers are Loquendo[42], LumenVox[43], CMU Sphinx[44], a speech recognition toolkit by the Carnegie Mellon University, or Festival[45], a general multi-lingual speech synthesis system developed at The Centre for Speech Technology Research[46] from the University of Edinburgh.

But unfortunately ASR and TTS technologies are still far from being accurate and understandable at human communication level. The requirement of establishing a fluent dialogue with the machine makes ASR and TTS technologies often being replace with more simple alternatives such as Keyboard/Dual Tone Multi-Frecuency (DTMF) input and audio recording respectively, which can be more precise and understandable.

Speech technology is still far from being accurate and understandable.

Input in telephone-based applications can be taken both from ASR and Dual-Tone Multi-frequency signals (DTMF) generated by the end-user device. However the later is preferred for some applications and use cases. This decision is mainly related with the effectiveness of speech recognition and factors such as the speaker's independence, his/her expertise, confusability of the speaker's vocabulary and grammar, speaking mode (rate and coarticulation), channel conditions and user tolerance of errors (Roe, 1995). We should also mention that the demands of using speech rather than keyboard entry may slow speech users more in the higher cognitive-load task of composition (Shneiderman, 2000).

As for the output side, pre-recorded human voices can be used as a more naturalistic alternative for managing static information. However, when managing dynamic information, speech synthesis is a more appropriate technology for automating the generation of voice prompts, as it can reproduce both static and dynamic audio contents.

[42] http://www.loquendo.com/
[43] http://www.lumenvox.com/
[44] http://cmusphinx.sourceforge.net/
[45] http://www.cstr.ed.ac.uk/projects/festival/
[46] http://www.cstr.ed.ac.uk/

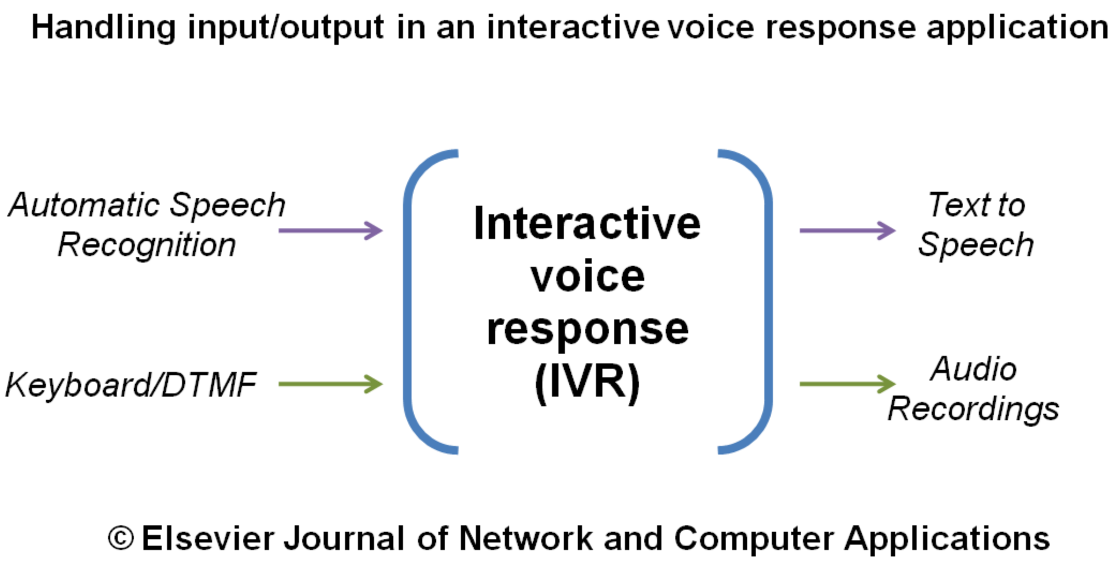


Illustration 19: Handling input/output in an interactive voice response application

While recognition accuracy is the key factor to measure the acoustical and environmental robustness of a speech recognition system (Acero & Stern, 1990), text to speech considers speech prosody and acoustic parameters such as word pronunciation, phonemisation, duration rules, tone and intonation (Schröder & Trouvain, 2003).

We also ought to acknowledge that most of the Automatic Speech Recognition and Text to Speech solutions we have evaluated focus on the English language. Hence, in countries in which English is not the primary language, multi-language support is an additional decision factor when choosing the adequate ASR and TTS technologies.

In the described context, speech-enabled systems have had limited success and only in very specific applications such as screen readers, dictation systems, in-car applications and telephone-based services (Schiller et al., 2004). In our professional experience, telephony has been the most successful commercial field for voice interaction, with plenty of examples of automated speech responses for customer services, along with other business or leisure interactive voice response services.

Enabling aural access to the Web

Nowadays websites use, above all, the visual channel to communicate content, functionality and navigation and interaction capabilities (Bolchini et al., 2006).

Nowadays websites use, above all, the visual channel.

A voice portal is the voice equivalent of a web portal. It uses spoken commands and voice responses to guide the user through the site. Voice portals have been deployed by Internet portal companies such as AOL[47] and Yahoo[48] and by dedicated companies such as Tellme Networks[49]. Latterly, the technology has been used in vCommerce (voice-commerce) applications, a term coined by Nuance Communications[50] which makes reference to the usage of speech technology over the telephone in commercial applications (Duggan & Deegan, 2003).

A more traditional approach to enable audio access to the web has been the use of assistive technologies which provide an audio interpretation of what it is displayed on the screen. Users using screen readers are forced to listen for long to irrelevant information on the page until being able to catch the relevant content (Ramakrishnan et al., 2004). This approach represents a major limitation in the attempt to carry out a naturalistic dialogue with the user. Screen readers can be delivered as desktop standalone applications, such as Jaws[51], Orca[52] and Apple's Voiceover[53], or they can be integrated in web browsers, such as the case of Opera Voice[54].

Audio interpretation of what it is displayed on the screen represents a major limitation to carry out a naturalistic dialogue with the user.

Alternatively, the release of technologies for delivery of voice communications over Internet Protocol networks (VoIP) in the second half of 1990s, allowed software developers to easily integrate voice systems with Internet applications, making it possible to provide visual and auditory access to the information by adding voice interfaces to the existing visual displays.

VoIP allows developers to easily integrate voice systems with Internet applications.

With a great proliferation of hardware and software artefacts, communication protocols and transmission technologies, many different approaches exist for the developing and deploy of voice over IP communication services. The Session Initiation Protocol (SIP), widely used for controlling multimedia communication sessions, along with the Real-time Transport Protocol (RTP), a transportation format for delivering audio and video over the Internet, can be mentioned among the most used VoIP technologies.

[47] http://www.aol.com/
[48] http://www.yahoo.com/
[49] http://www.tellme.com/
[50] http://www.nuance.com/
[51] http://www.techno-vision.co.uk/JAWS.htm
[52] http://live.gnome.org/Orca
[53] http://www.apple.com/accessibility/voiceover/
[54] http://www.opera.com/browser/tutorials/voice/

There is also still agreement on the adoption of standards for speech languages or APIs for the use of interactive voice response features and VoIP call centres, which limits the portability of solutions developed for a specific voice platform. Examples of this are the Sip Express Router (SER)[55] and Asterisk[56], which are used in call centre environments.

Voice over IP technologies encourage the convergence of telephony and web-based services (Ahuja & Ensor, 2004). Most current and envisioned VoIP services are converged services, integrating features and functions from multiple existing services. In the particular case of web-based e-learning, recent studies (Motiwalla, 2009) (Chevrin et al., 2006) propose to factorised e-learning systems into a collection of dedicated services to address the issues of multi-channel access and meet the needs of mobile learning communities.

Designing voice-based web applications

Through the visual channel, a website conveys various messages simultaneously, basically concerning content, interaction/navigation capabilities and the graphical layout of the elements of the page.

The success of a voice enabled web system hinges on acceptance by the people who use it.

Voice-based web applications can be the aural "counterpart" of existing interactive web applications, or a stand-alone application conceived and designed to meet specific purposes. Like all human-computer interaction technologies, the success of a voice enabled web system hinges on acceptance by the people who use it (Duggan, 2004).

The application of principles and guidelines for interaction design are essential.

The application of principles and guidelines for interaction design are essential in all human-computer interfaces (Fitzpatrick & Higgins, 1998). The International Standards Organisation (ISO) in the ISO 9241 – Part 11 standard (ISO 9241, 1998) and Human Factors and Ergonomics Society (HFES) in the HFES 200 standard (HFES 200, 2001) propose three core criteria, which are critical for the design of any human-computer interface:

- Effectiveness or usefulness is the accuracy and completeness which specified users can achieve specified goals in particular environments.

[55]http://sip-router.org/
[56]http://www.asterisk.org/

- Efficiency or usability describes the resources expended in relation to the accuracy and completeness of goals achieved.
- User satisfaction or desirability describes the comfort and acceptability of the work system to its users and other people affected by its use.

However, instead of providing design support, the industries have focused on providing implementation solutions to develop audio interfaces. Two main languages lead the market in the field of voice applications: Speech Application Language Tags (SALT) and VoiceXML.

SALT[57] is an XML based markup language that is used in HTML and XHTML pages to add speech output and interaction capabilities to web based applications. It is designed to allow multimodal and telephony-enabled access to information at the same time. SALT is developed by Microsoft, and plug-ins are available to enable Microsoft Internet Explorer[58] to access SALT enabled applications.

To facilitate a more universal solution, and as it happened in the 1990s with web applications, the World Wide Web Consortium (W3C)[59] has defined a series of recommendations, within the "Voice Browser" activity[60], for the definition of dialogues, speech recognition grammars, call control, etc. Voice Browser recommendations include the VoiceXML language[61] and other XML-based technologies which together attempt to create a homogeneous environment for voice applications and also enable better integration with other Internet-based software.

W3C Voice Browser attempts to create a homogeneous environment for voice applications in Internet.

VoiceXML has received important support from major players in the industry (Niklfeld et al., 2001) and development platforms and voice servers have been slowly incorporating Voice Browser technologies to describe human-machine dialogues, favouring the portability and interoperability of the software solutions that are being implemented. Examples of this are the Loquendo MRCP server[62], i6Net VXI* VoiceXML browser[63] and the Voxeo Prophecy platform[64].

These languages are crucial for the implementation of voice-based web applications, but they cannot solve the design issues needed to be considered for an effective user experience.

Voice languages are crucial, but they cannot solve design issues.

57 http://msdn.microsoft.com/en-us/library/ms994629.aspx
58 http://www.microsoft.com/windows/internet-explorer/
59 http://www.w3.org/
60 http://www.w3.org/Voice/
61 http://www.w3.org/TR/voicexml21/
62 http://www.loquendo.com/en/technology/speechsuite.htm
63 http://www.i6net.com/
64 http://www.voxeo.com/prophecy/

Conclusions

The challenge of using machines to mimic the capability of speaking naturally and responding properly to spoken language has intrigued engineers and scientists for centuries. The problem of speech synthesis and automatic speech recognition has been approached and progressively improved over the years.

However, there are still many technical difficulties in simulating a human conversation using human-machine interfaces.

Users are increasing their acceptance of applications that utilise spoken communication.

A number of situations exist in which spoken communication with machines may be advantageous, and users are increasing their acceptance of applications that utilise synthesised speech or voice recognition.

Traditionally, voice access to Internet-based information has been enabled through the use of assistive technology such as screen-readers. Assistive technologies provide only aural descriptions of the visual content, which represents a major limitation in the attempt to carry out a naturalistic dialogue with the user.

VoIP technologies allow enabling multi-channel access to Internet-based software.

Alternatively, voice over IP applications are hosted on the network, allowing integrating voice systems with any Internet-based application. Voice over IP technologies encourage the convergence of telephony and web-based services, and allow to address the issue of enabling multi-channel access to Internet-based software.

Despite the little agreement on the adoption of standards for voice-based applications, the World Wide Web Consortium has defined a series of recommendations based on markup languages, which together attempt to create a homogeneous environment for voice applications and also enable better integration with other Internet-based software.

A rethinking of the design principles to adapt visual to aural web applications is required.

However, in order to design effective multi-channel applications, a rethinking of the design principles to adapt visual to aural web applications is required.

V

Adaptive Educational Hypermedia

Previous chapters have focused on reviewing current knowledge and research directions for developing e-learning systems, as well as the level of maturity of voice-enabled Internet applications. The aim of this chapter is to offer an introduction to adaptive educational hypermedia and the methods and techniques that may be applicable to adapt user interaction.

This section reviews the evolution of adaptive hypermedia, and summarises the main existing approaches for adapting the presentation and navigation of traditional web-based hypermedia systems accordingly to user, usage and environmental factors.

Introduction

The research in adaptive educational hypermedia systems started in the 1990s. The first research proposals, made between years 1990 and 1996, can be categorised in two research streams. The area of intelligent tutoring systems (ITS) tried to combine traditional student modelling and adaptation processes to find new applications to ITS (Brusilovsky et al., 1993). Meanwhile, another stream of research focused on educational hypermedia in an attempt to make e-learning system adapt to individual students (De Bra, 1996).

Adaptive hypermedia adapts visible aspects of the systems to the users' specific needs.

Brusilovsky (Brusilovsky, 1996) defines adaptive hypermedia systems (AHS) as all hypertext and hypermedia systems which reflect some features of the user in the user model and apply this model to adapt various visible aspects of the system to the user. The user model is an essential component for any system that aims at adapting content to the users' specific needs. In adaptive hypermedia applications the system keeps track of evolving aspects of the user, such as preferences and domain knowledge. Adaptive hypermedia systems build a model of the goals, preferences and knowledge of each individual user, and use this model throughout the interaction with the user, in order to adapt to the needs of that user (De Bra et al., 1999).

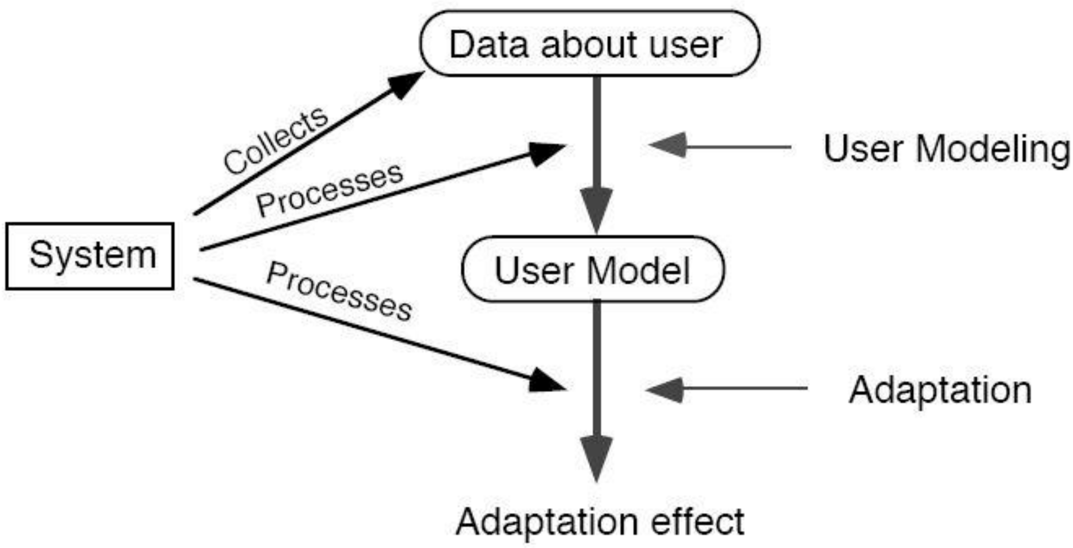


Illustration 19: Classic loop "user modeling - adaptation" in adaptive systems

Many adaptive hypermedia systems exist to date. The majority of them are used in educational applications, but some are used, for example, for on-line information systems or information retrieval systems (Wu et al., 2000). Hence, adaptive hypermedia systems have found application in different application areas (Brusilovsky et al., 1996):

- Educational hypermedia systems
- On-line information systems
- On-line help systems
- Information retrieval hypermedia
- Institutional hypermedia
- Personalised views

Educational hypermedia was one of the first application areas for adaptive hypermedia and is currently one of the most popular and well-investigated (Brusilovsky, 2001)) (De Bra & Calvi, 1998) (Weber & Specht 1997). Examples of existing adaptive educational environments are ACE (Specht & Oppermann, 1998), AHA (De Bra et al., 2002), Arthur (Gilbert & Han, 2002) and CAMELEON (Laroussi & Benahmed, 1998).

Educational hypermedia is one of the first application areas and one of the most popular and well-investigated.

Adaptive Educational Hypermedia caters to the needs of each individual student, adapting e-learning accordingly to individual needs or abilities. Traditionally, adaptation decision in adaptive systems was based on a user model that represents relevant aspects of the user such as preferences, knowledge and interests.

Adaptive hypermedia in e-learning refers to the organization of the information in the e-learning systems, integrating different paces of content (adaptive presentation) and navigation (adaptive navigation support) in order to be able to respond to diverse needs of the students and to avoid the cognitive overload (Paule, 2008). Users with different goals and knowledge may be interested in different pieces of information presented on a hypermedia page and may use different links for navigation. Furthermore, e-learning systems making a great deal of information available to the student can be too overwhelming cognitively and may lead students to quit the training process. Adaptive hypermedia addresses this need by introducing a customised hypermedia structure and allowing content to be displayed differently to diverse users depending on the needs of that particular user or a particular usage context.

Adapting to what?

The goal of adaptive hypermedia systems is to optimise the user experience by personalising the web content and navigational structure to suit the user individual requirements (Muntean & McManis, 2006).

One distinctive feature of any adaptive system is the user model that represents essential information about each user.

One distinctive feature of any adaptive system is the user model that represents essential information about each user. There are a number of features related to the user context and the user as an individual which are taken into account by adaptive systems. Traditionally, adaptation decision in adaptive systems was based on taking into account the various characteristics of their users, goals (Cliford, 2000), knowledge level (De Bra & Calvi, 1998), background, interest, preferences (Brusilovsky, 1996), stereotypes (Zakaria & Brailsford, 2002), cognitive preferences (Chen & Macredie, 2002) and learning styles (Stash et al., 2004). These characteristics can be evaluated and mapped into attribute-value pairs that represent the user model. Among the different characteristics, this research focuses on users' learning styles, and more particularly in the visual/aural dimensions.

Learning styles are characteristic cognitive, affective and psychological behaviours that serve as relatively stable indicators of how learners perceive, interact with, and respond to the learning environment (Keefe, 1979). Psychologists have proposed several complementary taxonomies of learning styles (Alonso et al., 2002) (Dunn & Dunn, 1978) (Sarasin, 1998) (Felder & Silverman, 1988) (Kolb, 1984).

The different categories proposed by the taxonomies have been mapped into questionnaires and indicators in an attempt to understand, recognise and group learning styles. When we focus on the visual/verbal learning style dimension, we can clearly see the distinction between a visual learner who learns and remembers best what is seen and a verbal learner who learns best when using audio explanations.

The Index of Learning Styles (ILS)[65] is an instrument designed to assess preferences on the four dimensions of the Felder-Silverman learning-style model. The Web-based version of the ILS is taken hundreds of thousands of times per year and has been used in a number of published studies, some of which include data reflecting on the reliability and validity of the instrument (Felder & Spurlin, 2005).

[65] http:///www.engr.ncsu.edu/learningstyles/ilsweb.html

From 1996, and influenced by the rapid increase in the use of the World Wide Web, adaptive hypermedia attracted the attention of a larger community of researchers. In 1999 Kobsa et al. (Kobsa et al., 1999) suggested to distinguish adaptation to user data, usage data, and environment data. User data comprise the traditional adaptation target, various characteristics of the users. Usage data comprise data about user interaction with the systems that cannot be resolved to user characteristics (but still can be used to make adaptation decisions). Environment data comprise all aspects of the user environment that are not related to the users themselves, such as the hardware, software or network characteristics.

Environment data comprises all aspects of the user environment that are not related to the users themselves.

Simple adaptation to the device usually involves selecting the type of material and media (i.e. image vs. audio) to present the content. I.e. in the emerging ubiquitous computing field, each student interacts with many embedded devices, which makes recommendable to complement visual-only interaction with other modes of communication. Moreover, the popularity of mobile phones makes it common, for at least one of the devices involved in the learning process, to have audio input/output capabilities. I.e. HIPS project (Oppermann & Specht, 1998) uses audio explanations to support nomadic activities (Alexander, 2004) adapted to the individual visitor of a museum. Nomadic adaptive systems require both a user model where the information according to user needs, knowledge and preference is evaluated and a usage model where the information about the user environment and client hardware and software is held up-to-date.

Usage model holds information about the user environment and client hardware and software.

What can be adapted in adaptive hypermedia

In adaptive hypermedia, the adaptation space is quite limited: there are not so many features which can be altered. Hypermedia consists of a set of nodes or hyperdocuments (web pages) connected by links. Each page contains some local information and a number of links to related pages. What can be adapted in adaptive hypermedia are the content of regular pages (content-level adaptation) and the links from regular pages, index pages, and maps (link-level adaptation) (Brusilovsky, 1996).

From the previous classification, we can distinguish between two different classes of hypermedia adaptation: adaptive presentation and adaptive navigation support, which are related with content-level and link-level adaptation respectively.

Adaptive presentation adapts the content of a web page.

In adaptive presentation the main idea consists of adapting the content of the web page which is accessed by a particular user to his/her learning style, current level of knowledge, goals, etc. The content of a regular web page in hypermedia systems, may not only contain text (as in classic hypertext systems), but also a set of various multimedia items. This characteristic provides us of a new classification into adaptive text presentations and adaptive multimedia presentations.

Adaptive navigation helps users to find their paths in hyperspace.

Adaptive navigation support is concern with helping users to find their paths in hyperspace by adapting the way of presenting links to goals, knowledge, and other characteristics of an individual user. The elements susceptible of being adapted are:

- Local non-contextual links
- Contextual links or “real hypertext” links
- Links from index and content pages
- Links on local maps and links on global hyperspace maps

This classification has been widely referenced and used in the study of techniques and methods to enable presentation and navigation adaptation.

Methods and techniques of adaptive hypermedia

Adaptation methods and techniques describe how to adapt information in adaptive hypermedia systems.

Adaptation methods and techniques describe how to adapt information in adaptive hypermedia systems. It is possible to establish a distinction between different approaches for performing the adaptation, and thus, adaptation methods and techniques are often used to classify adaptive hypermedia systems.

Adaptation techniques refer to methods of providing adaptation in existing AH systems. These techniques are a part of the implementation level of an AH system. Each technique can be characterised by a specific kind of knowledge representation and by a specific adaptation algorithm.

Adaptation methods are defined as generalisations of existing adaptation techniques. Each method is based on a clear adaptation idea which can be presented at the conceptual level. The same conceptual method can be implemented by different techniques. At the same time, some techniques are used to implement several methods using the same knowledge representation (Brusilovsky, 2001).

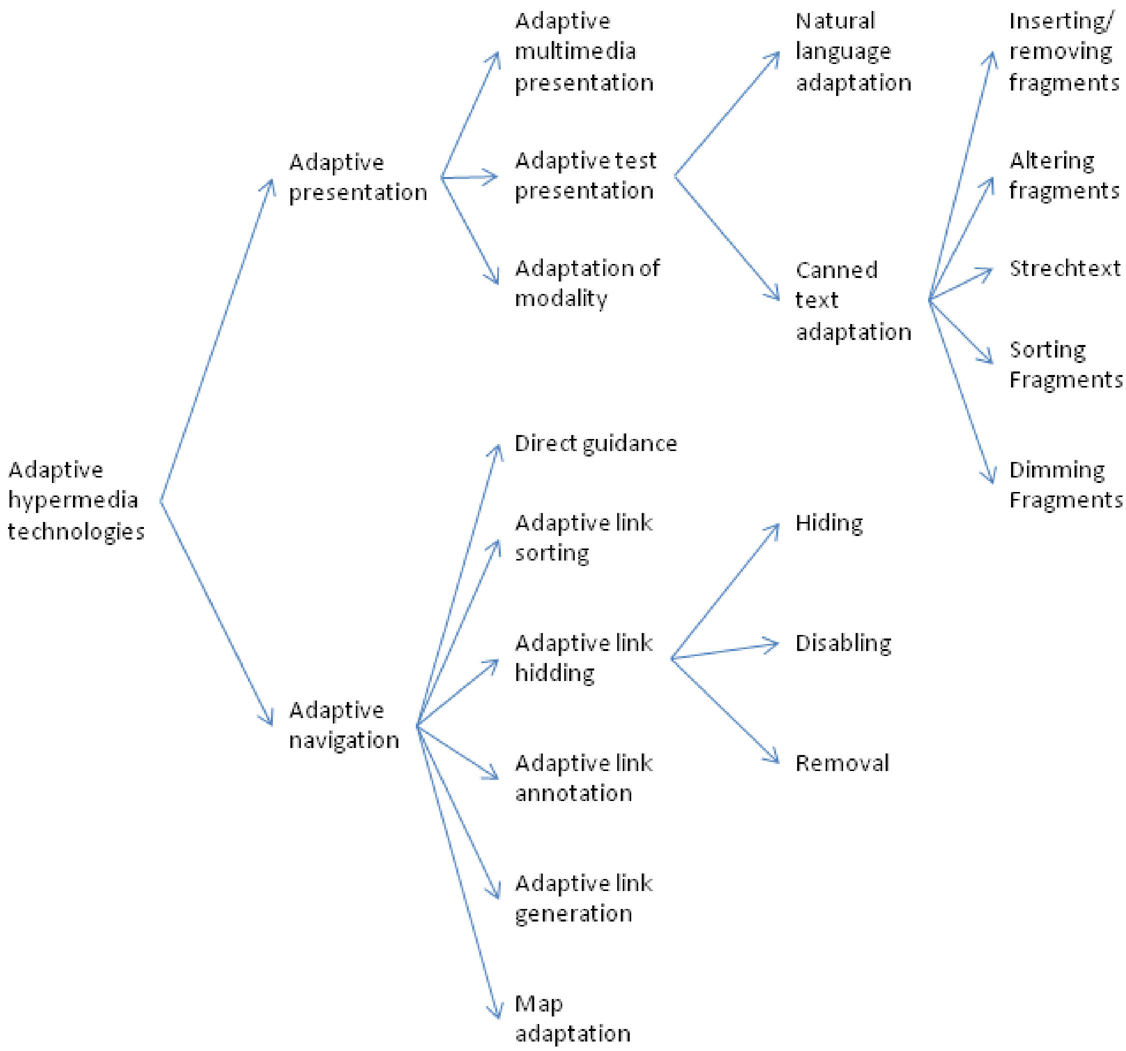


Illustration 20: The updated taxonomy of adaptive hypermedia technologies

Brusilovsky's taxonomy originally provided a mechanism for classifying the various adaptive hypermedia systems at the time. Since then, more systems have been developed, some of which fit in to the existing taxonomy, and others that have forced extensions to the taxonomy making it more lengthy and complex. I.e. Bailey et al. (Bailey et al., 2002) propose a complementary taxonomy with significant changes and a focus on media and natural language adaptation. Sequencing methods can be used to join information fragments together and conserve the progression of the narrative flow.

The association of adaptation methods with adaptation techniques can provide practitioners with useful insight information on how to perform adaptation and determine categories of classification of adaptive hypermedia systems:

Methods and techniques of adaptive presentation					
Technique **Method**	Stretchtext	Conditional Text	Fragment Variants	Page Variants	Frame-based technique
Additional Explanations	X	X			X
Prerequisite Explanations	X	X			X
Comparative Explanations	X	X			X
Explanations Variants		X	X	X	X
Sorting					X

Table 1: Methods and techniques of adaptive presentation

Methods and techniques of adaptive navigation							
Technique **Method**	Direct guidance	Sorting of links	Hiding of links	Annotation of links	Generation	Map Adaptation	Passive Navigation
Additional Explanations	X	X			X		X
Prerequisite Explanations	X	X	X	X	X	X	X
Comparative Explanations			X	X	X	X	
Explanations Variants		X	X	X	X	X	

Table 2: Methods and techniques of adaptive navigation

Adaptation methods and techniques are a common ground for researchers and developers of educational adaptive hypermedia. More research is being undertaken and many examples on how hypermedia systems can adapt to the users and their environments can be found in the scientific literature on this subject.

Conclusions

Research in adaptive educational hypermedia suggests the importance of considering the various characteristics of users and devices for the design of learning processes. Adaptive Hypermedia provides users with new means of interacting with information by adapting the presentation and navigation of traditional web-based hypermedia systems accordingly to user and environmental data. This involves providing not only relevant information, but presenting it in an appealing format as well. The methods and techniques of adaptive hypermedia can also be applied in order to create audio experiences for learners. Adapting both the interface and content format accordingly to user, usage and environmental factors, allows hypermedia systems to address users' specific needs and learning styles, and may optimise the overall performance and operational effectiveness of audio-driven applications.

Adaptive Hypermedia provides users with new means of interacting with information.

VI

Case Studies

Previous chapters have focused on reviewing the evolution, current status and future perspectives in the various fields concerned. This chapter presents case studies, conducted across a period of two years, which are used to gain insight into the best way to design and develop a voice interactive classroom. This section is aimed to help to understand research limitations and opportunities for the study, and properly capture the requirements of the solution.

Introduction

Research studies help to understand and clarify the boundaries and conditions in the development of voice interactive learning.

Previous chapter describes the evolution and level of development in the various fields of knowledge and technology associated with this research. This chapter presents a series of research studies, conducted between years 2007 and 2009, which help to understand and clarify the boundaries and conditions in the development of voice interactive learning applications. From both an understanding-oriented and action-oriented perspectives, this section describes the evolution of the process for setting up, designing, developing, documenting and evaluating prototypes.

The subjects in this study cannot be understood without the influence of my previous professional and research experience. I shall therefore attempt to summarise them briefly before further describing the evolution of the prototypes.

Illustration 21: III Virtual Spanish American Congress of Pathological Anatomy

In September 1999 I joined the research team at the Image Processing Service (University of Oviedo) where I was invited to participate in my first e-learning research projects, the educational system "Welcoming Environment for Lifelong Learning in Pathology" and the "III Virtual Spanish American Congress of Pathological Anatomy". The later determined also the orientation of my final year project entitled "Virtual Congress Management System" from which I was awarded with a mention of merit in June 2000.

Besides meeting the common requirements for an e-learning platform (institutional information, user and content management, communication facilities, etc.) and the various novel and specific requirements for allowing the celebration of an on-line congress (review and publishing process, scheduled events, etc.), we introduced a rest area with shoulder events (news and touristic information, photography contest, etc) to create the sense of being present at the congress. The enthusiasm and magnificent team work provided an outstanding starting point for my research and professional activity.

Illustration 23: Nokia social networking

After this initial research, I moved into a professional environment in Barcelona where I particularly focused on mobile development. In this period I participated in the Internet-based social network project entitled "Nokia 3G Community", which made use of the new standards for mobile telecommunications to provide a social experience through the mobile phone. Besides analysing and designing the various social networking aspects, the early stage in Wireless Application Protocol (WAP) and Java 2 Platform Micro Edition (J2ME) development set an additional technical difficulty to the already challenging task of presenting visual information on devices with limited screen sizes and entering text with a reduced keyboard. It was also my first participation in a project entirely developed and documented using the English language.

The experience in Barcelona aroused my subsequent interest in researching how mobile communications and software architectures can be combined to provide more flexible access to Internet resources and enable new modes of interaction in e-learning.

With the aforementioned goal in mind, I completed a new version of the "Virtual Congress Management System" in the year 2004. This version combined mobile technologies and service orientation to provide mobile access to features already present in the original software application. During the first half of the same year, I also resumed my collaboration with the team at the Image Processing Service to set up and design a system for making medical guides available from mobile Palm[66] and Microsoft Pocket PC devices.

Illustration 24: Medical guides for PALM and Pocket PC

Later, I moved to London, where I was invited to join a telecommunications company in the role of software developer and voice over IP consultant. In London I undertook the design and development of IP Private Branch Exchange (IP-PBX) and IP telephone-based systems using free software exclusively. In addition to study Interactive Voice Response (IVR) processes in depth, I became more familiar with a large range of free-licensed software, including operating systems, system management and monitoring tools, web development environments and, of course, voice communication technologies.

In autumn 2006 I went back to Asturias to join the European doctoral program in computer science and began my academic career as a lecturer in computer science and telecommunications at the University of Oviedo. During the next years I combined my Ph.D. studies and lectures at the University of Oviedo with external consultancy for several companies in the United Kingdom involved in VoIP and IVR development.

During my first doctoral course, I had the opportunity to meet the members of the various research groups and get a better understanding of the process and overall conditions for my research at the University of Oviedo. In summer 2007, after discussing the option with academic colleagues, I decided to make use of all the previous background and experience for researching the use of audio interfaces and software architectures in e-learning. The following describes the initial case studies conducted between years 2007 and 2009.

[66] http://www.palm.com/

Personalised multi-channel learning using feed syndication

One of the subjects I was asked to lecture during my first academic year (2006-2007) was web programming, an optional subject which was given in the third course of the computer science degree. After analysing the lecturing method and results from previous academic years, and together with professors María del Puerto Paule and Manuel Antonio Martín, we took the decision of changing the method of the subject in order to improve the level of satisfaction, develop professional competences and improve the level of analysis (Bloom et al., 1956) of the students.

Considering that an active method was more suitable for our purpose, and following previous successful experiences from other universities (Bridges & Hallinger, 1997), we decided to perform an adaptation of Problem-Based Learning (PBL) for programming subjects. In 1982 Barrows, one of the fathers of the Problem-Based Learning, defined it as "a learning method based on the principle of using problems as starting point for the acquisition and integration of new knowledge". In PBL, teachers are responsible for providing the problems and act as facilitators during the learning process, but the knowledge is actively constructed by the student (Ben-Ari, 1998).

In the second term of the academic year 2006-2007, we asked the students of web programming to form teams and develop a web emulator for a voice-based application. The introduction of freedom for the election of the problem and techniques to solve it, generated a large number of ideas during the laboratory classes. I.e. students developed small prototypes for simulating a voice-enabled automated teller machine or a voice-interactive telephony response system for ordering pizzas. One of the most technologically challenging ideas consisted in developing a voice-based RSS reader named Feedo. Isaac Rubio, one of the members of the student team who proposed Feedo, decided to continue this line of experimentation for his final year project.

Originally developed by Netscape in 1999, RSS (which can stand for RDF Site Summary, Rich Site Summary, or Really Simple Syndication) is an XML-based format that allows web developers to describe and syndicate web site content (Harmmersley, 2003). Since its creation, RSS has been used on the Internet as a means to retrieve information from various web sites, not linked to one another, to be read in one specific application the user can easily access. Using pull technology, the end user no longer needs to go through all the relevant web sites in order to obtain the latest updates or interesting information. By subscribing to feeds available on the web, information can be selected according to personal preferences and can be delivered at the user's convenience.

The main reasons for this syndication feature to arise were the need to save time when looking for information on the web and the possibility to give users more control over the information retrieval process, personalising the information.

However, as in the case of other web-based applications, feed readers have been made to communicate with users by visual means, using text-based content and showing that new updates are available by putting the feed title in bold or using another typographical style option which visually stands out. This limits the audience which can be reached by RSS readers as text does not fit well in all situations.

Visual communication limits the audience who can be reached by RSS readers.

Coming to realise that feed readers were not made to communicate to users who required using audio while accessing material, we started to work on our first prototype with the aim to enable audio users to have equal access to RSS. Feedo was designed to communicate feeds by using both visual means and speech interaction, and initially created and used in the context of our university for educational purposes.

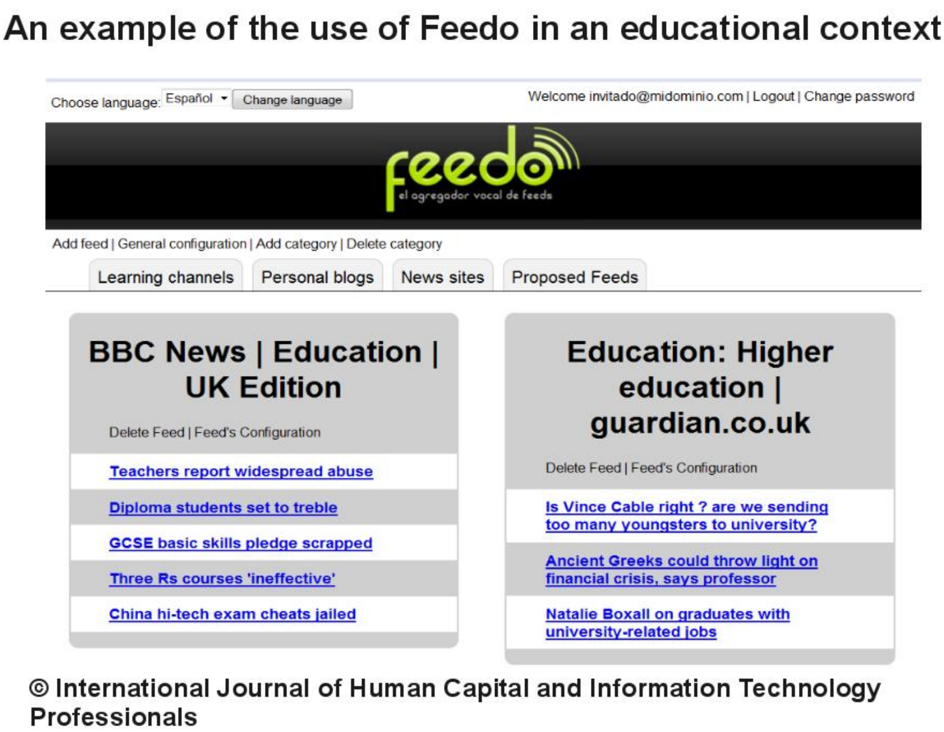


Illustration 22: An example of the use of Feedo in an educational setting

Feedo is a personalised web and voice-based feed aggregator.

Feedo is a personalised web and voice-based feed aggregator which allows teachers and students to classify their feeds into categories and read or listen to them using text or voice. Feedo has three main features, user personalisation, feed retrieval and a multi-channel adaptation.

Feedo allows users to create their own personal learning environment (PLE) (van Harmelen, 2006) based on categories and resources of their choice. Default categories can be managed using an administrative profile and they can include items such as 'Learning Channels', 'Personal Blogs' or 'News Sites'. Teachers and students can add and delete categories and feeds to create an educational environment of their preference. Both categories and resources are database persistent. Resources are stored using the specific URL for the desired source and updated to keep information about already visited feeds.

Any information update sent out through RSS syndication can be automatically retrieved by a feed reader. Users can then use the reader to check out the title and a short summary of the information and, in the case the user is interested, use the link to go to the original source and read the complete information. RSS feed retrieval requires a software tool called RSS parser, which performs the operation of checking updates from a given URL and download a determined number of RSS items. The current version of Feedo uses Magpie PHP RSS[67].

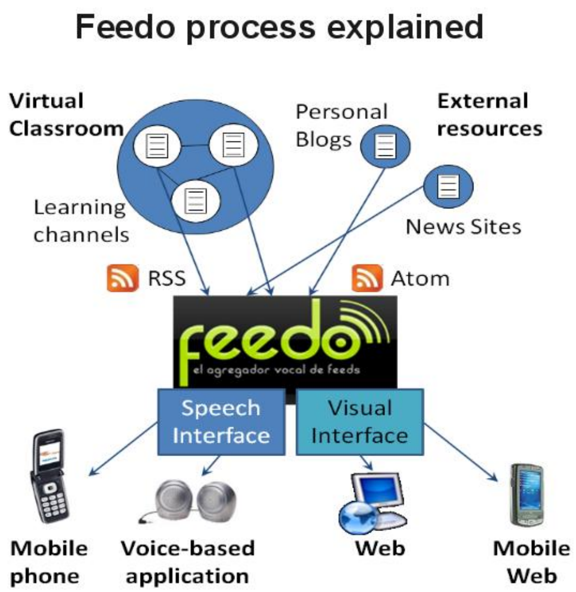


Illustration 23: Feedo process explained

[67]http://magpierss.sourceforge.net/

Feedo complements the standard behaviour of a feed reader with the possibility to listen to the information using an audio device, such as a voice over IP terminal or a traditional telephone, enabling audio access to the information offered through RSS syndication.

Multi-channel adaptation, consisting in enabling web and voice access, was the most challenging task of the project. Feedo considered three approximations for the solution, the use of the architectural software pattern Model-View-Controller (MVC) (Reenskaug, 1979), XSL transformations[68] and web services[69].

The use of the pattern Model-View-Controller isolates program logic from the user interface, giving a clear separation between data manipulation and the application's appearance. This characteristic was addressed with the support of Symfony Web PHP Framework[70], which simplifies the use of an MVC approach in software development.

Since feeds are provided in XML format with a well-defined structure, it is possible to apply XSL transformations to obtain appearances adapted to the access device. For example, given a resource in RSS format and using XSL transformations, we can obtain an HTML version of the information which can be read by web browsers or a VoiceXML version which can be used by voice-based clients. However, this approach has not yet been put in practice.

The state-of-the-art study influenced the last stage of the development when Feedo's functionality was mapped into web services to raise the level of interoperability of the solution. The web service API was written in REST, and used from a voice application written using Asterisk control language for dialogues and call flow description.

Feedo's functionality was mapped into web services to raise the level of interoperability.

By nature, Symfony's framework allows developers to create REST services quite easily, thanks to an architecture which associates web services with URLs. Symfony's architecture also helped to edit, internationalise and refactor (Fowler et al., 1999) the application. PhpEdit[71] and PhpDocumentor[72] were used to edit and auto-document Feedo's source code. The project used Metrica 3, a Spanish methodology for information systems. Hence, deliverables included Metrica 3 documents, as well as user and installation manuals. Feedo was hosted on Isolde[73], a Linux machine used by our research group which includes useful web and VoIP resources.

[68]http://www.w3.org/TR/xslt/
[69]http://www.w3.org/2002/ws/
[70]http://www.symfonyproject.org/
[71]http://www.phpedit.com/
[72]http://www.phpdoc.org/
[73]http://www.pulso.uniovi.es/

The testing was performed by placing VoIP calls using X-lite[74] and Linphone[75] softphones. Feedo followed a test-driven development process aimed by Symfony unit tests.

Notwithstanding the mentioned benefits of using a software development framework such as Symfony, we ought to acknowledge that understanding and mastering the different framework tools and components proved to have a steep learning curve.

Feedo reduced some of the RSS limitations by making a sensible use of technology.

We consider that Feedo reduced some of RSS limitations by making a sensible use of technology.

Feedo offers the possibility to select default sources of information. I.e. teachers can make use of this feature to offer students those sources which are considered most adequate for the assignment, so that students have them ready for consultation. The classification of information into categories also allows users to access a larger quantity of information without creating a cognitive overload.

In the area of accessibility, Feedo offers a significant improvement when compared to contemporary RSS readers. Although the idea of RSS being read by using voice emerged almost immediately at the same time as RSS itself, it is not a feature integrated in RSS readers. Feedo improves the accessibility of feeds by using speech interaction technologies. The incorporation of the use of voice means the reader can be adapted to situations in which voice needs to be used as a means of communication.

Web-based visual sources make it very difficult to separate information from its representation.

Nevertheless, using speech interaction also points directly to some of the limitations RSS has. The original source of the information is most often web-based and visually driven, making it very difficult to separate information from its representation. Feedo makes use of RSS technology to access the title and the summary through web and speech interaction, but it doesn't make possible to consult the full article.

Specific platform languages break the portability of voice dialogues to other platforms.

Feedo also raised one of the issues commonly associated with voice interaction, the platform and language for creating voice user interfaces. As we have indicated, the voice dialogues developed for Feedo were written using Asterisk language which creates a dependency with the platform and breaks the portability of dialogues to other platforms. Our research group addressed this problem in a later study on the use of VoiceXML.

[74]http://www.counterpath.com/x-lite.html
[75]http://www.linphone.org/

Using W3C recommendations of VoiceXML to access web-based learning

In summer 2008, a few months prior to the conclusion of our first research study for this dissertation and derived from its results, we decided to set up a second study on the use of W3C recommendations of VoiceXML as means of overcoming the limitations of using unusual or non-standard languages to describe voice dialogues.

VoiceXML helps overcoming the limitations of using unusual or non-standard languages to describe voice dialogues.

The main goal of this study was to learn more and analyse the convenience of using a standard XML-based language for the description of audio interfaces, preserving the portability of the dialogues across different voice platforms. Student Moisés Riestra undertook this research and performed a series of experiments using different voice platforms and web-based e-learning applications.

The project required to design, develop, document and evaluate five prototypes of increasing complexity. By organising the study in this manner, the necessary expertise to create a voice interactive learning application is developed along the time path. The following describes briefly the five prototypes.

Prototype 1: Introduction to voice interactive learning

The goal pursued by the first prototype was to introduce the development of interactive voice response applications in the academic realm. The development consisted of implementing a very limited functionality to allow a student to consult information using a telephone-based system. By making a voice over IP call, the student can check his/her grades for a class work. This prototype did not require using any actual solution or factual data. Authentication was based on a fictitious ID numeric code and the grades were manually introduced in a database. Asterisk was the choice for the voice platform and the dialogues were expressed following the Asterisk control language.

Prototype 2: Introduction to VoiceXML

On completion of the first prototype, and following the same idea, we set up the design of a first VoiceXML dialogue system. The prototype had to reproduce both the speech interface and the actions performed in the first prototype. For this experiment we used the same voice platform, Asterisk, but in this case we enabled the use of XML-based dialogues by using OpenVXI[76], a portable open source library that interprets the VoiceXML dialog markup language, and Voiceglue[77], a software tool that allows the integration of Asterisk and OpenVXI, and hence creates a VoiceXML interpreter based on Asterisk.

Prototype 3: Alternative voice platform

The dialogues in the second prototype were expressed in VoiceXML, a portable markup language for the definition of voice interfaces. In order to demonstrate the portability of the dialogues, we required to use an alternative voice platform for the third experiment. After considering different alternatives, JvoiceXML, a free VoiceXML interpreter for Java, was selected to take the place of the previous Asterisk, OpenVXI and Voiceglue combination.

VoiceXML dialogues could not be entirely ported due to differences in the support for the VoiceXML 2.1

Besides differences in the development and deployment environment as well as in the programming language for coding the actions, we found that the dialogues could not be entirely ported due to differences in the support given for the VoiceXML version 2.1. This situation forced us to replace also some of the dialogue instructions.

Prototype 4: Multi-channel e-learning platform

Moodle was chosen to enable multi-channel access in a real e-learning setting.

The three previous prototypes didn't make use of any e-learning feature or factual data as they were intended for training purposes. In this experience we proposed the use of Moodle to use voice interaction in a real e-learning setting. In this experiment we enabled multi-channel access to Moodle, giving the student the option to consult his/her grades using the traditional access through a web browser or alternatively using a VoIP audio device. Any change in the grades was immediately reflected both in the visual and audio interfaces.

Besides having to study the insides of the chosen e-learning platform in order to gain audio access to it, this experiment also arouse difficulties in the association of traditional e-learning features with the corresponding voice interactive version. A simple example can be found in considering the aural authentication process, which

[76]http://sourceforge.net/projects/openvxi/
[77]http://www.voiceglue.org/

may want to enable input from a DTMF (touch-tone signals via a telephone) source as to avoid the use of an unsafe speak aloud process. This implies the use of numeric values for both the username and password fields, a representation that may be inapplicable or inappropriate for the corresponding fields in the e-learning platform.

This experiment became the core of the present study and a reference for the subsequent research, as it set a precedent in the creation of voice dialogues for e-learning platforms using VoiceXML. This process was gradually refined over the course of development.

Prototype 5: VoiceXML Feedo

To reaffirm the technique developed, we set up one last prototype with the aim of designing and implementing VoiceXML dialogues to provide audio access to a different e-learning system. Making use of our previous experience, the software application chosen for this development was the educational feed reader Feedo. Feedo, which was already a multi-channel application, was developed using Asterisk's control language, which broke the portability of the voice interface. We embraced this experiment as an opportunity to upgrade Feedo and give it a new portable interface.

In order to try different components, we introduced the use of i6Net VXI* VoiceXML browser[78], which proved easy to install and reliable in use. VXI* can be used altogether with Asterisk and various Text to Speech and Automatic Speech Recognition solutions, including a few which provide support for the use of the Spanish language. This characteristic makes it highly convenient for our developments and an important factor to choose it in following studies.

It is worth to mention that, in addition to the use of new audio technologies, this study adopted several other changes in the method for creating and evaluating the prototypes.

This study adopted several changes in the method for creating and evaluating the prototypes.

Dialogues were designed and discussed using text-based material. In the absence of an integrated software framework, a combination of several software tools supported the development of the prototypes, including Gedit[79] for editing, PhpDocumentor for auto-generated documentation and SimpleTest[80] for unit testing (Myers, 1979). Following a suitable object-oriented methodology, deliverables included components descriptions along with unified modelling language (UML) diagrams and installation, execution, user and programmer manuals.

[78]http://www.i6net.com/products/vxi/
[79]http://projects.gnome.org/gedit/
[80]http://www.simpletest.org/

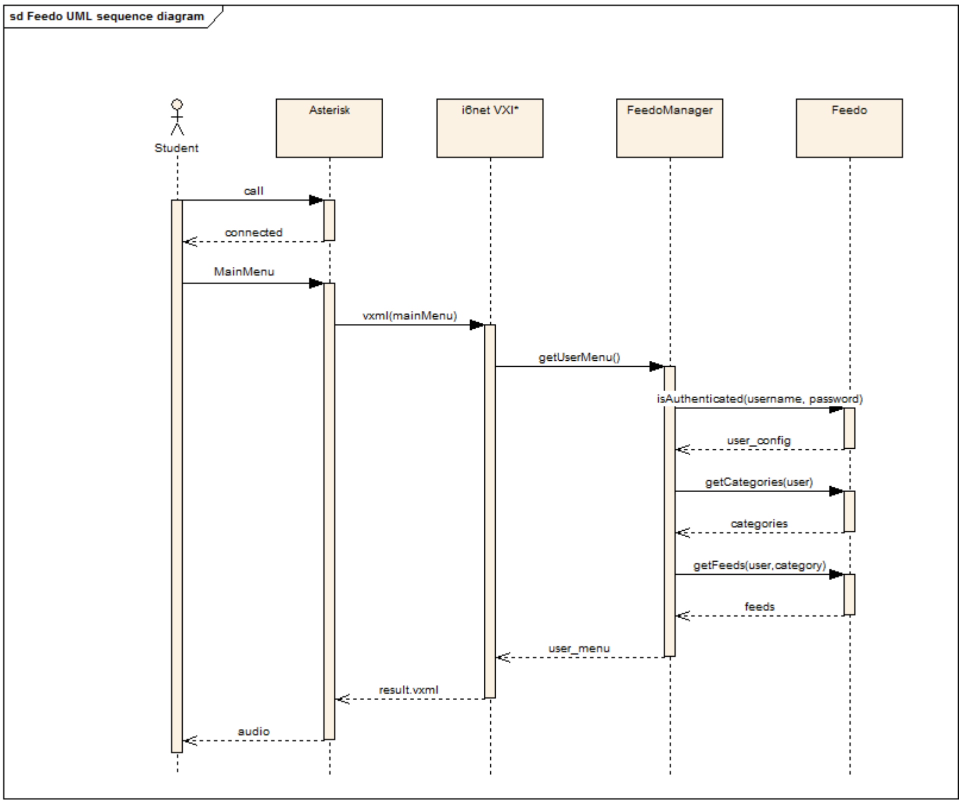


Illustration 24: Feedo UML sequence diagram

Questionnaires were used to assess qualitative parameters associated with usability and user satisfaction.

One of the most significant additions was the questionnaires to assess qualitative parameters associated with user satisfaction, which was included in all the prototypes and conducted with students at the completion of the developments as means to complement the more technical unit, integration and system testing.

The following list resumes the techniques used for measuring the usability and user satisfaction level for the prototypes:

- Design of structured audio guided tasks in concordance with the prototypes
- Write a general single response questionnaire to evaluate the proficiency in the use of computer and audio applications.
- Write a more specific questionnaire to measure the level of usability and user satisfaction during the assigned task.
- Annotate the time to perform the assigned task and observed errors.
- Follow the "think aloud" method during testing and write down users' reactions and comments.
- Include an extra comment field for both users and evaluators.

<table>
<tr><th colspan="6">Questionnaire to measure the level of usability and user satisfaction</th></tr>
<tr><td colspan="2">Ease of use</td><td>Always</td><td>Frequently</td><td>Occasionally</td><td>Never</td></tr>
<tr><td colspan="2">¿Is the application easy to use?</td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="2">With the given audio messages, Do you know what you have to do?</td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="2">Functionality</td><td>Always</td><td>Frequently</td><td>Occasionally</td><td>Never</td></tr>
<tr><td colspan="2">¿Does each task work in the way you expect?</td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="2">Is time elapsed too long?</td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="6">Interface Quality - Voice aspects</td></tr>
<tr><td></td><td>Too high pitched</td><td>High pitched</td><td>Regular</td><td>Low pitched</td><td>Too low pitched</td></tr>
<tr><td>The tone of messages is</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td>Too high</td><td>high</td><td>Regular</td><td>Low</td><td>Too low</td></tr>
<tr><td>The volume is</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td colspan="2">Quick</td><td>Regular</td><td colspan="2">Slow</td></tr>
<tr><td>The vocalization is</td><td colspan="2"></td><td></td><td colspan="2"></td></tr>
<tr><td colspan="6">Dialogues design</td></tr>
<tr><td colspan="3"></td><td>Yes</td><td>No</td><td>Sometimes</td></tr>
<tr><td colspan="3">Can you follow the dialogues easily?</td><td></td><td></td><td></td></tr>
<tr><td colspan="3">Do you find the dialogues design appropriate?</td><td></td><td></td><td></td></tr>
<tr><td colspan="3">Do you consider the dialogues are well structured?</td><td></td><td></td><td></td></tr>
<tr><td colspan="6">Observations</td></tr>
<tr><td colspan="6">User comments</td></tr>
</table>

Table 3: Questionnaire to measure the level of usability and user satisfaction

In sum, the five prototypes developed for this final year project introduced a number of methodological and practical refinements with respect to our previous study. By following this approach we complemented our research work and obtained the necessary experience and understanding of the development of voice interactive learning solutions. This experience contributed to improve the understanding of research limitations and opportunities, providing the grounding needed for making a transition to a more innovative approach.

This experience provided the grounding needed for making a transition to a more innovative approach.

Conclusions

This chapter presents a series of research studies, conducted between years 2007 and 2009, which help to understand and clarify the limits and opportunities for the design and development of voice interactive learning applications.

The research, which was influenced by previous experience, was addressed by two final year projects at the University of Oviedo.

In the first study Isaac Rubio developed Feedo, a personalised web and voice-based feed aggregator which allows teachers and students to classify their feeds into categories and read or listen to them using text or voice.

Feedo overcomes some of the limitations of current RSS but it was constrained from using a visually-based source of information and a non-standard language for describing audio interfaces.

In the second study Moisés Riestra developed a series of prototypes to progressively approach the development of voice interactive learning applications using VoiceXML.

VoiceXML is supported by multiple voice solutions, but it is not infallible and the success of the portability depends on the compliance of those solutions with specifications.

The second study also introduced a number of methodological and practical refinements with respect to Feedo.

These research experiences allowed us to gain insight into the conditions and best way to design and develop a voice interactive classroom.

These research experiences allowed our team to gain insight into the conditions and best way to design and develop a voice interactive classroom, and helped us to capture the requirements for the solution.

VII

Solution Requirements

Previous chapters present the studies conducted across a period of two years which helped us to understand the limits and possibilities of our research, and permitted us to state the requirements for the solution, which are presented within this chapter.

From a conceptual point of view, our previous research and professional experience suggested the convenience of enabling audio access to e-learning, but one of the problems appeared to be the lack of a more precise description of the e-learning scenarios which can benefit from the usage of voice interaction. This matter, which was particularly arisen during conversations with Dr. Raul Izquierdo in Oviedo and Dr. Marcus Specht in Heerlen, was regarded as one of the goals for the dissertation. The enumeration of the scenarios (Chapter VIII) was finally addressed during a train journey with Remko van Dort in the Netherlands.

The rest of the goals had a more technical orientation; they were identified during the initial stages of this thesis (Chapter II, III, IV and V) and in the course of development of the first case studies (Chapter VI), and addressed in a second stage of prototyping (Chapter IX).

Establish the nature and characteristics of the problem and its solution required a comprehensive and thoughtful study.

Establish the nature and characteristics of the problem required a comprehensive and thoughtful study. Furthermore, as it is described in the state-of-the-art, previous attempts to enable voice access to existing e-learning platforms were not successful. Thus, it was required the adoption of a research method to achieve a more accurate description of the problem and its solution.

We decided to combine an exploratory and constructive research aimed to lead us to better describe the problem and finally be able to propose and adopt a solution to fill the lack of multi-channel access to web-based learning platforms.

The synthesis of the various issues concerning the problem, as well as the enumeration and description of solution requirements, were addressed iteratively and progressively supported by research findings and results obtained from case studies.

Information on the use of technologies, and their perceived performance during the research, can give practitioners a first insight into technical aspects of the implementation of voice-enabled learning solutions.

Although our previous experience provided a starting point in the adoption of techniques for the creation of audio-driven systems, the nature of this dissertation required to explore more in depth a number of technologies and approaches for the creation of voice interactive classrooms.

Complementing existing e-learning visual interfaces with audio interactions requires the use of e-learning platforms, voice interpreters and advanced voice modules, as well as other side technologies such as communication and transmission systems.

The exploration of the aforementioned technologies implies to select and discard material. The process of prototyping was influenced by our previous knowledge and expertise, and relied on the state-of-the-art in web-based e-learning platforms, which indicates a strong preference in this field for free licensed solutions.

This situation led us to decide to give priority to the adoption of free software developments with the purpose of guarantee the compatibility of our proposals with current solutions. However, it is important to remark that, in our opinion, this should not exclude in any case the consideration and use of more commercial approaches, especially in cases where free applications do not provide a required feature or perform lower than a commercial alternative.

Our proposed procedure was carried out as a series of iterative use cases which help to analyse problems and the best means of solving them. A number of theoretical and practical concerns arose during this analysis.

One of the main problems consists on the isolation of the first ad-hoc solutions with respect to the systems in use at academic and educational institutions, thereby making it difficult to adopt a more generic approach to allow the integration and enable interoperability among different voice-enabled learning features and e-learning platform components.

Another of the main concerns of the study is the adaptation of visual interfaces into voice dialogues in such a way that, despite their different nature, we can consider them equivalent to one another.

From the outline of our first voice interactive applications, to the voice-based feed aggregator Feedo and the first prototypes using Moodle, the evolution of the research use cases provided valuable clues and directions towards the definition of the solution requirements.

The comprehensive review of scientific literature, along with the results from the research activities conducted during a period of two years, finally led us towards a description of the requirements of a software architecture intended to enable multi-channel access to web-based learning platforms. A second stage in the development of research use cases, this time carried out using our proposed software architecture, was required to demonstrate the challenging task of developing integrated and interoperable voice interactive learning applications.

Our research method led us towards a description of the solution requirements.

Taking into account the previous observations, the following describes, in a more structured manner, the requirements for our solution.

Definition of a reference software architecture

This thesis is devoted to the definition of a reference software architecture to support building cross-platform and multi-channel applications for Internet-based learning systems.

After stressing the importance of the previous item as the main requisite to be achieved in order to provide a feasible and practical oriented solution, the rest of the requirements are described below.

Identify e-learning scenarios which can benefit from the usage of voice interaction

The convenience of creating a model to enable audio access to existing e-learning needs to find a motivation in a clear vision of what it means to learn. Thus, there is a requirement for describing the situation and circumstances in which e-learners can benefit from this research.

Propose the adoption of technologies to fill the lack of multi-channel access to web-based learning

Besides being the instrument used to create the prototypes, the different technologies which provide voice access to research use cases have to be discriminated and described to provide a preliminary and non-formal guide for practitioners on the implementation of voice-enabled learning solutions.

Devise, design and develop speech-enabled software applications for Internet-based learning

Understanding the nature of the problem proposed in this dissertation requires a previous pursuit of devising, designing and developing approaches to support the integration of voice interactive solutions in e-learning settings. Research use cases are proposed to specify and outline prototypes and, in some specific cases, develop a more complete solution to run-through the most practical aspects of the experimental procedure.

Enable multi-channel access to existing Internet-based e-learning platforms

The state-of-the-art describe a number of e-learning platforms and indicates the way academic and educational institutions adopt them for delivering course contents and learning activities at a distance. Integrate voice interactive learning into current solutions implies to enable a new access mode and adopt software mechanisms to link spoken interfaces with existing e-learning platforms features.

Enable the creation of voice interactive features compliant with existing e-learning and voice platforms

The proposed software architecture ought to be capable of enabling the creation of voice interactive features which can be linked seamlessly to a number of existing e-learning and voice platforms. This process is addressed by creating a service-oriented middleware supported by the use of e-learning specifications and recommendations by the World Wide Web Consortium.

Adapt the user interface of learning management systems for the audio communication channel

Audio interfaces should not result from a sequential parsing of visual content but from a process of adapting visual interfaces into more naturalistic audio dialogues. This process is aimed by the use of user design guidelines and methods and techniques of adaptive educational hypermedia.

Design and develop speech-enabled software applications following the proposed solution

Speech-enabled software applications, covering different features present in current e-learning systems, are an expected outcome from the research studies. It is suggested to approach this matter by setting up, designing and developing a new series of research case studies considering the various features of the proposed solution. Additionally, the prototypes based on the proposed solution, and optionally some of the prototypes developed in an earlier stage, can be made available for the scientific community at the completion of the experimentation.

VIII

Proposed Solution: Voice Interactive Classroom

This chapter describes thoroughly the proposed solution to the problem of designing a software architecture to enable cross-platform multi-channel access to Internet-based learning. The architecture follows a middleware approach which encourages adherence to web-based e-learning frameworks and recommendations for developing interactive voice response applications, as well as user design guidelines. The solution hereby presented appears on a special issue on "Middleware Trends on Network Applications" published by Elsevier, and it is already being used in a number of academic projects.

This chapter extends the information provided by the mentioned article, adding further explanations for design decisions and discussing the alternatives in more detail. In addition, some extensions of this research which have not yet been published are also mentioned.

Voice Interactive Classroom

As it is stated in the previous chapter, web-based e-learning relies on learning management systems, such as Moodle, Blackboard and Sakai, and standards and specifications proposed by web-based e-learning frameworks like IMS Abstract Framework, The Open Knowledge Initiative and The e-Framework.

Learning management systems and e-learning frameworks have focused on visual communications.
.

Learning management systems and e-learning frameworks have been developed focusing on a web environment that was created for visual communication. Although they don't include modules or specific techniques for enabling speech access to Internet-based learning, in our opinion, voice interaction has added value in a world that is being progressively virtualised and subjected to new pervasive ways of accessing and manipulating information.

The growth and increasing acceptance of applications that utilise synthesised speech or voice recognition suggests that, for most on-line learning programs, the question is no longer deciding whether to adopt speech, but rather determining which applications are most suitable for speech and developing a speech strategy that most effectively complements and integrates with an overall e-learning strategy.

By using vocal interactions the on-line learning process can be better adapted to new e-learning scenarios.

Speech technologies open the possibility to combine vocal interactions with interactive e-learning tools, so that the on-line learning process can be better adapted to new e-learning scenarios. In particular, the following five scenarios make it either necessary or highly recommended to complement visual learning with audio interactions:

- Learning processes in which sound is the main source of information, as in the case of music and language learning (Patel, 2008).
- Students with perceptual auditory learning styles or those who prefer to learn by using both their visual and auditory senses (Barbe & Milone, 1981).
- Hands-busy, eyes-busy and mobility-required situations, e.g., in routine physical activities like walking and driving (Shneiderman, 2000).
- The device has speech input/output capabilities but lack of usable displays or keyboards (Lai, 2000).
- A disability does not allow a person to make use of a visual display or input/output devices, for instance when users have visual or motor impairments (Cohen and Oviatt, 1995).

However, we consider that current solutions to combine visual web-based e-learning platforms and voice interactions, including the use of assistive technologies such as screen readers, are neither convenient nor sufficient. Instead of performing appropriate visual to audio adaptations, assistive technologies provide only aural descriptions of the visual content, covering the multiple elements on a screen from left to right and from the top of the page to the bottom. Furthermore, the first attempts of embedding VoiceXML support into e-learning platforms, like MoodleVoice[81], failed due to the inefficiency of assistive technologies.

Provide only aural descriptions of the visual content is not adequate.

At this point we ought to remark that projects like MoodleVoice and our early prototype Feedo, which has been explained in Chapter VI, were ad-hoc solutions consisting of developing functionalities for a particular e-learning application, which could not be ported to others. This precludes the possibility of sharing resources and information among different e-learning systems.

Ad-hoc solutions preclude the possibility of sharing resources and information.

Our commitment to enable new e-learning scenarios the students can benefit from, along with the aforementioned situations, motivated us to design a software architecture that makes it possible to use simple, intuitive and naturalistic voice dialogues with functionalities already present in the various e-learning platforms. The following subsections describe our proposals and contributions in further depth.

Software Architectural Design

Traditional passive e-learning is being replaced by more dynamic and active e-learning. Active e-learning applies the broad range of technologies of the Internet to achieve pedagogic scenarios otherwise inaccessible to traditional forms of learning. Technologies themselves do not directly cause learning to occur, but can afford certain tasks that themselves may result in learning or give rise to certain learning benefits (Dalgarno & Lee, 2010). To support the diversity in technologies, software architectures for web-based e-learning platforms are evolving from monolithic to service-oriented systems. Steps towards service-oriented e-learning have been defined by IMS Abstract Framework, The Open Knowledge Initiative and The e-Framework. Future e-learning platforms will support federated exchange (information and control), various levels of interoperability (intra-domain and inter-domain) and service composition (orchestration and choreography) (Dagger et al., 2007).

Software architectures are evolving to support the diversity in e-learning technologies.

[81] http://docs.moodle.org/en/Development:Voice

In a middleware architecture, federated exchange, interoperability and service composition are balanced by a layer which connects software components. Voice Interactive Classroom adopts this architectural vision and proposes a service-oriented middleware to assemble a composite of voice modules and e-learning resources.

Owing to the fact that visual mode is already present in e-learning management systems, we have focused our attention on adding voice interactions to e-learning in order to provide auditory as well as visual access to web-based e-learning environments.

Voice Interactive is a layered architectural design where a middleware provides the support to assemble voice services upon a set of heterogeneous e-learning platforms.

Voice Interactive Classroom is a layered architectural design where a middleware provides the support to assemble voice services upon a set of heterogeneous e-learning platforms. The state-of-the-art study gave us the background for the definition and composition of the services. Instead of proposing a new specification, we encourage the adherence to one of the established e-learning frameworks. Particularly, we have chosen the Open Service Interface Definitions by The Open Knowledge Initiative to help dealing with the complexity of having different e-learning platforms.

OKI OSIDs are added on top of e-learning platforms to provide a contract-first design for learning features, linking learning services and e-learning platforms components. Accordingly to service-orientation principles, the architectural design of the Voice Interactive Classroom distinguish three main layers, presentation & navigation, application and system, each one with well defined responsibilities, as it is following explained.

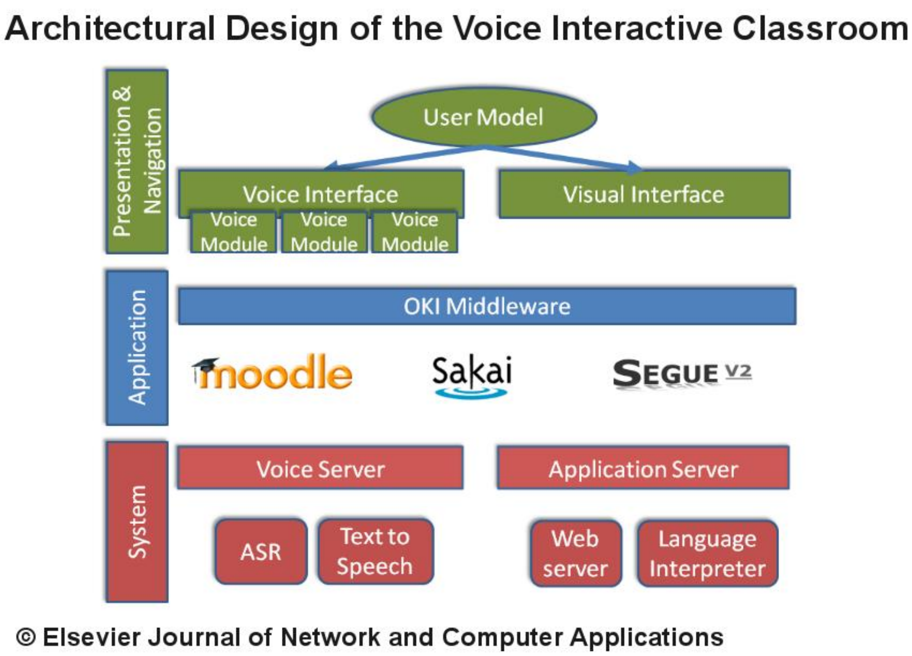


Illustration 25: Architectural design of the Voice Interactive Classroom

Presentation & navigation layer

This layer is responsible for determining the mode of presentation of the learning management system, visual or vocal, as well as adapting the presentation and navigation according to the user and usage models (Brusilovsky, 2001b) (Kobsa et al., 1999). In our proposal, the user model represents the perceptual learning style preference, visual or aural, and the access channel is one of the environment parameter in the usage model, in such a way that, when the user accesses the application from multiple devices such as a PC/laptop, a mobile device or and IP phone, the access mode parameter is inferred from the header of the transport protocol in the TCP/IP architecture. Subsequently, the system is capable of identifying the appropriate communication channel and adapting the presentation and navigation accordingly.

This layer adapts the presentation and navigation according to the user and usage models.

In case the communication channel is visual, the system renders a visual interface, which is the default interface of the learning management system. Otherwise, when the communication channel is verbal, it is necessary to perform adaptations both in the presentation and in the navigation of the audio interface.

The way the information is structured and organised differs from visual to verbal interactions. While visual interfaces are displayed on a screen where the size limits the amount of data presented and the time to read and react is unlimited, speech interfaces try to mimic naturalistic human dialogues, limiting the amount of information that can be retained as well as the response time.

In order to make an adequate conversion from the given LMS visual interfaces to audio dialogues and accommodate individual learning style differences, we propose to take into account user design guidelines, including the following three principles from The Universal Design for Learning:

- Perceptible information

  The design communicates necessary information effectively to the user, regardless of ambient conditions or the user's sensory abilities.

- Simple and intuitive use

  Use of the design is easy to understand, regardless of the user's experience, knowledge, language skills or current concentration level.

- Equitable use

  The design is useful and marketable to people with diverse abilities.

We aim to make visual and aural interactions equivalent to one another.

Even though visual and voice interfaces show differences in presentation and navigation, and -with reference to the UDL principles of “perceptible information”, “simple and intuitive use” and “equitable use”- we aim to create naturalistic dialogues and make visual and audio interactions equivalent to one another to provide the same functionality in a user-friendly manner, in our experience, this goal is not always easy to achieve.

While it is possible to use different formats to define the interface of the audio dialogues, the state-of-the-art study highlights the convenience of using W3C's recommendations of VoiceXML. VoiceXML is a XML-based representation which allows voice applications to be developed and deployed in an analogous way to (X)HTML. As in the case of (X)HTML and Web browsers, VoiceXML can be interpreted by multiple voice platforms, which ensures the portability of the voice modules.

In a visual interface, the presentation is generated from a set of (X)HTML pages (dynamic or static), which are linked by a navigation that can be carried out using a graph structure where the user select which route to take according to the visual options available. In contrast, in voice interaction, the presentation is made via voice dialogues, in our case represented in VoiceXML, and the navigation is determined by a dialogue flow, which also makes use of a graph structure. However, in this case, the routes are selected from the options offered in audio prompts. This implies that in order to create a naturalistic dialogue flow and avoid information overloads, both the amount of information presented and the number of navigation options in audio interfaces have to be restricted.

As it is stayed in the state-of-the-art, methods and techniques of adaptive hypermedia can be applied in order to create audio experiences for learners. Adapting both the interface and content format accordingly to user, usage and environmental factors, allows hypermedia systems to address users’ specific needs and learning styles, and may optimise the overall performance and operational effectiveness of audio-driven applications.

Adaptive hypermedia (AH) systems build models of the individual user, where the information according to user needs, knowledge and preference is evaluated, and of the usage, where the information about the user environment and client hardware and software is held up-to-date, and apply them for adaptation.

In our case, we propose the use of a minimum of two of the adaptive hypermedia techniques for adapting visual interactions into voice dialogues, the conditional inclusion of fragments and link removal.

In order to avoid information overloads, the technique conditional inclusion of fragments can be used to reduce the amount of information included in the spoken interface.

At the same time, link removal allows to restrict the number of navigation options available in voice dialogues and give them a more naturalistic and simplistic flow. Other techniques of adaptive hypermedia can also be considered, such as reordering information, providing explanation variants and sorting of links.

Once a first version of the dialogues is defined, a user-centred iterative design process allows evaluating and refining aural interactions to make them more user-friendly and better suit learners' requirements. This process is further described in next chapter.

Besides the user interface, the presentation and navigation layer has to provide the sequence of actions to perform when an option is chosen. These actions determine the interaction with the e-learning system as well as the information that is to be reading out on the next dialogue-step. In our architecture, actions are expressed as a sequence of OKI OSIDs services requests and responses, in such a way that the logic for the actions is the same independently from the e-learning platform running underneath. This approach assures the independence from any particular e-learning solution and the re-usability of the components designed for the presentation and navigation layer.

Tips for the presentation and navigation layer
1. Design dialogues Use a neutral XML-based representation Follow user design guidelines Adapt presentation & navigation using AH: • conditional inclusion of fragments • link removal • consider other techniques Refine dialogues using user-centred approach
2. Define actions Use neutral OSID requests & responses

Table 4: Tips for the presentation and navigation layer

Application layer

This layer allows voice services to interoperate with the various e-learning platforms.

In our proposed architecture, the application layer is placed between voice dialogues and system resources, and it is responsible for allowing voice services to interoperate with the various e-learning platforms. This middleware layer includes the e-learning platforms as well as the interfaces and inner implementation of a service-oriented middleware based on OKI Open Service Interfaces. The middleware raises the level of abstraction at which e-learning services are defined. This scheme makes it possible to use any underlying e-learning platform and achieve one of the goals of this dissertation, make the architecture cross-platform compliant.

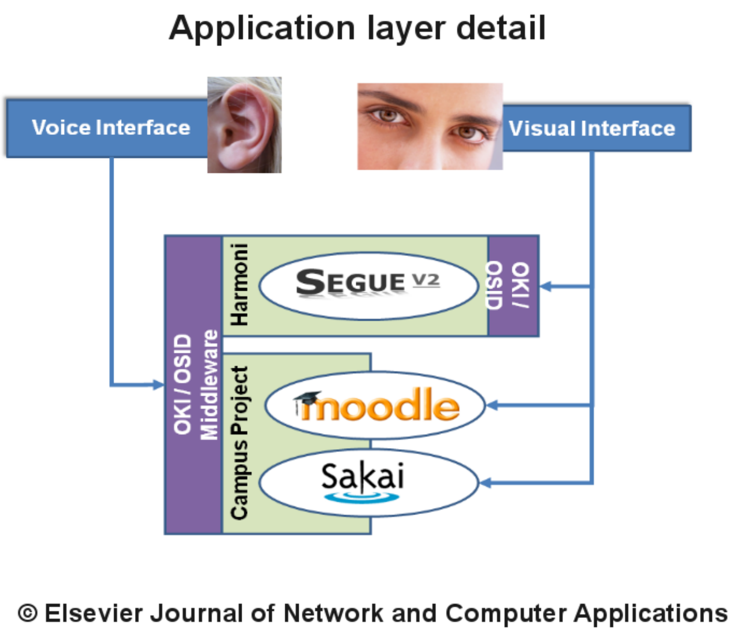


Illustration 29: Application layer detail

We propose to firstly adopt the Voice Interactive Classroom architecture to build voice interactive learning solutions upon Moodle, Sakai and Segue. The reasons we have found for prioritise projects on these e-learning platforms are the following:

- Moodle

Although this e-learning platform is not compliant with OKI OSIDs, it is currently the one in use for the virtual classroom at the University of Oviedo. It is also the preferred e-learning platform in Spanish universities.

- Sakai

  It is a full service-oriented web-based e-learning platform built partially on the principles of IMS Abstract Framework and OKI.

- Segue

  This LMS by the University of Middlebury is fully compliant with OKI OSID interoperability standards.

The process to create the middleware based on OKI OSIDs requires firstly understanding the specifications in order to identify the service interfaces needed for a particular functionality. It is important to acknowledge that although OSIDs specifications cover major sections and common functionalities in e-learning platforms, each solution has its own singularities and hence, depending on the particular task to be achieved, few services might be identified, if any. In such cases, it is recommendable to use as many standard services as possible and keep the OSID programming style when defining new services is unavoidable.

It should also be noted that, besides the merit of understanding an e-learning framework and identify the appropriate services, the completion of this layer requires to find or perform adequate implementations of every identified service for each selected e-learning platform(s). Notwithstanding the clear advantage of reusing previous implementations, we ought to take into account the availability of previous solutions, as well as the various e-learning platforms and programming languages in which they are implemented, and in some occasions, even differences among versions of the same platform. These factors can radically change the time-consumption and effort for the accomplishment of the application layer.

Tips for the application layer
1. Select e-learning platform(s) Prioritise projects and tasks Consider existing OKI solutions
2. Identify services Learn how to use OSID specifications Translate functionalities into OSID sequences
3. Implement services Re-use existing implementations when possible

Table 5: Tips for the application layer

System layer

This layer provides support to the deployment and execution of web and voice-based components.

The system layer is required to provide support to the deployment and execution of web and voice-based components. While e-learning platforms are deployed and executed in application servers, voice interactive applications rely on voice servers.

Application servers manage communication and back-end transactions, they include a web server to handle the HTPP requests/responses and language interpreters for different programming languages. We propose the use of the Apache Software Foundation projects[82], which include the necessary support for the web server, as well as PHP and Java language interpreters.

In addition, voice servers are capable to handle aural processes and output audio resources; they provide command interpreters as well as Automatic Speech Recognition (ASR) and Text to Speech (TTS) modules for voice interaction. We propose to take into account parameters such as accuracy, tone, intonation and language support for the choice of ASR and TTS modules.

In our proposal, voice dialogues follow the recommendations of VoiceXML, a characteristic that makes them compliant with multiple voice servers. Following the results of the state-of-the-art study and our own research and professional experience, we propose to develop telephone-based solutions supported by Asterisk, a software implementation of a telephone private branch exchange which supports the use of VoiceXML along with several input and output audio technologies.

Tips for the system layer
1. Select an application server Communication & back-end transactions Programming language interpretation
2. Select a voice server XML-based dialogues Adequate audio input/output capabilities

Table 6: Tips for the system layer

[82]http://www.apache.org

Voice modules

Voice modules define the audio functionality available within e-learning management systems. Our research experience shows that each organisation develop its own set of LMS functionalities according to its particular needs, taking into account parameters such as the cost and outcomes of these developments in the academic or educational setting. Voice modules should therefore be able to meet the needs of specific academic environments and on-line learning communities.

Our research team has cooperated with the Department of Innovation to identify, set up and develop a set of voice modules which are aimed to add voice interactive support to the existing web-based virtual classroom at the University of Oviedo. The current backlog brings together more than ten voice modules distributed among several LMS sections.

Voice modules backlog	
LMS Section	**Voice Module**
Personal Area	My courses
	Grades/Grading
	RSS reader
	Calendar/Events
Learning Content	Institutional information
	Degrees and courses
	Content reader
Evaluation	Quizzes & surveys
Communication	Audioforum
	Audiochat/Conference
	Telephony & voicemail
Cooperation	Audioblogs/Podcast
	Audiowikis

Table 7: Voice modules backlog

We ought to remark that the various web-based e-learning components are not equally suitable to be translated into voice-based instructional modules. Thus, it is important to take into account two factors when prioritising the development of voice modules:

- Implement first the most used e-learning components at the academic or educational setting.
- Give priority to components which are more suitable (usable and technically feasible) to be translated into voice modules.

From the voice modules backlog, and following an iterative design process, we have first prototyped and evaluated the modules for consulting grades, checking calendar/events and filling out quizzes and surveys. These modules have been assembled following the Voice Interactive Classroom architecture.

Deployment diagram of the Voice Interactive Classroom

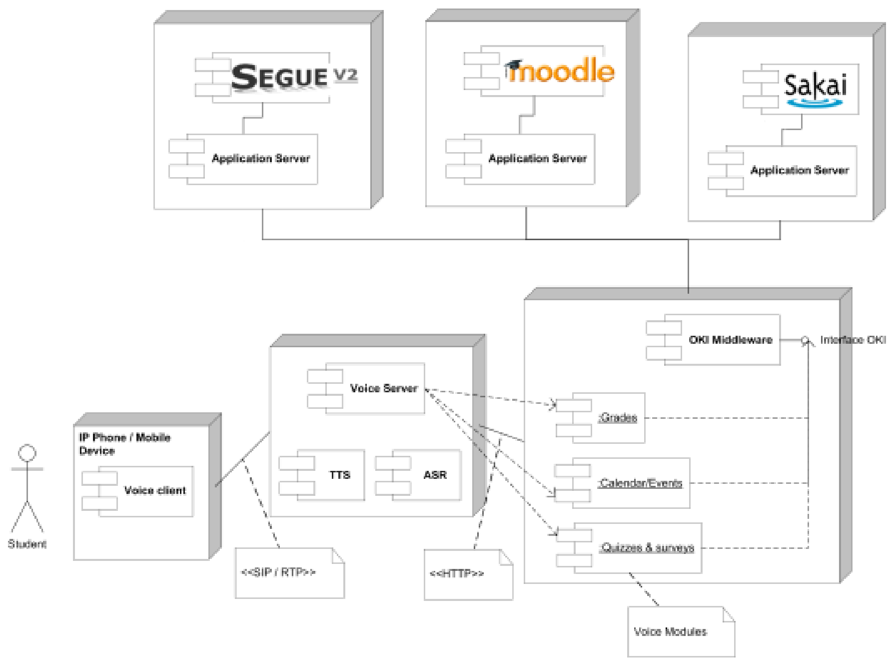


Illustration 30: Deployment diagram of the Voice Interactive Classroom

IX

Design of Case Studies

Previous chapter describes the proposed solution, a middleware architecture to enable cross-platform multi-channel access to Internet-based learning. This chapter presents recommendations and discusses the conditions for the design of case studies based on the Voice Interactive Classroom. It is intended to provide students and practitioners with indications they can use to plan, implement and evaluate case studies based on the proposed solution.

Introduction

Previous chapter describes the proposed solution, a middleware architecture named Voice Interactive Classroom which allows enabling cross-platform and multi-channel access to e-learning platforms under the premises described in Chapter VII.

This chapter provides an insight into the process for designing, developing, documenting and evaluating research case studies based on the proposed architecture. Since the results cannot be assessed in the absence of previous experiences, this section does not intend to provide a formal method to construct or evaluate voice interactive classrooms. Nonetheless, in order to provide students and practitioners with indications they can use to plan, implement and track prototypes, we consider necessary at this stage to present recommendations and discuss the conditions for the design of case studies based on the proposed architecture.

In addition to considering the findings of the present study, the design of case studies following presented in this chapter is influenced by and complements previous research studies, which have been described in Chapter VI.

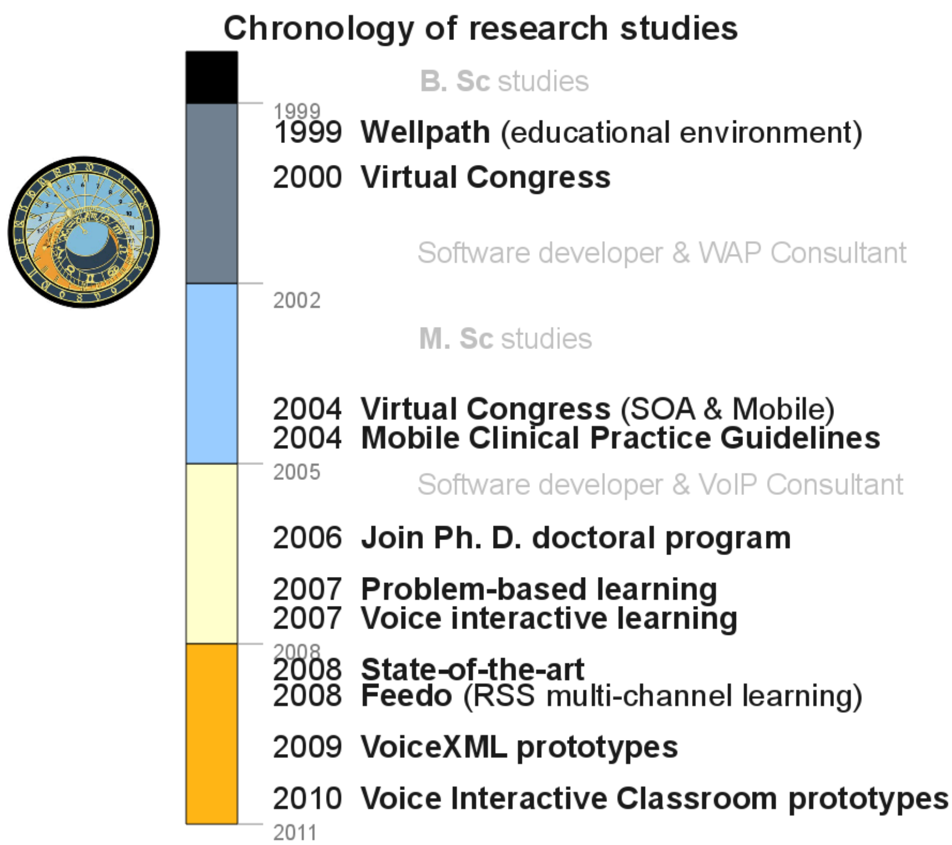


Illustration 31: Chronology of research studies

By summer 2009, we had already introduced and discussed our ideas and the findings of our research at our University and in national and European congresses. It was then that I found myself in the position of writing a first draft of the solution and the design of case studies based on our proposed architecture. Voice Interactive Classroom was defined not only in terms of functional and non-functional characteristics of a software system, but as a complete presentation of the research and motivation, as well as the goals and expected outcomes of this research. The document I drafted was first discussed internally before being shared with the Centre for Learning Sciences and Technologies at the Open University of the Netherlands, destination of my research visit in the autumn of the same year.

The draft was also used to introduce our next research study to the students selected to undertake the development of case studies based on the Voice Interactive Classroom, which were aimed to demonstrate the incidence and benefits of the proposed solution.

The development of case studies was aimed to demonstrate the incidence and benefits of the proposed solution.

In the following we describe the design of case studies, an evolving document where the reader will find both constructive and exploratory elements to support strategies for the development of voice interactive classrooms. Likewise, it is important to acknowledge that construction and exploration in our research are not linear but iterative, as the knowledge and findings from each study is used to identify and replace elements that perform below expectations.

Application scenarios of the case study

Research case studies aim to provide knowledge and experience in the area of their domain. The usefulness of a case study lies in its ability to contribute to the progress of a study and propose an advance in the state of development in a given field.

The usefulness of a case study lies in its ability to contribute to the progress of a study and propose an advance in the state of development in a given field.

Our first consideration in the design of case studies was to be able to demonstrate the benefits of the proposed solution, and show how it gives a step forward towards enabling cross-platform multi-channel access to Internet-based learning.

Furthermore, we should take into consideration that this thesis is devoted to adapt traditional visual-based e-learning to innovative audio-based e-learning scenarios.

Therefore, a research case study based on the Voice Interactive Classroom should additionally give a justification of how and why the solution could be used and explain its contribution to enable new e-learning scenarios.

In Chapter VIII five scenarios are proposed that make it either necessary or highly recommended to complement visual learning with audio interactions. These scenarios offer a valuable resource in the consideration of the usefulness of a proposed case study.

e-learning scenarios for audio-based learning
Learning processes in which sound is the main source of information
Students with perceptual auditory learning styles or those who prefer to learn by using both their visual and auditory senses
Hands-busy, eyes-busy and mobility-required situations
The device has speech input/output capabilities but lack of usable displays or keyboards
A disability does not allow a person to make use of a visual display or input/output devices, for instance when users have visual or motor impairments

Table 8: Proposed e-learning scenarios for audio-based learning

The choice of one of the above listed or an alternative scenario finds a strong motivation in the target of the case study and its application scope.

Generally, the following parameters influence the selection of application scenarios for the case studies:

- Application setting: academic/educational, institutional or commercial.
- Characteristics of the users it is intended to serve: learning styles, preferences and needs.
- Characteristics of the access device: PC, mobile device or landline.
- Situations that require the use of a particular access mode, such as visiting a city, a museum, or car navigator applications.

Although the choice of one scenario provides a more clear focus of the case study, it is not in discordance with enabling alternative uses. At the same time, it is often possible to consider a combination of two or more scenarios for the case studies.

Choice of technologies

The choice of technologies for the case studies has to consider several different variables which are going to affect the overall performance and behaviour of the prototype, being the most relevant the underlying e-learning platform, which has to satisfy the requirements of the learning process and which is directly related with the size of the middleware development effort.

The choice of technologies considers variables that affect the performance and behavior of the prototype.

Choice of e-learning platform

Several factors can be taken into consideration for the choice of an e-learning platform:

1. License costs can be a major consideration for the long-term success of the learning environment. The state-of-the-art shows a number of commercial and free software initiatives.
2. The popularity of the solution, which can determine its ability to integrate with existing software systems.
3. E-learning platforms in use at the surrounding institutional or commercial environment can be better supported or easier to update. E.g. e-learning platforms in use at academic institutions in the same country or region.
4. The reluctance of teachers and students to switch platforms makes it more likely for the academic and educational institutions to maintain a current solution rather than switching to a different platform.
5. When using e-learning frameworks, such as OKI OSIDs, the compliance with the specifications and the existence of previous implementations can determine the size of the middleware development effort.

Stressing the e-learning platform as the most relevant consideration, there are a number of technologies, in relation with the system layer in the Voice Interactive classroom, which have to be considered. They can be classified in two major groups, those related with the application server and those which are associated with the voice server.

Choice of web technologies

The application server can be intimately related with the e-learning platform as it is required to provide support for communications and back-end transactions, as well as handle protocol requests and responses, according to the needs of the e-learning platform and the programming language it has been written with.

There are a number of alternatives which can be considered and use in conjunction to provide a suitable web support. In our professional experience, examples of these alternatives are:

1. Microsoft's Internet Information Server[83], Active Server Pages (ASP)[84] or .NET development technologies[85], and SQL Server database[86].
2. BEA Weblogic[87] or Sun's Java System Application Server[88], Java[89] and Oracle database[90].
3. Apache software foundation projects, PHP and MySQL database[91].

The choice and assimilation of web technologies is generally influenced by the decision strategy used to adopt new technology and the institution or company experience.

Choice of voice technologies

The voice server can be initially separated from the e-learning platform, but we would also like to acknowledge the importance of the posterior integration between the voice and the e-learning systems.

Voice servers have a clear orientation to telephone-based systems and devices. The list of choices for the voice server includes Microsoft Speech Technologies[92], IBM WebSphere voice server[93], Loquendo MRCP Server, Voxeo Prophecy and Asterisk IP-PBX.

As the state-of-the-art indicates, voice interactive applications commonly require the use of automatic speech recognition and speech synthesis technologies. Among the choices for ASR and TTS modules we can mention Loquendo, LumenVox and Festival. Price range can vary accordingly to the characteristics of the product and the list of features required. Thus, besides the economical factor, it is also necessary here to consider the target of the case study and its application scope.

83 http://www.iis.net/
84 http://msdn.microsoft.com/en-us/library/aa286483.aspx
85 http://www.microsoft.com/net/
86 www.microsoft.com/sqlserver/
87 http://www.oracle.com/bea/
88 http://developers.sun.com/appserver/
89 http://www.java.com/
90 http://www.oracle.com/us/products/database/
91 http://www.mysql.com/
92 http://www.microsoft.com/speech/
93 http://www-01.ibm.com/software/pervasive/voice_server/

Software development methodology

Once an initial selection of the various technologies in the system is made, an agile iterative process for prototyping inspired on Scrum (Schwaber, 1995) (Schwaber & Beedle, 2001) and Extreme programming (Beck & Fowler, 2000) supports the construction of voice modules for the voice interactive classroom.

The list of voice modules presented in Chapter IX can be taken as a reference to set up a initial set of voice modules to be implemented. In Scrum, this planning is likely to be captured in a product backlog or sprint backlog. In this case we will refer to it as voice modules backlog.

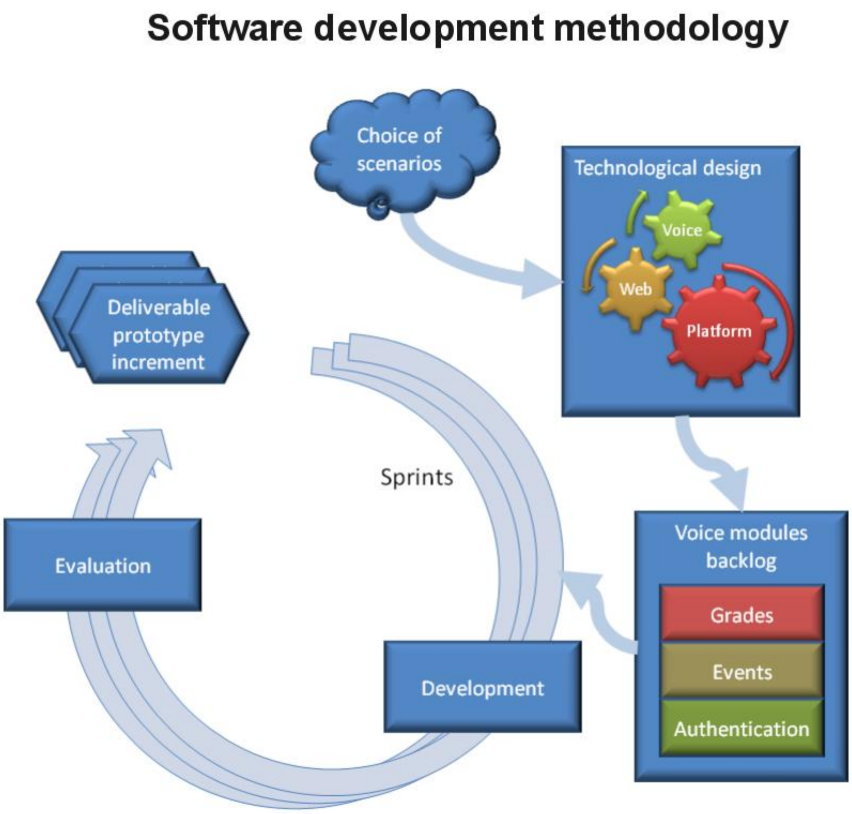


Illustration 32: Software development methodology

As it is shown in the figure above, prototyping and evaluation are part of an iterative and cyclic process. The evaluation provides the necessary feedback to support an iterative process for refining and improving the prototypes. The focus on delivering an increment of research or business value, along with the absence of unnecessary development overhead, are decisive factors for choosing Scrum.

Prototyping

Prototyping is a process that enables the developer to create a model of the software to be built.

Prototyping is a process that enables the developer to create a model of the software to be built and can take one of three forms: a paper prototype that enables the user to understand aspects of the interaction with the system, a working prototype that implements some subset of the desired software, or an existing program that performs all or part of the functions desired but has other features to be improved upon in the new development effort. (Pressman, 1987) (Macaulay et al., 1990).

We consider the following steps in the creation of the prototypes:

- Voice modules backlog.
- Design of voice dialogues.
- Implementation of service interfaces.
- Documentation and software practices.

Voice modules backlog

The selection of voice modules have to consider the particularities of the research or professional activity.

The initial selection and prototyping of voice modules have to consider the particularities of the research or professional development activity, and be prioritised accordingly. Some of the parameters that can be taking into account are the cost, technical and operational success factors, and the outcomes of these developments in the academic or educational setting.

Design of voice dialogues

In order to create effective, efficient and desirable voice interfaces, it is instructive to consider the key HCI design factors for voice enabled web systems (Duggan, 2004), as well as previous experiences with spoken dialog systems (Boyce, 2000) (Vaittinen, 2003) (Bolchini at al., 2006). This is particularly recommended during the first iteration in the design of the dialogues, when user's feedback from evaluation is not yet available.

Paper prototyping (Snyder, 2003) and voice dialogue diagramming techniques are also introduced in this stage to better discuss and provide a graphical representation of dialogue flow. As a result, interaction with the system is proposed to be described through paper prototypes, voice dialogue diagrams and conversational text-based representations.

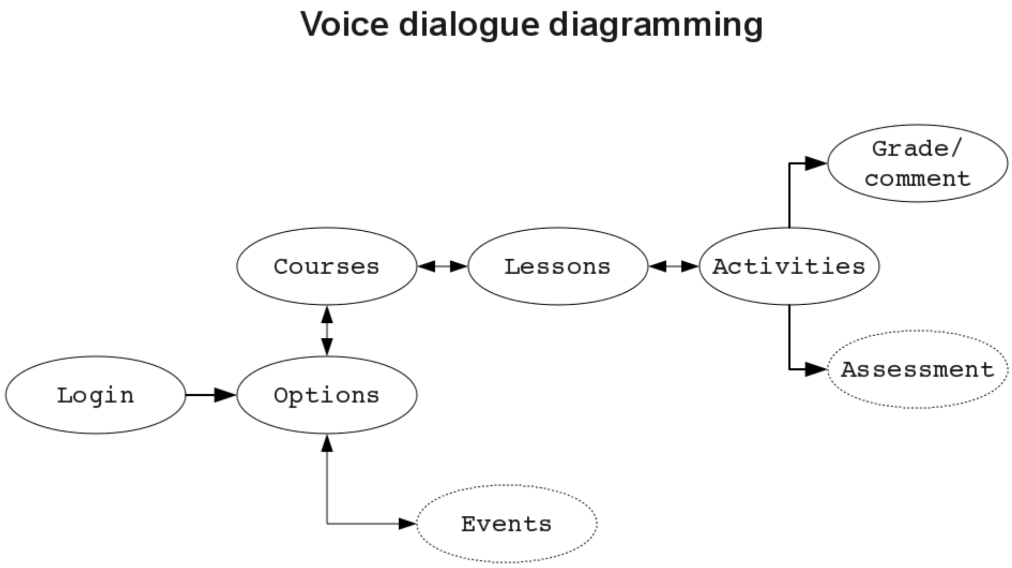


Illustration 33: Voice dialogues diagramming

The design of voice dialogues makes use of the standard description language VoiceXML. VoiceXML files contain static and dynamic elements. On design-time, a general dialogue structure is written in a VoiceXML skeleton or template from which dynamic data elements are replaced with tags. On run-time, tags are filled up accordingly to user preferences, usage context and system actions, giving as result the final VoiceXML document.

The design of voice dialogues makes use of the standard description language VoiceXML.

	Design-time VoiceXML	Run-time VoiceXML
Dialogue interactions	<vxml xml:lang=”LANGUAGE”>	<vxml xml:lang=”en-GB”>
	<prompt> COURSE_AUDIO </prompt>	<prompt> Please select the course number </prompt>
	VAR_DEFAULT	<var name==”course” expr=”5”>
	FORM_HEADER	<form id=”course” >
	OPTIONS	<option dtmf=”1” value=”1”>
	<filled> IF_CONDITION </filled>	<filled> <if cond=”mainvar==”1”> <goto next=”#course1”/> </filled>
System actions	<submit next=URL />	<submit next=”http://vic.org/course.php?lang=en_GB” />

Table 9: Design-time and Run-time VoiceXML

Implementation of service interfaces

Voice dialogues connect to the middleware using the logic described by OKI OSIDs.

Portability and logical independence from the e-learning platform is addressed at this stage. The design of the Voice Interactive Classroom allows voice dialogues to connect to the middleware using the logic described by the Open Knowledge Initiative Open Service Interface Definitions. It is required for each voice module to identify and make use, or implement when needed, the corresponding OKI Open Service Interface Definitions.

Re-using previous implementations is a major advantage in service orientation and it is highly recommended at this point, as it can significantly reduce the development effort and save time.

In addition, the level of expertise with the inner implementation of the e-learning platform should be considered, as developers might be required to provide a new implementation of the OKI OSIDs by linking service interface definitions with features provided by the underlying e-learning platform.

Documentation and software practices

Documentation and software practices are among the aspects which should be taken into consideration.

Documentation and software practices are among the aspects which should be taken into consideration in any software development. The following elements can help supporting the development of the prototypes:

Documentation

Software documentation and manuals contain valuable information on the development process, software architecture and usage of the system, as well as information related to the application domain. When documenting, it is always recommended to follow a methodology according to the characteristics and goals pursued by the project, taking also into consideration the various backgrounds of the members of the development team.

Auto-generated documentation

Auto-generated documentation helps summarising aspects of a software project using different formatting styles. Software tools support different automatic documentation generation, with Javadoc, a mechanism for generating Java API documentation, being the most popular at the present time. Similar alternatives can be used for software projects using a programming language different than Java. Examples of these are phpDocumentor and Rdoc, for the PHP and Ruby programming languages respectively.

Organised directory hierarchy

We recommend the use of an organised directory hierarchy to store and deploy voice dialogues, OKI OSIDs definitions and implementations, as well as the rest of software elements and resources involved in the process of prototyping.

We recommend the use of an organised directory hierarchy for prototyping.

The following table shows the proposed directory hierarchy and the description of each directory listed. We would like to remark that this directory hierarchy should be taken as an initial recommendation rather than an imposition. We encourage students and practitioners to improve this proposal to better accommodate the various elements involved and contribute to the evolution of this proposal.

Proposed Directory Hierarchy			
Vic **Voice Interactive Classroom root directory**			
	Asterisk Configuration files to allow using Voice Interactive Classroom with the Asterisk IP-PBX		
	Audio Internationalised audio files		
	Config General configuration files		
	Daemons Background processes for maintenance and updating purposes		
	Doc General documentation		
	Languages Internationalised messages		
	Platforms Learning management systems specific files		
		[LMS] I.e. Moodle, Sakai, Segue	
			Config Configuration files for a LMS (if required)
			Doc Documentation for a LMS (including auto-generated documentation)
			Log System activity register
			Modules Voice applications (dynamic dialog generation)
			Testing Unit, integration and system testing
	Vxml Static dialogues in VoiceXML format		
	Wrappers Libraries required to link the features with the network, operating system, Text to Speech, voice recognition systems, etc.		

Table 10: Proposed directory hierarchy

Testing

Software testing assures computer code does what it was designed to do and does not do anything unintended.

Software testing is a process, or a series of processes, designed to make sure computer code does what it was designed to do and that it does not do anything unintended. Software should be predictable and consistent, offering no surprises to users (Myers, 1979). Testing plays an important part in any software development project, and a difficult one, since it is directly influenced by the evolution of programming languages, operating systems and hardware platforms. In order to address unit, integration and system testing we recommend making use of software testing tools and documenting both the process and its results.

Project visibility

We aim to make the results publicly available for the scientific community and society.

In our aim to make the results of this thesis publicly available for the scientific community and society, we encouraged students to use a free software license for the prototypes and publish them in the open source software development web site Sourceforge.net, under the project entitled Voice Interactive Classroom[94].

Evaluation

An e-learning application should be evaluated considering both its usability and its didactic effectiveness (Ardito et al., 2006). Although the scientific literature describes methods and processes for usability evaluation in audio and e-learning settings, they have not been related to one another. Our process for evaluating the prototypes suggests taking into account a combination of usability aspects in e-learning (Ardito et al., 2006) and speech-enabled applications (Mellor et al., 1996). It applies both quantitative and qualitative measures to evaluate the effectiveness and efficiency of the prototypes, as well as the satisfaction of e-learners.

Evaluation captures both technical and pedagogical usability issues.

The evaluation process follows a combination of usability engineering and a user-centred approach to capture both technical and pedagogical usability issues. It starts from the ideas introduced in Chapter VI to extend and refine the measures used in the evaluation. We introduce in the following measures to evaluate the effectiveness and efficiency of the prototypes and e-learners' satisfaction.

[94] http://victhesis.sourceforge.net/

Effectiveness

Effectiveness is the accuracy and completeness with which users achieve certain goals (ISO 9241, 1998). There are several indicators of effectiveness, including quality of solution and error rates (Frøkjær et al., 2000). We use error rates as primary indicator for effectiveness. This quantitative measurement can be calculated using the number of turns in the perfect understanding situation (TurnsPerfect is a constant number predefined by the dialog interface design) and the number of turns in the practice or task (TurnsPractice is obtained in practice from logs files and audio recorded files where the e-learner's voice is recorded) (Ai & Weng, 2008).

At the same time, we should acknowledge that one of the main concerns in voice interactive systems is word error rate scores on speech recognition tasks. Accuracy in speech recognition affects the effectiveness of the interaction and it is a key factor in the way users perceive the system. We suggest making use of system logs and monitor students activity while using the prototypes to register ASR error rate scores.

Efficiency

Efficiency is the relation between the accuracy and completeness with which users achieve certain goals, and the resources expended in achieving them. Indicators of efficiency include task completion time and learning time (Frøkjær et al., 2000). Efficiency in our approach is addressed by measuring the task completion time. We propose to make use of three different data sources to obtain the task completion time: log files, recorded audio files, and information gathered from voice software such as softphones or IP Phones.

Satisfaction

Satisfaction of e-learner can be defined as the interface ability to interact with a user in a pleasant way, which can be assessed as a summary affective response of varying intensity that follows e-learning activities (Giese & Gote, 2000). Surveys are a common method to get e-learners feedback and assert qualitative parameters. This study customises an empirically validated survey instrument developed by Wang (Wang, 2003) for measuring learner satisfaction in asynchronous e-learning systems.

For the answers of the questionnaire, we suggest using a four point Likert scale without mid-point (1-2: highly disagree/disagree; 3-4: agree/highly agree). By intentionally avoiding a neutral or ambivalent category, we help in the interpretation of the results.

Questionnaire to measure learners' satisfaction		
Item code	**Item description**	**Access mode**
Q1	The e-learning system provides content that exactly fits your needs.	Visual
		Audio DTM Tones
		Audio Automatic Speech Recognition
Q2	The e-learning system provides useful content	Visual
		Audio DTM Tones
		Audio Automatic Speech Recognition
Q3	The e-learning system provides sufficient content	Visual
		Audio DTM Tones
		Audio Automatic Speech Recognition
Q4	The e-learning system provides up-to-date content	Visual
		Audio DTM Tones
		Audio Automatic Speech Recognition
Q5	The e-learning system is easy to use	Visual
		Audio DTM Tones
		Audio Automatic Speech Recognition
Q6	The e-learning system makes it easy for you to find the content you need	Visual
		Audio DTM Tones
		Audio Automatic Speech Recognition
Q7	The e-learning system provides content which is easy to understand	Visual
		Audio DTM Tones
		Audio Automatic Speech Recognition
Q8	The e-learning system is user-friendly	Visual
		Audio DTM Tones
		Audio Automatic Speech Recognition
Q9	System operations are stable	Visual
		Audio DTM Tones
		Audio Automatic Speech Recognition
Q10	Voice Interaction gives added value to e-learning platform	Not applicable

Table 11: Questionnaire to measure learners' satisfaction (adapted from Wang)

Evaluation completes one iteration in the methodology proposed for developing the prototypes.

The process to evaluate the effectiveness and efficiency of the prototypes and e-learners' satisfaction completes one iteration in the methodology proposed for developing the prototypes. At this point a deliverable prototype increment is made available. From the results obtained in the assessment, the development team must consider whether another iteration is required to refine the prototype.

X

Demonstration: Case Studies Revisited

This chapter completes the contributions of this dissertation. It is aimed to demonstrate the benefits of the proposed solution, and describe how it contributes to enable cross-platform multi-channel access to Internet-based learning.

The approach followed for such demonstration consists on describing a series of case studies conducted between years 2009 and 2010 to address the experimental aspects derived from the proposed solution. This section extends the case studies described in Chapter VI, but this time using the Voice Interactive Classroom architecture to enable voice access to the e-learning platforms Moodle and Sakai.

Introduction

Chapter VI started the contributions of this dissertation by presenting a series of research studies, conducted between years 2007 and 2009, which helped us to understand the preliminary level of knowledge in the development of voice interactive learning applications.

In the following we complete those contributions by describing a new series of case studies conducted in the following academic year (2009-2010). These research case studies address the experimental aspects of the solution, which has been described in Chapter VIII, and the design of case studies described in Chapter IX. They are intended to demonstrate the benefits of using the Voice Interactive Classroom architecture, and how it contributes to enable cross-platform multi-channel access to Internet-based learning.

A first draft summarising Chapter VIII and Chapter IX was used to introduce the research to our colleagues at the Centre for Learning Sciences and Technologies, as well as to students Manuel Pérez and Moisés Riestra, who were selected to undertake the prototyping and evaluation using the e-learning platforms Sakai and Moodle respectively.

From this draft, and taking into account the research findings and experience from previous studies, we set up the first experiments at our University. The need to reduce the voice modules backlog to three or four features to fit within the scope of an academic project, made us carefully select the experiments. Besides taking into account the match between voice modules and e-learning components in Sakai and Moodle, we had to consider the technical complexity of the proposed developments, and prioritise the features which, according to the Department of Innovation at our university, were the most used in the virtual classroom or the most attractive for the use of voice interactions. Taking into account the aforementioned factors, we decided to conduct the following three case studies:

- Authentication and authorisation
- Checking grades
- Checking calendar notes/events

Next sections describe in more detail the decision aspects and the development of the three case studies.

Application scenario of the case studies

From the five e-learning scenarios described in Chapter IX, we selected the use of devices with speech input/output capabilities. Particularly, we focused in case studies to provide students with VoIP access to e-learning platforms. This choice does not prevent the case studies to be used in other scenarios (e.g. students who prefer to use both their visual and auditory senses).

From the five e-learning scenarios, we selected the use of devices with speech input/output capabilities.

The choice found a strong motivation in our academic context, where we have students who make use of a visually-based e-learning platform, Moodle, but they are not being offered with the possibility to access the platform by aural means from landlines or mobile devices. It is our particular interest to address this issue by enabling an innovative voice over IP telephone access to our virtual classroom.

Enabling telephone access requires complementing the existing visual mode with aural interactions, hence meeting the goal of enabling multi-channel access to Internet-based learning.

Additionally, we must keep in mind that this research is also concerned with providing cross-platform access to e-learning. This goal makes it necessary to conduct the case studies at least upon two e-learning platforms with different technological characteristics (software architectures, web technologies, programming languages, etc.). In the following section we explain the choice of e-learning platforms, as well as the associated technologies needed for the deployment of voice interactive classrooms.

Choice of technologies

Choice of e-learning platforms

According to the state-of-the-art, license cost is a major consideration for the long-term success of the learning environment. At the same time, economic costs along with the reluctance of teachers and student to switch platforms makes it more feasible for academic and educational institutions to maintain current solutions rather than switching to different platforms. This made us decide to select Moodle, the e-learning platform currently in use at our University, and the most popular in Spanish universities, as our first choice for the case studies.

Choice of e-learning platforms
Moodle: the e-learning platform in use for the virtual classroom at the University of Oviedo and the preferred e-learning platform in Spanish universities
Sakai: full service-oriented web-based e-learning platform built partially on the principles of IMS Abstract Framework and OKI

Table 12: Choice of e-learning platforms

A second e-learning platform was required to demonstrate the multi-platform feature of the solution. The different technical characteristics of the e-learning platforms, including software architectures, related web technologies and the programming languages they have been written with, as well as their popularity, were factors for taking this decision. Despite not being totally compliant with OKI standards, Sakai contains many technical differences with respect to Moodle, and has certain level of popularity. For the mentioned reasons, we considered Sakai very convenient for our second choice.

The choice of an e-learning platform has a number of technology side-effects.

The choice of an e-learning platform has a number of technology side-effects and it affects the rest of the choices of technologies for the application and voice servers.

There is currently no separation between the e-learning platform and the web technologies which have been used to develop it. I.e. while e-learning platforms such as Moodle or Segue have been implemented using PHP, Sakai has used Java. Hence, by choosing Moodle and Sakai for the case studies, we need to consider differences in both the internal structures and programming mechanisms; and in this way the studies are forced to adopt two different software development approaches in concordance with the construction of PHP and Java-based software applications.

Choice of web technologies

Although it is possible to consider other alternatives for the application server, the choice of the Apache Software Foundations projects results in a natural one in the context of Moodle, PHP and the associated free-software technologies, and for this reason we selected them in conjunction with MySQL as the database manager.

From this decision, we addressed the possibility of supporting Sakai within the same application server by using the Apache Tomcat extension to recognise Java Server Pages and Java Servlets.

Choice of voice technologies

The importance of the integration between the e-learning platform and voice technologies in our academic context, led us to decide to give priority to the adoption of free software developments with the purpose of guarantee the compatibility of our proposals with the e-learning setting. However, our previous experiences with the use of free software were not completely satisfactory, and we preferred that some of the components in the voice server were of a commercial nature.

Asterisk is a well-known open source IP-PBX which I was very experienced with, and therefore a very convenient choice for the operational success and completion of the prototypes. The choice of a command interpreter, as well as ASR and TTS modules for voice interaction, was influenced by the choice of Asterisk, and based in the experience from previous case studies, where free-software components showed unsuccessful or performed on a lower level than commercial alternatives.

Recognition and text to speech engines must be tailored for particular regions to be effective (Markowitz, 1996). i6Net[95] is a Spanish company which takes into consideration ASR and TTS modules which include support for the use of the Spanish languages. i6Net VXI* VoiceXML browser[96] was our choice for the VoiceXML language interpreter. At the same time, the orientation to telephony applications, as well as the availability and good performance for the use of Spanish languages, made us select the ASR and TTS modules from the Spanish company Verbio[97]. Both i6Net and Verbio kindly offered us a free version of their products for research purposes.

Prototyping & evaluation

In the initial selection of the scope to be explored, we prioritised the features which, according to the Department of Innovation at our university, were the most used in the virtual classroom or the most attractive for the use of voice interactions. In this choice, we also took into account the complexity of the audio technology involved, as well as the compliance of the features with OKI specifications and e-learning platform components. As a result, we decided to conduct three case studies (Table 13).

95 http://www.i6net.com/
96 http://www.i6net.com/products/vxi/
97 http://www.verbio.com/

Voice modules backlog	
Type of service	**Voice module**
Educational services	Grades/Grading
	Calendar/Events
Common service	Authentication & authorisation

Table 13: Voice modules backlog

These case studies were considered adequate to address the experimental aspects of the Voice Interactive Classroom, enabling voice access to the e-learning platforms Sakai and Moodle. In the following we describe in more detail the process for developing and evaluating each one of the prototypes.

Authentication and authorisation

The first case study is concerned with the design of a voice-enabled authentication and authorisation feature. We should acknowledge that, although this case study is used as an introductory procedure, its simplicity does not carry with it a loss in realism, since authentication and authorisation are basic mechanisms in web applications and they give support to a number of other modules, including checking grades and checking calendar notes/events, which are the subject of next sections.

A common aspect in the design of case studies consists in using voice dialogues to define the interaction with the user and the Spanish and English languages to internationalise the application.

A review of the authentication and authorisation processes in Sakai and Moodle provides a first understanding of these features and a starting point for designing the user interaction with the system.

The inner implementation of the authentication and authorisation services in Sakai and Moodle differs in programming languages (Java and PHP respectively), authentication types (some are coincident, i.e. username/password accounts and LDAP), internal data structures and programming mechanisms. In this case study we selected the most common authentication type, consisting in asking the user for a username and password and entering through system accounts.

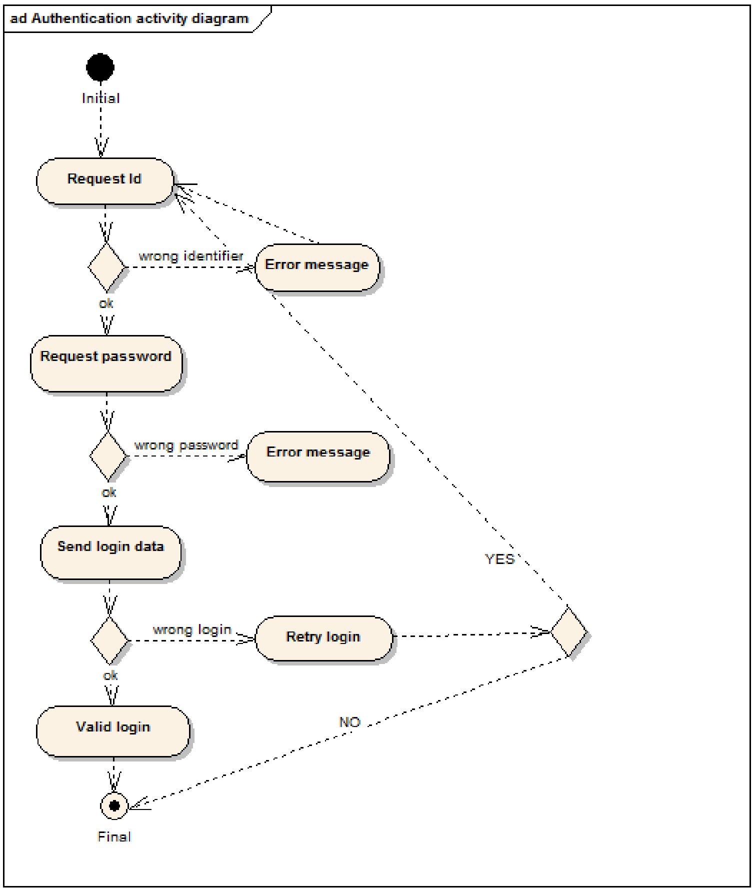


Illustration 27: Authentication activity diagram in Sakai

In the proposed architecture, the development of the authentication and authorisation processes relies in using and implementing, when necessary, the suitable OKI Open Service Interface Definitions.

The focus of the Authentication OSID is authenticating and mapping users to the AgentId representing them in the system. The Authentication service is designed to allow for support of more than one Authentication process. For example, the same Authentication implementation could support authenticating users through user id and password, certificates, or any other method. In our case, we have decided to authenticate users through a traditional username and password access control mechanism.

OKI OSIDs cover services that are likely to involve restricted content and operations. The Authorization OSID provides a way to define who is authorised to perform an operation in the system.

Authorisations associate agents (who), which represent the user or another actor in the system, with functions (what) and qualifiers (when/where). Before performing an operation, services will want to know if there is an authorisation that covers the operation. If the agent is authorised, the operation can be attempted. If not, the operation is not allowed.

The implementation of abstract classes and primitive-type iterators is supported by an open service interface definition named Shared. This OSID provides basic primitives, and receives the name of shared because many of its services are widely used across other OSIDs.

An examination of an Open Service Interface Definition (OSID) usually begins with the Manager. Managers provide the way to create the objects that implement the principal interfaces in the service. In this case, the classes AuthenticationManager and AuthorizationManager contain the methods required to perform the authentication and authorisation.

Consequently, it is possible to infer the OKI service and interfaces involved in the development of this component:

1. Authentication
    - AuthenticationManager
    - AuthenticationException

2. Authorization
    - AuthorizationManager
    - Authorization
    - AuthorizationException
    - AuthorizationIterator
    - Function
    - FunctionIterator
    - Qualifier
    - QualifierIterator

3. Agent
    - AgentManager
    - Agent
    - AgentException
    - AgentIterator
    - Group

4. Shared
    - Id
    - Properties
    - PropertiesIterator
    - Type
    - TypeIterator

In keeping with the philosophy of re-using when possible, we found that the authentication and authorisation components already had an implementation for both Sakai and Moodle, given by project Campus. Thus, it was only necessary to assemble and reuse the code provided by project Campus to have the authentication and authorisation services available for standalone use or to support any other service.

Authentication and authorisation components already had an OKI implementation for Sakai and Moodle.

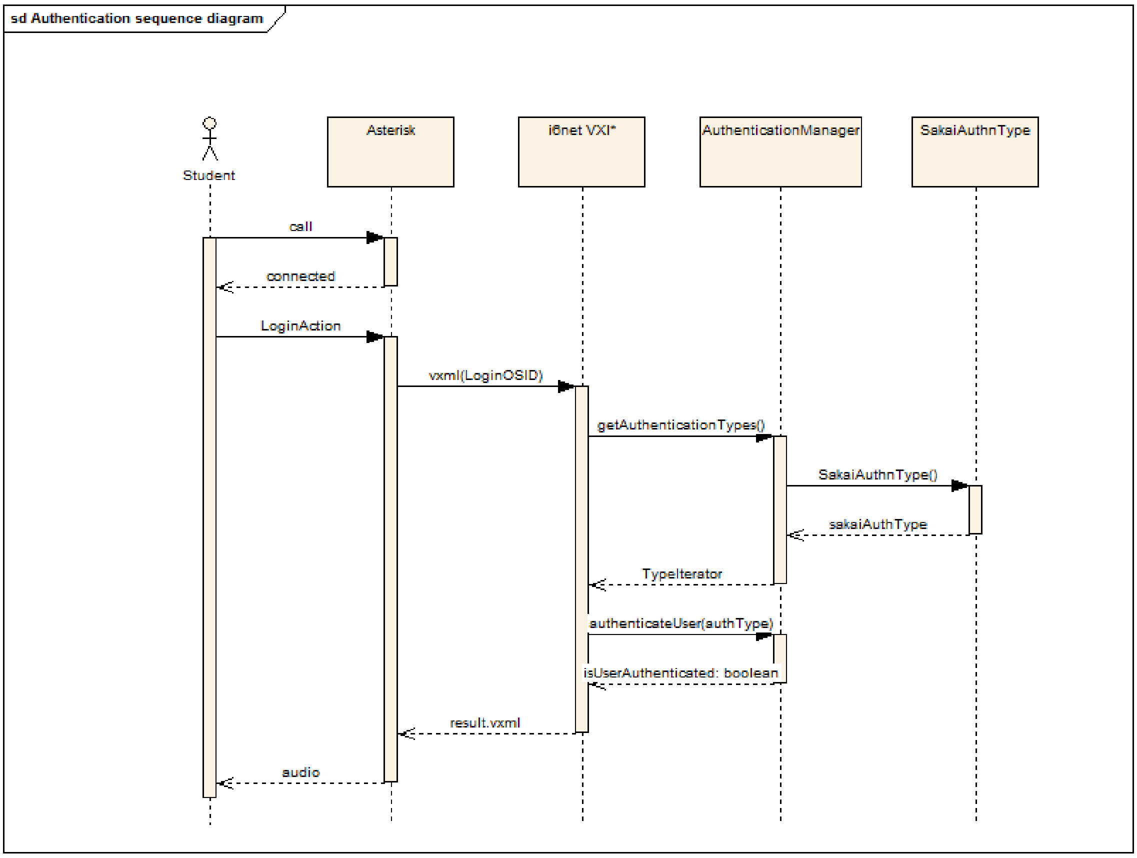


Illustration 35: Authentication sequence diagram

This case study provided little feedback during the evaluation, mainly due to its simplicity and small number of user interactions. Results showed how usability parameters were very similar to those in the visual interface. However, this case study is too narrow to draw definitive conclusions.

Checking grades

In the second case study we developed a voice-enabled feature to allow students to check their grades in a particular activity or course. Sakai and Moodle present differences in their features to manage courses and display grades. It was hence necessary to analyse the approached proposed by OKI before proposing a design for the user interaction.

In the OKI OSID model, checking grades makes reference to how students (agents) perform in the different courses. This makes it necessary to consider two additional OKI services for the construction of this component, CourseManagement and Grading.

The CourseManagement Open Service Interface Definition primarily supports creating and managing a catalogue of courses. Although similar in structure, e-learning platforms present variations in the hierarchy and elements of a catalogue of courses. In OKI this catalogue is organised into:

- CanonicalCourses, which are general and exist across terms.
- CourseOfferings for CanonicalCourses, which occur in a specific term.
- CourseSections for CourseOfferings, which have a meeting location, schedule, student roster, etc.

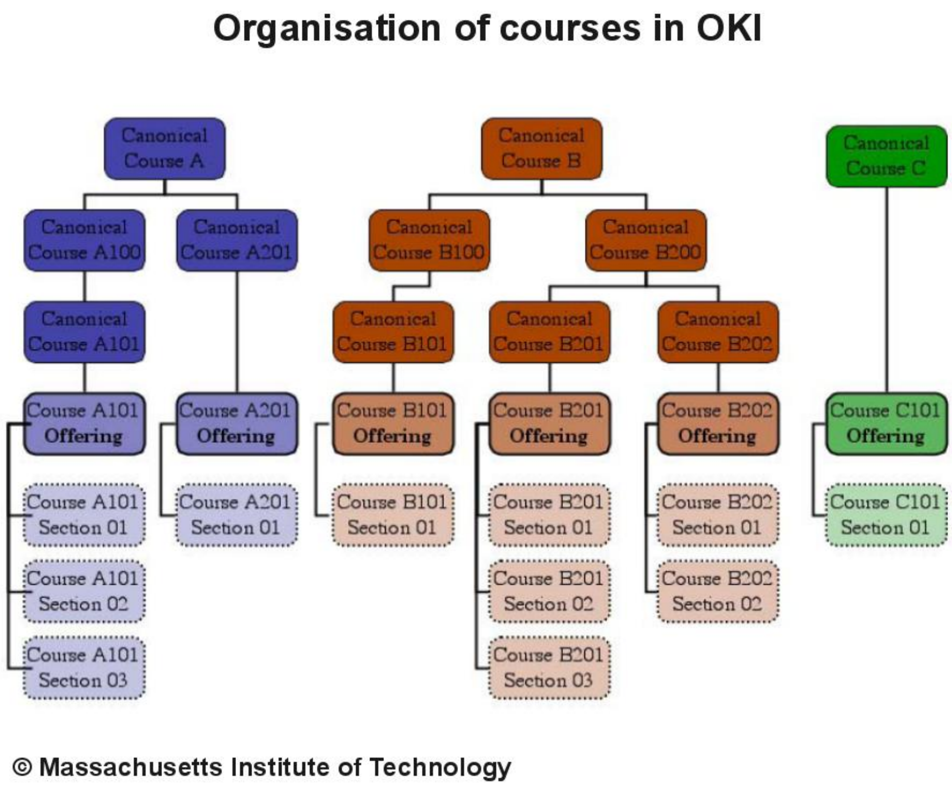


Illustration 29: Organisation of courses in OKI

A GradableObject holds information about a specific CourseSection. The Grading OSID supports characterising, storing and retrieving Grades. A grade is specified with four elements: a Grade-Value, GradeType, GradeScale, and ScoringDefinition. These four elements provide a general and flexible way to characterise a grade, which can be used to provide an audio output of the information.

In order to associate the information about grades, students and activities/courses, the service also provides support for managing GradeRecords, which join information about the grade, the agent (student) whose grade it is, and the object that was graded.

The initial design of the voice dialogues can be inferred from the navigation structure and various elements described in OKI's CourseManagement and Grading.

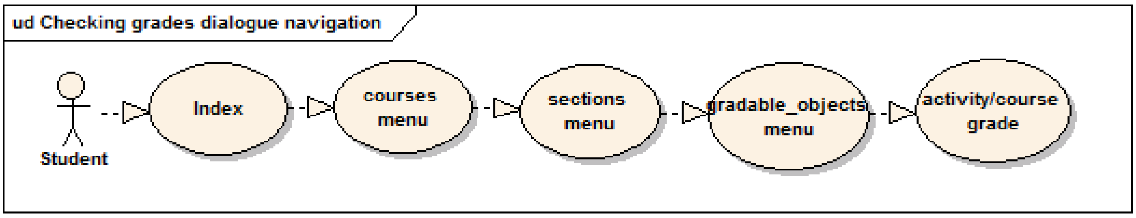


Illustration 37: Checking grades dialogue navigation

The development of the checking grades voice module involves the use of the following additional OKI services and interfaces:

5. CourseManagement
   - CourseManagementManager
   - CourseOffering
   - CourseOfferingIterator
   - CourseSection
   - CourseSectionIterator
   - CourseGradeRecord
   - CourseGradeRecordIterator

6. Grading
   - GradingManager
   - GradableObject
   - GradableObjectIterator
   - GradeRecord
   - GradeRecordIterator

However, in contrast to what happened in the authentication and authorisation component, Course Management and Grading OSIDs were not available from the Campus or any other project at the time of this development. The lack of previous developments made it necessary to provide the interfaces with two new implementations for Sakai and Moodle. These implementations did not cover all the possible methods or interfaces, as it was not required for the completion of the prototype.

CourseManagement and Grading OSIDs were not available for re-use, making it necessary to implement them for Sakai and Moodle.

A closer look into the code shows both the contrasts and similarities in the implementations. Table 11 illustrates the programming style differences as well as the coincident logic between the two implementations. Even when re-usability is not possible and a new implementation of the OKI services is required, the use of OKI OSIDs helps to generalise and unify the logic of e-learning platforms features.

Even when re-usability is not possible, OKI OSIDs help to generalize and unify the logic of e-learning platforms features.

Checking grades source code, Sakai/Moodle comparison	
Sakai: GradingManager.GetGradableObject	**Moodle: GradingManager.GetGradableObject**
public GradableObjectIterator getGradableObjects(Id courseSectionId, Id externalReferenceId) throws GradingException {	function getGradableObjects ($courseSectionId, $externalReferenceId) {
// initialize variables GradebookService gs = (GradebookService) ComponentManager .get ("org.sakaiproject.service.gradebook.GradebookService");	//initialize variables $idSection = ''; $idCourse = ''; if($courseSectionId == null || $externalReferenceId == null) { throw new GradingException(NULL_ARGUMENT); }
// obtaining the gradebook String gradebookUid; try { gradebookUid = courseSectionId.getIdString(); } catch (SharedException e) { throw new GradingException(e.getMessage()); }	//obtaining the gradebook $idSection = $courseSectionId->getIdString(); $idCourse = $externalReferenceId->getIdString(); //get a section with $idSection and $idCourse $section = get_course_section($idSection,$idCourse);
// obtaining assignments for this gradebook List<Assignment> list = null; try { list = gs.getAssignments(gradebookUid); } catch (GradebookNotFoundException e) { return null;}	// obtaining assignments for this gradebook $gradable_items = explode(",",$section->sequence); if(count($gradable_items)==0){ return null; }
// create a GradableObject and add it to a list GradableObject go = new org.unioviproject.component.osid.grading.GradableObject(courseSectionId, assignmentName, assignmentName, extRefId, assignmentId, pointsPossible, dueDateLong, released); // add it to a list listGO.add(0, go);	// create a GradableObject and add it to a list $gradable = get_record('grade_items','courseid', $gradable_item->course,'iteminstance', $gradable_item>instance, 'itemmodule',$module->name); $id = new okiserver_osid_shared_Id($gradable->id); $gradable_object = new org_vic_osid_grading_GradableObjectImpl($id); array_push($gradable_objects_array,$gradable_object);
// return gradable objects list return (new org.unioviproject.component.osid.grading.GradableObjectIterator (listGO));	// return gradable objects list return new org_vic_osid_grading_GradableObjectIteratorImpl($gradable_objects_array);

Table 14: Checking grades source code, Sakai/Moodle comparison

The interaction with the user was correct and students were highly satisfied with the voice feature.

The evaluation of this case study yielded to meaningful results. Efficiency measures showed that the completion time for the aural task is 56% longer than the visual equivalent. This increment is caused by the intrinsic sequentiality of aural navigation. Effectiveness indicated that the interaction with the user was correct, with an error rate below 10% and ASR accuracy of 95%. In the measure of learners' satisfaction, all the students indicated they highly agreed with the content, presentation and navigation of the voice feature.

Checking calendar notes/events

A third case study completes the objectives and features planned for this stage of the research process. This study proposes the development of a voice-enabled feature to allow students to check calendar notes, announcements or calendar updates made by themselves, their supervisors or the teacher responsible of a subject they are enrolled in.

To provide students with a certain grade of flexibility regarding the events they can listen to, the navigational structure of the voice dialogues considered several date ranges and options (Table 15).

Listen to events aural template		
Reading order	**Number of subjects**	**Period consideration** (from request moment)
1	all enrolled	next 24 hours
2	all enrolled	next 7 days
3	all enrolled	next 30 days
4	one subject	next 30 days

Table 15: Listen to events aural template

For each one of the options, events are arranged in time order, so that the event which takes place sooner is reproduced first. Information regarding calendar notes or events can include the subject, time period and location.

In OKI, calendar events require to manage additional information to associate activities and their schedule with agents and the courses they are enrolled upon. This feature is provided by the OKI service Scheduling.

The Scheduling Open Service Interface Definition provides a means of associating agents with specific activities (ScheduleItems) that have a specific start and end date and time. For each ScheduleItem, the Agent or Agents involved have a Status that reflects their level of commitment to the activity. There are provisions for both browsing through ScheduleItems and for finding the times when a set of Agents are available.

The Scheduling OSID is linked with the CourseManagement OSID and additionally provides a way for an application to integrate or use a calendar system.

The development of the checking calendar/events voice module involves the use of the following additional OKI service and interfaces:

7. Scheduling
   - SchedulingManager
   - ScheduleItem
   - ScheduleItemIterator

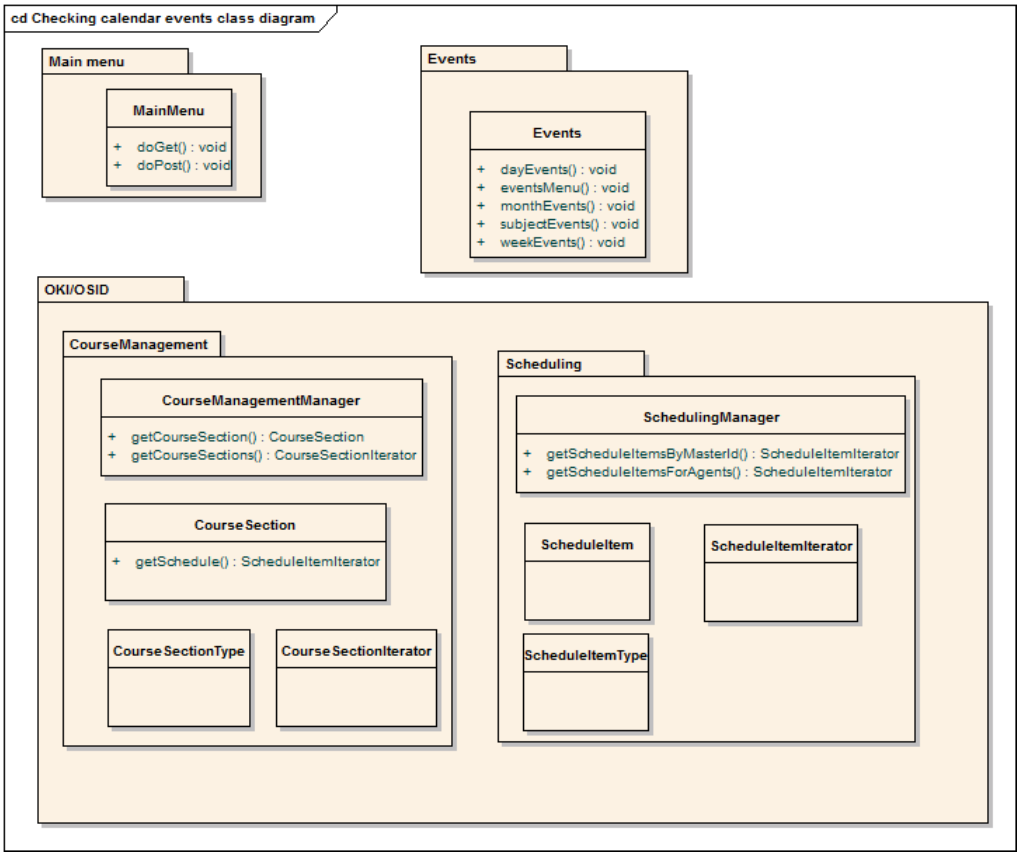


Illustration 38: Checking calendar events class diagram

OKI Scheduling service was not available for re-use.

As occurred with the previous prototype, the OKI Scheduling service was not available for re-use and required a new implementation of the interfaces for Sakai and Moodle, resulting in the subsequent effort and time consumption.

Evaluation scores were slightly lower than in the previous case.

Results from the evaluation complemented previous findings. Events case study is similar to that for checking grades, but with different query information and less number of dialogue turns. Efficiency measures showed that completion time for the aural task was 23% longer than the equivalent visual interface. Effectiveness indicated that the interaction with students was correct, scoring an error rate of approximately 10% and ASR accuracy of 90%. In this case, satisfaction scores were slightly lower than in the previous case; with students indicating they highly agreed with 89% of the content and interaction characteristics.

Conclusions

The development of a series of case studies successfully enabled us to adapt visual interfaces to voice dialogues and to provide aural access to e-learning features already present in Sakai and Moodle.

Case studies enabled us to provide aural access to Sakai and Moodle.

At the same time, the described design and development of case studies provide students and practitioners with useful procedures and guidelines that they can use to plan, implement and evaluate their own voice-based applications.

Despite the case studies give a decisive step forward towards technology independence and interoperability, there are aspects that are still dependent on the characteristics of the underlying software architecture. This is the case of the connection between services and e-learning platforms, which is decisively shaped by the characteristics of the underlying e-learning platforms.

In this middleware approach, re-usability is a crucial factor for the development of prototypes. Identifying the OSIDs needed for a particular feature and implementing them proved to be time consuming tasks. Although we were able to accommodate some previous OSID implementations, OKI adoption was still in an early stage and many of the service interfaces required new implementations.

In this middleware approach, re-usability is a crucial factor.

The following table summarises the OKI development for the three voice modules and helps to understand the level of re-usability achieved by using Open Service Interface Definitions in the year 2009.

OKI OSIDs re-usability in 2009		
Service	**Interface**	**Re-used**
Agent	AgentManager	✓
	Agent	✓
	AgentException	✓
	AgentIterator	✓
	Group	✓
Authentication	AuthenticationManager	✓
	AuthenticationException	✓
Authorization	AuthorizationManager	✓
	Authorization	✓
	AuthorizationException	✓
	AuthorizationIterator	✓

OKI OSIDs re-usability in 2009		
Service	**Interface**	**Re-used**
	Function	✓
	FunctionIterator	✓
	Qualifier	✓
	QualifierIterator	✓
CourseManagement	CourseManagementManager	✗
	CourseOffering	✗
	CourseOfferingIterator	✗
	CourseSection	✗
	CourseSectionIterator	✗
	CourseGradeRecord	✗
	CourseGradeRecordIterator	✗
Grading	GradingManager	✗
	GradableObject	✗
	GradableObjectIterator	✗
	GradeRecord	✗
	GradeRecordIterator	✗
Scheduling	SchedulingManager	✗
	ScheduleItem	✗
	ScheduleItemIterator	✗
Shared	Id	✓
	Properties	✓
	PropertiesIterator	✓
	Type	✓
	TypeIterator	✓
	Level of re-usability	**20/35 57%**

Table 16: OKI OSIDs re-usability in 2009

In the year 2009 the percentage of services that could be reused for the planned components was 57%. While most of the basic OKI services were already available for Sakai and Moodle, more specific services were yet to be released. With the release of the version 1.4.1 of the Campus project in August 2010, the level of re-usability for the given prototypes raised up to 100%. Although specific e-learning platforms, features and services are still required to provide new implementations of the OSID interfaces, this increase can be taken as an indicator of the progress and interest arisen by this service-oriented approach.

The results from the evaluation are very positive. The students who participated in these experiences highly appreciated the possibility of using speech interaction to access learning management systems, and they were very supportive of the use of audio technologies in the proposed scenarios.

Students highly appreciated using speech interaction and were very supportive of the use of audio technologies.

The successful development of case studies using the e-learning platforms Sakai and Moodle completes the experimental aspects of this dissertation, and demonstrates that the Voice Interactive Classroom architecture, along with the associated proposals of this dissertation, can be used to enable cross-platform multi-channel access to Internet-based learning.

XI

Scientific Communication & Validation

This chapter describes the scientific communication and validation of this dissertation, which has been achieved by setting up and following an active strategy for scientific outreach. Activities include planning, coordinating and supervising final year projects, master projects and master dissertations, writing and presenting articles and communications in scientific meetings, and the publication of this research study in a prestigious scientific journal.

"Difficulties are rewarded when a prestigious scientific journal accepts one of our articles for publication. This means a group of experts has considered our work represents an advance for science, and that is the top aspiration".

"Las dificultades se ven recompensadas cuando una revista científica de prestigio acepta uno de nuestros artículos para su publicación. Eso significa que un grupo de expertos ha considerado que ese trabajo representa un avance para la ciencia, y eso es lo máximo a lo que se puede aspirar".

Dr. Carlos López Otín
Universidad de Oviedo

Introduction

A validation strategy required to communicate and publish this research in scientific conferences and journals.

From the beginning of the dissertation, we agreed to complement our proposals and empirical findings with a validation strategy based in producing valuable innovations and enhance the scientific quality of this dissertation by means of communications and publications in conferences and journals. In order to achieve the expected outcomes within a specific time, we set up a scheme of incremental difficult in publishing the results.

On a first institutional level, we planned, coordinated and supervised projects involving research activities from computer science students at different levels. Along with following the planning and indications regarding the development of exploratory case studies and prototypes, the students were asked to write, present and defend their research with other members of our academic institution. This work produced outputs such as final year projects, master projects and master dissertations defended at the University of Oviedo.

Communications and articles presented in scientific meetings provide a first review by a group of experts, and a great opportunity to compare our research with related work and discuss future research directions with other research groups. During the research and elaboration of this dissertation, we actively participated and contributed with papers presented in several national and international congresses.

The final and most challenging achievement consisted in the inclusion of this research in the scientific literature in the field of computer science and the reward of a publication in a prestigious scientific journal. Following the directives and recommendations by The National Agency for Quality Assessment and Accreditation (ANECA)[98], we carried out a preliminary selection from a list of journals from Thomson Reuters' Journal Citation Reports[99]. A rigorous peer review process by a group of experts in the field assures the research accepted for publication meets high standards in terms of theoretical and empirical rigour and originality. At the same time, papers published in this manner, makes the research accessible to the scientific community and more generally, to society as a whole.

[98]http://www.aneca.es/
[99]http://thomsonreuters.com/products_services/science/science_products/a-z/journal_citation_reports

Final year projects, master projects and master dissertations.

Isaac Rubio and his team fellows at the laboratory of Internet programming started working on the first project in February 2007. The development of Feedo, the personalised multi-channel RSS aggregator oriented to e-learning settings continued until December 2008, when Isaac Rubio was awarded B.Sc. in Computer Science for his final year project, obtaining a grade of 9 out of 10.

On the 7th of July 2008 I defended with honours my master of advanced studies dissertation entitled “Software frameworks for developing electronic learning platforms. From monolithic to flexible systems”, a research work that provided the basis for the state-of-the-art, and pointed out the main concerns of this dissertation.

The second project was initiated by Moisés Riestra in the same month of 2008. This study, which analysed the use of W3C recommendations of VoiceXML in audio-based e-learning settings, was carried out within a period of one year. Moisés Riestra was awarded B.Sc. in Computer Science for his final year project in July 2009, obtaining a grade of 9 out of 10.

In order to address simultaneously the complementary case studies on the use of our proposed software architecture upon the e-learning platforms Sakai and Moodle, we decided to overlap these projects within the period from July 2009 to July 2010. Manuel Pérez performed the study for Sakai and was awarded B.Sc. in Computer Science for his final year project in February 2010, obtaining a grade of 9.5 out of 10. Moisés Riestra, who undertook the Moodle version and proposed a further study on the evaluation of the usability, was awarded M.Sc. in Computer Science for his master dissertation in July 2010, obtaining a grade of 9 out of 10.

Additionally, we ought to mention two research studies which have been intentionally omitted for not being direct part of this dissertation, but which are in relation with its contributions and the proposals for further research derived from it.

A study by Diego Rosado comprised our experience in adapting the agile development methodology Scrum to the development characteristics of final year projects in computer science. This approach helped greatly to understand and support the process of rapid prototype development of the voice interactive classrooms. This project was initiated in summer 2008 and finished in March 2010, when Diego Rosado was awarded B.Sc. in Computer Science obtaining a grade of 9.5 out of 10.

In autumn 2009, and influenced by my research visit in the Netherlands, we set up a collaboration study with members of the Centre for Learning Sciences and Technologies on the use of augmented-reality and audio-based interactions in mobile learning settings. At the University of Oviedo, this study was entitled "UniDroid, Audio Augmented Reality at University" and it was addressed from September 2009 to July 2010, when Pablo Muñoz was awarded M.Sc. in Computer Science for his master project, obtaining a grade of 9 out of 10.

Active participation in scientific meetings

In the period from October 2008 to October 2010 we participated in seven scientific meetings.

In the period from October 2008 to October 2010, we participated in seven scientific meetings where we presented the work derived from this dissertation.

In 2008, a first article derived from my master in advanced studies dissertation, and entitled "Presente y futuro del desarrollo de plataformas web de elearning en educación superior", was accepted at the V Spanish Multidisciplinary Symposium on Design and Evaluation of Reusable Educational Content (SPDECE 2008). On the 20th of October 2010, I gave a talk on this work at the Pontifical University of Salamanca, Spain.

A second research work entitled "A Case study of the adaptation of Problem-Based Learning for programming subjects", which explains the academic experience that facilitated and motivated the development of our first prototype Feedo, was selected for the ACM SIGCSE (spSIGCSE), First Workshop on Methods and Cases in Computing Education (MCCE'08), an event celebrated after SPDECE08 and in the same location. On the 22nd of October 2010 I gave a presentation on this work at the Pontifical University of Salamanca, Spain. A paper extending this talk was selected for publication in the International Journal of Teaching and Case Studies.

In 2009 a research study on personalised multi-channel learning using RSS, and entitled "Personalized web and voice-based learning using feed syndication", was accepted at the IASK E-Activity and Leading Technologies International Conference (E-ALT2009). On the 23rd of June 2009 I gave a talk on this work in Seville, Spain. A paper extending this talk was selected for publication in the Journal of Human Capital and Information Technology Professionals.

In the same year, we wrote two more articles to explain the topic of an unfinished master dissertation by student Héctor Parajón. This work proposes a structural and semantic analysis of the Web to provide aural access to Internet resources. The papers, entitled "Hacia una navegación web basada en diálogos de voz" and "Acceso web por voz: un enfoque orientado al diálogo" were accepted, and Dr. María del Puerto Paule and Héctor Parajón gave talks on these works at the Conferencia Ibero-Americana InterTIC 2009 (Seville, June 2009) and X Congreso Internacional de Interacción Persona Ordenador (Barcelona, September 2009), respectively.

During my research visit at the Centre for Learning Sciences and Technologies in 2009, I was invited to participate in the JTEL Winter School on Advanced Learning Technologies 2010, which was held in Innsbruck, Austria from the 1st to the 6th of February 2010. Besides attending the various talks, workshops and activities, I participated in a pecha kucha session where I attempted to describe, in 180 seconds, the basics of this dissertation.

When the projects were coming to conclusion, a paper describing the master dissertation by Moisés Riestra and entitled "Plataformas de e-learning multicanales: Caso de estudio sobre Moodle" was accepted for the VII Spanish Multidisciplinary Symposium on Design and Evaluation of Reusable Educational Content (SPDECE 2010). On the 1st of July 2010, Moisés Riestra gave a talk on this work in Cádiz, Spain, and I was privileged to be invited to act as chair for the session.

At the present time, we continue the dissemination of this research. At the time of writing this dissertation, we have been invited to present our research at the 4th Spanish Sakai Congress, which will be celebrated in November 2010 in Barcelona, and the 2nd International Symposium on Middleware and Network Applications, which will be held in Las Vegas in April 2011.

At the present time we continue disseminating this research.

Research visit

From the 20th of September 2009 to the 20th of December 2009, and as part of the planning for applying for the European doctorate mention, I carried out a research visit at the Centre for Learning Sciences and Technologies, Open University of the Netherlands, where I was privileged to join Dr. Rob Koper, Dr. Marcus Specht and their research team.

In 2009 I carried out a research visit at the Centre for Learning Sciences and Technologies.

The purpose of this visit was to introduce the research work conducted by our research team, give a final impulse to this thesis and consider possible research synergies.

Initial formal and informal meetings already provided a good understanding on the research areas that were the interest of both parties, and helped to encourage communication and interaction between our research teams.

We reviewed this dissertation and consolidated its theoretical background. Dr. Marcus Specht (CELSTEC) and Dr. Raul Izquierdo (UNIOVI) proposed to explain the contexts in which users could benefit from this research by reviewing and better describe the scenarios for audio-based learning. From this suggestion and with the collaboration of Remko van Dort, we defined the five scenarios for voice interactive learning as they have been described throughout this dissertation.

Writing was one of the main focuses in the course of the research visit.

Writing was one of the main focuses in the course of the research visit. Two papers were written and sent for revision during this period. A first article, which extended the study on RSS-based learning using audio, is a forthcoming publication in the International Journal of Human Capital and Information Technology Professionals. The second article summarises the proposals and findings of this dissertation and is described in detail in next section.

During the research period we also discussed several collaboration lines for introducing voice interactive learning in different environments and projects at the Celstec.

These proposals had in common the requirement consisting in introducing speech-enabled technology, and particularly Text to Speech (TTS) and Automatic Speech Recognition (ASR). Thus, we decided to evaluate different free and commercial tools, as well as libraries which could provide speech synthesis and recognition in multiple languages. The tools evaluated included Sphinx, Verbio, Loquendo and Android SDK.

From the proposed collaboration lines, we gave preference to research on augmented reality for mobile Android. From this idea, we proposed the master project entitled "UniDroid, Audio Augmented Reality at University".

We also participated together in the proposal of European projects which focus on the supporting of young researchers in the field of mobile and contextualised learning.

Publication in a journal

One of the main concerns during my research visit at the Open University of the Netherlands was to put together and summarise the findings and proposals from this dissertation and write them for publication in a prestigious scientific journal.

After considering several journals, we decided to submit a manuscript for the Special Issue of Journal of Network and Computer Applications on "Middleware Trends for Network Applications". Elsevier Journal of Network and Computer Applications is a scientific journal indexed in the Thomson Reuters' Journal Citation Reports (JCR) with a 5-year impact factor of 1.255 (2009). The decision of publishing our research in this journal was decisively influenced by one of the topics of the special issue, "service oriented middleware", which in our opinion fitted our research very well.

We decided to submit a manuscript for the Special Issue of Journal of Network and Computer Applications on "Middleware Trends for Network Applications".

Illustration 39: Special Issue of Journal of Network and Computer Applications on "Middleware Trends for Network Applications"

We wrote the first version of the paper from the ideas of previous manuscripts and in a month time, from the 29th of September 2009 to the 31st of October 2009, date of the submission of the manuscript for peer review to Elsevier editorial office. During this period I was living in Maastricht and working in Heerlen.

We received the favourable response from Elsevier on the 11th of January 2010, along with the reviewers' comments and suggested revisions that needed further consideration. The revision of the manuscript was addressed during the next month. As dates were coincident with the JTEL Winter School 2010, part of this writing was done in Innsbruck, Austria. We submitted the revised version to the editorial on the 10th of February 2010.

We received the final acceptance on the 18th of March 2010; and with it the scientific recognition to the research which had started more than two and a half years before.

We received the final acceptance in March 2010; and with it the scientific recognition to this research.

In words of Dr. Jameela Al-Jaroodi and Dr. Nader Mohamed, editors of the special issue, the article demonstrates the use of middleware to facilitate specific applications in interactive learning. "Alvarez et al. describe how middleware can be used to facilitate voice interactive classrooms. This approach allows the middleware to support different types of applications used for interactive e-learning by allowing for the incorporation of voice-based interfaces. Generally, enabling speech controlled software is done within a specific application and cannot be ported to another. However, in this article, a middleware approach is used to provide this feature for various web-based learning applications."

Citations

The aforementioned research studies and publications have been arranged in alphabetical order of authors' surnames and date of publication, and can be cited as follows:

Víctor Manuel Álvarez García, María del Puerto Paule Ruiz, Remko van Dort, Juan Ramón Pérez Pérez. RSS-based learning using audio. Journal of Human Capital and Information Technology Professionals. Forthcoming publication. doi:10.4018/jhcitp.2010100105.

Víctor Manuel Álvarez García, María del Puerto Paule Ruiz, Juan Ramón Pérez Pérez, Manuel Antonio Martín García. A Case study of the adaptation of problem-based learning for programming subjects. International Journal of Teaching and Case Studies. Forthcoming publication.

Víctor Manuel Álvarez García, María del Puerto Paule Ruiz, Juan Ramón Pérez Pérez. Voice interactive classroom, a service-oriented software architecture for speech-enabled learning. Elsevier Journal of Network and Computer Applications, Volume 33, Issue 5, Pages 603-610, September 2010.

Víctor Manuel Álvarez García, María del Puerto Paule Ruiz, Remko van Dort, Juan Ramón Pérez Pérez. Personalized web and voice-based learning using feed syndication. Proceedings of the IASK E-Activity and Leading Technologies International Conference (E-ALT2009), Volume 1, Pages 90-96. Sevilla (Spain), 22-24 June 2009.

Víctor Manuel Álvarez García, María del Puerto Paule Ruiz, Juan Ramón Pérez Pérez. A case study of the adaptation of problem-based learning for programming subjects. Spanish Chapter of the ACM Special Interest on Computer Science Education ACM SIGCSE, Volume 1, Pages 33-39. Salamanca (Spain), 22 October 2008.

Víctor Manuel Álvarez García, María del Puerto Paule Ruiz, Juan Ramón Pérez Pérez, Ignacio Gutiérrez Menéndez. Presente y futuro del desarrollo de plataformas web de elearning en educación superior. V Simposio Plurisdisciplinar sobre Diseño y Evaluación de Contenidos Educativos Reutilizables (SPDECE 2008), Volume 1. Salamanca (Spain), 20-21 October 2008.

Víctor Manuel Álvarez García. Software frameworks for developing electronic learning platforms. From monolithic to flexible systems. Master of Advanced Studies dissertation, Department of Computer Science, University of Oviedo, July 2008.

Pablo Muñoz Andrade. Unidroid: Realidad aumentada por audio para escenarios de e-learning. Master project, Department of Computer Science, University of Oviedo, July 2010.

Héctor Parajón Sánchez, María del Puerto Paule Ruiz, Víctor Manuel Álvarez García, Juan Ramón Pérez Pérez. Acceso web por voz: un enfoque orientado al diálogo. X Congreso Internacional de Interacción Persona Ordenador. Barcelona (Spain) 7-9 September 2009.

Héctor Parajón Sánchez, María del Puerto Paule Ruiz, Víctor Manuel Álvarez García, Juan Ramón Pérez Pérez. Hacia una navegación web basada en diálogos de voz. Conferencia Ibero-Americana InterTIC 2009, Volume 1, Pages 269-274. Sevilla (Spain), 22-24 June 2009.

Manuel Pérez González. Prototipo para la implementación de servicios de OKI accesibles mediante voz. Final year project, Department of Computer Science, University of Oviedo, February 2010.

Moisés Riestra González. Plataformas de e-learning multimodales e interoperables: caso de estudio sobre Moodle. Master dissertation, Department of Computer Science, University of Oviedo, July 2010.

Moisés Riestra González, María del Puerto Paule Ruiz, Víctor Manuel Álvarez García, Juan Ramón Pérez Pérez. Plataformas de e-learning multicanales: Caso de estudio sobre Moodle. VII Simposio Plurisdisciplinar sobre Diseño y Evaluación de Contenidos Educativos Reutilizables (SPEDECE 2010), Volume 1. Cádiz (Spain), 30 June - 2 July 2010.

Moisés Riestra González. Acceso a recursos de Internet mediante servicios de voz utilizando el estándar del W3C VoiceXML. Final year project, Department of Computer Science, University of Oviedo, July 2009.

Diego Rosado López. Herramienta de soporte a la planificación ágil de proyectos fin de carrera basado en la metodología "Scrum". Final year project, Department of Computer Science, University of Oviedo, March 2010.

XII

Conclusions & Further Research

The last chapter of this thesis summarises, in a compressed form, which are in our opinion its major contributions. This section also describes briefly the areas of research that we believe will lead to further progress in the near future.

Major contributions

As on-line learning is becoming more and more prevalent, the rising interest in this type of education has motivated the appearance of numerous platforms as well as new software frameworks specifically designed for e-learning, but not necessarily linked to one another. This situation has created a heterogeneous set of software approaches to support the development of on-line learning environments. E-learning technologies have traditionally based the interaction with the user mainly on visual means, limiting the number of scenarios in which e-learning processes can take place. This limitation has become more significant recently because pervasive and mobile computing devices are often used in hands-busy, eyes-busy settings, and they usually lack of usable displays and keyboards. However, other e-learning scenarios should also be taken into consideration. In this research we have identified five scenarios that make it either necessary or highly recommended to complement visual learning with audio interaction. This allows us to accommodate individual learning styles differences and enable innovative learning scenarios for students.

We consider that current solutions to combine visual web-based e-learning platforms and voice interaction, including the use of guidelines for web application accessibility and assistive technologies, fail to support adequately the overall design process, providing only aural descriptions of the visual content. The way the information is structured and organised differs from visual to auditory interactions. This thesis describes how visual and audio-based e-learning can be achieved by making an adequate conversion from the given LMS visual interfaces to audio dialogues. The adaptation from visual-only learning into naturalistic voice dialogues can be addressed by applying methods and techniques of adaptive hypermedia. In addition to following guidelines for interaction design, this thesis introduces a user-centred iterative design process that helps to make audio interactions more user-friendly and better suit learners' requirements.

Furthermore, by using a service-oriented middleware supported by OKI Open Service Interface Definitions, voice dialogues can easily be re-used and integrated within a heterogeneous set of e-learning platforms. This thesis describes how a middleware can be constructed and used to accommodate interoperable voice interactive classrooms. This approach results in a software architectural design which supports different types of voice-enabled applications used for interactive e-learning.

Generally, enabling speech controlled software is done within a specific application and cannot be ported to another. However, in this thesis, a middleware approach is used to provide this feature for various web-based learning applications. E-learning frameworks help to deal with this diversity, but understanding the specifications and implementing a robust set of service interfaces can be laborious and time consuming task. The level of interoperability that can be achieved relies on the externalisation of educational services and on the ongoing progress in the adherence to service specifications such as IMS Abstract Framework and OKI OSIDs.

Further research

There are a number of conceptual as well as technical issues which have arisen from this dissertation and are in the following proposed for further exploration.

The description on the design of case studies provides a general understanding on the evolution of the research and the enhancement in prototyping and evaluation, but does not intend to propose a complete development procedure. At the same time, it is clear that the methodology applied during the development of case studies influences the completeness and quality of the voice interactive classrooms. For these reasons, we suggest the procedure for building voice-enabled applications based on the proposed architecture to be revised and further extended. We consider that address this matter would help students and practitioners to reproduce more accurately the experiences hereby presented, and contribute to enlarge the number of possible solutions.

Voice interactive classrooms need extensive user evaluation. Although the scientific literature describes methods and processes for usability evaluation in audio and e-learning settings, they have not been related to one another. The novelty of this thesis prevents the availability of previous models for measuring effectiveness, efficiency and students satisfaction in multi-channel (visual & audio) e-learning settings. For such research, we suggest taking into account existing methods, adapting the evaluation to assess usability and application-specific measures in voice interactive classrooms.

While this research makes a significant contribution towards achieving independence from technologies, there are still aspects that depend on the characteristics of the underlying software architecture.

This is the case of the connection between learning services and e-learning platforms, which is decisively shaped by the web technologies and program languages the e-learning platforms have been written with. We suggest addressing this technical difficulty by following a more dynamic approach allowing splitting out design-time technology-neutral representations from run-time more specific code for a particular web technology or programming language.

Another aspect that could be dynamised is the adaptation of the visual-based learning process into voice dialogues, which is currently addressed at design-time, and which could be automated to be responsive to real-time changes in the e-learning platforms.

This research shows that the various audio features available within e-learning management systems are not equally easy to address, and this is partly caused by the heterogeneity and current disagreements in the adherence to standards for web-based resources and learning technologies. In our opinion, the study of the structural and semantic nature of the Web, as well as agreements in the adherence to interoperability specifications and semantic ontologies, can contribute to model common e-learning domains and also help to facilitate aural access to Internet resources in the future. Additionally, semantics can also play an important role in the evolution of speech recognition, since current automatic speech recognition is mainly based in the use of a predefined set of short commands instead of longer natural language sentences.

Finally, we consider that voice interactive learning is not restricted to e-learning platforms, but can also find applications in new exciting fields, such as the integration of novel context-aware and mobile systems with educational approaches for supporting learning processes. Although the research on this topic is out of the scope of this thesis, my research visit at the Centre for Learning Sciences and Technologies in the year 2009, along with the participation in the JTEL Winter School 2010, have already led us to initiate research projects to explore innovative technologies for contextualised and mobile learning in the context of the European Network of Excellence called STELLAR[100], which brings together a number of European research organizations in the field of Technology Enhanced Learning.

[100] http://www.stellar.net.eu

XIII

Conclusiones e Investigación Futura

El último capítulo de esta tesis resume las que en nuestra opinión son sus principales aportaciones. Esta sección también describe brevemente las áreas de investigación que en nuestro juicio contribuirán a su progreso en el futuro.

Principales aportaciones

A medida que el aprendizaje en línea se vuelve más frecuente, el creciente interés en este tipo de educación ha motivado la aparición de numerosas plataformas así como frameworks de software específicamente diseñados para el aprendizaje electrónico, pero no necesariamente vinculados entre sí. Esta situación ha creado un conjunto heterogéneo de enfoques para permitir el desarrollo de entornos de aprendizaje en línea. Las tecnologías de aprendizaje electrónico han basado tradicionalmente la interacción con el usuario en el uso de la visión, lo que limita el número de escenarios en los que los procesos de aprendizaje se pueden llevar a cabo. Esta limitación se ha vuelto más significativa recientemente debido a que los dispositivos móviles y los dispositivos pervasivos son utilizados habitualmente en situaciones en las que los ojos y las manos están ocupados, y en las que no se dispone de pantallas y teclados utilizables. Sin embargo, también se han de tener en cuenta otros escenarios de aprendizaje electrónico. Esta investigación ha identificado cinco escenarios que hacen necesario o altamente recomendable complementar el aprendizaje visual con la interacción por voz. Esto nos permite dar cabida a las diferencias individuales en los estilos de aprendizaje y habilitar escenarios de aprendizaje innovadores para los estudiantes.

Consideramos que las soluciones actuales para combinar las plataformas de aprendizaje basadas en web y la interacción vocal, incluyendo el uso de guías de accesibilidad web y las tecnologías asistenciales, son inadecuadas para dar soporte al proceso global de diseño, puesto que solamente proporcionan descripciones auditivas del contenido visual. La manera en que la información se estructura y organiza difiere de las interacciones visuales a las vocales. Esta tesis describe cómo es posible conseguir un aprendizaje electrónico visual y auditivo realizando conversiones adecuadas de las interfaces visuales de los LMS a diálogos de voz. La adaptación del aprendizaje basado solamente en la visión a diálogos naturales puede ser abordada mediante la aplicación de métodos y técnicas de hipermedia adaptativa. Además de seguir las guías para el diseño de interacciones, esta tesis introduce un proceso de diseño iterativo y centrado en el usuario que permite que las interacciones auditivas sean más fáciles de utilizar y estén mejor adaptadas a los requisitos de los estudiantes.

Además, usando un middleware orientado a servicios que se apoya en la especificación OKI Open Service Interface Definitions, los diálogos de voz pueden ser reutilizados e integrados de manera sencilla en un conjunto heterogéneo de plataformas de aprendizaje.

Esta tesis describe como se puede construir un middleware que permite la creación de aulas de voz interactivas. Este enfoque da lugar a un diseño de arquitectura software que permite desarrollar diferentes tipos de aplicaciones dirigidas por voz usadas en el aprendizaje interactivo.

Generalmente, el software controlado por voz es utilizado en aplicaciones específicas y no puede ser portado a otras. Sin embargo, en esta tesis un enfoque middleware proporciona esta característica para diversas aplicaciones de aprendizaje basadas en la web. Los frameworks de aprendizaje electrónico ayudan a tratar esta diversidad, pero entender las especificaciones e implementar un conjunto robusto de interfaces de servicios puede ser una tarea laboriosa y difícil. El nivel de interoperabilidad que se puede alcanzar depende de la externalización de los servicios educativos y del progreso en la adopción de especificaciones de servicios tales como IMS Abstract Framework y OKI OSIDs.

Investigación futura

A partir de esta tesis doctoral han surgido una serie de cuestiones conceptuales y técnicas que son propuestas a continuación para ser exploradas en el futuro.

La descripción del diseño de casos de uso proporciona una comprensión general sobre la evolución de la investigación y la mejora en los procesos de prototipado y evaluación, pero no pretende proponer un procedimiento completo de desarrollo. Al mismo tiempo, resulta claro que la metodología aplicada durante el desarrollo de casos de estudio influye en la completitud y calidad de las aulas de voz interactivas. Por estos motivos, sugerimos que el procedimiento para construir aplicaciones dirigidas por voz y basadas en la arquitectura propuesta sea revisado y extendido. Consideramos que abordar esta cuestión puede ayudar a los estudiantes y profesionales a reproducir de manera más precisa las experiencias presentadas y contribuir a ampliar el número de posibles soluciones.

Las aulas de voz interactivas necesitan una mayor evaluación. Aunque la literatura científica describe métodos y procesos para la evaluación de la usabilidad en entornos auditivos y de aprendizaje electrónico, no se han relacionado entre sí. La originalidad de esta tesis evita la disponibilidad de modelos previos para la medición de la eficacia, eficiencia y satisfacción de los estudiantes en entornos de e-learning multicanales (visuales y auditivos). Para esta investigación, sugerimos que se tengan en cuenta los métodos existentes para adaptar la evaluación y determinar medidas de usabilidad específicas en aulas de voz interactivas.

Aunque esta investigación realiza una contribución significativa para conseguir la independencia de las tecnologías, aún hay aspectos que dependen de las características de la arquitectura software subyacente. Este es el caso de la conexión entre los servicios educativos y las plataformas de aprendizaje, que está claramente condicionada por las tecnologías web y los lenguajes de programación usados por las mismas. Sugerimos abordar esta dificultad técnica siguiendo un enfoque más dinámico que permita separar representaciones neutras en tiempo de diseño frente a código más específico en tiempo de ejecución para una tecnología web o lenguaje de programación.

Otro aspecto que podría ser dinamizado es la adaptación del proceso visual a diálogos de voz, que actualmente es abordado en tiempo de diseño, y que podría ser automatizado para responder a los cambios en tiempo real en las plataformas de aprendizaje.

Esta investigación muestra que las diversas funcionalidades de audio disponibles en las plataformas de aprendizaje no son igualmente sencillas de abordar, y esto se debe en parte a la heterogeneidad y actuales desacuerdos en la adopción de estándares para los recursos web y las tecnologías de aprendizaje. En nuestra opinión, el estudio de la estructura y semántica de la Web, así como los acuerdos en la adopción de especificaciones de interoperabilidad y ontologías semánticas, pueden contribuir a modelar dominios comunes de aprendizaje y ayudar a proveer un acceso auditivo a los recursos de Internet en el futuro. Además, la semántica también puede jugar un papel importante en la evolución del reconocimiento automático del habla, dado que en la actualidad está principalmente basado en el uso de un conjunto predefinido de comandos cortos en lugar de sentencias más largas en lenguaje natural.

Finalmente, consideramos que el aprendizaje interactivo por voz no se limita a las plataformas de aprendizaje, sino que puede encontrar aplicación en nuevos campos, como la integración de los sistemas sensibles al contexto y móviles con los enfoques educativos para apoyar los procesos de aprendizaje. Aunque la investigación en esta materia se encuentra fuera del ámbito de esta tesis, mi estancia de investigación en el Centre for Learning Sciences and Technologies en el año 2009, junto con mi participación en la JTEL Winter School 2010, nos han llevado a iniciar proyectos de investigación para explorar este tipo de tecnologías en el marco de la red europea de excelencia STELLAR, que reúne a una serie de organizaciones europeas de investigación en el campo de la tecnología educativa.

References

(Acero & Stern, 1990) Acero A, Stern R. Environmental robustness in automatic speech recognition. IEEE International Conference on Acoustics, Speech, and Signal Processing, 1999, p.849-852.

(Ahuja & Ensor, 2004) Ahuja SR, Ensor R. VoIP: What is it good for?. Queue, 2004;2(6):48-55.

(Ai & Weng, 2008) Ai H, Weng, F. User Simulation as Testing for Spoken Dialog Systems. Proceedings of the 9th SIGdial Workshop on Discourse and Dialogue, 2008, p.164–171.

(Alexander, 2004) Alexander B. Going nomadic: mobile learning in higher education. EDUCASE review 39, 2004;5:28-35.

(Alonso et al., 2002) Alonso C, Gallego D, Honey P. Los Estilos de aprendizaje: procedimientos de diagnóstico y mejora. Bilbao, Spain: Ediciones Mensajero, 2002.

(Ardito et al., 2006) Ardito C, Costabile MF, Marsico M, Lanzilotti R, Roselli S, Rossano V. An approach to usability evaluation of e-learning applications. Universal Access in the Information Society, 2006;4:270-283.

(Barbe & Milone, 1981) Brusilovsky P, Milone MN. What we know about modality strengths. Educational Leadership 1981:378–80.

(Beck & Fowler, 2000) Beck K, Fowler M. Planning Extreme Programming. Addisson-Wesley, 2000.

(Ben-Ari, 1998) Ben-Ari M. Constructivism in computer science education. Proceedings of the twenty-ninth Technical Symposium on Computer Science Education. ACM/SIGCSE, 1998;30(1):257-261.

(Bernsen et al, 1998) Bernsen NO, Dybkjar H, Dybkjar L. Designing Interactive Speech Systems: From First Ideas to User Testing. Springer-Verlag, Germany, 1998.

(Bloom et al., 1956) Bloom BS, Englehart MD, Hill WH, Furst EJ, Krathwohl DR. Taxonomy of educational objectives: Handbook 1: The cognitive domain, McKey, New York, 1956.

(Bolchini et al., 2006) Bolchini D, Colazzo S, Paolini P, Vitali D. Designing aural information architectures. Proceedings of the 24th anual conference on design of communication (SIGDOC'06). ACM Press, 2006.

(Boyce, 2000) Boyce SJ. Natural spoken dialogue systems for telephony applications. Communications of the ACM, 2000: 43(9); 29-34.

(Bridges & Hallinger, 1997) Bridges EM, Hallinger P. Using problem based learning to prepare educational leaders. Peabody Journal of Education, 1997; 72: 131-146.

(Brusilovsky et al., 1993) Brusilovsky P, Leonid P, Mikhail Z. Towards an Adaptive Hypermedia Component for an Intelligent Learning Environment. Springer-Verlag Human-Computer Interaction. Lecture Notes in Computer Science, 1993; 753: 348-358.

(Brusilovsky, 1996) Brusilovsky P. Methods and techniques of adaptive hypermedia. User Modeling and User-Adapted Interaction, 1996; 6(2-3): 87-129.

(Brusilovsky, 2001) Brusilovsky P. Adaptive Educational Hypermedia. Proceedings of the Tenth International PEG Conference, Tampere, Finland, 2001a, p.8–12.

(Chen & Macredie, 2002) Chen SY, Macredie RD. Cognitive styles and hypermedia navigation: development of a learning model. Journal of the American Society for Information Science and Technology, 2002; 53(1): 3-15.

(Chevrin et al., 2006) Chevrin V, Derycke A, Rouillard J. Project UBI-Learn: an intermediation infrastructure multi-channel access to future LMS. Proceedings of the Advanced International Conference on Telecommunications and International Conference on Internet and Web Applications and Services (AICT/ICIW 2006), IEEE Press, La Guadeloupe, France, 2006.

(Cohen and Oviatt, 1995) Cohen PR, Oviatt SL.The Role of Voice Input for Human-Machine Communication. In: Proceedings of the National Academy of Science, Washington, 1995;92(22):9921–9927.

(Cliford, 2000) Cliford R. Adaptive Hypermedia for Music Instruction. Proceedings of the Seventh international technological directions in music learning conference, 2000.

(Dagger et al., 2007) Dagger D, O'Connor A, Lawless S, Walsh E, Wade V. Service-Oriented E-Learning Platforms: From Monolithic Systems to Flexible Services. IEEE Internet Computing 2007; 11(3): 28–35.

(Dalgarno & Lee, 2010) Dalgarno B, Lee MJW. What are the learning affordances of 3-D virtual environments?. British Journal of Educational Technology, 2010;41(1):10-32.

(Davis et al., 1952) Davis KH, Biddulph R, Balashek S. Automatic Recognition of Spoken Digits. Journal of the Acoustical Society of America, 1952;24(6): 627-642.

(De Bra, 1996) De Bra P. Teaching Hypertext and Hypermedia through the Web. Journal of Universal Computer Science, 1996; 2(12): 797-804.

(De Bra & Calvi, 1998) De Bra P, Calvi L. AHA! An open adaptive hypermedia architecture. The New Review of Hypermedia and Multimedia, 1998; 4: 115-39.

(De Bra et al., 1999) De Bra P, Brusilovsky P, Houben GJ. Adaptive Hypermedia: From Systems to Framework. ACM Computer Surveys, 1999; 31:4.

(De Bra et al., 2002) De Bra P, Aerts A, Smits D, Stash N. AHA! Version 2.0, more adaptation flexibility for authors. Proceedings of the Wold Conference on e-learning (E-Learn 2002), 2002, p.240-246.

(Dougiamas & Taylor, 2003) Dougiamas M, Taylor P. Moodle: Using Learning Communities to Create an Open Source Course Management System, in World Conference on Educational Multimedia, Hypermedia and Telecommunications. Honolulu, Hawaii, 2003.

(Dudley et al., 1939) Dudley H, Riez RR, Watkins SA. A synthetic speaker. Elsevier Journal of the Franklin Institute, 1939; 227: 739-746.

(Dudley & Tarnoczy, 1950) Dudley H, Tarnoczy TH. The speaking machine of Wolfgang von Kempelen. Journal of the Acoustical Society of America, 1950; 22: 151-166.

(Duggan & Deegan, 2003) Duggan B, Deegan M. Considerations in the usage of text to speech (tts) in the creation of natural sounding voice enabled web systems. Proceedings of the 1st International sympoosium on Information and Communication Technologies, Trinity College, Dublin, 2003.

(Duggan, 2004) Duggan B. Creating effective, efficient and desirable voice enabled web interfaces. Proceedings of the 8th ERCIM Workshop "User Interfaces for All", June 2004.

(Dunn & Dunn, 1978) Dunn R, Dunn K. Teaching Students through their individual learning styles: a practical approach. Reston, VA: Reston Publishing; 1978.

(Felder & Silverman, 1988) Felder RM, Silverman LK. Learning and teaching styles in engineering education. Journal of Engineering Education, 1988; 78(7): 674–81.

(Felder & Spurlin, 2005) Felder RM, Spurlin J. Applications, Reliability and Validity of the Index of Learning Styles. International Journal of Engineering Education, 2005;21(1):103-112.

(Fielding, 2000) Fielding RT. Architectural styles and the design of network-based software architectures. PhD Thesis dissertation, University of California, Irvine, 2000.

(Fitzpatrick & Higgins, 1998) Fitzpatrick R, Higgins C. Usable software and its attributes: A synthesis of software quality, European Community law and human-computer interaction. People and Computers XIII. Proceedings of the HCI98 Conference, Springer, London, 1998.

(Følstad et al., 2001) Følstad A, Bae Brandtzæg P, Heim J. Usability analysis and evaluation of mobile ICT. Proceedings of the 8th International Symposium on Human Factors in Telecommunication (HFT 2001), Bergen, Norway, 2001.

(Forgie & Forgie, 1959) Forgie JW, Forgie CD. Results obtained from a vowel recognition computer program. Journal of the Acoustical Society of America, 1959; 31(11): 1480-1489.

(Fowler et al., 1999) Fowler M, Beck K, Brant J, Opdyke W, Roberts D. Refactoring: Improving the Design of Existing Code. Addison-Wesley, 1999.

(Frøkjær et al., 2000) Frøkjær E, Hertzum M, Hornbæk K. Measuring Usability: Are effectiveness, efficiency, and satisfaction really correlated?. Proceedings of the ACM CHI Conference on Human Factors in Computing Systems, Hague, The Netherlands, 2000, p. 345-352.

(Gamma et al., 1995) Gamma E, Helm R, johnson R, Vlissides J. Design Patterns: Elements of Reusable Object-Oriented Software. Addison-Wesley, Reading, Massachusetts, 1995.

(Giese & Gote, 2000) Giese JL, Gote JA. Defining consumer satisfaction. Academy of Marketing Science Review, 2000.

(Gilbert & Han, 2002) Gilbert JE, Han CY. Arthur: A personalized instructional system. Journal of Computing In Higher Education, 2002;14(1):113-129.

(Goldberg et al., 1996) Goldberg MW, Salari S, Swoboda P. World wide web - course tool: An environment for building www-based courses. Computer Networks and ISDN Systems, 1996; 28: 1219–1231.

(Hammersley, 2003) Hammersley B. Content Syndication with RSS. O'Reilly & Associates, 2003.

(HFES 200, 2001) HFES ANSI 200: Human Factors Engineering of Software User Interfaces. Human Factors and Ergonomics Society, 2001.

(ISO 9241, 1998) ISO 9241: Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability. International Standardisation Organisation (ISO), 1998.

(Laroussi & Benahmed, 1998) Laroussi M, Benahmed M. Providing an adaptive learning through the Web case of CAMELEON: Computer Aided MEdium for LEarning on Networks. Proceedings of the 4th International Conference on Computer Aided Learning and Instruction in Science and Engineering, Alvegard Editors, Goteborg, Sweden, 1998, p.411-416.

(Herrell, 2006) Herrell E. Evaluating speech self-service architectures. Forrester Research, 2006.

(Juang & Rabiner, 2005) Juang BH, Rabiner LR. Automatic speech recognition - a brief history of the technology. Elsevier Encyclopedia of Language and Linguistics, Second Edition, 2005.

(Kamm & Walker, 1997) Kamm C, Walker MA. Design and Evaluation of Spoken Dialog Systems. Proceedings of the 1997 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU). Santa Barbara, California, 1997.

(Keefe, 1979) Keefe JW. Learning Style: An Overview. Student Learning Styles: Diagnosing and Prescribing Programs. Reston, Virginia: National Association of Secondary School Principals, 1979.

(Klamma et al., 2007) Klamma R, Chatti MA, Duval E, Hummel H, Hvannberg EH, Kravcik M, Law E, Naeve A, Scott P. Social Software for Life-long Learning. Educational Technology & Society, 2007;10(3):72-83.

(Kobsa et al., 1999) Kobsa A, Koenemann J, Pohl W. Personalized hypermedia presentation techniques for improving on-line customer relationships. Technical report No. 66 GMD, German National Research Center for Information Technology, St. Augustin, Germany, 1999.

(Kolb, 1984) Kolb DA. Experiential learning experience as the source of learning and development. Englewood Cliffs, NJ:PrenticeHall; 1984.

(Koper & van Es, 2004) Koper R, van Es R. Modeling units of learning from a pedagogical perspective. Online education using learning objects (open and flexible learning). Routledge/Falmer, London, 2004.

(Kratzenstein, 1782) Kratzenstein CG. Sur la naissance de la formation des voyelles. Journal de Physique, 1782; 21: 358-380.

(Lai, 2000) Lai J. Conversational Interfaces. Communications of the ACM, 2000; 43(9): 24–27.

(Lippmann, 1997) Lippmann RP. Speech recognition by machines and humans. Elsevier Speech Communication, 1997; 22: 1-15.

(Luyten & Coninx, 2001) Luyten K, Coninx K. An XML-based runtime user interface description language for mobile computing devices. Proceedings of the eigth International Workshop on Design, Specification, and Verification of Interactive Systems (DSV-IS'2001). Glasgow, 2001, p. 20-29.

(Macaulay et al., 1990) Macaulay L, Fowler C, Kirby M, Hutt A. USTM: a new approach to requirements specification. Elsevier Interacting with Computers, 1990; 2(1): 92-118.

(March & Smith, 1995) March ST, Smith GF. Design and natural science research on information technology. Elsevier Decision Support Systems 1995; 15: 251-266.

(Mane et al., 1996) Mané A, Boyce S, Karis D, Yankelovich N. Designing the User Interface for Speech Recognition Applications. SIGCHI Bulletin, 1996; 28(4): 29-34.

(Masie, 2002) Masie E. Blended learning: the magic is in the mix. The ASTD E-Learning Handbook. NewYork: McGraw-Hill; 2002. p.58–63.

(Markowitz, 1996) Markowitz J. Using Speech Recognition. Prentice Hall, Prentice-Hall, Upper Saddle River, New Jersey, 2001.

(Mellor et al., 1996) Mellor BA, Barber C, Tunley C. Evaluating automatic speech recognition as a component of a multi-input device human-computer interface. Proceedings of the International Conference on Spoken Language Processing, Philadelphia, USA, 1996.

(Meyer, 2002) Meyer KA. Quality in distance education: Focus on on-line learning. ASHE–ERIC Higher Education Report 2002; 29: 1–121.

(Mikropoulos & Natsis, 2010) Mikropoulos TA, Natsis A. Educational virtual environments: A ten-year review of empirical research (1999-2009). Elsevier Computers & Education 2010; doi: 10.1016/j.compedu.2010.10.020.

(Moore, 2003) Moore RK. A comparison of the data requirements of automatic speech recognition systems and human listeners. Proceedings of the 8th European conference on speech communication and technology (EURO-SPEECH'03). Geneva, 2003, p. 2582-2584.

(Motiwalla, 2009) Motiwalla LF. A voice-enabled interactive services (VoIS) architecture for e-learning. International Journal on Advances in Life Sciences, 2009; 1(4): 122-133.

(Muntean & McManis, 2006) Muntean CH, McManis J. Fine grained content-based adaptation mechanism for providing high end-user quality of experience with adaptive hypermedia systems. Proceedings of the 15th International Conference on World Wide Web, ACM, New York, 2006, p.53-62.

(Myers, 1979) Myers JG. The Art of Software Testing. Wiely Interscience, New York, 1979.

(Niklfeld et al., 2001) Niklfeld G, Finan R,Pucher M. Architecture for adaptive multimodal dialog systems based on VoiceXML. Proceedings of the seventh European conference on speech communication and technology (Eurospeech'01), Aalborg, Denmark, 2001.

(Ogaya & Yano, 2004) Ogata H, Yano Y. Context-aware support for computer-supported ubiquitous learning. Proceedings of IEEE International Workshop on Wireless and Mobile Technologies in Education. IEEE Computer Society; 2004, p.27–34.

(Olson & Belar, 1956) Olson HF, Belar H. Phonetic Typewriter. Journal of the Acoustical Society of America, 1956; 28(6): 1072-1081.

(O'Malley et al., 2003) O'Malley C, Vavoula G, Glew J, Taylor J, Sharples M, Lefrere P. Guidelines for learning/teaching/tutoring in a mobile environment. Mobilearn project deliverable; 2003.

(Oppermann & Specht, 1998) Oppermann R, Specht M. Adaptive support for a mobile museum guide. Proceedings fo the Conference on Interactive Applications of Mobile Computing (IMB'98), Rostock, Germany, 1998.

(Patel, 2008) Patel AD. Music, language, and the brain. NewYork: Oxford University Press; 2008.

(Paule, 2008) Paule MP. Fernández MJ. Ortín F. Pérez JR. Adaptation in current e-learning systems. Elsevier Computer Standards and Interfaces, 2008; 30(1-2): 62-70.

(Pressman, 1987) Pressman RS. Software Engineering. A Practitioner's Approach. McGraw-Hill, New York, 1987.

(Ramakrishnan et al., 2004) Ramakrishnan I, Stent A, Yang G. Hearsay: Enabling Audio Browsing in Hypertext Content. Proceedings of the 13th International Conference on World Wide Web, 2004, p.80-89.

(Rodríguez, 2009) Rodríguez A. Plataforma de e-learning Sakai integrada con el gestor documental Alfresco. Final Year Project, Department of Computer Science, University of Oviedo, 2009.

(Roe, 1995) Roe DB. Deployment of human-machine dialogue systems. Proceedings of the national academy of science. Washington, 1995; 92: 10017-10022.

(Rose & Meyer, 2002) Rose HD, Meyer A. Teaching every student in the digital age: universal design for learning. Association for Supervision and Curriculum Development (ASCD); 2002.

(Rubio, 2008) Rubio I. Feedo, agregador RSS vocal. Final Year Project, Department of Computer Science, University of Oviedo, 2008.

(Rumbaugh et al., 1991) Rumbaugh J, Blaha M, Premerlani W, Eddy F, Lorensen W. Object-Oriented modeling and design. Prentice Hall, Englewood Cliffs, New Jersey, 1991.

(Sarasin, 1998) Sarasin LC. Learning style perspectives: impact in the classroom. Madison, Wisconsin: Atwood Publishing; 1998.

(Schiller et al., 2004) Schiller CA, Gan YM, Kemper C, Tydlitat B. Exploring new ways of enabling e-Government services for the Smart Home with speech interaction. Proceedings of the International conference on smart home and health telematics (ICOST'2004), Singapore, 2004.

(Schröder & Trouvain, 2003) Schröder M, Trouvain J. The German text-to-speech synthesis system MARY: A tool for research, development and teaching. International Journal of Speech Technology, 2003;6:365-377.

(Schwaber, 1995) Schwaber K. Scrum Development Process. Proceedings of the Conference on Object-Oriented Programming Systems, Languages, and Applications (OOPSLA'95) Workshop on Business Object Design and Implementation, 1995.

(Schwaber & Beedle, 2001) Schwaber K, Beedle M. Agile Software Development with Scrum. Prentice-Hall, Upper Saddle River, New Jersey, 2001.

(Shneiderman, 2000) Shneiderman B. The limits of speech recognition. Communications of the ACM 2000; 43(9): 63–65.

(Snyder, 2003) Snyder C. Paper prototyping: The fast and easy way to design and refine user interfaces. Morgan Kaufmann Publishers, London, 2003.

(Specht & Oppermann, 1998) Specht M, Oppermann R. ACE – Adaptive Courseware Environment. The New Review of Hypermedia and Multimedia, 1998;4:141-161.

(Stash et al., 2004) Stash N, Cristea A, De Bra P. Authoring of Learning Styles in Adaptive Hypermedia: problems and solutions. Proceedings of the 13th international world wide web conference, 2004, p.114-123.

(Vaittinen, 2003) Vaitinnen T. Guideline-supported user-centered design of multimodal speech-enabled TV-guide Software frameworks for developing electronic learning platforms. Department of Computer Science and Engineering, Helsinki University of Technology, 2003.

(van Harmelen, 2006) van Harmelen M. Personal Learning Environments. Sixth IEEE International Conference on Advanced Learning Technologies (ICALT06); 2006, p. 815-816.

(Wang, 2003) Wang Y. Assessment of learner satisfaction with asynchronous electronic learning systems. Elsevier Information and Management, 2003; 41: 75-86.

(Weber & Specht, 1997) Weber G, Specht M. User modelling and adaptive navigation support in WWW-based tutoring systems. Proceedings of the 6th International Conference on User Modeling, 1997, p. 289-300.

(Wheatstone, 1879) Wheatstone C. The Scientific Papers of Sir Charles Wheatstone. Physical Society of London, 1879.

(Zakaria & Brailsford, 2002) Zakaria MR, Brailsford TJ. User modeling and adaptive educational hypermedia frameworks for education. New Review of Hypermedia and Multimedia 2002; 8: 83-97.

(Zubizarreta, 2004) Zubizarreta, J. The learning portfolio: Reflective practice for improving student learning. Anker Publishing, Bolton, Massachusetts, 2004.