



Universidad
Carlos III de Madrid

Departamento de Informática

PROYECTO DE FIN DE GRADO

Desarrollo de un sistema multimodal para la planificación de viajes

Autor: Jennifer Catalina Vaquero

Tutor: David Griol Barres

Leganés, febrero de 2013

Título: Desarrollo de un sistema multimodal para la planificación de viajes
Autor: Jennifer Catalina Vaquero
Director: David Griol Barres

EL TRIBUNAL

Presidente: _____

Vocal: _____

Secretario: _____

Realizado el acto de defensa y lectura del Proyecto Fin de Carrera el día __ de _____
de 20__ en Leganés, en la Escuela Politécnica Superior de la Universidad Carlos III de
Madrid, acuerda otorgarle la CALIFICACIÓN de

VOCAL

SECRETARIO

PRESIDENTE

Agradecimientos

Me gustaría utilizar estas líneas para expresar mi más sincero agradecimiento a todas las personas que han confiado en mí y que sin su apoyo no me hubiera sido posible finalizar este proyecto.

Gracias a toda mi familia por su infinito apoyo y por estar a mi lado en todo momento, fácil o difícil.

Gracias a mi hermana, por todos los momentos, consejos y ayuda. Porque ‘nadie dijo que iba a ser fácil’ y porque esto es el comienzo de lo mejor que está por venir.

Gracias a todos mis compañeros por estos cuatro años de carrera. En especial, gracias a mi compañera de prácticas por todos los días de risas y prácticas que hemos pasado en la universidad y porque al final, lo hemos conseguido.

Gracias al tutor de este proyecto, David Griol, por su dedicación, ayuda y ánimo durante todo el transcurso del mismo. Le agradezco todo lo que he aprendido de él y el darme la oportunidad de involucrarme en el ámbito de los sistemas de diálogo.

Este proyecto de fin de grado supone el final de una etapa que no hubiese podido completar sin la ayuda, el apoyo y los consejos de muchas personas. A todos ellos, gracias.

Resumen

El objetivo principal de este trabajo de fin de grado es el desarrollo de una aplicación cuya interfaz de usuario es multimodal. Una interfaz multimodal permite al usuario interactuar con una aplicación o sistema utilizando diferentes modalidades para proveer las entradas y/o conocer las salidas generadas por la aplicación. Su objetivo fundamental es hacer más natural y completa la interacción hombre - máquina. En el caso de la aplicación desarrollada para este trabajo de fin de grado los modos de acceso seleccionados son el visual y el oral.

Con respecto al modo de acceso visual, la aplicación cuenta con una página web que es accesible desde cualquier navegador mediante los interfaces tradicionales (teclado, ratón y pantalla). Con respecto al modo de acceso oral, la aplicación cuenta con un sistema de diálogo a modo de portal de voz de forma que toda la información disponible en la web sea accesible oralmente.

Con la combinación de ambas modalidades de interacción con la aplicación, se consigue una mejor accesibilidad y usabilidad para el usuario contribuyendo a eliminar barreras de acceso, aumentando el número de entornos en los que la aplicación es accesible y posibilitando el acceso a personas con discapacidades motoras o visuales.

Como aplicación práctica de esta tecnología, se ha desarrollado un planificador de viajes, en el cual, a partir del destino que indique el usuario, puede consultarse cómo llegar, los medios de transporte disponibles, alojamiento, atracciones turísticas, así como información meteorológica. Los contenidos que facilita la aplicación se almacenan íntegramente en una base de datos y se actualizan dinámicamente de forma que la información ofrecida sea actual y verídica.

Para el desarrollo de la aplicación se han utilizado diversas tecnologías con el objetivo de hacerla dinámica y flexible y proporcionar la mayor accesibilidad a la información. Así pues, se han utilizado diferentes lenguajes de programación (SQL, PHP, HTML, VoiceXML), bases de datos (MySQL), servidores web y plataformas IVR (Voxeo Evolution).

El proyecto se complementa con una evaluación de la aplicación por parte de un grupo de usuarios y un estudio estadístico sobre la efectividad de su uso.

Palabras clave: interfaz multimodal, sistemas de dialogo, portal web, portal de voz.

Abstract

The main objective of this Bachelor's project is the development of a multimodal interface application. A multimodal interface provides the user with multiple input and output modes of interacting with a system or application. Its aim is to make human – computer interaction more natural and complete. In the developed application for this Bachelor project the chosen input and output modes are visual and oral.

Regarding the visual input, the application provides the user with a web page which can be accessed in any browser and by conventional interfaces (keyboard, mouse and screen). Regarding the voice input, the application provides the user with oral interaction so that the information presented on the web page is also accessible by spoken dialog.

In that way, the application is more accessible and usable and this contributes to the attempt of removing access barriers increasing the number of computing environments in which the application can be accessed.

Using this technology, the aim of this Bachelor project is to develop a travel scheduler in which once the user has specified the city he or she wants to travel to, he or she can check how to arrive there, transports available, accommodation facilities and tourist attractions. Moreover, it is also possible to check the current weather of the desired city. All the information provided by the application is stored in a data base that is dynamically updated so that the information is up to date and truthful.

In order to develop the application, diverse technologies have been used so that the application is dynamic, flexible and accesible. Different programming languages as SQL, PHP, HTML, and VoiceXML have been used, as well as data bases (MySQL), web servers and IVR platforms (Voxeo Evolution).

The project also provides with an evaluation of the developed application by a group of users as well as a stadistic study about its efectivity.

Keywords: multimodal interfaces, spoken language systems, web page, voice portal

Índice general

1. INTRODUCCIÓN	1
1.1 Introducción	1
1.2 Objetivos	3
1.3 Fases del desarrollo	4
1.3.1 Planificación.....	4
1.3.2 Desarrollo.....	5
1.3.3 Documentación.....	6
1.4 Planificación Temporal	7
1.5 Medios empleados.....	1
1.6 Estructura de la memoria.....	2
2. ESTADO DEL ARTE	4
2.1 Sistemas de Diálogo	4
2.1.1 Introducción a los Sistemas de Diálogo.....	4
2.1.2 Arquitectura de un Sistema de Diálogo.....	5
2.1.3 Clasificación de los Sistemas de Diálogo.....	8
2.1.4 Sistemas de Diálogo Existentes: Ejemplos representativos	10
2.2 Estándar VoiceXML	12
2.2.1 Introducción al estándar VoiceXML	12
2.2.2 Conceptos Básicos.....	15
2.2.3 Elementos de VoiceXML.....	16
2.2.4 Constructores de Diálogo.....	17
2.2.5 Gramáticas	21
2.3 Plataformas IVR.....	22
2.3.1 Voxeo	23
2.3.2 Otras plataformas IVR.....	31
3. DESCRIPCIÓN GENERAL DEL SISTEMA	32
3.1 Presentación del sistema desarrollado.....	32
3.2 Tecnologías empleadas	34
3.2.1 Base de datos: MySQL.....	34
3.2.2 Servidor web: x10Hosting	37

3.2.3	<i>Intérprete VoiceXML: Voxeo</i>	38
4.	DESCRIPCIÓN DETALLADA DEL MÓDULO DE ACCESO VISUAL	40
4.1	Introducción	40
4.2	Descripción de los módulos del acceso visual	41
4.2.1	<i>Módulo 'Inicio'</i>	41
4.2.2	<i>Módulo 'Información General'</i>	42
4.2.3	<i>Módulo 'Cómo Llegar'</i>	44
4.2.4	<i>Módulo 'Transportes'</i>	45
4.2.5	<i>Módulo 'Alojamiento'</i>	46
4.2.6	<i>Módulo 'Atracciones Turísticas'</i>	48
4.2.7	<i>Módulo 'Condición Meteorológica'</i>	49
4.2.8	<i>Módulo 'Barra de Inicio'</i>	49
5.	DESCRIPCIÓN DETALLADA DEL MÓDULO DE ACCESO ORAL	52
5.1	Introducción	52
5.2	Gramáticas	54
5.2.1	<i>Gramáticas estáticas</i>	55
5.2.2	<i>Gramáticas dinámicas</i>	55
5.3	Descripción de los módulos del acceso oral	58
5.2.3	<i>Módulo 'Inicio'</i>	58
5.2.4	<i>Módulo 'Información General'</i>	59
5.2.5	<i>Menú</i>	60
6.	EVALUACIÓN DE LA APLICACIÓN	69
6.1	Metodología de evaluación	69
6.2	Resultados de la evaluación	73
7.	CONCLUSIONES Y TRABAJO FUTURO	78
7.1	Conclusiones	78
7.2	Trabajo futuro	80
8.	PRESUPUESTO	82
9.	GLOSARIO	85
10.	REFERENCIAS	87

Índice de figuras

<i>Figura 1 Diagrama de Gantt con la planificación temporal del proyecto</i>	1
<i>Figura 2 Arquitectura general de un sistema de diálogo</i>	5
<i>Figura 3 'Hello World' en VoiceXML</i>	13
<i>Figura 5 Ejemplo menú en VoiceXML</i>	21
<i>Figura 6 Vista general de una cuenta en Voxeo</i>	24
<i>Figura 7 Funcionalidad Quick Start Guide ofrecida por Voxeo</i>	24
<i>Figura 8 Funcionalidad Application Manager ofrecida por Voxeo</i>	25
<i>Figura 9 Configuración de una aplicación en Voxeo</i>	26
<i>Figura 10 Información de contacto de una aplicación</i>	26
<i>Figura 11 Funcionalidad Voxeo Designer ofrecida por Voxeo</i>	27
<i>Figura 12 Funcionalidad Application Analytics ofrecida por Voxeo</i>	27
<i>Figura 13 Funcionalidad Application Debugger ofrecida por Voxeo</i>	28
<i>Figura 14 Funcionalidad Device Monitoring ofrecida por Voxeo</i>	29
<i>Figura 15 Funcionalidad File, Logs, Reports ofrecida por Voxeo</i>	29
<i>Figura 16 Funcionalidad Support Tickets ofrecida por Voxeo</i>	30
<i>Figura 17 Funcionalidad Users & Contacts ofrecida por Voxeo</i>	30
<i>Figura 18 Funcionalidad Account Journal ofrecida por Voxeo</i>	31
<i>Figura 19 Información ofrecida por la aplicación</i>	33
<i>Figura 20 Arquitectura de la aplicación</i>	34
<i>Figura 21 Captura de pantalla de la aplicación phpMyAdmin</i>	35
<i>Figura 22 Modelo Relacional de la base de datos desarrollada para la aplicación</i>	36
<i>Figura 23 Funciones para el manejo de la base de datos</i>	37
<i>Figura 24 Localización de la URL inicial</i>	38
<i>Figura 25 Modulo de inicio desarrollado para la aplicación</i>	42
<i>Figura 26 Mensaje informativo de la aplicación</i>	42
<i>Figura 27 Query consulta</i>	43
<i>Figura 28 Módulo de información general desarrollado para la aplicación</i>	44
<i>Figura 29 Módulo Como llegar desarrollado para la aplicación</i>	45
<i>Figura 30 Módulo Transportes desarrollado para la aplicación</i>	46

<i>Figura 31 Módulo Alojamiento desarrollado para la aplicación</i>	<i>47</i>
<i>Figura 32 Módulo Atracciones Turísticas desarrollado para la aplicación.....</i>	<i>48</i>
<i>Figura 33 Módulo Condición Meteorológica desarrollado para la aplicación.....</i>	<i>49</i>
<i>Figura 34 Módulo Barra de inicio desarrollado para la aplicación.....</i>	<i>50</i>
<i>Figura 35 Listado de destinos.....</i>	<i>50</i>
<i>Figura 36 Funcionalidad Acerca de desarrollado para la aplicación.....</i>	<i>51</i>
<i>Figura 37 Módulos de la aplicación de voz.....</i>	<i>53</i>
<i>Figura 38 Ejemplo de gramática estática.....</i>	<i>55</i>
<i>Figura 39 Método crear_gramatica que genera una gramática dinámica</i>	<i>57</i>
<i>Figura 40 Flujo de datos y ficheros del módulo inicio</i>	<i>58</i>
<i>Figura 41 Caso de uso. Módulo Inicio</i>	<i>59</i>
<i>Figura 42 Caso de uso. Módulo Información general.....</i>	<i>60</i>
<i>Figura 43 Flujo de datos y ficheros del módulo menú</i>	<i>60</i>
<i>Figura 45 Caso de uso. Menú</i>	<i>61</i>
<i>Figura 46 Caso de uso. Módulo cómo llegar</i>	<i>63</i>
<i>Figura 47 Caso de uso. Módulo transportes</i>	<i>64</i>
<i>Figura 48 Caso de uso. Módulo alojamiento.....</i>	<i>66</i>
<i>Figura 49 Caso de uso. Módulo atracciones turísticas</i>	<i>67</i>
<i>Figura 50 Caso de uso. Módulo condición meteorológica.....</i>	<i>68</i>
<i>Figura 51 Cuestionario de evaluación</i>	<i>71</i>
<i>Figura 52 Formulario de evaluación de la aplicación.....</i>	<i>72</i>
<i>Figura 53 Resultados de la evaluación. Pregunta 1</i>	<i>73</i>
<i>Figura 54 Resultados de la evaluación. Pregunta 2.....</i>	<i>73</i>
<i>Figura 55 Resultados de la evaluación. Pregunta 3.....</i>	<i>73</i>
<i>Figura 56 Resultados de la evaluación. Pregunta 4.....</i>	<i>74</i>
<i>Figura 57 Resultados de la evaluación. Pregunta 5.....</i>	<i>74</i>
<i>Figura 58 Resultados de la evaluación. Pregunta 6.....</i>	<i>74</i>
<i>Figura 59 Resultados de la evaluación. Pregunta 7.....</i>	<i>75</i>
<i>Figura 60 Resultados de la evaluación. Pregunta 8.....</i>	<i>75</i>
<i>Figura 61 Resultados de la evaluación. Pregunta 9.....</i>	<i>75</i>
<i>Figura 62 Resultados de la evaluación. Pregunta 10.....</i>	<i>76</i>
<i>Figura 63 Resultados de la evaluación. Pregunta 11.....</i>	<i>76</i>
<i>Figura 64 Resultados de la evaluación. Pregunta 12.....</i>	<i>76</i>

Índice de tablas

<i>Tabla 1 Elementos de VoiceXML</i>	17
<i>Tabla 2 Atributos de un formulario</i>	18
<i>Tabla 3 Ítems de entrada en VoiceXML</i>	18
<i>Tabla 4 Ítems de control en VoiceXML</i>	19
<i>Tabla 5 Atributos de los ítems de formulario en VoiceXML</i>	19
<i>Tabla 6 Atributos del elemento menú en VoiceXML</i>	20
<i>Tabla 7 Atributos del elemento choice en VoiceXML</i>	21
<i>Tabla 8 Elementos SRGS (XML Form)</i>	22
<i>Tabla 9 Desglose de costes en recursos humanos</i>	83
<i>Tabla 10 Coste total del proyecto</i>	84

Capítulo 1

Introducción

1.1 Introducción

Hoy en día, gracias a los avances tecnológicos de las últimas décadas, es posible acceder a información en la red desde dispositivos inteligentes que facilitan este acceso de manera prácticamente instantánea, desde cualquier lugar y en cualquier momento. Con el objetivo de proporcionar la información a los usuarios de la forma más simple y eficiente surge la necesidad de diseñar interfaces usables que se adapten a las diferentes necesidades de los usuarios.

Se conoce como interfaces multimodales a aquellos dispositivos que utilizan varias modalidades de entrada y/o salida (por ejemplo, imagen, vídeo, texto, voz) para llevar a cabo la interacción con el usuario [1.1].

En las últimas décadas, la interacción multimodal ha tenido especial relevancia y ha sido ampliamente estudiada dentro del ámbito de la comunicación humano – máquina. Esto se debe principalmente a dos motivos:

- Las interfaces multimodales suponen una forma “más humana” de interactuar con el ordenador. Estas formas engloban el diálogo, gestos u otras modalidades que generalmente prefieren los usuarios ante las modalidades convencionales.
- Está demostrado que las interfaces multimodales ofrecen una mayor flexibilidad y eficacia que otras modalidades de interacción [1.2].

CAPÍTULO: Introducción

El presente trabajo de fin de grado pretende desarrollar una aplicación cuya interfaz sea multimodal y aborde dos modalidades de comunicación con sus usuarios, el modo visual mediante una interfaz web y el modo oral mediante una interfaz de voz. La aplicación se ha desarrollado para un dominio específico: un portal web y un portal de voz que ayudan a los usuarios a planificar sus viajes proporcionando diferentes tipos de información sobre el destino que el usuario vaya a visitar.

Actualmente, existen diferentes portales que ayudan a los usuarios a planificar sus viajes, buscar billetes de avión o tren así como información de hoteles y atracciones turísticas. Sin embargo, la mayoría de estos portales están focalizados en ofrecer un único servicio. Esto supone que los usuarios se ven obligados a consultar distintos portales y aplicaciones para obtener información completa del destino que van a visitar. Así pues, la aplicación desarrollada para este trabajo de fin de grado propone centralizar la información que se ha considerado más importante a la hora de planificar un viaje en un mismo portal web y además dotar a dicha aplicación de la accesibilidad mediante la voz.

La interfaz web presenta información a los usuarios de forma visual y es accesible desde cualquier navegador en cualquier momento que el usuario desee. Las páginas web se dividen en estáticas o dinámicas. Las páginas web estáticas están constituidas por contenidos estáticos a los que se le da formato con etiquetas HTML (*HyperText Markup Language*) [1.3]. En cambio, las páginas web dinámicas, tal y como su nombre indica, se generan de forma dinámica en tiempo de ejecución dependiendo de múltiples factores, como el paso de parámetros, cookies o recuperación de información de una base de datos. Estas últimas páginas, al poder contener aplicaciones dentro de la propia página web, son más interactivas con los usuarios. Ejemplos de páginas web dinámicas son las desarrolladas para foros y libros de visita, atención al cliente personalizada, pedidos online, etc.

Respecto a la estructura y organización de la información en las páginas web, existen directrices cuya intención es normalizar el diseño de forma que se facilite la visualización e interpretación del contenido. La iniciativa para la accesibilidad web (WAI, *Web Accessibility Initiative*) [1.4] es una rama de W3C (*World Wide Web Consortium*) [1.5] que se encarga de publicar guías de accesibilidad al contenido web que deben de seguir los desarrolladores web. En función del grado de accesibilidad que tenga cada web será catalogada con criterios mínimos de accesibilidad, extendidos, o accesibilidad máxima.

No obstante, dado que la forma de acceder a la información a través de la interfaz web es exclusivamente visual, surge la necesidad de buscar nuevas interfaces que faciliten el acceso a la información de forma más natural e intuitiva posibilitando que la información sea accesible desde nuevos entornos que se adapten a las necesidades de los usuarios. Se busca eliminar barreras de acceso de forma que los usuarios que por diversos motivos no puedan acceder a una interfaz visual dispongan de más alternativas.

Actualmente, las interfaces de voz son un componente muy importante de las interfaces multimodales, dado que una interfaz de voz sofisticada puede ser de gran utilidad para proporcionar acceso a la información a distintos tipos de usuario de una forma sencilla y natural. La incorporación del diálogo en un sistema informático, permite la interacción persona-máquina usando el habla para la solución de diversos problemas cotidianos [1.6]: comprar un billete de tren, reservar un vuelo, proporcionar información meteorológica, etc.

De este modo, los sistemas de diálogo tienen como objetivo facilitar la interacción natural mediante el habla entre una persona y un sistema informático [1.7]. Un sistema de diálogo ideal reconoce el habla sin restricciones y proporciona respuestas bien formadas, con sentido y voz completamente natural. Sin embargo, existen algunas limitaciones en los sistemas de diálogo, como pueden ser la limitación de la comprensión a un dominio restringido o problemas que pueda suponer el diálogo espontáneo por parte del usuario al interactuar con la máquina. Estas limitaciones son objeto de estudio e investigación hoy en día.

Para la implementación de los sistemas de dialogo, el W3C propone el estándar VoiceXML (*Voice Extensible Markup Language*) [1.8]. VoiceXML es un lenguaje con etiquetas que permite navegar mediante la voz como modo de interacción con el ordenador, en lugar del teclado, ratón u otros periféricos de interfaz persona-máquina. Algunas de las características con las que cuenta este lenguaje son el habla sintetizada, reproducción de audio digital, reconocimiento de entrada oral y reconocimiento de tonos DTMF (*Dual-Tone Multi-Frequency*), así como grabación de la entrada de voz y funciones de telefonía (reconocimiento del número del usuario, desconexión de llamadas, transferencia de llamadas, etc.)

VoiceXML es independiente de la plataforma por lo que permite portabilidad y transferencia entre aplicaciones heterogéneas. Esta característica es muy importante, ya que si se utiliza el lenguaje VoiceXML junto con otros estándares se puede conseguir una base para crear aplicaciones basadas en los sistemas de diálogo eficientes y eficaces en diferentes dominios.

1.2 Objetivos

El objetivo fundamental de este trabajo de fin de grado es, tal y como se ha mencionado anteriormente, la implementación de una aplicación cuya interfaz sea multimodal y permita a los usuarios acceder al contenido de dos formas diferentes:

- De forma visual mediante un portal web. Dicho portal web está alojado en un servidor web y puede ser accedido desde cualquier navegador siendo necesario un dispositivo electrónico con conexión a Internet.
- De forma oral mediante un portal de voz. Para ello se ha desarrollado un sistema de diálogo en el que los usuarios pueden interactuar mediante voz y/o señalización DTMF con la aplicación y acceder al mismo contenido que existe en el portal web.

Este objetivo principal se complementa con los siguientes:

- Analizar el estado del arte correspondiente a las aplicaciones multimodales existentes en la actualidad para poder así aplicar todas sus posibilidades en el desarrollo de la nueva aplicación

CAPÍTULO: Introducción

- Llevar a cabo un estudio de la información que se considera más útil para poder planificar un viaje e integrarla en la aplicación.
- Utilizar el potencial del desarrollo web de forma que el portal web sea accesible y de fácil uso y que los usuarios consideren que se trata de un método cómodo para obtener la información que desean.
- Utilizar el potencial de los sistemas de dialogo y su implementación de forma que la aplicación de voz sea útil y de fácil uso y que los usuarios consideren que se trata de un método cómodo para obtener la información que desean.
- Llevar a cabo un estudio detallado de los sistemas de diálogo incluyendo su arquitectura, criterios para su clasificación y principales funcionalidades.
- Completar un estudio detallado del estándar VoiceXML, así como de los demás lenguajes necesarios para implementar todas las funcionalidades de la aplicación de voz.
- Realizar un estudio del sistema de gestión de bases de datos relacional MySQL para que la base de datos que requiere la aplicación tenga el diseño más eficiente y se reduzca el tiempo requerido para el acceso a la información.
- Facilitar el acceso a la aplicación a usuarios con impedimentos físicos o visuales que no pueden acceder a la web y su única vía de acceso a la información es el acceso oral. También se pretende facilitar el acceso a la aplicación a personas que no dispongan de un dispositivo y/o conocimientos informáticos para acceder al portal web, como pueden ser personas de la tercera edad que generalmente están más acostumbradas al uso del teléfono.

1.3 Fases del desarrollo

El presente trabajo se ha llevado a cabo en diferentes etapas atendiendo a las fases de desarrollo que siguen los sistemas de software del mercado que se desarrollan actualmente. Dichas fases son la de planificación, desarrollo y documentación.

A continuación se detallan las fases citadas anteriormente.

1.3.1 Planificación

En las primeras etapas del proyecto se definieron los objetivos generales y el funcionamiento deseado para la aplicación. Además, se llevó a cabo un estudio de las tecnologías necesarias para llevar a cabo la implementación de la aplicación, así como una búsqueda exhaustiva de documentación relativa a las citadas tecnologías.

Así pues, gracias a artículos de divulgación científica, manuales y foros especializados, se completó una búsqueda y estudio de las siguientes tecnologías necesarias para el desarrollo de la aplicación y consecución de los objetivos:

- Gestores de bases de datos.
- Lenguajes de programación PHP, Javascript, hojas de estilo CSS.
- Servidores web.
- Sistemas de diálogo.
- Plataforma Voxeo.
- Lenguaje VoiceXML

En esta primera fase se han llevado a cabo diversas aplicaciones prácticas a modo de ejemplo para consolidar los conocimientos adquiridos de las tecnologías citadas y establecer una base sólida de la cual partir para realizar la aplicación.

1.3.2 Desarrollo

La fase de desarrollo se compone de distintas subfases que se describen a continuación:

1.3.2.1 Análisis y diseño

En esta fase se ha efectuado un análisis de las distintas funcionalidades en los diferentes módulos y submódulos del sistema. Así pues, se han analizado los requisitos indispensables que deben tener la aplicación web y la aplicación de voz para satisfacer los objetivos.

Se ha definido el diseño de la página web estableciendo el estilo y distribución de los elementos principales que la componen. La página web de inicio, que será la primera página que visite un usuario al acceder a la aplicación, muestra un formulario a través del cual se solicita la ciudad de la que se desea conocer información al mismo tiempo que se informa al usuario del funcionamiento de la aplicación y cómo puede acceder a la aplicación de voz. Esta página de inicio dispone de un menú que será visible en todas las restantes páginas del portal web a través del cual el usuario podrá volver al inicio, consultar la información relativa a la aplicación y acceder a un listado actualizado de los destinos que están dados de alta y por tanto disponibles en la aplicación. Una vez que el usuario ingresa el destino, el diseño y menús de páginas a través de las cuales puede navegar es similar de forma que la navegación por el portal sea sencilla.

La aplicación de voz se ha diseñado de forma análoga a la del portal web con el objetivo de que la navegación en la aplicación de voz sea idéntica a la del portal web y esto facilite que los usuarios no tengan dificultades en utilizar las dos interfaces simultáneamente o en momentos diferentes.

Por último, también se ha llevado a cabo el diseño de la base de datos que utiliza la aplicación. La base de datos es un elemento esencial para el funcionamiento de la aplicación por lo que se ha realizado un diseño óptimo de tablas y columnas de forma que las consultas a la base de datos sean rápidas y eficaces.

1.3.2.2 Desarrollo

Esta fase engloba la implementación de todos los módulos y submódulos de la aplicación, tanto para el interfaz web como para el de voz.

Además, se han creado las tablas y estructuras necesarias de la base de datos para dotar de contenidos a la aplicación y poder llevar a cabo las evaluaciones necesarias en condiciones reales de funcionamiento.

1.3.2.3 Evaluación

En esta fase se han llevado a cabo pruebas tanto unitarias como de integración.

Las pruebas unitarias han permitido determinar si cada módulo de la aplicación está listo y correctamente terminado. Las pruebas de integración han estado enfocadas a la interacción entre unidades.

Se ha prestado una especial atención y dedicación en la fase de pruebas para asegurar que la versión final está libre de errores y se trata de una versión estable.

Dentro de la fase de pruebas se incluye una evaluación de la aplicación mediante una encuesta de valoración subjetiva a un grupo de usuarios y el análisis de los resultados obtenidos para un conjunto de parámetros estadísticos.

1.3.3 Documentación

Durante el desarrollo de proyecto se ha ido documentando la información que se ha considerado importante para el mismo. Se ha redactado el presente documento de memoria del trabajo final de grado y se ha preparado la presentación.

1.4 Planificación Temporal

Una vez establecidos los objetivos del proyecto y las fases de desarrollo del mismo se ha realizado la planificación temporal con la ayuda de un diagrama de Gantt. Este tipo de diagrama muestra el tiempo de dedicación previsto para cada tarea a lo largo del tiempo previsto para la realización del proyecto.

Al comienzo y finalización del proyecto se han establecido dos hitos indicando la fecha de inicio y fin del mismo. Los hitos son tareas de duración nula que se utilizan para marcar puntos de control en el transcurso del proyecto.

El diagrama de Gantt resultante se puede observar en la Figura 1. En este diagrama se muestra una visión global de la planificación del proyecto, de sus distintas fases y de su duración.

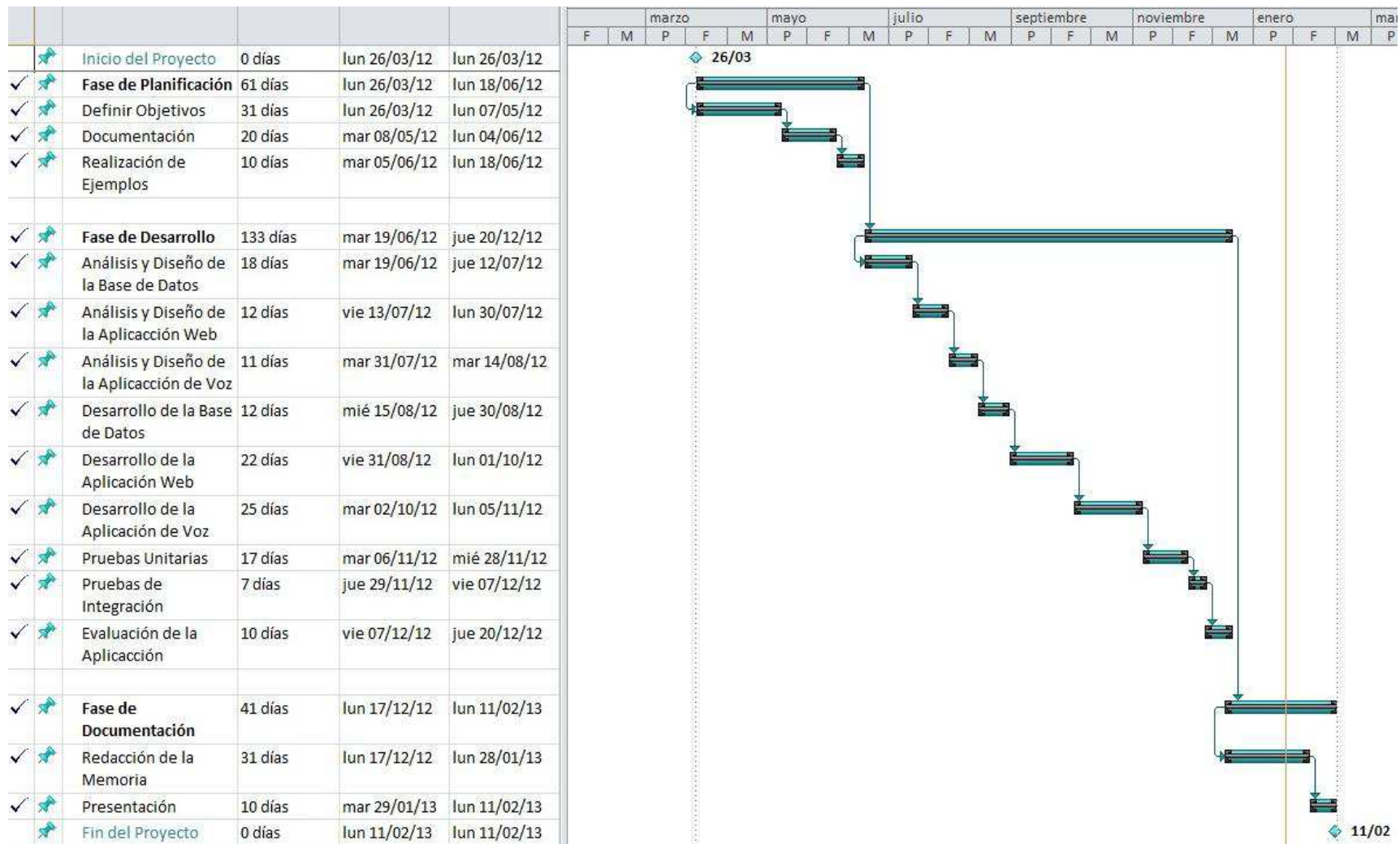


Figura 1 Diagrama de Gantt con la planificación temporal del proyecto

1.5 Medios empleados

En esta sección se presentan los medios que han sido necesarios para el desarrollo de la aplicación. Distinguiremos entre dispositivos hardware, software y lenguajes de programación.

- Dispositivos hardware

Los dispositivos hardware necesarios han sido un ordenador, auriculares con micrófono, periféricos habituales (ratón y teclado), así como un servidor web para albergar la aplicación.

- Software

Para generar todo el código se ha utilizado el editor de textos *Notepad ++*. También han sido muy importantes para la implementación de la aplicación la plataforma *Voxeo* [1.9] y la plataforma *phpMyAdmin* [1.10] para la gestión de la base de datos.

Por último, hay que destacar la importancia de la herramienta *Skype* [1.11] para la realización de las pruebas de la aplicación de voz, así como el paquete *Microsoft Office 2010* utilizado para la redacción de la memoria.

- Lenguajes de programación

Para el desarrollo de la aplicación web ha sido necesario usar lenguajes de programación web tales como PHP, HTML y hojas de estilo CSS para el estilo de la página web.

Para el desarrollo de la aplicación de voz se ha utilizado VoiceXML como lenguaje de programación principal. No obstante, para dotar a la aplicación de la funcionalidad deseada ha sido necesario integrar otros lenguajes como PHP y Javascript.

Por último, para el desarrollo de la base de datos se ha utilizado el gestor MySQL para la creación y gestión de la base de datos necesaria para la inmensa cantidad de datos de los que hace uso la aplicación.

Además de las categorías descritas anteriormente, ha sido necesario hacer uso de numerosa documentación para el correcto desarrollo. Dentro de esta documentación hay que destacar artículos y manuales acerca de los lenguajes de programación así como de las plataformas utilizadas.

1.6 Estructura de la memoria

Para facilitar la lectura de la memoria, se incluye a continuación un breve resumen de cada capítulo:

Capítulo 1: Introducción y objetivos. En este capítulo se establece el propósito y los objetivos de todo el contenido del proyecto de fin de grado. Además, se incluyen las fases de desarrollo, la planificación temporal, los medios empleados y la estructura de la memoria.

Capítulo 2: Estado del Arte. En este capítulo se hace un estudio completo de los sistemas de diálogo, se analiza en detalle el lenguaje estándar VoiceXML y se presenta la plataforma utilizada para la implementación de la aplicación, Voxeo.

Capítulo 3: Descripción general del sistema. En este capítulo se proporciona una visión global de la aplicación desarrollada y se analizan las tecnologías empleadas para ello.

Capítulo 4: Descripción detallada del módulo de acceso visual. En este capítulo se proporciona una descripción detallada de la interfaz y funcionalidad del portal web desarrollado.

Capítulo 5: Descripción detallada del módulo de acceso oral. En este capítulo se proporciona una descripción detallada de la interfaz y funcionalidad de los módulos que componen el portal de voz desarrollado.

Capítulo 6: Evaluación de la aplicación. En este capítulo se lleva a cabo una evaluación de la aplicación desarrollada prestando especial atención a la aplicación de voz. La evaluación se basa en las opiniones subjetivas recogidas en un cuestionario de un grupo de usuarios.

Capítulo 7: Conclusiones y trabajo futuro. Se exponen las principales ideas, cuestiones y conclusiones derivadas de la realización del proyecto, así como las posibles líneas de investigación que se podrían llevar a cabo partiendo de este proyecto.

Presupuesto. Este apartado contiene un análisis de los costes del diseño y desarrollo del proyecto, detallando el coste de personal y del material necesario para llevar a cabo su realización.

Glosario. En este apartado se recogen los principales términos y conceptos técnicos utilizados en la memoria, con el objetivo de facilitar su comprensión al lector.

Referencias. En este apartado se reflejan las citas bibliográficas que se han consultado para la realización tanto del proyecto como de la memoria.

Capítulo 2

Estado del Arte

2.1 Sistemas de Diálogo

2.1.1 Introducción a los Sistemas de Diálogo

Los sistemas de diálogo (SLS, *Spoken Language Systems*) [2.1] [2.2] son una tecnología concebida para facilitar la interacción natural mediante el habla entre una persona y un ordenador. Constituyen un intermediario entre un usuario y un sistema informático, que tiene la ventaja de no requerir un teclado o ratón ya que recurren al medio de comunicación propio de los humanos, la voz. Es por esto por lo que los sistemas de diálogo se suelen usar en servicios que se proporcionan a través del teléfono, aunque también se usan en otros casos en los que la lengua oral permite llevar a cabo una determinada tarea de forma flexible y eficaz.

Las capacidades ideales de un sistema de diálogo engloban el reconocimiento del habla espontánea sin dificultades, la comprensión de todo tipo de contenidos, la capacidad de proporcionar respuestas completas con sentido, gramaticalmente bien formadas y pragmáticamente adecuadas. Otras capacidades ideales de un sistema de diálogo son el uso de una voz completamente natural que no sea distinguible a la de un humano por parte de la máquina y una interacción multimodal a través del habla y la imagen.

No obstante, los sistemas de diálogo que podemos usar hoy en día están sujetos a las limitaciones del reconocimiento del habla y condicionados por la naturalidad del

habla sintetizada. Otra limitación es que las respuestas están condicionadas a un dominio o ámbito específico.

Sin embargo, el uso de los sistemas de diálogo se está extendiendo cada vez más en aplicaciones reales, tanto en el entorno de la administración pública como en el de la empresa privada. El principal motivo de esta extensión de su uso se debe a que los sistemas de diálogo permiten ofrecer un servicio telefónico continuado durante 24 horas a lo largo de los 365 días del año, y al mismo tiempo, mejoran las prestaciones que proporcionan al cliente al limitar el trabajo de los operadores humanos a los problemas complejos que requieren atención por parte de un operador humano [2.3].

Actualmente, el diseño de sistemas de diálogo orales es un campo muy estudiado y en el que se desarrolla actualmente una mayor investigación dentro de los campos del procesamiento del lenguaje natural y las tecnologías del habla. Desarrollar interfaces orales posibilita dos objetivos fundamentales, permiten el acceso a aplicaciones en entornos en los que no es posible utilizar interfaces convencionales y posibilitan el acceso a dichas aplicaciones a personas con discapacidades visuales o motoras [2.4].

2.1.2 Arquitectura de un Sistema de Diálogo

Para llevar a cabo sus funciones, un sistema de diálogo debe de estar formado por una serie de módulos. Cada módulo está especializado en un aspecto concreto de la interacción entre la persona y la máquina. La Figura 2 muestra la arquitectura de un sistema de diálogo basado en el procesamiento del habla y la descripción y principales características de cada módulo.

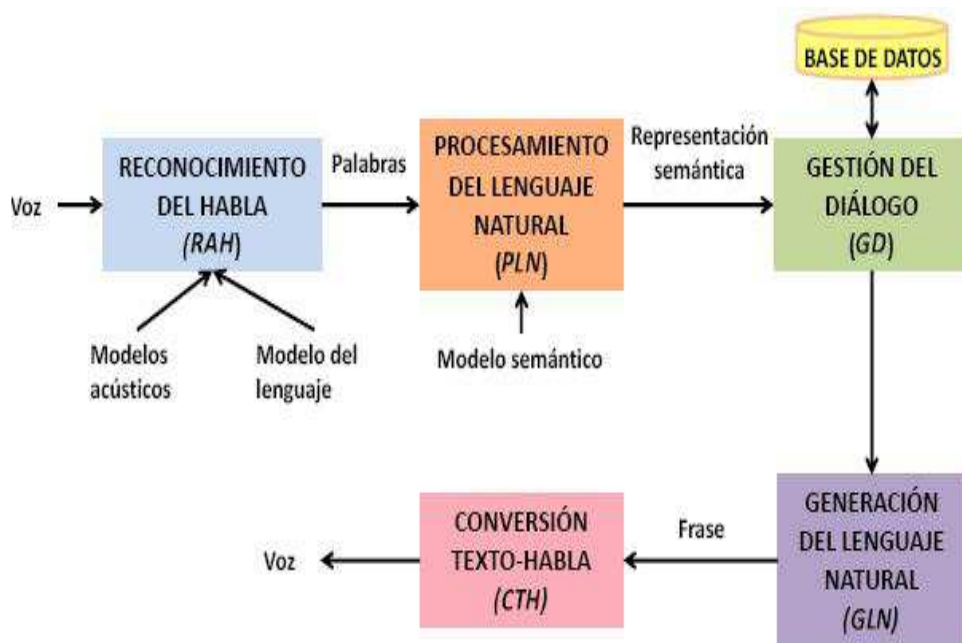


Figura 2 Arquitectura general de un sistema de diálogo

- Reconocimiento del habla

Este módulo es el encargado de reconocer la señal pronunciada por el usuario y proporcionar la secuencia de palabras reconocida cuya aparición sea más frecuente [2.4].

Se trata de un módulo con una alta dificultad en su desarrollo dado que la señal pronunciada por el usuario puede estar incompleta, ser ambigua o verse afectada por factores como el tono de voz, edad, sexo o acento.

Para poder ejecutar la tarea de reconocimiento del habla del interlocutor, el sistema utiliza diferentes metodologías. La más utilizada es la aproximación estadística. Esta metodología trata de encontrar una secuencia de palabras W dada una secuencia de ondas acústicas A . Esta secuencia W puede determinarse a partir de la siguiente expresión:

$$W = \max_W P(W | A)$$

Para deducir la probabilidad $P(W|A)$ se aplica la regla de Bayes, por lo que la expresión anterior se puede redefinir de la siguiente forma:

$$P(W | A) = \frac{P(A | W) P(W)}{P(A)}$$

Sin embargo, como la probabilidad de la secuencia de ondas acústicas $P(A)$ es la misma independientemente de la secuencia de palabras pronunciada en el proceso de maximización, esta probabilidad puede ser eliminada (la secuencia de palabras que da el máximo no varía). De esta forma se obtiene la fórmula fundamental del reconocimiento automático del habla:

$$W = \max_W P(A | W) P(W)$$

- Procesamiento del lenguaje natural

Este módulo es el encargado de extraer el significado semántico de las palabras reconocidas por el módulo anterior. Para poder llevar a cabo esta tarea, el módulo de procesamiento de lenguaje natural lleva a cabo análisis morfológicos, sintácticos, semánticos y pragmáticos.

El análisis morfológico consiste en detectar la relación que se establece entre las unidades mínimas que forman una palabra, como puede ser el reconocimiento de sufijos o prefijos. Este tipo de análisis mantiene una estrecha relación con el lexema de las palabras. El lexema contiene la raíz de las palabras con formas regulares, siendo el analizador morfológico el que se encarga de determinar si el género, número o flexión que componen el resto de la palabra son adecuados.

El análisis sintáctico tiene como función clasificar cada uno de los componentes sintácticos que aparecen en la oración y analizar cómo las palabras se combinan para formar construcciones gramaticalmente correctas. El resultado de este proceso consiste en generar la estructura correspondiente a las categorías sintácticas formadas por cada una de las unidades léxicas que aparecen en la oración.

El análisis semántico trata de extraer el significado de la estructura de la frase completa a partir de sus diferentes componentes. Dado que definir el significado no es una tarea sencilla, pueden dar lugar diversas interpretaciones. Así pues, es posible distinguir entre significado independiente y significado dependiente del contexto. El primero, hace referencia al significado que las palabras tienen por sí mismas sin considerar el significado adquirido según el uso en una determinada circunstancia. El segundo, hace referencia al componente significativo de una frase asociado a las circunstancias en que ésta se da.

El análisis pragmático añade información adicional al análisis del significado de la frase en función del contexto donde aparece. Se trata de uno de los niveles de análisis más complejos, cuya finalidad es incorporar al análisis semántico la aportación significativa que pueden hacer los participantes, la evolución del discurso o información presupuesta [2.5].

- Gestión de diálogo

El módulo de gestión de diálogo se trata del módulo más complejo ya que es el que más tareas realiza. Así pues, este módulo se encarga de considerar la interpretación semántica de la petición del usuario, la historia del proceso del diálogo, la información de la aplicación disponible así como el estado del sistema y a partir de esto decide cual va a ser la siguiente acción que debe tomar el sistema [2.4].

Para la implementación de los gestores de diálogo se utilizan diferentes metodologías. La metodología más sencilla se basa en modelar el diálogo como una máquina de estados finitos, donde el sistema utiliza las acciones de los usuarios para modelar las transiciones entre las respuestas del sistema. Por otra parte, la metodología basada en *frames* se basa en crear una estructura compuesta de campos (conceptos y atributos) que el sistema se encargará de ir cumplimentando según se va recibiendo la información del usuario.

Para llevar a cabo tareas más complejas se pueden utilizar metodologías basadas en planes. Este tipo de metodología se basa en el establecimiento de un objetivo común en el diálogo entre el usuario y el sistema. De esta forma, usuario y sistema cooperan para llevar a cabo el mencionado objetivo.

Por último, es posible implementar gestores de diálogo mediante metodologías basadas en agentes. Este tipo de metodología combina las ventajas de los modelos basados en máquinas de estados finitos y los modelos basados en *frames* y generalmente se utilizan si existe la necesidad de ejecutar operaciones en un dominio que cambia dinámicamente.

- Generación del lenguaje natural

Este módulo también se conoce bajo el nombre de módulo de generación de respuestas. Su objetivo es generar una frase, gramaticalmente correcta y en un lenguaje lo más cercano posible al lenguaje natural a partir de la respuesta recibida por parte del sistema. La frase gramaticalmente correcta que genera el módulo de generación de lenguaje natural debe transmitir el mensaje generado por el gestor del diálogo.

Las tareas fundamentales que se llevan a cabo en este módulo son la determinación de contenido, la planificación del documento, es decir, cómo se va a organizar el mismo, la generación de expresiones que describan cada referencia u objeto, la agregación y lexicalización de lo que se va a decir, y por último, este módulo también se encarga de dar forma lingüística al contenido resultante [2.6].

- Conversión Texto – Habla

Este módulo también se conoce bajo el nombre de módulo de síntesis y voz. La función que realiza este módulo es generar la señal de audio a partir de la respuesta que obtiene del sistema en forma de texto. La señal de audio generada es la respuesta final que obtiene el usuario.

En cuanto a la arquitectura del módulo de síntesis y voz, dicho módulo se compone de dos partes, un *front-end* y un *back-end*. El *front-end* desempeña dos tareas principales. En primer lugar, toma el texto y convierte partes problemáticas como números y abreviaturas en palabras equivalentes. A continuación asigna una transcripción fonética a cada palabra, y divide y marca el texto en unidades. El *back-end*, toma la representación lingüística simbólica y la convierte en sonido. El back-end se llama a menudo sintetizador.

2.1.3 Clasificación de los Sistemas de Diálogo

En función de cómo se produce la interacción entre el usuario y la máquina se pueden distinguir distintos tipos de diálogo [1.7].

- Sistemas de diálogos guiados. La interacción se realiza mediante alternancias cerradas entre pregunta y respuesta.
- Sistemas de diálogo cooperativos. Aceptan las interrupciones y negociaciones por parte del usuario, reparten de forma equilibrada el turno de palabra e incorporan mecanismos de detección de incoherencias gramaticales.

- Sistema de diálogos adaptativos. El sistema es capaz de aprender nuevas estrategias comunicativas en función del comportamiento del usuario. El usuario lleva la iniciativa de diálogo en cualquier instante de la interacción y se le permite que utilice lenguaje natural y no se delimitan los mensajes del sistema.

Existen otros criterios para clasificar los sistemas de diálogo [2.7]:

- Según la dependencia del hablante

- Dependiente del hablante. Un sistema de este tipo está desarrollado para funcionar únicamente con un solo hablante. Estos sistemas son fáciles de implementar, baratos y flexibles, pero generalmente no tan flexibles como los sistemas de diálogo independientes del hablante.
- Independiente del hablante. Un sistema de este tipo está desarrollado para funcionar con cualquier hablante de un determinado tipo. Estos sistemas son complejos de implementar y su implementación es más cara. Sin embargo, son más flexibles.

- Según el tipo de comunicación

- Unimodales. Son aquellos sistemas de diálogo que únicamente utilizan el habla.
- Multimodales. Son aquellos sistemas de diálogo que utilizan varios canales de comunicación para proporcionar las entradas y las salidas.

- Según el tipo de idioma

- Monolingüe. Permiten la interacción con el sistema mediante un único idioma.
- Multilingüe. Permiten la interacción con el sistema mediante varios idiomas.

- Según el tipo de discurso

- Reconocimiento de palabras aisladas. El patrón acústico de cada palabra del vocabulario se almacena como una secuencia temporal de características derivadas. El reconocimiento se lleva a cabo comparando el patrón acústico de la palabra a reconocer con los patrones almacenados.
- Reconocimiento de palabras conectadas. En estos sistemas la entrada hablada es una secuencia de palabras de un vocabulario específico, y el reconocimiento se lleva a cabo basándose en la coincidencia de palabras de referencia aisladas.
- Reconocimiento continuo. En estos sistemas las palabras están conectadas, es decir, no están separadas por pausas. El lenguaje continuo es más difícil de tratar debido a la variedad de efectos.

- Reconocimiento discreto. Estos sistemas funcionan con palabras simples, necesitando de una pausa entre la dicción de cada palabra.
- Según la adaptación
- Sistemas adaptativos. El sistema es capaz de aprender nuevas estrategias comunicativas en función del comportamiento del usuario.

2.1.4 Sistemas de Diálogo Existentes: Ejemplos representativos

En las siguientes líneas se presentan algunas de las aplicaciones de distintos ámbitos cuya interfaz está diseñada para interactuar con el usuario a modo de sistema de diálogo oral.

MASK (Multimodal-Multimedia Automated Service Kiosk) [2.5] Sistema desarrollado por el *Spoken Language Processing Group* (Francia). Proporciona horarios de tren, reservas, precios, etc. Los usuarios se pueden comunicar con el sistema mediante habla y una pantalla sensible al tacto, mientras que éste se puede comunicar con los usuarios mediante habla, gráficos, vídeo y sonido.

August [2.9] Sistema desarrollado en el *Centre for Speech Technology, Royal Institute of Technology* (Suecia), para proporcionar información acerca del autor sueco August Strindberg (a quien debe su nombre) y sobre restaurantes y otros servicios de la ciudad de Estocolmo. La interfaz presenta una figura animada con aspecto similar al autor sueco del cual recibe su nombre la aplicación. La comunicación con el usuario se lleva a cabo a través de diálogos sintéticos, expresiones faciales y movimientos de cabeza, acordes con los movimientos de cabeza de los usuarios captados a través de una cámara. La información adicional se muestra a través de nubes de pensamiento que manan de la cabeza de August.

Mercury [2.10] Se trata de una interfaz oral basada en el uso telefónico que proporciona información acerca de vuelos y sus precios. Mercury permite a los usuarios reservar vuelos a más de 200 destinos en Estados Unidos y en el resto del mundo.

DARPA Communicator [2.11] Se trata de una interfaz conversacional inteligente para la reserva de billetes de avión desarrollada en Estados Unidos.

LARRI (Language-based Agent for Retrieval of Repair Information) [2.12] Se trata de un sistema de diálogo dedicado al soporte, mantenimiento y reparación de dominios. LARRI se basa en la tecnología utilizada por DARPA. Esta aplicación ha sido desarrollada por el grupo de ciencias de computación de la universidad de Carnegie Mellon.

Jupiter [2.13] Se trata de un sistema de diálogo oral diseñado para proporcionar al usuario la información meteorológica actualizada día a día mediante la vía telefónica. Ha sido desarrollado por el *Instituto Tecnológico de Massachusetts* (MIT), en Estados Unidos.

ITSPOKE [2.14] Se trata de un sistema de diálogo desarrollado para el ámbito académico. El sistema conversa con los estudiantes y les proporciona información así como correcciones de sus errores de aprendizaje.

VOCALIZA [2.15] Sistema de diálogo que se utiliza para terapias de voz en español y ayuda a los terapeutas a tratar a pacientes con distintas patologías y carencias en su habilidad lingüística.

VICO (Virtual Intelligent Co-Driver) [2.16] Permite la interacción natural entre los usuarios y dispositivos digitales. El sistema de diálogo está basado en el ámbito del automóvil y proporciona a los usuarios una interfaz usable con lenguaje natural así como servicios innovadores como pueden ser navegación, planificación de rutas, posibilidad de reservar hoteles o diseñar rutas turísticas.

AthosMail [2.17] Se trata de un sistema de diálogo multilingüaje que permite a los usuarios la lectura de correos electrónicos.

COGNIRON [2.18] Se trata de un robot inteligente cognitivo.

LET'S GO! [2.19] Se trata de un sistema de diálogo orientado principalmente a usuarios extranjeros y a usuarios de la tercera edad. Su objetivo es proporcionar información acerca de los autobuses de Pittsburgh.

DisCoH [2.20] Sistema de diálogo oral que proporciona información sobre congresos a los usuarios.

Siri [2.21] Se trata de un asistente de voz desarrollado por la compañía Apple que permite a los usuarios comunicar mediante la voz la acción que quiere realizar con su dispositivo.

2.2 Estándar VoiceXML

2.2.1 Introducción al estándar VoiceXML

El origen de VoiceXML data de 1995. Fue diseñado como un lenguaje de diálogo basado en XML (*eXtensible Markup Language*) [1.8] cuyo objetivo era simplificar el proceso de desarrollo de aplicaciones basadas en el reconocimiento de la voz dentro de un proyecto de la compañía AT&T llamado PML (*Phone Markup Language*). Actualmente, junto a las compañías Lucent y Motorola se sigue trabajando en PML como lenguaje propio de diseño.

En 1998, el W3C organizó una conferencia acerca de los navegadores de voz. En ese momento, AT&T y Lucent tenían diferentes variantes del original PML, mientras que Motorola e IBM estaban trabajando en el desarrollo de VoxML y SpeechML respectivamente. Muchos de los asistentes de la mencionada conferencia estaban también desarrollando lenguajes similares para el diseño de diálogos, por ejemplo, HP trabajaba en TalkML y PipeBeach en VoiceHTML.

AT&T, IBM, Lucent y Motorola decidieron formar el Foro de VoiceXML con el objetivo de aunar sus esfuerzos. La misión de este foro era definir y estandarizar un lenguaje que los desarrolladores pudieran usar en sus aplicaciones de voz. Se eligió XML como la base de este lenguaje.

En el 2000, el Foro de VoiceXML lanzó al público VoiceXML 1.0. Poco tiempo después, VoiceXML 1.0 se presentó a W3C como la base para la creación de un nuevo estándar internacional.

VoiceXML 2.0 es el resultado de la colaboración conjunta de las empresas, miembros de W3C, otros grupos de trabajo del W3C y el público.

VoiceXML es un lenguaje diseñado para crear sistemas de diálogo mediante la voz que incluyen voz sintetizada, reconocimiento de entrada por voz o DMTF, grabación de diálogos, funciones de telefonía e iniciativa mixta. Su objetivo principal es usar las ventajas y características del desarrollo web en el desarrollo de aplicaciones de voz interactivas.

La Figura 3 muestra un ejemplo sencillo del uso del lenguaje.

```

<?xml version="1.0" xml:lang="es-ES" encoding="UTF-8"?>
  <vxml version="2.0">
    <form>
      <block> ;Hola Mundo! </block>
    </form>
  </vxml>

```

Figura 3 'Hello World' en VoiceXML

Este ejemplo consta de un formulario que sintetiza y reproduce la expresión “¡Hola Mundo!” al usuario.

La arquitectura definida para una aplicación desarrollada en voiceXML contiene los componentes que se pueden observar en la Figura 4.

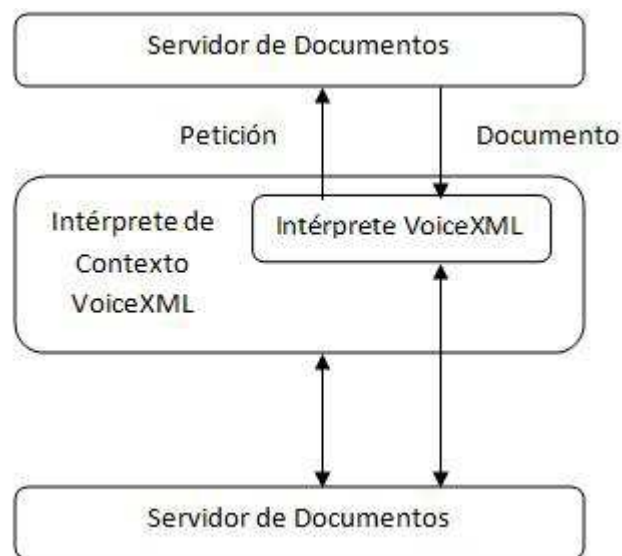


Figura 4 Arquitectura de una aplicación VoiceXML

El servidor de documentos (un servidor web, por ejemplo) procesa las peticiones de la aplicación cliente, el intérprete VoiceXML, a través del intérprete de contexto. La respuesta del servidor es un documento con formato VoiceXML que es procesado por el intérprete VoiceXML. El intérprete de contexto VoiceXML puede monitorizar entradas de usuario en paralelo con el intérprete de VoiceXML.

La plataforma de implementación está controlada por el intérprete de contexto VoiceXML y por el intérprete de VoiceXML. Dicha plataforma genera eventos en respuesta a las acciones del usuario (orales o en forma de acciones como desconectar) y los eventos del sistema (contador de limitadores temporales). Algunos de estos eventos

son transmitidos por el Intérprete a sí mismo, tal como se especifica en el documento de VoiceXML, mientras que otros son transmitidos por el Contexto VoiceXML.

El objetivo principal de VoiceXML es llevar todo el potencial del desarrollo web existente a las aplicaciones basadas en sistemas de diálogos. Este lenguaje permite la integración de servicios de voz con servicios de datos utilizando el paradigma del cliente-servidor. Se considera un servicio de voz como una secuencia de diálogo interactiva entre un usuario y una plataforma de implementación. Los servidores de documentos proporcionan los diálogos que pueden ser externos a la plataforma de implementación. Los servidores de documentos deben mantener la lógica general del servicio, realizar operaciones en la base de datos y sistemas heredados, y además producir cuadros de diálogo. Un documento VoiceXML especifica cada interacción de diálogo que se llevó a cabo por el intérprete. La entrada del usuario afecta a la interpretación del diálogo y se recoge en las solicitudes presentadas a un servidor de documentos. El Servidor de documentos responde con otro documento VoiceXML para proseguir con la sesión del usuario que continúa con otros diálogos.

Por tanto, VoiceXML es un lenguaje de marcas con las siguientes características fundamentales:

- Minimiza las interacciones cliente – servidor mediante la especificación de múltiples interacciones por documento.
- Protege los detalles específicos de la plataforma.
- Separa el código referente a la interacción entre la máquina y el usuario (en VoiceXML) de la lógica de servicio.
- Promueve la portabilidad del servicio en diferentes plataformas.
- Es fácil de usar para interacciones simples, y a su vez proporciona características avanzadas para soportar diálogos complejos.

En referencia a la funcionalidad disponible, VoiceXML es un lenguaje que describe la interacción persona– máquina que ofrecen los sistemas de respuesta de voz y que incluye:

- Generación de síntesis de voz (text-to-speech).
- Reproducción de archivos de audio.
- Reconocimiento de la entrada oral.
- Registro de entrada oral.
- Control del flujo del diálogo.
- Funciones de telefonía como transferencia de llamadas y desconexión.

2.2.2 Conceptos Básicos

Un documento VoiceXML (o un conjunto de ellos) puede interpretarse como una máquina de estados finita. El usuario siempre está en estado de conversación o de diálogo. Cada diálogo determina cómo será el siguiente y la transición que tendrá lugar. Dichas transiciones se especifican usando URIs (*Uniform Resource Identifier*) [2.22] que determinan el siguiente documento y diálogo a utilizar. Si la URI no hace referencia a ningún documento, se asume que debe transicionar al documento actual, del mismo modo, si la URI no hace referencia a un diálogo, asume que es el primer diálogo existente en el documento al que debe efectuar la transición. La ejecución se termina cuando un diálogo no especifica un sucesor o si existe algún elemento que explícitamente termine la conversación.

A continuación se explican una serie de conceptos básicos referentes a VoiceXML:

- Diálogos y Subdiálogos

Existen dos tipos de diálogos: formularios y menús. Los formularios definen la interacción que recoge los valores de un conjunto de variables mientras que los menús presentan al usuario diferentes opciones, y en función de la opción elegida por el usuario, efectúa una transición hacia el diálogo relacionado con la opción elegida.

Un subdiálogo se puede considerar como una llamada a una función, que proporciona un mecanismo para invocar una nueva interacción y volver al formulario original.

- Sesiones

Una sesión comienza al mismo tiempo que el usuario inicia la interacción con el interprete de contexto VoiceXML. La sesión se mantiene activa mientras que se cargan y procesan los documentos y termina cuando el usuario, el documento o el intérprete de contexto lo solicitan.

- Aplicaciones

Una aplicación es un conjunto de documentos que comparten el mismo documento raíz. Dicho documento se carga cada vez que el usuario interacciona con un documento de la aplicación. Además, el documento raíz permanece cargado mientras que se llevan a cabo las transiciones entre un documento y otro y únicamente se descarga cuando el usuario pasa a un documento que no pertenece a la aplicación.

Mientras que el documento raíz está activo, las variables y gramáticas que contenga dicho documento están disponibles.

- Gramáticas

Cada diálogo puede tener una o más gramáticas de voz o de DTMF asociadas.

En las aplicaciones que implementan un sistema de diálogo guiado, las gramáticas están activas únicamente cuando el usuario se encuentra en el diálogo que tiene lugar.

En las aplicaciones que implementan un sistema de diálogo mixto, las gramáticas de diferentes diálogos al que tiene lugar pueden estar activas. En esta situación si el usuario dice algo que coincida con la entrada de una gramática de otro diálogo, se ejecuta una transición a ese diálogo. La iniciativa mixta agrega flexibilidad y potencia para las aplicaciones de voz.

- Eventos

VoiceXML proporciona un mecanismo de rellenado de formularios para el manejo de entradas de usuario "normales" y define un mecanismo para controlar los eventos no cubiertos por este procedimiento.

Los eventos son lanzados por la plataforma debido a diferentes circunstancias, como por ejemplo, cuando el usuario no responde, la respuesta no es la correcta o esperada o si el usuario solicita ayuda. El intérprete también produce eventos si existen errores semánticos en los documentos.

El control de eventos puede especificarse en cualquier nivel y se aplica a todos los niveles inferiores.

- Enlaces

Los enlaces especifican transiciones a otros puntos del documento, otro documento dentro de la aplicación, u otro documento de otra aplicación.

2.2.3 Elementos de VoiceXML

La Tabla 1 muestra los elementos de VoiceXML.

Elemento	Funcionalidad
<assign>	Asigna un valor a una variable.
<audio>	Reproduce audio dentro de un mensaje.
<block>	Contenedor (no interactivo) de código ejecutable.
<catch>	Captura un evento.
<choice>	Define un elemento (ítem) del menú.
<clear>	Borra uno o más ítems del formulario.
<disconnect>	Desconecta una sesión.
<else>	Elemento utilizado en la estructura <if>.
<elseif>	Elemento utilizado en la estructura <if>.
<enumerate>	Notación abreviada para enumerar las opciones en un menú.
<error>	Captura un evento error.
<exit>	Finaliza la sesión.
<field>	Declara un campo de entrada en un formulario.
<filled>	Acción a ejecutar cuando se rellenan los campos.
<form>	Cuadro de diálogo para presentar información y recopilar datos.

<goto>	Transición a otro diálogo en el mismo o diferente documento.
<grammar>	Especifica un reconocimiento de voz o una gramática DTMF.
<help>	Captura un evento ayuda.
<if>	Lógica condicional.
<initial>	Declara código inicial antes de entrar en un formulario (iniciativa mixta).
<link>	Especifica una transición válida para todos los diálogos en el ámbito de la aplicación.
<log>	Genera un mensaje de depuración.
<menu>	Cuadro de diálogo para elegir entre varias alternativas.
<meta>	Define un elemento de metadata en formato nombre/valor.
<metadata>	Define información metadata mediante el esquema metadata.
<noinput>	Captura el evento noinput.
<nomatch>	Captura el evento nomatch.
<object>	Define extensiones a medida.
<option>	Especifica una opción en un <field>.
<param>	Parámetro en <object> o <subdialog>.
<prompt>	Salida de audio para el usuario.
<property>	Configuración de la plataforma de control de la aplicación.
<record>	Graba una muestra de audio.
<reprompt>	Reproduce un prompt cuando un campo es revisitado después de un evento.
<return>	Retorno desde un sub-diálogo.
<script>	Especifica un bloque de código ECMAScript.
<subdialog>	Invoca a un diálogo como un subdiálogo del actual.
<submit>	Presenta valores a otro documento del servidor.
<throw>	Lanza un evento.
<transfer>	Transfiere la llamada a otro destino.
<value>	Inserta el valor de una expresión en un prompt.
<var>	Declara una variable.
<vxml>	Elemento de mayor jerarquía en cada documento VoiceXML.

Tabla 1 Elementos de VoiceXML

2.2.4 Constructores de Diálogo

2.2.4.1 Formularios

Los formularios son un componente clave de los documentos VoiceXML. Un formulario contiene:

- Un conjunto de elementos de formulario (ítems de formulario).
- Declaración de variables que no son elementos del formulario.
- Control de eventos.
- Acciones, o bloques de lógica que se ejecutan cuando las variables toman un valor a partir de la entrada del usuario.

Los formularios son interpretados por un algoritmo de interpretación de formularios implícito (FIA, *Form Interpretation Algorithm*). El FIA tiene un bucle principal que selecciona uno a uno los ítems del formulario y los va visitando. El ítem seleccionado es el primero en el documento cuya condición de guarda no es satisfecha.

El FIA finaliza cuando interpreta una sentencia de transferencia de control o con una etiqueta de *<exit>* implícita al no quedar ítems que seleccionar en el formulario.

La Tabla 2 muestra los atributos que tiene un formulario.

Atributo	Descripción
Id	Nombre del formulario. Si se especifica, el formulario se puede referenciar dentro del documento o de otro documento.
scope	Alcance por defecto de las gramáticas del formulario.

Tabla 2 Atributos de un formulario

Tal y como se ha mencionado anteriormente, los formularios contienen elementos de formulario (ítems de formulario). Estos elementos pueden ser de dos tipos:

- Ítems de entrada

Estos ítems tienen prompts para decirle al usuario qué decir o qué teclear, gramáticas que definen las entradas permitidas y manejadores que procesan cualquier evento resultante. Los ítems de entrada se muestran en la Tabla 3.

Ítem	Descripción
<field>	Input de entrada cuyo valor se obtiene a través de gramáticas de voz y DTMF.
<record>	Input de entrada cuyo valor de entrada es un fragmento de audio que graba la entrada del usuario. (por ejemplo, un mensaje en un contestador de voz)
<transfer>	Input de entrada que realiza una transferencia a otro número de teléfono.
<object>	Este input de entrada invoca a un objeto de una plataforma específica a partir de varios parámetros. El objeto de retorno es un objeto ECMAScript.
<subdialog>	Input de entrada que invoca otro diálogo en el documento activo o invoca a otro documento VoiceXML. El objeto de retorno es un objeto ECMAScript.

Tabla 3 Ítems de entrada en VoiceXML

- Ítems de control

En la Tabla 4 podemos observar los ítems de control existentes:

Item	Descripción
<block>	Contienen un bloque de código a ejecutar que no solicita la

	entrada de datos.
<initial>	Controla la interacción inicial en un formulario de interacción mixta.

Tabla 4 Ítems de control en VoiceXML

Cada ítem del formulario tiene asociado una variable, la cual por defecto está inicializada a *undefined* cuando se entra en el formulario, y una condición de guarda, la cual establece si ese ítem puede ser elegido o no por el FIA. En la tabla 5 se muestran los atributos comunes que tienen todos los ítems de formulario.

Atributo	Descripción
name	Nombre de la variable del ítem del formulario con ámbito de diálogo que contendrá el valor del ítem del formulario.
expr	El valor inicial de la variable del ítem del formulario. Su valor por defecto es <i>undefined</i> .
cond	Expresión a evaluar junto con la variable del ítem del formulario. Si no se especifica, por defecto será <i>true</i> .

Tabla 5 Atributos de los ítems de formulario en VoiceXML

Una vez analizados los elementos que forman los formularios es importante distinguir los tipos de formularios existentes en función del tipo de interacción. Así pues, existen dos tipos de formulario atendiendo al criterio definido anteriormente.

- Formularios dirigidos

Se trata de formularios en los que los ítems son ejecutados en orden secuencial con el fin de implementar una interacción dirigida por la máquina.

- Formularios de iniciativa mixta

Se trata de formularios en los que, ambos, la máquina y el usuario dirigen la conversación. Deben tener uno o más niveles de gramáticas, definiendo unas gramáticas globales para el formulario y otras específicas para cada parte a rellenar del formulario.

2.2.4.2 Menús

Los menús son estructuras que a partir de una serie de opciones solicitan al usuario hacer una elección. En función de la elección, hace transiciones a diferentes partes del documento. Pueden tener su gramática activa aunque el usuario esté ejecutando otro diálogo. Es posible incluir gramáticas DMTF en los menús.

Los menús están formados por distintos elementos. Estos elementos se analizan a continuación.

- Menú

Este elemento identifica al menú y determina el ámbito de sus gramáticas. Los principales atributos de este elemento se muestran en la Tabla 6.

Atributo	Descripción
id	Identificador del menú.
scope	Ámbito de aplicación de la gramática del menú.
dtmf	Si se establece a true, se permite que las primeras nueve alternativas reciban los valores DTMF “1”, “2”, etc a pesar de que no se haya indicado implícitamente. El valor por defecto es false.
accept	Si se establece a "exact" (valor por defecto), el texto de los elementos de elección en el menú define la frase exacta a ser reconocida. Si se establece a "approximate", se permite que el usuario diga una frase aproximada al valor del elemento de elección.

Tabla 6 Atributos del elemento menú en VoiceXML

- Elemento Choice

Este elemento puede ser utilizado con diferentes fines.

- Para especificar una gramática de voz.
- Para especificar una gramática DTMF.
- Para formar una cadena con el contenido y utilizarse en el elemento `<enumerate>`. Este elemento proporciona una descripción generada automáticamente de las opciones disponibles que tiene el usuario.

En la Tabla 7 se describen los atributos del elemento *choice*.

Atributo	Descripción
dtmf	Secuencia DTMF para esta elección.
accept	Anula la configuración de accept en <menu> para esta elección particular. Si se establece a "exact" (valor por defecto), el texto de los elementos de elección en el menú define la frase exacta a ser reconocida. Si se establece a "approximate", se permite que el usuario diga una frase aproximada al valor del elemento de elección.
next	URI del siguiente cuadro diálogo o documento.
expr	Especifica una expresión a evaluar como URI para la transición.
event	Especifica un evento que será lanzado.
eventexpr	Expresión ECMAScript que evalúa el nombre del evento que se producirá.
message	Cadena de mensaje que proporciona contexto adicional sobre el evento que será lanzado.

messageexpr	Expresión de ECMAScript para evaluar la cadena del mensaje.
--------------------	---

Tabla 7 Atributos del elemento choice en VoiceXML

La Figura 5 muestra un ejemplo genérico de un elemento menú. Este menú permite al usuario elegir entre 'Opción 1' y 'Opción 2'. En función de la elección del usuario, se realizará una transición a los formularios correspondientes a cada opción, es decir, a 'URI1' o 'URI2'. El atributo *dtmf* indica que para este menú está activa una gramática DTMF a la vez que lo está la gramática de voz. Por tanto, el usuario podrá realizar su elección mediante la voz o mediante el teclado numérico del teléfono.

```

<menu id="opciones" dtmf="true">
  <prompt>
    Elige una de las siguientes opciones: <enumerate/>
  </prompt>
  <choice next="#URI1"> Opción 1 </choice>
  <choice next="#URI2"> Opción 2 </choice>
</menu>

```

Figura 5 Ejemplo menú en VoiceXML

2.2.5 Gramáticas

Las gramáticas son elementos a partir de los cuales el motor del discurso determina qué palabras o frases debe reconocer como respuesta del usuario. Por tanto, se trata de un elemento fundamental para el desarrollo de aplicaciones de voz. El ámbito del documento en el que está activa una gramática depende de la misma. Además, es posible establecer jerarquías en las gramáticas mediante pesos.

2.2.5.1 Gramáticas de Voz

El elemento `<grammar>` se utiliza para establecer una gramática de voz que especifique el conjunto de palabras que el usuario debe pronunciar para ejecutar una acción o facilitar una información. Si el usuario proporciona una secuencia que coincida con algún elemento de la gramática, se devolverá la correspondiente interpretación semántica.

La Tabla 8 recoge los elementos definidos en el *XML Form* de la W3C y que están disponibles en VoiceXML 2.0

Elemento	Descripción
<code><grammar></code>	Elemento raíz de una gramática XML.

<meta>	Declaración del encabezado del contenido meta de un HTTP equivalente.
<metadata>	Declaración del encabezado del contenido de metadatos XML.
<lexicon>	Declaración del encabezado de un léxico de pronunciación.
<rule>	Declara una expansión de regla con nombre para una gramática.
<token>	Define una palabra u otra entidad que puede servir de entrada.
<ruleref>	Hace referencia a una regla definida localmente o externamente.
<item>	Define una expansión con repetición y probabilidad opcional.
< one-of >	Define un conjunto de expansiones de reglas alternativas.
<example>	Elemento contenido dentro de una definición de la regla que proporciona un ejemplo de entrada que coincide con la regla.
<tag>	Define una cadena arbitraria que es incluida en línea en una expansión que puede utilizarse para la interpretación semántica.

Tabla 8 Elementos SRGS (XML Form)

2.2.5.2 Gramáticas DTMF

Una gramática DTMF se distingue de una gramática de voz por el atributo “mode” del elemento *<grammar>*. Las gramáticas DTMF y las gramáticas de voz pueden ser gramáticas de línea o externas y los atributos “type” y “scope” se utilizan de igual forma para ambas.

Para trabajar con este tipo de gramáticas se necesita que las plataformas VoiceXML admitan el formato XML de gramática DTMF.

2.3 Plataformas IVR

IVR (*Interactive Voice Response*) es la tecnología que automatiza la interacción entre el humano y la máquina. Históricamente, las soluciones IVR utilizaban pre-grabado de voz y menús para presentar la información, y los usuarios proporcionaban las entradas al sistema mediante tonos del teclado del teléfono. Las soluciones IVR modernas permiten la entrada y las respuestas a través de las palabras habladas proporcionando así servicios de reconocimiento de voz. Estos sistemas también ofrecen la posibilidad de reproducir y grabar mensajes, así como funciones de telefonía.

La tecnología IVR es comúnmente implementada en empresas que reciben grandes cantidades de llamadas con el fin de reducir la necesidad de operadores.

Esta tecnología involucra el uso de otras tecnologías para proporcionar mejores servicios. Las tecnologías involucradas son las siguientes.

- DTMF (Dual Tone Multi Frequency)
- TTS (Text to Speech)

- ASR (Automatic Speech Recognition)

2.3.1 Voxeo

Voxeo Corporation [2.23] es una de las empresas líderes en el sector de los sistemas de diálogo y en concreto en plataformas IVR. Proporciona soluciones a través de voz sobre IP (VoIP) gracias a sus servidores y sistemas de desarrollo, apoyados en las tecnologías de VoiceXML y CCXML [2.24].

Voxeo cuenta con la plataforma Voxeo Evolution. Dicha plataforma proporciona a los desarrolladores un servicio completo para diseñar, crear, probar y depurar aplicaciones IVR de forma gratuita. Además, la plataforma cuenta con foros de soporte dónde se puede interactuar con la Corporación Voxeo y con otros miembros de la comunidad.

El funcionamiento de Voxeo es similar al funcionamiento de un servidor web. En los servidores web, los navegadores realizan peticiones a los servidores por medio de protocolos de Internet. Los servidores web alojan las páginas estáticas que devuelven al navegador. En Voxeo, las aplicaciones de voz siguen el mismo esquema que los servidores web convencionales. Sin embargo, se utilizan lenguajes de marcado de voz para escribir la página. De esta forma, lo que se devuelve al usuario no es una página web convencional sino audio.

2.3.2.1 Voxeo Evolution. Funcionamiento y Características

Para poder acceder a los servicios que proporciona Voxeo Evolution es necesario estar registrado y disponer de una cuenta de usuario.

Una vez que se dispone de la cuenta se puede comenzar a utilizar todas las funcionalidades que Voxeo Evolution proporciona. La vista general de la cuenta personal es la que se puede observar en la Figura 6.

Account > Overview

ACCOUNT FEATURES

- Quick Start Guide**
If you're new to Voxeo Evolution or voice applications, you'll definitely want to read our handy Quick Start Guide for instructions and development tips.
- Application Manager**
The application manager contains settings for all of your voice and messaging applications, including URLs, mapped phone numbers, and outbound dialing tokens.
- Voxeo Designer**
Use Voxeo Designer to build voice applications without any coding at all. Simply create a project, and let the web-based visual Voxeo Designer handle the rest!
- Application Analytics**
Use application analytics to view detailed reports and statistics for your applications.
- Application Debugger**
Having trouble getting your application to work? Watch your application activity in real-time so you can see exactly where the problems occur.
- Device Monitoring**
The Monitoring Manager controls which devices (or Web sites) you want to monitor and the specific actions that should be taken based upon their status.
- Files, Logs, & Reports**
If you'd like to host your voice application files on our servers, use the file manager to transfer your files. The file manager also includes a scratchpad where you can write your scripts online. Additionally, the file manager stores call logs and other archived reports.
- Support Tickets**
Create and monitor tickets for account and application issues.
- Users & Contacts**
The Contact Manager updates your account profile (including your login password) and controls how you and your colleagues receive notifications from Voxeo Evolution.
- Account Journal**
The Account Journal stores a permanent log of changes to your contacts and applications.

Welcome, jenn.cva! AccountID: 113195

On this page you'll find descriptions of the various Voxeo Evolution account features, as well as up-to-date information regarding your account and network status.

OPEN ACCOUNT TICKETS

Open New Account Ticket | RSS

no items were found

New Updates No New Updates

Figura 6 Vista general de una cuenta en Voxeo

Desde esta vista general es posible acceder a las diferentes funcionalidades que Voxeo Evolution ofrece. A continuación se ofrece una descripción de cada una de estas funcionalidades.

- Quick Start Guide

Permite consultar una guía de iniciación a la plataforma Voxeo. En la Figura 7 se puede observar el esquema de funcionamiento general de la plataforma. A continuación, se explican los pasos necesarios para poder crear una aplicación de voz en Voxeo.

Documentation > Quick Start Guide

How does it work?

Web pages consist of HTML that is rendered into a visible page by a Web browser. Similarly, a voice application consists of XML (VoiceXML, CCXML, or CallXML) which becomes an interactive voice application when processed by the Voxeo Corporation VoiceCenter network. All you need to do is write the application's XML, map it to a phone number in the Voxeo Application Manager, and give it a call.

As you can see in the diagram below, our servers fetch your voice application's XML, turning that XML into an interactive voice application for your callers.

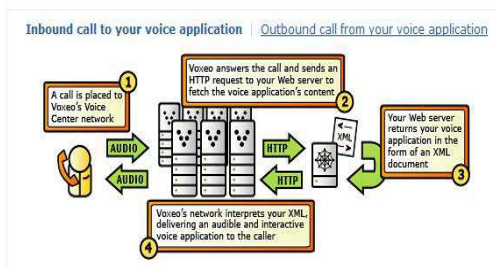


Figura 7 Funcionalidad Quick Start Guide ofrecida por Voxeo

- Application Manager

Esta herramienta contiene la configuración de las aplicaciones de voz y mensajería y permite crear nuevas aplicaciones. En la Figura 8 se muestra la apariencia de esta herramienta.

Account > Applications

EXISTING APPLICATIONS - APPLICATION VIEW

Application View | Number View

Application Name	Application Type	Deployment Type
PlanificadorDeViajes	Prophecy 11 VoiceXML, Premium ASR/TTS	Staging
SubDialog	Prophecy 11 VoiceXML, Premium ASR/TTS	Staging
Subgrammar	Prophecy 11 VoiceXML, Premium ASR/TTS	Staging
VoiceRecognition	Prophecy 11 VoiceXML, Premium ASR/TTS	Staging

Add Application

Figura 8 Funcionalidad Application Manager ofrecida por Voxeo

Si accedemos a una de las aplicaciones creadas podemos observar la configuración y la información de contacto relativa a la aplicación tal y como se muestra en las Figuras 9 y 10.

Account > Applications > PlanificadorDeViajes

APPLICATION SETTINGS

Application Settings | Contact Methods

*** Application Name:**
PlanificadorDeViajes

*** What forms of communication will this application support?**

Voice phone calls
 Text messaging
 Both

*** Voice Application Type:**

Deployment	Region	App Type	ASR/TTS	Platform
Development	USA	CCXML CallXML Designer VoiceObjects VoiceXML	DTMF-Only Nuance 9 Premium ASR Premium ASR/TT:	Prophecy 11 VoiceXML

Selected application type: Staging, Prophecy 11 VoiceXML, Premium ASR/TTS

*** Voice URL:** file manager | view file
<http://planificador.elementfx.com/appvoz/index.php>

+ Add a failover URL

Update Application | Delete Application

Figura 9 Configuración de una aplicación en Voxeo

Account > Applications > PlanificadorDeViajes

APPLICATION SETTINGS

Application Settings | Contact Methods

Phone Numbers & Addresses

The following contact numbers and addresses are mapped to your application.

Number Type	Number
<input type="checkbox"/> International (Voice Only) - Spain	+34 91 1237648
<input type="checkbox"/> iNum Number (Voice Only)	+883510001362782
USA Toll Free PIN Access (Voice Only)	(800) 289-5570 then PIN: 9996176912
USA Domestic PIN Access (Voice Only)	(407) 386-2174 then PIN: 9996176912
Skype VoIP	+990009369996176912
SIP VoIP	sip:9996176912@sip.voxeo.net
Phono Number	sip:9996176912@sip.voxeo.net

Move | Delete

Use the following form to add a new number to this application. Please note that international numbers may have [country-specific restrictions](#).

Phone Number:
 United States (+1) | Select a region to add a phone number... | Add

Outbound Dialing Tokens

Call Start Tokens, also known as Outbound Dialing Tokens, allow you to initiate phone calls with an HTTP fetch. For example, you could use a Call Start Token to place a phone call by clicking a button on a web page.

No Call Start Tokens are linked to this application. Contact support@voxeo.com to request a token.

Figura 10 Información de contacto de una aplicación

- Voxeo Designer

Esta herramienta permite crear aplicaciones de voz de forma gráfica mediante un editor y sin necesidad de codificación. Este editor se muestra en la Figura 11.

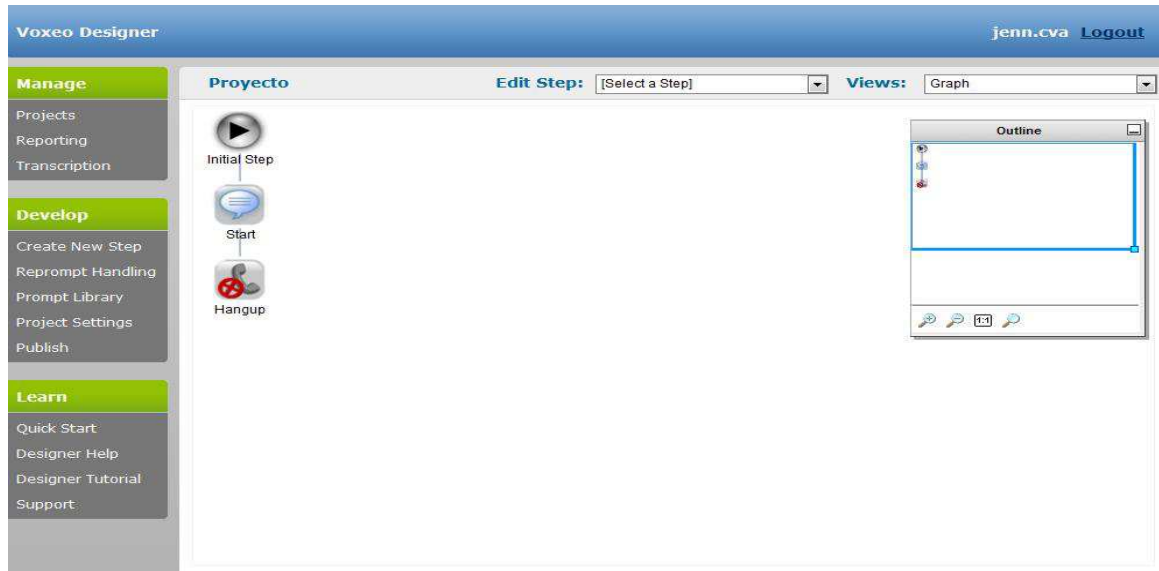


Figura 11 Funcionalidad Voxeo Designer ofrecida por Voxeo

- Application Analytics

Aplicación para ver informes detallados y estadísticas de las aplicaciones implementadas. Los informes y estadísticas se pueden generar para unas fechas determinadas. En la Figura 12 se muestra el formulario ofrecido por Voxeo para la generación de informes.

Account > Application Analytics

VIEW ANALYTICS DATA FOR YOUR APPLICATIONS

Use the following form to set the parameters for your analytics report. Reports will show statistics for the services mapped to your [VoiceObjects On-Demand](#) applications.

Start Date: End Date:

Service Name: Report Type:

Please note; application analytics will only exist for [VoiceObjects On-Demand](#) applications. Since you do not presently have any of these applications, you will be unable to run an analytics report.

Figura 12 Funcionalidad Application Analytics ofrecida por Voxeo

- Application Debugger

Esta herramienta permite depurar las aplicaciones en tiempo real para localizar errores existentes.

Para observar el estado de la aplicación, una vez que se tiene el depurador abierto, se marca el número de teléfono asignado y se llama a la aplicación. Si se produce algún problema se puede utilizar el soporte incorporado, enviando una copia de los registros del depurador al equipo de soporte de Voxeo.

En la Figura 13 se muestra una de las pantallas de esta herramienta.

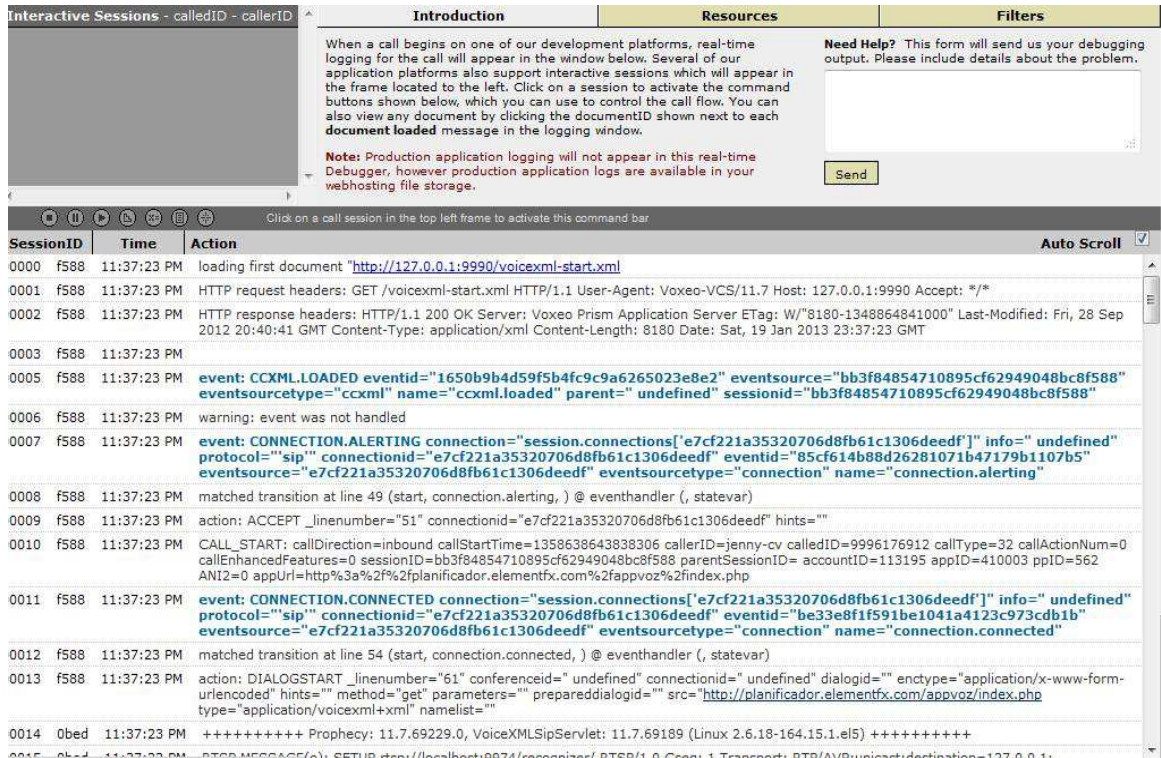


Figura 13 Funcionalidad Application Debugger ofrecida por Voxeo

- Device Monitoring

La herramienta Device Monitoring controla qué dispositivos (o sitios Web) se desean supervisar y las acciones específicas que se deben realizar basadas en su estado. Esta herramienta ha de activarse para la cuenta de usuario registrada en Voxeo y conlleva un coste adicional. En la Figura 14 se puede ver que dispositivos pueden ser monitorizados y su coste.

Account > Device Monitoring

Your account has been activated for device monitoring!

EXISTING DEVICES

No devices are configured. Use the form below to add your first device.

Add a New Device

Device test protocol	Annual fee: 1 test every 5 minutes	Annual fee: 1 test every minute
Web - HTTPb connection	\$150	\$850
Web - HTTP with content testing	\$425	\$1,800
Web - HTTPS with content testing	\$425	\$1,800
Email - SMTP	\$150	\$850
Email - POP3	\$150	\$850
Email - IMAP4	\$425	\$850
File - FTP	\$425	\$1,800
IP - Ping test	\$150	\$850
IP - TCP connection to a single port	\$150	\$850
Custom - Custom test script	\$960	\$3,000

Each account can setup one free 5-minute HTTPb monitored device. To add additional devices and protocols, please [contact our sales team](#).

Introduction

The Monitoring Manager controls which devices (or services) you want to monitor and the specific actions that should be taken based upon status.

Additionally, you can review individual monitoring attempts and historical alerts.

ACTIVE ALERTS

No active alerts.

EXTREME SUPPORT

If you have any questions or problems, contact our 24x7 Extreme Support team at support@voxeo.com

Figura 14 Funcionalidad Device Monitoring ofrecida por Voxeo

- Files, Logs, & Reports

Si se desean alojar los archivos de aplicación de voz en los servidores de Voxeo, se utiliza el administrador de archivos para transferir estos archivos. Aquí también se almacenan los registros de llamadas y otros informes que pueden ser de utilidad para el desarrollador. En la Figura 15 se puede observar la interfaz de esta funcionalidad.

Account > Files, Logs, Reports

EXISTING FILES

Note: All voice application files and directories must be in the [root/www](#) path.

root

Name	Size	Date Modified	Actions
logs		1/26/2012	
recordings		1/26/2012	
reports		1/26/2012	
www		1/26/2012	

ADD NEW FILE OR DIRECTORY

Upload File:

New Directory:

New File:

New file names must have one of the following extensions: **.txt**, **.grammar**, or **.*xml** (any extension ending in xml)

Figura 15 Funcionalidad File, Logs, Reports ofrecida por Voxeo

- Support Tickets

Con la herramienta Support Tickets se crean y controlan los tickets para cuestiones de cuenta y de aplicación. En la Figura 16 se muestra la interfaz proporcionada por Voxeo para crear un nuevo ticket.

Account > Support Tickets

SUPPORT TICKETS : POST FORM

Ticket Home > Account Tickets

Subject:

Message:

Attach File (optional):

Notify me when this thread is updated

Close This Ticket?

Figura 16 Funcionalidad Support Tickets ofrecida por Voxeo

- Users & Contacts

En esta aplicación se gestiona el perfil de la cuenta (incluyendo la contraseña de inicio de sesión) y se controla cómo los usuarios reciben notificaciones desde Voxeo Evolution. En la Figura 17 se muestra el formulario para actualizar los datos de la cuenta de usuario.

Account > Contact Manager > jenn.cva

VOXEO ACCOUNT PROFILE

* First Name: Jennifer	* Phone: +34635441364
* Last Name: Catalina Vaquero	* Password:
Company:	* Confirm Password:
Job Title:	* E-mail: 100080512@alumnos.
Address:	
City:	Send To This Email Address:
State/Region:	<input checked="" type="checkbox"/> Account Journal Updates
Postal Code:	<input checked="" type="checkbox"/> Account Tickets
Country:	<input checked="" type="checkbox"/> Forum Posts
	<input checked="" type="checkbox"/> Global Alert Tickets
	<input checked="" type="checkbox"/> Global Maintenance Tickets
	<input type="checkbox"/> Newsletters
	<input type="checkbox"/> Wiki Updates
	<input type="button" value="Update"/>

Figura 17 Funcionalidad Users & Contacts ofrecida por Voxeo

- Account Journal

Herramienta que almacena un registro permanente de los cambios de los contactos y aplicaciones del usuario. Dicha herramienta se puede observar en la Figura 18.

The screenshot shows the 'Account Journal' interface. At the top, there's a navigation bar with 'Account > Account Journal'. Below it, a blue header reads 'MOST RECENT JOURNAL ENTRIES'. To the right of this header are 'RSS' and 'Search' links. The main content area displays two entries, both dated '10/2/2012 12:10 PM (EDT)'. Each entry is titled 'Application Edited: PlanificadorDeViajes'. The first entry states: 'Your application named "PlanificadorDeViajes" was edited by Jennifer Catalina Vaquero (jenn.cva). Modified information is shown below: The first Voice URL was changed from <http://planificador.elementfx.com/appvoz/ampliarComoLegar.php> to <http://planificador.elementfx.com/appvoz/index.php>. Posted by jenn.cva in account:jenn.cva'. The second entry states: 'Your application named "PlanificadorDeViajes" was edited by Jennifer Catalina Vaquero (jenn.cva). Modified information is shown below: The first Voice URL was changed from <http://planificador.elementfx.com/appvoz/index.php> to <http://planificador.elementfx.com/appvoz/ampliarComoLegar.php>. Posted by jenn.cva in account:jenn.cva'. On the right side, there's an 'Introduction' section with a notepad icon. It explains that the account journal keeps a permanent log of account activities and that an entry is created for application updates, account updates, and contact updates. Below this is a calendar for 'FEBRUARY 2013' showing dates from 1 to 28. At the bottom right, there's a 'FEBRUARY 2013 ARCHIVES' section stating 'This month has no journal entries'.

Figura 18 Funcionalidad Account Journal ofrecida por Voxeo

2.3.2 Otras plataformas IVR

Además de la plataforma Voxeo, analizada anteriormente, existen otras plataformas IVR que ofrecen servicios similares y se mencionan a continuación.

CallFire [2.25] Proporciona una solución IVR que permite crear aplicaciones telefónicas tanto simples como avanzadas. Permite enlazar la aplicación implementada a un número de teléfono local. También proporciona servicios de mensajería.

IVR Technology Group [2.26] Algunos de los servicios que ofrece son almacenamiento IVR de aplicaciones de voz propias, mensajería, servicios de pago mediante teléfono, verificación de la voz, transferencia de llamadas, etc.

SiatVox [2.27] Esta plataforma ha sido desarrollada íntegramente por *Prosodie Ibérica*. Cuenta con servicios de reconocimiento natural y síntesis de voz, y permite integrarse fácilmente con soluciones *Call Center* de clientes externos.

Loquendo [2.28] La solución IVR de Loquendo permite la realización de servicios interactivos self-service basados en IVR y portales de voz, para que las empresas se comuniquen de forma eficaz con sus propios clientes por teléfono (tanto red fija como móvil), o a través de Internet.

Capítulo 3

Descripción general del sistema

En este capítulo se presenta una descripción detallada del sistema desarrollado y se explican las características fundamentales de cada uno de los módulos así como su funcionalidad. Por último, se analizan las diferentes tecnologías utilizadas para la implementación de la aplicación.

3.1 Presentación del sistema desarrollado

El planificador de viajes desarrollado ofrece a los usuarios dos modalidades de acceso:

- Acceso visual mediante una página web dinámica.
- Acceso mediante la voz.

La Figura 19 muestra los servicios ofrecidos por esta aplicación multimodal.



Figura 19 Información ofrecida por la aplicación

La interacción con el sistema comienza cuando el usuario accede a la aplicación de una de las siguientes formas:

- Mediante la aplicación de voz, a través de una llamada, bien la línea telefónica o bien mediante cualquier cliente de voz sobre IP (VoIP) como por ejemplo, Skype.
- Mediante la aplicación web, cuando el usuario accede a la URL en la que está disponible la aplicación desde cualquier navegador.

La aplicación está alojada en un servidor web, de forma que cuando un usuario accede a la aplicación se realizan peticiones a dicho servidor web. La información que se proporciona se encuentra almacenada en una base de datos.

Adicionalmente, es necesario para la aplicación de voz un intérprete de VoiceXML que se encargue, por un lado de responder las llamadas de los usuarios y, por otro, de interpretar los documentos VoiceXML para ofrecer el servicio al usuario. Para ello se ha utilizado la plataforma Voxeo.

En ambas modalidades, todos los ficheros se interrelacionan de tal forma que el usuario en su llamada o en su consulta a la web puede navegar por cualquiera de los servicios ofrecidos, pudiendo realizar más de una consulta en cada interacción con el sistema.

La arquitectura de la aplicación se presenta en la Figura 20, donde se observa cómo el usuario interactúa con la aplicación y con todos los elementos que la componen.

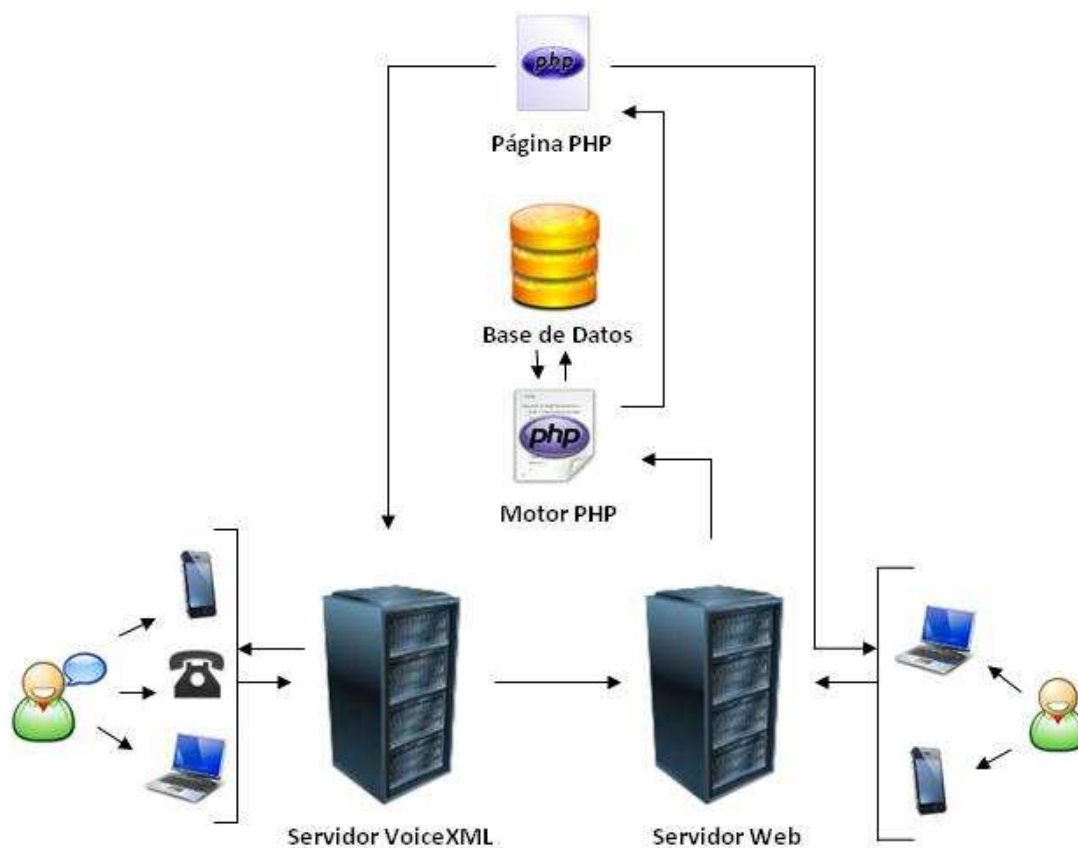


Figura 20 Arquitectura de la aplicación

3.2 Tecnologías empleadas

3.2.1 Base de datos: MySQL

MySQL es un sistema de gestión de bases de datos relacionales. El modelo relacional está basado en la lógica de predicados y en la teoría de conjuntos y presenta numerosas ventajas como garantizar la integridad referencial, evitar la duplicidad de registros y favorecer la normalización.

Otras de las características que presenta MySQL es que se trata de un sistema multihilo, es decir, permite el acceso a la información de forma concurrente y segura, así como un sistema multiusuario ya que permite proveer el servicio a varios usuarios simultáneamente.

MySQL es, desde enero de 2008, una subsidiaria de Sun Microsystems y ésta a su vez de Oracle Corporation desde abril de 2009.

MySQL se ofrece como software libre en un esquema de licenciamiento dual, es decir, se ofrece bajo la licencia pública general de GNU y para las empresas que deseen

incorporar MySQL en productos privados se exige disponer de una licencia adecuada que les permita su uso.

Se trata además del sistema de base de datos con uso más generalizado en aplicaciones web, en diferentes plataformas y sistemas operativos. Es por esto, junto con las numerosas ventajas que presenta este gestor de bases de datos (escalabilidad, conectividad, permisión de uso de numerosos tipos de datos y longitudes para columnas, etc) por lo que se ha decidido utilizar este sistema de gestión de bases de datos para la realización de este proyecto de fin de grado.

Además, se ha hecho uso de la herramienta de administración de bases de datos *phpMyAdmin* [1.10]. Este administrador es una herramienta de distribución libre escrita en PHP cuyo principal objetivo es proporcionar a los usuarios una interfaz visual, fácil e intuitiva de usar para la gestión de las bases de datos. Esta interfaz fácil e intuitiva puede observarse en la Figura 21.

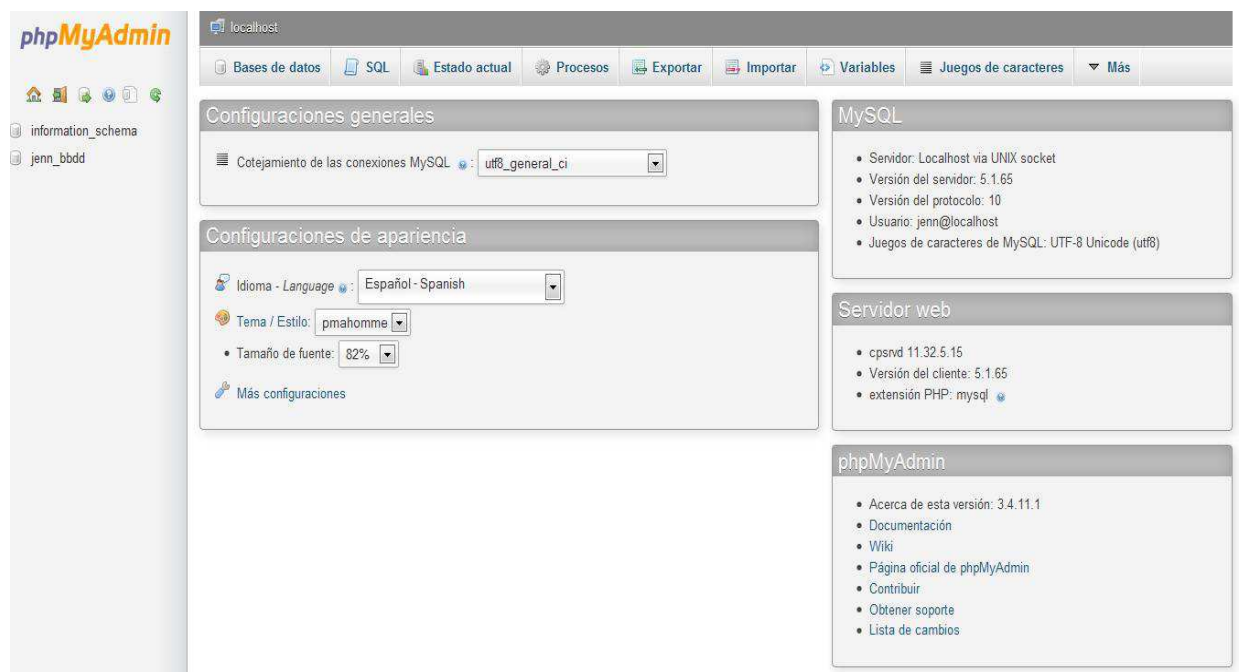


Figura 21 Captura de pantalla de la aplicación phpMyAdmin

Las operaciones frecuentes sobre una base de datos (insertar datos, tablas, relaciones, índices, permisos de usuarios) se pueden llevar a cabo mediante la interfaz visual. No obstante, también es posible ejecutar queries SQL.

phpMyAdmin cuenta con una amplia documentación de consulta, es multiplataforma y multilinguaje.

El modelo relacional de la base de datos utilizada en la aplicación ha sido objeto de un estudio profundo con el fin de diseñar un modelo que a la vez que se adapte a las necesidades de la aplicación, permita que las consultas a la base de datos sean eficaces y rápidas.

En la Figura 22 se puede observar el modelo relacional de la base de datos de la aplicación.

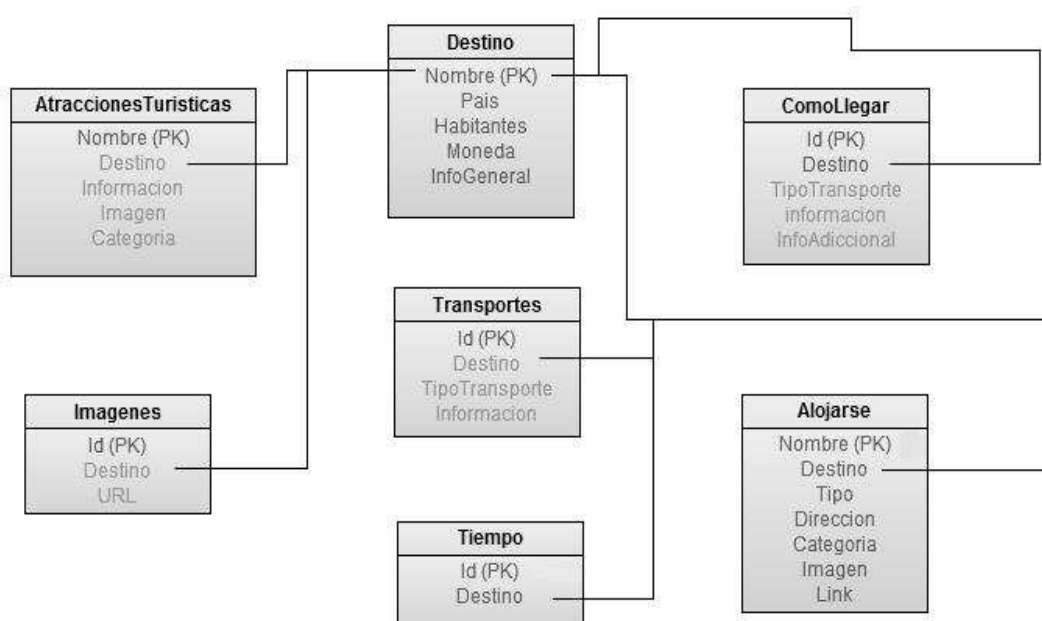


Figura 22 Modelo Relacional de la base de datos desarrollada para la aplicación

Las claves primarias de cada tabla se indican mediante (PK). Además, tal y como se puede observar en la Figura 22 se ha hecho uso de índices para mejorar y agilizar las consultas.

En cuanto a la codificación de los contenidos de las tablas se ha utilizado el formato UTF-8 (8-bit Unicode Transformation Format). UTF-8 es una norma de transmisión de longitud variable para caracteres codificados utilizando Unicode. UTF-8 usa grupos de bytes para representar el estándar Unicode para los alfabetos de muchos de los lenguajes del mundo. Es por este motivo por lo que se ha decidido hacer uso de este formato de codificación, ya que no solo se pretende ofrecer al usuario la posibilidad de consultar información acerca de destinos nacionales, sino que también se pretende que la aplicación sea completa y disponga de información de ciudades de todo el mundo. Así pues, para poder considerar caracteres que no existen en el castellano es necesario hacer uso de un sistema de codificación que los soporte.

Para realizar la conexión a la base de datos se dispone de una clase específica que realiza la conexión, así como las operaciones más comunes sobre la misma en el fichero *mysql.php*

Las funciones más importantes de esta clase son las que se pueden observar en la Figura 23.

```

/* Función que realiza la conexión a la base de datos */
function connect(){

    $this->conexion =
    mysql_connect("localhost","jenn_user","~d7PC33F18?
    ") or die($this->error_conexion);
    return $this->conexion;

}

/* Funcion que selecciona la base de datos a utilizar */
function select($db){

    $result = mysql_select_db($db, $this->conexion);
    if($result){
        $this->data_base = $db;
        return true;
    }else{
        return false;
    }
}

/* Función para cerrar la conexión a la base de datos */
function close(){

    mysql_close($this->conexion);

}

```

Figura 23 Funciones para el manejo de la base de datos

De esta forma, la conexión a la base de datos se puede realizar desde cualquier fichero de la aplicación invocando al método *connect ()* en primer lugar, *select ()*, en segundo lugar y por último, y tras haber realizado las operaciones pertinentes en la base de datos, se invoca a *close()* para finalizar la conexión con la base de datos.

3.2.2 Servidor web: x10Hosting

Para alojar los ficheros de la aplicación así como la base de datos se ha utilizado el servidor gratuito x10Hosting [3.1].

De entre las numerosas características que presenta este servidor en su versión gratuita, se destacan las siguientes.

- Espacio ilimitado.
- Transferencia de archivos ilimitada.
- cPanel 11

- Soporte de PHP 5, MySQL y scripts de instalación propia como Joomla, Wordpress, phpBB o SMF.
- Posibilidad de agregar dominios y subdominios.
- Cuentas FTP
- Soporte 24/7

3.2.3 Intérprete VoiceXML: Voxeo

El intérprete de lenguaje elegido para el desarrollo de la aplicación es la plataforma IVR detallada en la sección 2.3.2, Voxeo. Esta herramienta proporciona la infraestructura y componentes necesarios para poder crear y usar aplicaciones de voz.

Voxeo permite el acceso a la aplicación de voz creada a partir de distintos medios ya que proporciona un número de teléfono local (es posible especificar el país y región en la que se va a utilizar la aplicación, y por tanto el número que Voxeo proporcionará será válido en el país / región indicados), así como un número para poder realizar llamadas a través de Skype. De esta forma, facilita el acceso a la aplicación tanto para poder utilizarla una vez finalizada como para realizar las pruebas necesarias a la hora de la implementación y depuración de sus funcionalidades.

El número asignado a la aplicación tiene, internamente, asignado una URL en la que se indica la localización del primer fichero que debe cargar la aplicación cuando un usuario efectúe una llamada. Dicho fichero (al igual que el resto de los que forman la aplicación) puede estar almacenado en la misma plataforma Voxeo, o como en el caso de la aplicación desarrollada para este proyecto, en un servidor web. En la Figura 24 se puede observar la URL indicada (campo **Voice URL*) para esta aplicación en la configuración de la misma.

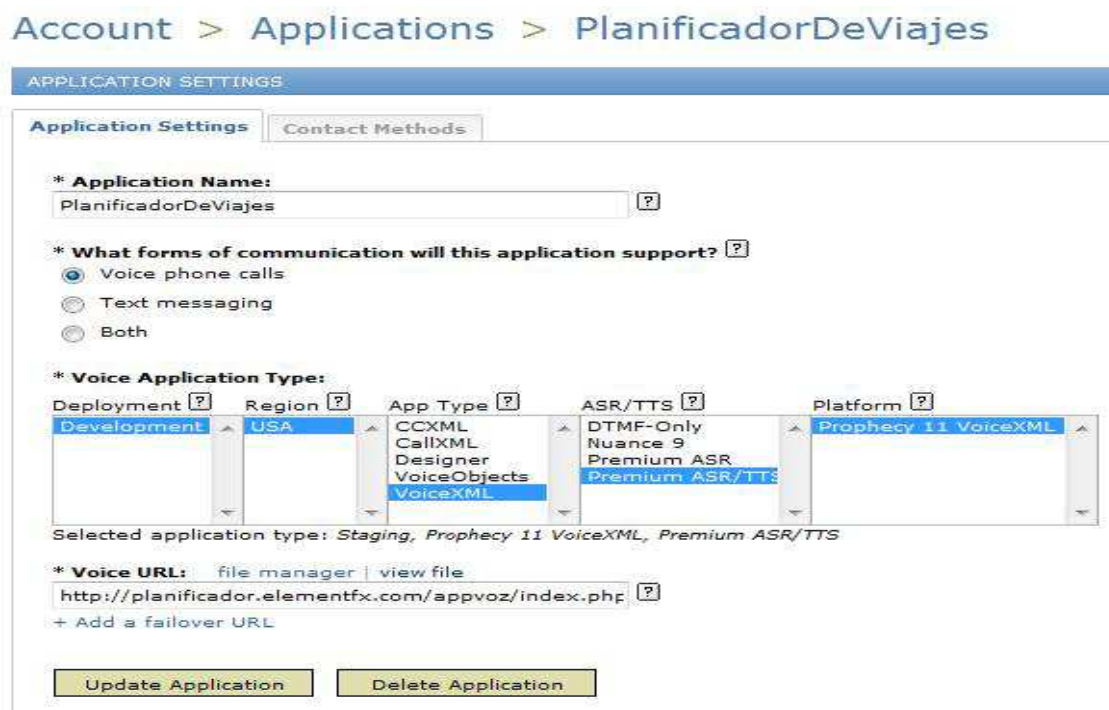


Figura 24 Localización de la URL inicial

3.2 Tecnologías empleadas

Tal y como se puede ver en la Figura 24, la aplicación se ha desarrollado utilizando la tecnología *Prophecy II VoiceXML*. Dicha tecnología ha permitido desarrollar la aplicación para su uso en idioma castellano.

Por último, la plataforma Voxeo posee un sistema de soporte rápido y eficaz, que incluye foros, tickets de soporte y una documentación muy completa.

Capítulo 4

Descripción detallada del módulo de acceso visual

En este capítulo se presenta una descripción detallada del módulo visual de la aplicación desarrollada y se explican las características fundamentales de dicho módulo así como su funcionalidad.

4.1 Introducción

La página web diseñada para presentar la información de modo visual en la aplicación se ha implementado utilizando los lenguajes de programación que se han considerado oportunos para que dicha página web ofrezca la funcionalidad y calidad necesaria para el correcto funcionamiento de la aplicación. Así pues, se ha hecho uso de lenguajes de programación web como PHP, HTML y Javascript.

El contenido de la página web se actualiza dinámicamente de forma que el contenido se genera a partir de la información introducida por el usuario en un formulario (en este caso, a partir del destino que indica el usuario). La información a mostrar se recupera de la base de datos.

La página web está alojada en el servidor x10Hosting y puede ser visitada mediante la siguiente URL: <http://planificador.elementfx.com/index.html>

En cuanto al estilo de la página web, se han utilizado hojas de estilo CSS. La idea principal del uso de CSS es separar la estructura de un documento de su presentación. Así pues, la información de estilo se presenta en un documento separado simplificando el código de los ficheros de contenido de la página web.

Una hoja de estilo CSS consiste en una serie de reglas. Cada regla consiste en uno o más selectores y un bloque de estilos con los estilos a aplicar para los elementos del documento que cumplan con el selector que les precede.

Las reglas definidas en la hoja de estilo CSS, se han diseñado de forma que el estilo se mantenga a lo largo de las páginas que componen el portal. Así pues, en todas las páginas se ha establecido un apartado que contiene la información relativa a la página actual, una barra de menú, y el menú relativo a los diferentes servicios de consulta que ofrece la aplicación. Estos elementos se presentan de la misma forma en todas las páginas con el objetivo de simplificar la navegabilidad en la web. De esta forma, es posible mantener un estilo uniforme en todas las páginas que forman el portal web para que la interfaz de usuario cumpla con los principios de usabilidad y accesibilidad y se mejore la interacción entre el portal web y el usuario.

En la siguiente sección y subsecciones se muestra el diseño y estilo utilizado en la página web junto con la funcionalidad de la misma.

4.2 Descripción de los módulos del acceso visual

4.2.1 Módulo ‘Inicio’

La página principal del portal web es la mostrada en la Figura 25. El estilo que presenta esta página se mantiene en el resto de páginas tal y como se ha mencionado anteriormente. En esta página de bienvenida se presenta un formulario a través del cual el usuario debe indicar el nombre de la ciudad de la cual desea consultar información.



Figura 25 Modulo de inicio desarrollado para la aplicación

A partir del valor recogido en el formulario se comprueba si la ciudad indicada está disponible en la base de datos, si lo está se redirige al usuario a la página que contiene la información general del destino y el menú de opciones disponibles. En el caso de que no esté dada de alta la ciudad solicitada en la base de datos, se mostrará al usuario un mensaje informativo que se puede observar en la Figura 26.



Figura 26 Mensaje informativo de la aplicación

4.2.2 Módulo 'Información General'

Una vez que el usuario ha indicado el nombre de la ciudad, se guarda dicho nombre en una variable a partir de la cual se realiza la consulta a la base de datos para obtener la información que se va a mostrar a continuación.

La query mediante la cual se realiza la consulta a la base de datos que se realiza es la mostrada en la Figura 27.

```
$consulta1 = sprintf("SELECT nombre, pais, habitantes, moneda,
infoGeneral FROM destino WHERE lower(nombre) LIKE '%s'",
mysql_real_escape_string($destino));
```

Figura 27 Query consulta

Tal y como se puede ver en la query de la Figura 27, se utiliza la función de PHP *sprintf* para indicar el formato del texto con el que se va a realizar la consulta. Dado que se espera que la entrada del usuario sea una cadena de texto, se especifica mediante ‘%s’, que el contenido de la variable donde almacenamos la entrada del usuario debe de ser un string. Además, se utiliza la función *mysql_real_escape_string()* para añadir más seguridad a la consulta, evitando que el usuario introduzca caracteres ilegales o, que incluso pueda realizar un ataque SQL Injection a la base de datos modificando su contenido o en casos peores, eliminándolo. Dada la importancia de mantener la integridad del contenido de la base de datos y establecer un nivel de seguridad en la aplicación web, todas las queries utilizadas para hacer consultas se realizan de este modo.

En la Figura 28, se muestra la página correspondiente al modulo de información general. Tal y como se puede observar en la figura, desde esta página es posible acceder a toda la información disponible del destino a través del menú situado a la derecha del apartado del contenido principal.



Figura 28 Módulo de información general desarrollado para la aplicación

4.2.3 Módulo ‘Cómo Llegar’

La Figura 29 muestra la interfaz visual correspondiente a este módulo. Para cada destino dado de alta en la aplicación, se han seleccionado los medios de transportes más adecuados para llegar a él, y mediante una consulta a la base de datos se obtiene toda la información necesaria que se muestra al usuario.

Planificador de Viajes

Inicio Destinos Acerca de

Madrid – Cómo llegar

Avión

Los accesos a Madrid por avión son muy privilegiados, de hecho, el aeropuerto de Madrid-Barajas es uno de los cinco más importantes de Europa tras su ampliación.

Accesos al aeropuerto

Por carretera por las principales vías de acceso y circunvalación de la ciudad de Madrid a través de la A-1 (autovía Madrid-Burgos) tomando la salida 17 en dirección aeropuerto, A-2 (carretera de Barcelona) cogiendo la salida hacia el aeropuerto en el punto kilométrico 12 o desde la M-40 (carretera de circunvalación de Madrid): las salidas 3 y 9 conectan respectivamente con la M-11 y la M-14, las cuales llegan directamente a las terminales.

En autobús se accede a través de la línea Exprés Aeropuerto de Barajas, un servicio que garantiza la conexión entre el aeropuerto y la capital las 24 horas del día, los 365 días del año. El precio del billete es tan sólo de 2 euros.

La línea 8 del Metro, Nuevos Ministerios-Barajas, comunica directamente todas las terminales del aeropuerto con el centro de la capital. El servicio vale dos euros, incluido el "suplemento aeropuerto". También es posible llegar al aeropuerto mediante la Red de Ferrocarriles Españoles (RENFE) a través de la línea C-1 que une la estación de Príncipe Pío y la T4.

Carretera

La privilegiada ubicación geográfica de la Comunidad de Madrid y especialmente de su capital, en el centro de la península ibérica, unida al carácter radial de las redes de carreteras y ferrocarriles españoles, proporciona una excelente conexión por tierra con otras comunidades así como con los países limítrofes.

Siete son las grandes carreteras de carácter estatal que unen a la ciudad de Madrid con el resto de urbes importantes españolas. Asimismo, la propia capital dispone de una serie de autovías, que facilitan la comunicación entre la amplia gama metropolitana y el resto de los términos municipales de la región madrileña.

La estructura básica de la red de carreteras españolas es radial, con centro en Madrid. Existen seis carreteras principales: A-1, Madrid-Burgos-Irún; A-2, Madrid-Barcelona-La Junquera; A-3, Madrid-Valencia; A-4, Madrid-Sevilla; A-5, Madrid-Badajoz y A-6, Madrid-A Coruña.

Accediendo por cualquiera de las seis vías principales encontrará tres anillos sucesivos denominados M-50, M-40 y M-30, que le facilitarán el acceso o el tránsito por la capital.

Tren

La empresa estatal RENFE presta servicios de viajeros y de mercancías en toda España. Numerosos ferrocarriles permiten los más diversos trayectos por el país. Los países vecinos, Portugal y Francia, y todos los países de la UE también están conectados con Madrid a través del ferrocarril.

Tel.: (+34)902 320 320

La ciudad de Madrid está conectada con once ciudades españolas a través del producto estrella de RENFE: la alta velocidad. Siete son las líneas que prestan el servicio en la actualidad: los AVE en funcionamiento van hasta Valladolid, Barcelona, Sevilla, Málaga, Valencia (a partir de diciembre 2010), Lleida y Huesca, con paradas en Ciudad Real, Puertollano, Córdoba, Guadalajara, Calatayud y Zaragoza. La calidad de su servicio es excelente.

© 2012 Planificador de Viajes. Proyecto de Fin de Grado realizado por Jennifer Catalina Vaquero

Información Práctica
Madrid
Cómo llegar
Transportes
Alojamiento
Atracciones Turísticas
Tiempo

Figura 29 Módulo Como llegar desarrollado para la aplicación

4.2.4 Módulo ‘Transportes’

La Figura 27, muestra la interfaz visual correspondiente a este módulo. Para cada destino dado de alta en la aplicación, se han seleccionado los transportes más útiles, y

CAPÍTULO: Descripción detallada del módulo de acceso visual

mediante una consulta a la base de datos se obtiene toda la información necesaria que se muestra al usuario.

Planificador de Viajes

Inicio Destinos Acerca de

Madrid – Transportes

Abono Transportes Turístico

Es un título de transportes de carácter personal, que permite a su titular efectuar un número ilimitado de desplazamientos en el conjunto de servicios de transporte regular de viajeros que operan en el interior de la zona elegida, con las excepciones que se indiquen.
Para su utilización, el usuario deberá estar en posesión de un cupón con la zona y el periodo de validez correspondiente, en el que se indicará, en el momento de la adquisición, el número del DNI o Pasaporte del titular. No serán válidos los cupones que no contengan dicha información.

Puntos de venta

- Oficinas del Consorcio Regional de Transportes
- Todas las estaciones de la red de Metro
- Centro de Atención al Viajero en la estación de Metro de Aeropuerto T1-T2-T3-T4
- En determinados estancos y kioscos de prensa de Madrid
- Centro de Información de Turismo de Madrid (Plaza Mayor, 27)
- Oficinas de Turismo de la Comunidad de Madrid
- Principales Tour Operadores internacionales y nacionales
- Principales agencias de viaje españolas y hoteles de la ciudad
- En lugares de interés turístico
- Internet : www.neoturismo.com

Autobús

El servicio regular de autobuses en el municipio de Madrid corre por cuenta de la Empresa Municipal de Transportes (EMT). Dispone de 216 líneas que ponen en contacto todas las calles de Madrid, 171 líneas son convencionales diurnas, 7 son líneas que prestan servicio en los entornos de algunos campus universitarios, 26 son líneas nocturnas convencionales y 12 son líneas nocturnas de 'Metrobúhos' (que sustituyen a la red de Metro las noches de fines de semana y vísperas de festivos).

Tel.: (+34)91 406 88 00

Cercanías

El cercanías de Renfe permite una comunicación eficaz dentro de la Comunidad y un acceso rápido a las comunidades limítrofes de Castilla y León y Castilla-La Mancha, además de llevar a cabo recorridos de interés a lo largo de los pueblos de la región. Desde numerosos puntos de esta red de cercanías es posible conectar con la red de metro, así como con los principales municipios de la Comunidad de Madrid.

Tel.: (+34)902 320 320

Metro

Metro de Madrid es un medio de transporte fundamental para la ciudad: rápido, puntual, eficaz y moderno, el transporte estrella.
Este transporte suburbano consta de tres metros ligeros y de 12 líneas que permiten moverte por la capital de manera ágil, desde las seis de la mañana hasta las 1.30 de la madrugada ininterrumpidamente.

Tel.: (+34)902 444 403

Taxi

Los más importantes municipios disponen de su propia red de taxis, los de mayor uso son los de la capital, blancos con una banda diagonal granate. Circulan todos los días y es fácil encontrarlos en las calles. Al mismo tiempo, los taxi tienen unas tarifas asequibles, si las comparamos con otras capitales europeas. Si en Madrid un recorrido desde el aeropuerto hasta el centro puede valer alrededor de 25 euros.

Radio Taxi: (+34)91 447 51 80

Información Práctica

- Madrid
- Cómo llegar
- Transportes
- Alojamiento
- Atracciones Turísticas
- Tiempo

© 2012 Planificador de Viajes. Proyecto de Fin de Grado realizado por Jennifer Catalina Vaquero

Figura 30 Módulo Transportes desarrollado para la aplicación

4.2.5 Módulo 'Alojamiento'

En este módulo, el usuario puede encontrar diferentes opciones de alojamiento en el destino seleccionado en función de sus necesidades.

En la Figura 31, se puede observar la página correspondiente a este módulo. Para esta página se ha implementado la funcionalidad de paginación para mantener el estilo de la interfaz ya que es posible que exista un gran número de alojamientos que mostrar.

Para implementar esta funcionalidad se ha establecido un índice y un valor de registros a mostrar que se utiliza para filtrar el número de resultados que devuelve la consulta a la base de datos. El valor de registros a mostrar se ha fijado a 4 de forma que cada página muestra 4 resultados.

Planificador de Viajes

Inicio Destinos Acerca de

Madrid – Alojamiento

Ganivet
 Tipo: Hotel
 El Hotel Ganivet goza de una ubicación céntrica, a 150 metros de la estación de metro de Puerta de Toledo de Madrid y a 10 minutos a pie de la Plaza Mayor.
 Categoría: 3 estrellas
[Sitio web](#)

Gran Versailles
 Tipo: Hotel
 El Gran Hotel Versailles está situado en una zona tranquila del centro de Madrid, junto a la estación de metro Alonso Martínez.
 Categoría: 4 estrellas
[Sitio web](#)

Gran Via-Sol
 Tipo: Apartamento
 Estos elegantes apartamentos se encuentran en el centro, a menos de 3 minutos a pie de la Gran Vía y de la Puerta del Sol.
 Categoría: Apartamentos
[Sitio web](#)

Hotel Ritz Madrid
 Tipo: Hotel
 El Ritz está situado en el centro del barrio del Retiro, a pocos pasos de los museos del Prado y Thyssen-Bornemisza, el teatro de la Zarzuela y el parque del Retiro. La zona financiera también está a poca distancia a pie.
 Categoría: 5 estrellas
[Sitio web](#)

1 2 3 Siguiente >

Información Práctica
 Madrid
 Cómo llegar
 Transportes
 Alojamiento
 Atracciones Turísticas
 Tiempo

© 2012 Planificador de Viajes. Proyecto de Fin de Grado realizado por Jennifer Catalina Vaquero

Figura 31 Módulo Alojamiento desarrollado para la aplicación

4.2.6 Módulo ‘Atracciones Turísticas’

En este módulo, el usuario puede encontrar diferentes atracciones turísticas en el destino seleccionado en función de sus gustos o intereses.

Al igual que en el módulo ‘Alojamiento’ (Sección 4.5.2), se ha implementado la funcionalidad de paginación con el objetivo de presentar la información más clara al usuario ya que es posible que exista un elevado número de atracciones turísticas en el mismo destino.

En la Figura 32 se puede observar la página correspondiente a este módulo.

Planificador de Viajes

Inicio Destinos Acerca de

Madrid – Atracciones Turísticas

Información Práctica
Madrid
Cómo llegar
Transportes
Alojamiento
Atracciones Turísticas
Tiempo

Palacio Real

El Palacio Real fue en su día la residencia oficial de los monarcas pero hoy, es un lugar de interés turístico, que acoge también actos de Estado, como encuentros diplomáticos y eventos oficiales.

Palacio Real de Aranjuez

Entre los ríos Tajo y Jarama se erige este Palacio Real de de estilo herreriano. El Palacio alberga una muestra sobre la vida cotidiana de las familias reales. Es obligatoria la visita a los Jardines de palacio, conocido como Jardín de la Isla, un espacio abierto, muy cuidado y que cuenta con un canal y numerosas fuentes.

Parque del Retiro

El Retiro o el ‘pulmón de la ciudad’ es uno de los principales parques de la ciudad y cuenta con una infinidad de estatuas, fuentes y monumentos conmemorativos que han ido poblando los jardines y lo han convertido en un museo de escultura al aire libre.

San Lorenzo de El Escorial

Ubicado a 40 kilómetros de Madrid, en plena Sierra de Guadarrama, el Real Sitio de San Lorenzo de El Escorial de estilo herreriano es un monumento patrimonio de la Humanidad y un lugar que no hay que dejar de visitar, por su emplazamiento y construcción.

< Anterior 1 2 3 Siguiente >

© 2012 Planificador de Viajes. Proyecto de Fin de Grado realizado por Jennifer Catalina Vaquero.

Figura 32 Módulo Atracciones Turísticas desarrollado para la aplicación

4.2.7 Módulo ‘Condición Meteorológica’

En este módulo se presenta la condición meteorológica así como la temperatura actual del destino seleccionado. Para ello, se ha hecho uso de la API proporcionada por Yahoo. Así pues, a través de una URL fija:

<http://weather.yahooapis.com/forecastrss?&u=c&w=>

a la cual es necesario concatenar el código asignado a la ciudad en cuestión, Yahoo devuelve un documento del cual gracias a otras clases auxiliares (*simplepie.inc* y *simplepie_yahoo_weather.inc*) es posible obtener la información necesaria para proporcionar el servicio. Dado que la información que ofrece Yahoo es en inglés, se ha efectuado la traducción equivalente a español para la aplicación.

El código único que Yahoo asigna a cada ciudad se recupera de la base de datos de la aplicación. Si este código no estuviera almacenado en la base de datos, se informa al usuario que la información no está disponible.

La Figura 33 muestra la página con la información relativa a este módulo.



Figura 33 Módulo Condición Meteorológica desarrollado para la aplicación

4.2.8 Módulo ‘Barra de Inicio’

Este módulo se corresponde con la barra horizontal de inicio que se encuentra presente en todas las páginas del portal web para dotar al portal de flexibilidad y navegabilidad. Dicha barra se puede observar en la Figura 34.



Figura 34 Módulo Barra de inicio desarrollado para la aplicación

Las tres opciones que presenta la barra se analizan a continuación.

- Inicio

Permite al usuario volver al módulo de inicio de la aplicación (Figura 25) en cualquier momento de forma que se pueda realizar una nueva búsqueda.

- Destinos

Muestra al usuario un listado paginado con el total de destinos disponibles en la aplicación (Figura 35). Desde esta lista, el usuario puede elegir el destino que le interesa y acceder a la información del mismo modo que lo haría si introdujera el nombre de la ciudad en el campo de búsqueda del módulo de inicio.



Figura 35 Listado de destinos

- Acerca de

Esta página permite al usuario consultar la información básica relativa a la aplicación. Así pues, se informa al usuario de las modalidades de acceso y la forma de acceder a la aplicación, tanto visualmente como oralmente. Esta funcionalidad se puede observar en la Figura 36.



Figura 36 Funcionalidad Acerca de desarrollado para la aplicación

Capítulo 5

Descripción detallada del módulo de acceso oral

En este capítulo se presenta una descripción detallada del acceso oral de la aplicación desarrollada y se explican las características fundamentales de dicho módulo así como las funcionalidades que ofrece. La explicación se complementa con casos de uso reales de la aplicación para ilustrar el funcionamiento de la misma.

5.1 Introducción

El acceso oral desarrollado implementa uno de los objetivos fundamentales del presente proyecto de fin de grado, esto es, dota a un portal web con la posibilidad de acceder al mismo mediante la voz. Así pues, el diseño del módulo de voz se ha llevado a cabo mediante diferentes módulos que son análogos a los vistos anteriormente en la sección 4.2 con el objetivo de ofrecer al usuario toda la información disponible en la página web a través de la voz. En la Figura 37 podemos observar la división de módulos en la aplicación de voz.

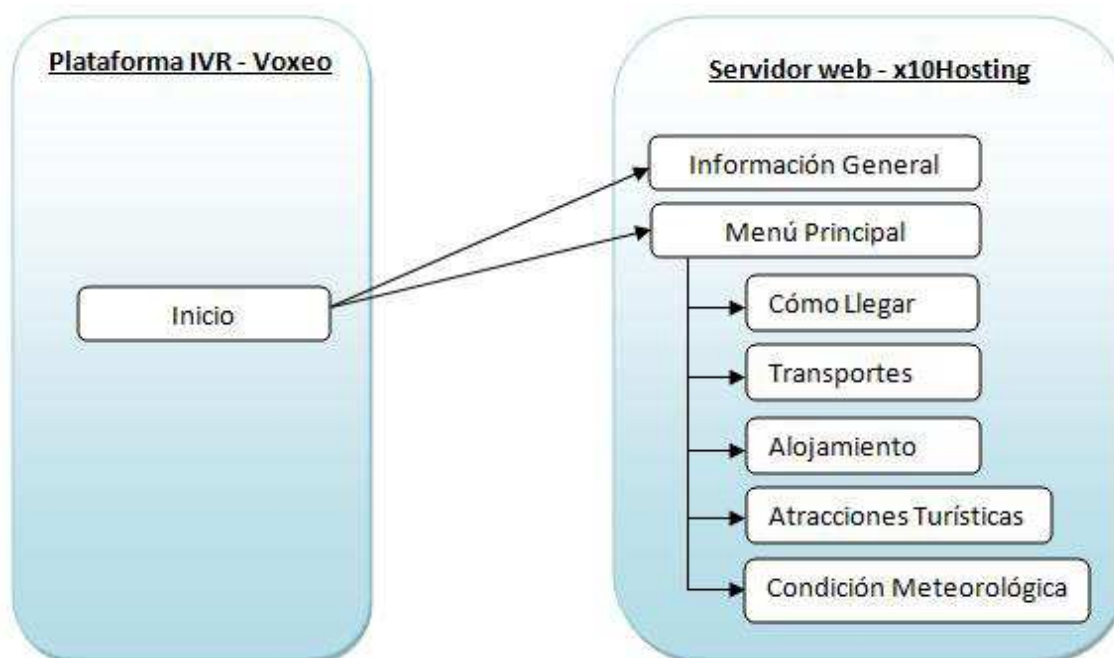


Figura 37 Módulos de la aplicación de voz

Los módulos que se pueden observar en la Figura 37 se explicarán con detalle en la sección 5.3 del presente capítulo.

En VoiceXML se pueden establecer valores de diferentes propiedades mediante la etiqueta `<property>`. Estas propiedades se utilizan para definir los valores que afectan al comportamiento de la plataforma.

Para la implementación de esta aplicación se ha hecho uso de diversas propiedades y eventos que se explican a continuación:

- Grammarfetchhint

Esta propiedad especifica cuándo una gramática externa debe ser buscada. Su valor por defecto, *prefetch*, significa que la gramática será buscada siempre que la página VoiceXML se cargue. En cambio, el valor *safe* establece que sólo se buscará la gramática cuando ésta sea utilizada. En la aplicación se utiliza su valor por defecto para que el usuario no experimente retrasos en el uso de la aplicación si la gramática a cargar contiene muchas entradas.

- Confidencelevel

Esta propiedad permite ajustar la exactitud que la expresión del usuario debe tener para ser reconocida por la gramática activa. Su valor por defecto es de '0.45'. Si se disminuye mucho este valor (a valores cercanos a 0), se reconocería cualquier entrada, mientras que si se ajusta a '1.0', ninguna expresión del usuario se correspondería con entrada en la gramática activa. El valor que se ha fijado para esta propiedad en los ficheros VoiceXML de la aplicación es de '0.3'.

- Sensitivity

Esta propiedad permite ajustar la sensibilidad del intérprete frente a la entrada nula. Su valor por defecto es de '0.5'. Si se aumenta este valor, el intérprete será más sensible, mientras que si se disminuye el intérprete será menos sensible. Esta propiedad puede ser útil si el usuario final accede a la aplicación en un entorno muy ruidoso. En el portal de voz implementado se ajusta este valor a '0.3'.

Por último, es importante destacar que el intérprete VoiceXML proporciona capturadores implícitos por defecto para los eventos *noinput*, *nomatch* y *help*. En la aplicación desarrollada y en función del módulo que esté activo interesa tratar estos tres eventos de forma específica y mostrar un mensaje determinado.

- Evento noinput

Este evento ocurre cuando el usuario no responde dentro del intervalo del tiempo de espera establecido. En la aplicación se ha tratado este evento informando al usuario de que no se ha recibido ninguna información y solicitándole que la proporcione de nuevo.

- Evento nomatch

Este evento se lanza cuando el usuario dice algo pero no es reconocido por ninguna de las gramáticas activas. En la aplicación se ha tratado este evento informando al usuario de que no se ha entendido la información y solicitándole que la proporcione de nuevo

- Evento help

Este evento ocurre en el caso en que el usuario pida ayuda. En la aplicación se ha tratado este evento de forma que si el usuario solicita ayuda se le proporcione información acerca de la aplicación y su uso, tal y como proporcionaba en el apartado 'Acerca de' del portal web. (Sección 4.2.8).

5.2 Gramáticas

Tal y como se indica en la sección 2.2.5, las gramáticas son un elemento fundamental en las aplicaciones de voz. En la aplicación de voz implementada se ha hecho uso de gramáticas de voz y DTMF de forma que la aplicación sea más accesible y proporcione diferentes medios de entrada a los usuarios.

Las gramáticas codificadas son de tipo XML ya que es el formato estándar definido por el W3C y, por tanto, es soportado en cualquier plataforma VoiceXML. Este formato permite una mayor flexibilidad en cuanto a la estructura de la gramática y es más fácil de depurar.

A lo largo de la aplicación se diferencia entre gramáticas estáticas y gramáticas dinámicas en función de las necesidades. Las diferencias entre los diferentes tipos de gramáticas se explican a continuación.

5.2.1 Gramáticas estáticas

Las gramáticas estáticas contienen información que no varía a lo largo del tiempo. Generalmente este tipo de gramáticas no presentan un gran número de entradas y pueden declararse en el documento *vmxl* o declararse como un documento externo e invocarse desde el documento *vxml* en el momento que se desee cargar la gramática.

Un ejemplo de este tipo de gramática se encuentra en el fichero *index.php*. La Figura 38 muestra un ejemplo de gramática estática que se activa cuando se pide confirmación al usuario acerca de si la ciudad que ha expresado es correcta o no.

```

<grammar xmlns="http://www.w3.org/2001/06/grammar" xml:lang="es-
    es" version="1.0" root = "MYRULE">

  <rule id="MYRULE">
    <one-of>
      <item>si<tag>out.MySlot="si";</tag> </item>
      <item>correcto<tag>out.MySlot="correcto";</tag></item>
      <item>no<tag>out.MySlot="no";</tag> </item>
      <item>incorrecto<tag>out.MySlot="incorrecto";</tag></item>
    </one-of>
  </rule>

</grammar>

```

Figura 38 Ejemplo de gramática estática

En toda la aplicación hay activas gramáticas estáticas que permiten al usuario volver al menú principal o al inicio para seleccionar un nuevo destino en cualquier momento. De este modo, se permite una mayor navegabilidad en la aplicación para que sea más fácil y cómoda de usar.

5.2.2 Gramáticas dinámicas

Las gramáticas dinámicas contienen información que varía en el tiempo. Al contrario de las gramáticas estáticas, este tipo de gramáticas suele contener un número elevado de entradas (ítems), y si se declararan en el documento *vxml* habría que modificar dicho documento de forma manual en cortos periodos de tiempo.

En la aplicación se ha hecho uso de gramáticas dinámicas en aquellos casos en los que la gramática puede variar en función del contenido actual de la base de datos. Así pues, las gramáticas dinámicas en la aplicación son aquellas que contienen los nombres

CAPÍTULO: Descripción detallada del módulo de acceso oral

de las ciudades disponibles y los nombres de transportes, alojamiento y atracciones turísticas asociados a cada ciudad.

Con el objetivo de que estas gramáticas se correspondan con el contenido de la base de datos y estén actualizadas en todo momento, el procedimiento que se sigue es eliminar la gramática existente (si la hubiera) en el momento en que sea necesario hacer uso de una gramática dinámica y a continuación crearla de nuevo en consonancia con el contenido de la base de datos.

Para ello, en el fichero *generarGramatica.php* se han implementado una serie de métodos a través de los cuales es posible llevar a cabo el procedimiento anteriormente explicado.

El método *borrar_gramatica* permite borrar una gramática existente. A partir del nombre que recibe por parámetro comprueba si existe un fichero con esa denominación, y si es así, lo elimina mediante el método *unlink*.

En la Figura 39 se muestra el método *crear_gramatica*. Este método recibe dos parámetros, el nombre que tendrá la gramática y un array con los ítems que contendrá la gramática que va a ser creada. El array de ítems se forma dinámicamente en función del contenido de la base de datos antes de invocar a este método. Después, mediante el método PHP *fopen* se crea un fichero nuevo en el servidor en el cual se va escribiendo el contenido de la gramática mediante el método *fwrite*. Al escribir los ítems de la gramática se eliminan las tildes de las palabras que las contengan en la base de datos mediante un método auxiliar, *retornarStringValido*. Una vez que la operación se ha finalizado se cierra el fichero mediante *fclose*.

```

    /*Método que crea una gramatica dinamicamente. Para
    ello, abre un fichero y lo va escribiendo línea a línea con el
    contenido correspondiente */

    function crear_gramatica($nombre_gramatica,
    $array_nombres){
        $gramatica=$nombre_gramatica;
        $fp =fopen($gramatica,"a");
        $contenido='<?xml version= "1.0"?>';
        $write= fwrite($fp, $contenido);
        fwrite($fp, chr(13).chr(10));

        $contenido='<grammar
xmlns="http://www.w3.org/2001/06/grammar" xml:lang="es-es"
version="1.0" root="MYRULE">';
        $write= fwrite($fp, $contenido);
        fwrite($fp, chr(13).chr(10));
        $contenido='<rule id="MYRULE">';
        $write= fwrite($fp, $contenido);
        fwrite($fp, chr(13).chr(10));
        $contenido='<one-of>';
        $write= fwrite($fp, $contenido);
        fwrite($fp, chr(13).chr(10));

        $contenido='<item>inicio<tag>out.MySlot="inicio";</tag><
/item>'; /*en todas las gramáticas generadas dinámicamente
añadimos la entrada de inicio. */
        $write= fwrite($fp, $contenido);
        fwrite($fp, chr(13).chr(10));

        foreach ($array_nombres as $opciones){
            $contenido='<item>';
            $write= fwrite($fp, $contenido);
            $opciones = retornarStringValido($opciones);
            fwrite($fp, $opciones);
            $contenido='<tag>out.MySlot="';
            $write= fwrite($fp, $contenido);
            $opciones =
retornarStringValido($opciones);
            fwrite($fp, $opciones);
            $contenido='";</tag> </item>';
            $write= fwrite($fp, $contenido);
            fwrite($fp, chr(13).chr(10));

        }

        $contenido='</one-of>';
        $write= fwrite($fp, $contenido);
        fwrite($fp, chr(13).chr(10));
        $contenido='</rule>';
        $write= fwrite($fp, $contenido);
        fwrite($fp, chr(13).chr(10));
        $contenido='</grammar>';
        $write= fputs($fp, $contenido);
        fclose($fp);
    }
}

```

Figura 39 Método crear_gramatica que genera una gramática dinámica

5.3 Descripción de los módulos del acceso oral

5.2.3 Módulo ‘Inicio’

Este módulo se corresponde con el fichero *index.php* que está alojado en el servidor web y es el fichero al cual se establece la referencia en Voxeo como fichero principal de la aplicación de voz.

Las acciones principales que realiza este módulo son reproducir un mensaje de saludo y solicitar al usuario el nombre de la ciudad en la que está interesado.

El saludo se implementa mediante un script que, en función de la hora en que el usuario realice la llamada, reproducirá “Buenos días”, “Buenas tardes” o “Buenas noches” seguido de “Bienvenido a Planificador de Viajes”.

A continuación se genera una gramática dinámica que contiene todos los nombres de los destinos disponibles y se solicita al usuario el nombre de la ciudad. Una vez que el usuario ha indicado dicho nombre, se le solicita confirmación del mismo y por último se da la opción al usuario de escuchar información general acerca del destino o de acceder al menú principal.

El flujo de datos y ficheros utilizados por este módulo se puede observar en la Figura 40.

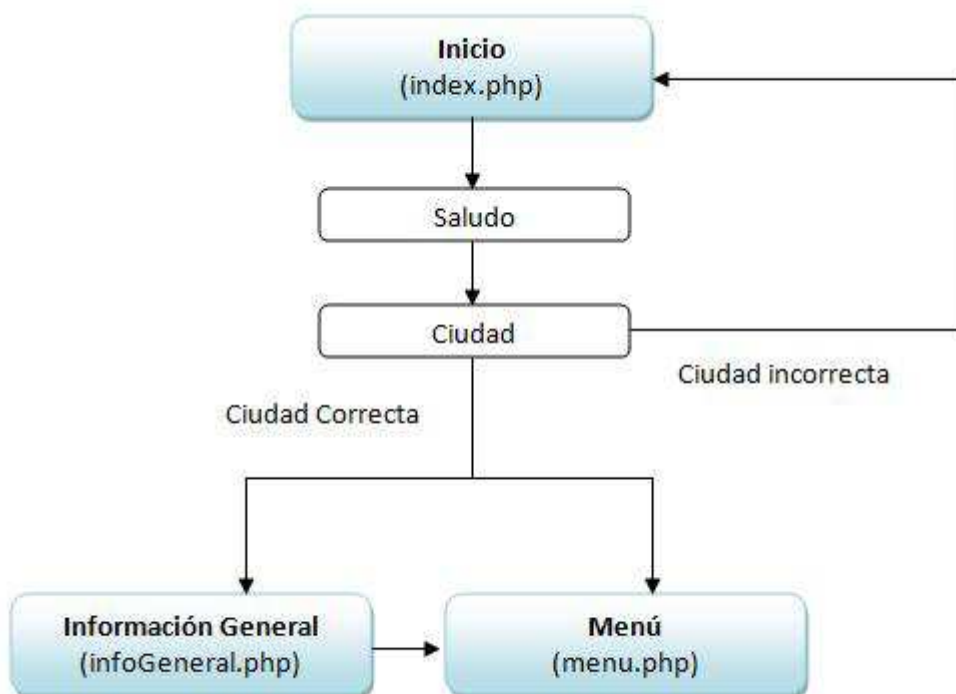


Figura 40 Flujo de datos y ficheros del módulo inicio

Un ejemplo de uso de este módulo se puede observar en la Figura 41 donde ‘S’ representa al sistema y ‘U’ representa al usuario.


```
S: Buenos días, bienvenido a planificador de viajes. Por favor, di el nombre de la ciudad de la cual deseas conocer información.  
U: Madrid.  
S: He entendido: Madrid, ¿es correcto?  
U: Si.  
S: Si desea escuchar información general acerca del destino diga información general. Si desea ir al menú, diga menú.
```

Figura 41 Caso de uso. Módulo Inicio

5.2.4 Módulo ‘Información General’

En este módulo, el usuario puede acceder a información general acerca de la ciudad seleccionada.

Tal y como se ha visto en la Figura 40, si el usuario tras indicar la ciudad desea escuchar información general de la misma, se realizará una transición al documento *infoGeneral.php* y se accederá a este módulo. En este fichero se realiza una consulta a la base de datos de la aplicación para obtener toda la información de la tabla ‘destino’ y transmitírsela al usuario. La información proporcionada incluye historia de la ciudad, moneda, país, y número de habitantes. Dado que el número de habitantes se almacena en la base de datos como un número entero, es necesario utilizar el elemento VoiceXML `<say-as>` de forma que el texto se interprete como un número y el usuario lo oiga como tal.

Una vez que el usuario ha escuchado esta información, se realiza una transición al menú principal de la aplicación (*menú.php*) para mostrar al usuario el resto de categorías que puede consultar.

En la Figura 42 se muestra un caso de uso de este módulo.

S: Si desea escuchar información general acerca del destino diga información general. Si desea ir al menú, diga menú.

U: Información general.

S: Madrid es la capital de España y de la Comunidad de Madrid, comunidad autónoma uniprovincial. También conocida como la Villa y Corte, es la ciudad más grande y poblada del país. Es la tercera ciudad más poblada de la Unión Europea –por detrás de Berlín y Londres– y la tercera área metropolitana, por detrás de las de París y Londres. Como capital del Estado, Madrid alberga las sedes del gobierno, Cortes Generales, ministerios, instituciones y organismos asociados, así como la residencia oficial de los Reyes de España y del Presidente del Gobierno. Madrid tiene 3265038 habitantes y su moneda es el Euro.

Figura 42 Caso de uso. Módulo Información general

5.2.5 Menú

Este módulo permite al usuario acceder oralmente a la información que se podía consultar visualmente en la página web en los módulos ‘Cómo llegar’ (Sección 4.2.3), ‘Transportes’ (Sección 4.2.4), ‘Alojamiento’ (Sección 4.2.5), ‘Atracciones Turísticas’ (Sección 4.2.6), ‘Condición Meteorológica’ (Sección 4.2.7).

El flujo de datos y ficheros utilizados por este módulo se puede observar en la Figura 43.

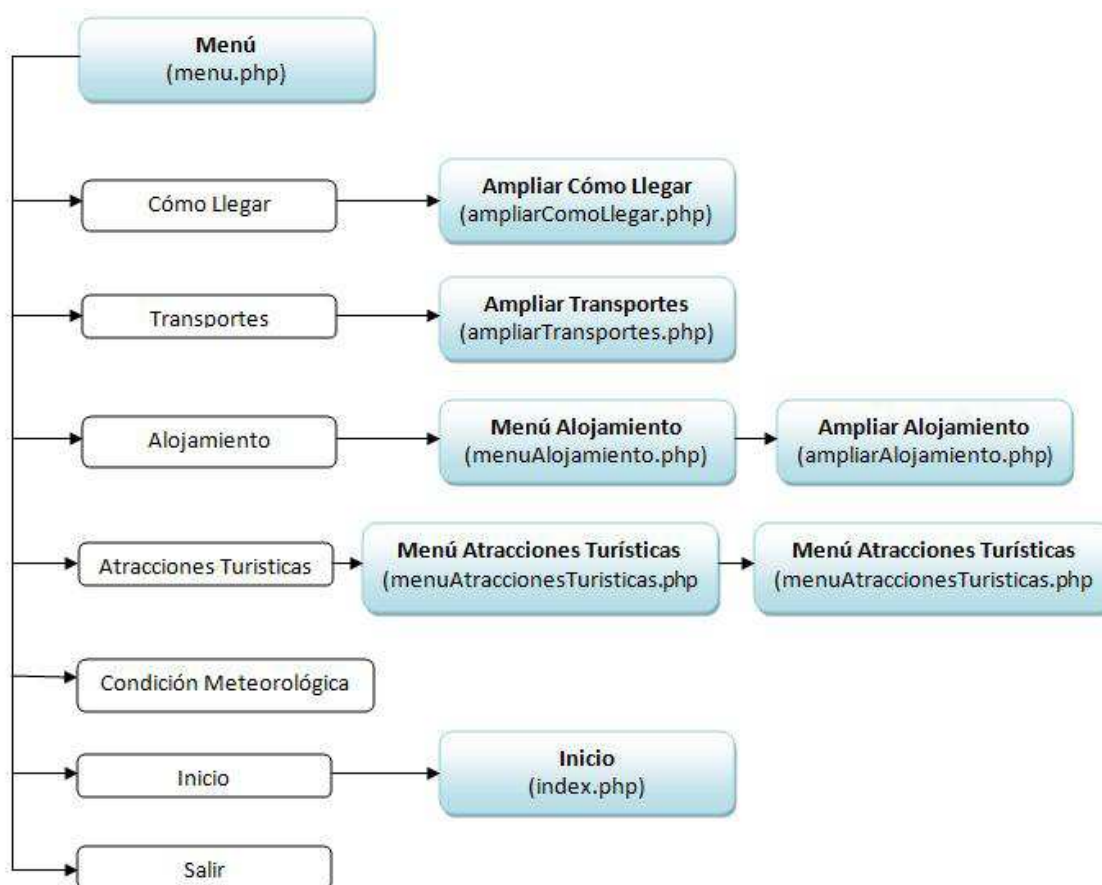


Figura 43 Flujo de datos y ficheros del módulo menú

El módulo correspondiente al menú principal se ha implementado mediante el elemento menú que proporciona VoiceXML. Se ha establecido el valor de la propiedad *dtmf* de dicha estructura como *true* de forma que el usuario pueda acceder a los ítems del menú tanto de forma oral como utilizando el teclado del teléfono, es decir, marcando un número determinado para cada ítem. Esta estructura genera dos gramáticas, una gramática de voz con las entradas ‘cómo llegar’, ‘transportes’, ‘alojamiento’, ‘atracciones turísticas’, ‘tiempo’, ‘inicio’ y ‘salir’, y una gramática DTMF que asigna valores a partir de 1 a los elementos que forman la gramática de voz. Por tanto, la finalidad del fichero *menú.php* es únicamente proporcionar la estructura necesaria para transicionar al usuario al módulo que desee y que de esta forma pueda acceder a la información.

La estructura del menú utilizado se puede observar en la Figura 44.

```

<menu id="menu" scope="dialog" dtmf="true">

  <prompt>
    Elige una de las siguientes opciones: <enumerate/>
  </prompt>

  <choice next="#comollegar"> Cómo llegar </choice>
  <choice next="#transportes"> Transportes </choice>
  <choice next="#alojamiento"> Alojamiento </choice>
  <choice next="#atraccionesTurísticas"> Atracciones
    Turísticas </choice>
  <choice next="#tiempo"> Tiempo </choice>
  <choice next="#inicio"> Inicio </choice>
  <choice next="#salir"> Salir </choice>

</menu>

```

Figura 44 Estructura del menú principal

En la Figura 45 se muestra un caso de uso de este módulo.

```

S: Elige una de las siguientes opciones:
Di cómo llegar o pulsa 1
Di transportes o pulsa 2
Di alojamientos o pulsa 3
Di atracciones turísticas o pulsa 4
Di tiempo o pulsa 5
Di inicio o pulsa 6
Di salir o pulsa 7

```

Figura 45 Caso de uso. Menú

En las siguientes subsecciones se explicará el contenido y funcionalidad de cada módulo accesible a través del menú principal.

5.2.5.1 Módulo ‘Cómo Llegar’

El módulo ‘Cómo Llegar’ proporciona al usuario información acerca de los medios de transporte más adecuados para llegar al destino indicado.

Dado que los transportes pueden variar en función del destino se realiza una consulta a la base de datos para recuperar dichos nombres y enumerárselos al usuario de forma que pueda escoger el medio de transporte del cual desea ampliar información. En la enumeración de los transportes es necesario incluir pausas mediante el elemento VoiceXML `<break>` de forma que la enumeración no se produzca de forma muy rápida y el usuario pueda escuchar todas las opciones correctamente.

En este módulo es necesario hacer uso de una gramática dinámica que contenga los nombres de los transportes asociados al destino que consulta el usuario ya que, como se ha indicado anteriormente, dichos nombres pueden variar en función del destino. De este modo, una vez enumerados los nombres al usuario se activará la gramática dinámica que contenga estos nombres y permita reconocer la entrada del usuario.

Por último, se realiza una transición al fichero *ampliarComoLlegar.php*. En este fichero, a partir del nombre del transporte indicado por el usuario se realiza una consulta a la base de datos en la que se obtiene la información referente al citado medio de transporte y se proporciona dicha información al usuario. A continuación, se ofrece la posibilidad al usuario de elegir otro medio de transporte del listado o volver al menú principal para acceder a otro tipo de información diferente.

En la Figura 46 se puede observar un ejemplo de caso de uso de este módulo.

S: Los medios de transporte más adecuados para llegar a Madrid son: avión, carretera, tren. Por favor, diga el medio de transporte del cual desea ampliar información.

U: Carretera

S: La privilegiada ubicación geográfica de la Comunidad de Madrid y especialmente de su capital, en el centro de la península ibérica, unida al carácter radial de las redes de carreteras y ferrocarriles españoles, proporciona una excelente conexión por tierra con otras comunidades así como con los países limítrofes.

Siete son las grandes carreteras de carácter estatal que unen a la ciudad de Madrid con el resto de urbes importantes españolas. Asimismo, la propia capital dispone de una serie de autovías, que facilitan la comunicación entre la amplia gama metropolitana y el resto de los términos municipales de la región madrileña.

La estructura básica de la red de carreteras españolas es radial, con centro en Madrid. Existen seis carreteras principales: A-1, Madrid-Burgos-Irún; A-2, Madrid-Barcelona-La Junquera; A-3, Madrid-Valencia; A-4, Madrid-Sevilla; A-5, Madrid-Badajoz y A-6, Madrid-A Coruña.

Accediendo por cualquiera de las seis vías principales encontrará tres anillos sucesivos denominados M-50, M-40 y M-30, que le facilitarán el acceso o el tránsito por la capital.

Si desea escuchar información acerca de otro medio de transporte diga otro. Si desea ir al menú, diga menú.

Figura 46 Caso de uso. Módulo cómo llegar

5.2.5.2 Módulo ‘Transportes’

Este módulo proporciona información acerca de los transportes más importantes y más útiles para desplazarse por la ciudad indicada por el usuario.

La arquitectura de este módulo es análoga a la del módulo ‘Cómo Llegar’ (Sección 5.2.5.1). Así pues, dado que los transportes pueden variar en función del destino se realiza una consulta a la base de datos para recuperar dichos nombres y enumerárselos al usuario de forma que pueda escoger el medio de transporte del cual desea ampliar información. En la enumeración de los transportes es necesario incluir pausas mediante el elemento VoiceXML `<break>` de forma que la enumeración no se produzca de forma muy rápida y el usuario pueda escuchar todas las opciones correctamente.

En este módulo es necesario hacer uso de una gramática dinámica que contenga los nombres de los transportes asociados al destino que consulta el usuario ya que, como se ha indicado anteriormente, dichos nombres pueden variar en función del destino. De

CAPÍTULO: Descripción detallada del módulo de acceso oral

este modo, una vez enumerados los nombres al usuario se activará la gramática dinámica que contenga estos nombres y permita reconocer la entrada del usuario.

Por último, se realiza una transición al fichero *ampliarTransportes.php*. En este fichero, a partir del nombre del transporte indicado por el usuario se realiza una consulta a la base de datos en la que se obtiene la información y se proporciona dicha información al usuario. Una vez proporcionada la información, se ofrece la posibilidad al usuario de elegir otro medio de transporte del listado o volver al menú principal para acceder a otro tipo de información diferente.

La Figura 47 muestra un ejemplo de caso de uso de este módulo.

```
S: Los medios de transporte más adecuados en Madrid son: abono
transporte turístico, autobús, cercanías, Metro, taxi. Por favor,
diga el medio de transporte del cual desea ampliar información.

U: Metro

S: Metro de Madrid es un medio de transporte fundamental para la
ciudad: rápido, puntual, eficaz y moderno, el transporte
estrella. Este transporte suburbano consta de tres metros ligeros
y de 12 líneas que permiten moverte por la capital de manera
ágil, desde las seis de la mañana hasta las 1.30 de la madrugada
ininterrumpidamente.

Tel.: (+34)902 444 403

Si desea escuchar información acerca de otro medio de transporte
diga otro. Si desea ir al menú, diga menú.
```

Figura 47 Caso de uso. Módulo transportes

5.2.5.3 Módulo ‘Alojamiento’

Este módulo proporciona información al usuario acerca de las opciones de alojamiento en el destino. Además, ofrece información detallada de cada uno de los posibles alojamientos.

Para lograr una mejor estructuración de la información y que ésta se pueda proporcionar al usuario de forma clara, se ha implementado un menú que contemple las posibles opciones de alojamiento de forma que en primer lugar el usuario elija el tipo de alojamiento en el que está interesado y a continuación pueda elegir el alojamiento concreto del cual desea ampliar información.

Los tipos de alojamiento contemplados son los siguientes:

- Hoteles de una estrella.

- Hoteles de dos estrellas.
- Hoteles de tres estrellas.
- Hoteles de cuatro estrellas.
- Hoteles de cinco estrellas.
- Apartamentos.
- Hostales.

Así pues, desde el fichero *menú.php* se realiza una transición transparente al usuario al fichero *menuAlojamiento.php*. En este último fichero se implementa la estructura menú que contiene como ítems los tipos de alojamiento citados anteriormente.

En función del tipo de alojamiento que el usuario proporcione a la aplicación, se realiza una consulta a la base de datos para obtener los nombres de los alojamientos correspondientes a la categoría indicada y se enumeran los nombres resultados de la consulta al usuario para que éste pueda elegir el que más le interese. En la enumeración de los nombres de los alojamientos que cumplen la condición de pertenecer a la categoría indicada por el usuario es necesario incluir pausas mediante el elemento VoiceXML `<break>` de forma que la enumeración no se produzca de forma muy rápida y el usuario pueda escuchar todas las opciones correctamente.

Dado que los nombres de los alojamientos varían en función del destino, es necesario hacer uso de una gramática dinámica que esté formada por los nombres de los alojamientos de la categoría indicada. Esta gramática se activará una vez se hayan enumerado los nombres a los usuarios para que la entrada del usuario sea reconocida.

Por último, se realiza una transición al fichero *ampliarAlojamiento.php*. En este fichero, a partir del nombre del alojamiento indicado por el usuario se realiza una consulta a la base de datos en la que se obtiene la información y se proporciona dicha información al usuario. Una vez realizado esto, se ofrece la posibilidad al usuario de elegir otro alojamiento o volver al menú principal para acceder a otro tipo de información diferente.

En la Figura 48 se puede observar un caso de uso de este módulo.

S: Elige la categoría del alojamiento:
Di una estrella o pulsa 1
Di dos estrellas o pulsa 2
Di tres estrellas o pulsa 3
Di cuatro estrellas o pulsa 4
Di cinco estrellas o pulsa 5
Di apartamentos o pulsa 6
Di hostales o pulsa 7
Di volver al menú principal o pulsa 8

U: cuatro estrellas.

S: En Madrid se pueden encontrar los siguientes alojamientos de cuatro estrellas: Gran Versailles, Petit Palace Mayor Plaza, Radisson Blu Hotel Madrid Prado, Silken Puerta Castilla. Por favor, diga el nombre del hotel del cual desea ampliar información.

U: Silken Puerta Castilla.

S: Este hotel se encuentra al lado de las torres Kio, en el paseo de la Castellana, a solo 2 paradas en metro del estadio del Real Madrid, el Santiago Bernabéu.

Si desea escuchar información acerca de otro alojamiento diga otro. Si desea ir al menú principal, diga menú.

Figura 48 Caso de uso. Módulo alojamiento

5.2.5.4 Módulo ‘Atracciones Turísticas’

Este módulo proporciona información al usuario acerca de las atracciones turísticas en el destino. Además, ofrece información detallada de cada una de las atracciones turísticas presentes en la base de datos de la aplicación.

Para lograr una mejor estructuración de la información y que ésta se pueda proporcionar al usuario de forma clara, se ha implementado un menú que contemple las posibles opciones de atracción turística de forma que el usuario elija el tipo de atracción en el que está interesado y a continuación pueda elegir la atracción concreta de la cual desea ampliar información.

Los tipos de atracción turística contemplados son los siguientes:

- Parques.
- Iglesias y Catedrales.
- Construcciones y Estructuras.
- Galerías de arte y Museos.
- Parques Temáticos.

5.3 Descripción de los módulos del acceso oral

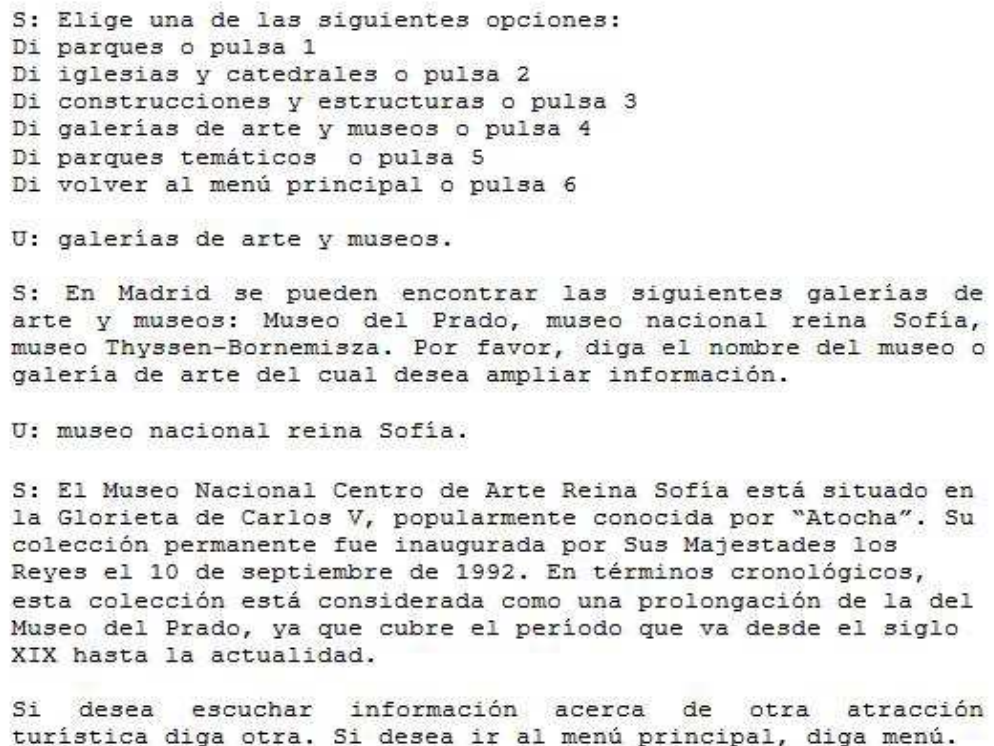
Así pues, desde el fichero *menú.php* se realiza una transición transparente al usuario al fichero *menuAtraccionesTuristicas.php*. En este último fichero se implementa la estructura menú que contiene como ítems los tipos de atracción turística citados anteriormente.

En función del tipo de atracción que el usuario proporcione a la aplicación, se realiza una consulta a la base de datos para obtener los nombres de las atracciones turísticas correspondientes a la categoría indicada y se enumeran los nombres resultados de la consulta al usuario para que este pueda elegir la que más le interese.

Dado que los nombres de las atracciones varían en función del destino, es necesario hacer uso de una gramática dinámica que esté formada por los nombres de las atracciones de la categoría indicada. Esta gramática se activará una vez se hayan enumerado los nombres a los usuarios para que la entrada del usuario sea reconocida.

Por último, se realiza una transición al fichero *ampliarAtraccionTuristica.php*. En este fichero, a partir del nombre de la atracción indicada por el usuario se realiza una consulta a la base de datos en la que se obtiene la información y se proporciona dicha información al usuario. Una vez realizado esto, se ofrece la posibilidad al usuario de elegir otra atracción turística o de volver al menú principal para acceder a otro tipo de información diferente.

En la Figura 49 se puede observar un caso de uso de este módulo.



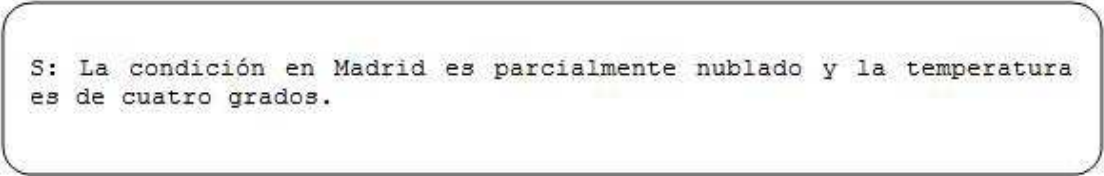
```
S: Elige una de las siguientes opciones:  
Di parques o pulsa 1  
Di iglesias y catedrales o pulsa 2  
Di construcciones y estructuras o pulsa 3  
Di galerías de arte y museos o pulsa 4  
Di parques temáticos o pulsa 5  
Di volver al menú principal o pulsa 6  
  
U: galerías de arte y museos.  
  
S: En Madrid se pueden encontrar las siguientes galerías de arte y museos: Museo del Prado, museo nacional reina Sofía, museo Thyssen-Bornemisza. Por favor, diga el nombre del museo o galería de arte del cual desea ampliar información.  
  
U: museo nacional reina Sofía.  
  
S: El Museo Nacional Centro de Arte Reina Sofía está situado en la Glorieta de Carlos V, popularmente conocida por "Atocha". Su colección permanente fue inaugurada por Sus Majestades los Reyes el 10 de septiembre de 1992. En términos cronológicos, esta colección está considerada como una prolongación de la del Museo del Prado, ya que cubre el período que va desde el siglo XIX hasta la actualidad.  
  
Si desea escuchar información acerca de otra atracción turística diga otra. Si desea ir al menú principal, diga menú.
```

Figura 49 Caso de uso. Módulo atracciones turísticas

5.2.5.5 Módulo ‘Condición Meteorológica’

Este módulo proporciona al usuario la información correspondiente a la condición meteorológica así como la temperatura actual del destino seleccionado. La implementación de este módulo se ha realizado de la misma forma que en el módulo visual por lo que se ha hecho uso de la API proporcionada por Yahoo. Así pues, a través de una URL fija a la cual es necesario concatenar el código asignado a la ciudad en cuestión, Yahoo devuelve un documento del cual gracias a otras clases auxiliares es posible obtener la información necesaria para proporcionar el servicio.

En la Figura 50 se puede observar un caso de uso de este módulo.



S: La condición en Madrid es parcialmente nublado y la temperatura es de cuatro grados.

Figura 50 Caso de uso. Módulo condición meteorológica

Capítulo 6

Evaluación de la aplicación

En este capítulo se describe la evaluación que se ha llevado a cabo de la aplicación desarrollada para este proyecto de fin de grado. Para ello, se ha desarrollado una página web con un cuestionario que recoge la valoración subjetiva de los usuarios que han utilizado la aplicación. Los resultados se han analizado para obtener, de esta forma, la completa evaluación del portal de voz.

6.1 Metodología de evaluación

Con el objetivo de optimizar los resultados de uso de la aplicación y minimizar costes es necesario realizar una evaluación de la aplicación, estudiando su rendimiento y su usabilidad.

La evaluación de la aplicación se ha llevado a cabo valorando la calidad de la misma. Para ello, se ha realizado un cuestionario utilizando la tecnología Google Forms que recoge la opinión subjetiva y el grado de satisfacción de los usuarios, obteniendo así una evaluación cualitativa de la percepción del sistema por parte de los usuarios.

En primer lugar, se pide al usuario que proporcione información sobre su grado de conocimiento acerca de nuevas tecnologías y uso de sistemas de diálogo para tener una idea del perfil de los usuarios que utilizan la aplicación.

A continuación, se realizan preguntas acerca del uso de la aplicación. Los principales aspectos que se han querido analizar son el grado en el cual el usuario valora

CAPÍTULO: Evaluación de la aplicación

que es entendido por el sistema y entiende los mensajes del mismo, la velocidad percibida de la de interacción, las seguridad de lo que se debe hacer en cada momento, la similitud en el comportamiento con un humano y el nivel de satisfacción con el sistema global. Además, se analiza las posibles diferencias de uso que pueda percibir el usuario al utilizar la aplicación web y la aplicación de voz.

Por último, se da la posibilidad al usuario de proporcionar información acerca de los aspectos que ha considerado más interesantes así como de los que mejoraría.

El cuestionario elaborado consta de 12 preguntas tal y como se puede observar en la Figura 51.

Puntúe en una escala de 1 a 5 su conocimiento previo acerca de las nuevas tecnologías para el acceso a información (1= 'Bajo', 5= 'Alto').

Puntúe en una escala de 1 a 5 su experiencia previa usando Interfaces de voz o sistemas similares (1= 'Bajo', 5= 'Alto').

¿Cuántas veces ha usado anteriormente Interfaces de Voz?

- Nunca
- Pocas veces
- A veces
- Bastantes veces
- Muchas veces

¿Qué tal le ha entendido el sistema?

- Muy mal
- Mal
- Regular
- Bien
- Muy bien

¿Cómo ha entendido los mensajes generados por el sistema?

- Muy mal
- Mal
- Regular
- Bien
- Muy bien

En su opinión la interacción fue...

- Muy Lenta
- Lenta
- Adecuada
- Rápida
- Muy rápida

¿Ha sido fácil obtener la información que usted solicitaba?
No, ha sido imposible
Sí, pero con gran dificultad
Sí, pero con ciertas dificultades
Sí, ha sido fácil
Sí, ha sido muy fácil

Establezca el nivel de dificultad del sistema para usted.
Muy difícil
Difícil
Normal
Fácil
Muy fácil

¿Cree usted que el sistema se comportó de manera similar a como lo haría un humano?
Nunca
Casi nunca
A medias
Casi siempre
Siempre

En términos generales, ¿está usted satisfecho con el sistema?
No, nada
Poco satisfecho
Satisfecho
Bastante satisfecho
Muy satisfecho

Con respecto al uso del portal web y el portal de voz...
Ha sido más fácil utilizar el portal web
Ha sido más fácil utilizar el portal de voz
No ha habido diferencia entre uno y otro.

¿Qué contenidos le han parecido más interesantes en la aplicación? ¿Cuáles mejoraría?

Figura 51 Cuestionario de evaluación

Tal y como se ha mencionado anteriormente, el cuestionario se ha realizado utilizando la tecnología *Google Forms* con el objetivo de facilitar el registro de las opiniones de los usuarios. El cuestionario puede ser accedido desde la siguiente URL: https://docs.google.com/forms/d/1qF_ZTszuppwtArpldQpgQXL649iKkNV2IINxYNIYL8/viewform

En la Figura 52 se puede observar el sitio web del formulario.

Cuestionario de Evaluación de la aplicación 'Planificador de viajes'

Puntúe en una escala de 1 a 5 su conocimiento previo acerca de las nuevas tecnologías para el acceso a información

1 2 3 4 5

Bajo Alto

Puntúe en una escala de 1 a 5 su experiencia previa usando Interfaces de voz o sistemas similares

1 2 3 4 5

Bajo Alto

¿Cuántas veces ha usado anteriormente Interfaces de Voz?

¿Qué tal le ha entendido el sistema?

¿Cómo ha entendido los mensajes generados por el sistema?

En su opinión la interacción fue...

¿Ha sido fácil obtener la información que usted solicitaba?

Establezca el nivel de dificultad del sistema para usted

¿Cree usted que el sistema se comportó de manera similar a como lo haría un humano?

En términos generales, ¿está usted satisfecho con el sistema?

Con respecto al uso del portal web y el portal de voz...

¿Qué contenidos le han parecido más interesantes en la aplicación? ¿Cuáles mejoraría?

Submit

Figura 52 Formulario de evaluación de la aplicación

6.2 Resultados de la evaluación

El test de valoración ha sido realizado por 20 usuarios a los que se les explicó las funcionalidades principales de la aplicación y se les pidió que completaran el cuestionario tras haber utilizado la aplicación. Los usuarios eligieron libremente las acciones a realizar y los módulos y submódulos a los que accedieron fueron variados. A continuación, se detallan los resultados obtenidos de la evaluación.

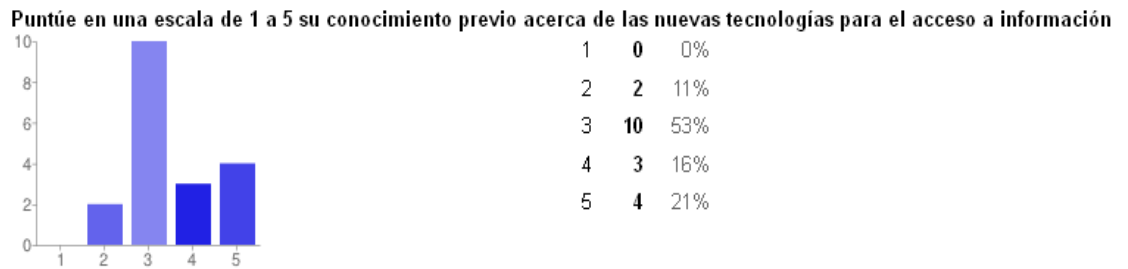


Figura 53 Resultados de la evaluación. Pregunta 1

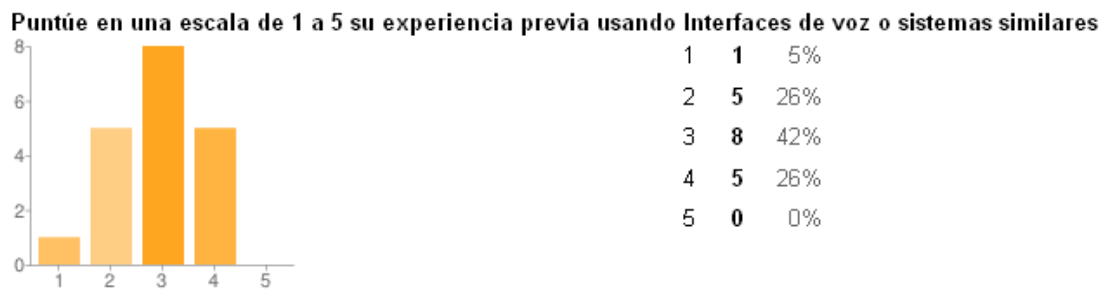
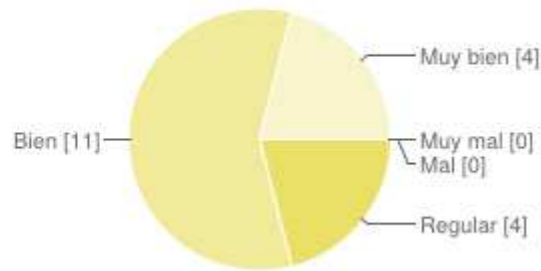


Figura 54 Resultados de la evaluación. Pregunta 2



Figura 55 Resultados de la evaluación. Pregunta 3

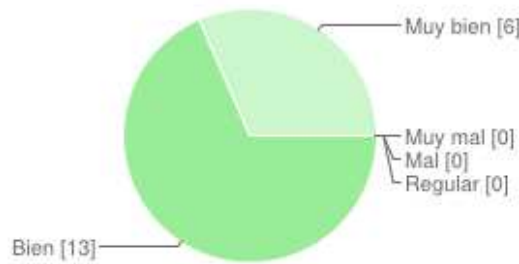
¿Qué tal le ha entendido el sistema?



Muy mal	0	0%
Mal	0	0%
Regular	4	21%
Bien	11	58%
Muy bien	4	21%

Figura 56 Resultados de la evaluación. Pregunta 4

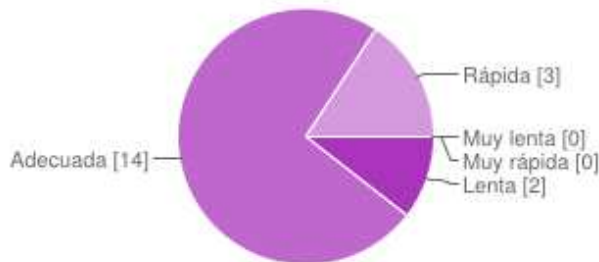
¿Cómo ha entendido los mensajes generados por el sistema?



Muy mal	0	0%
Mal	0	0%
Regular	0	0%
Bien	13	68%
Muy bien	6	32%

Figura 57 Resultados de la evaluación. Pregunta 5

En su opinión la interacción fue...



Muy lenta	0	0%
Lenta	2	11%
Adecuada	14	74%
Rápida	3	16%
Muy rápida	0	0%

Figura 58 Resultados de la evaluación. Pregunta 6

5.3 Descripción de los módulos del acceso oral

¿Ha sido fácil obtener la información que usted solicitaba?

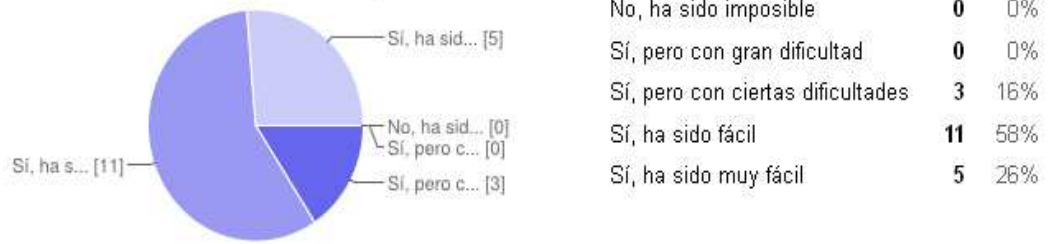


Figura 59 Resultados de la evaluación. Pregunta 7

Establezca el nivel de dificultad del sistema para usted

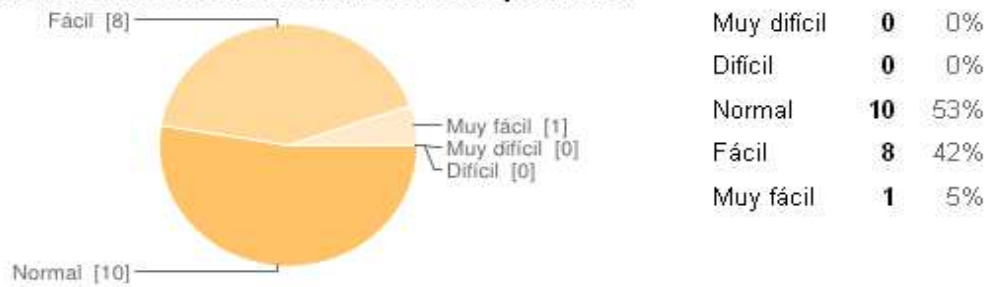


Figura 60 Resultados de la evaluación. Pregunta 8

¿Cree usted que el sistema se comportó de manera similar a como lo haría un humano?



Figura 61 Resultados de la evaluación. Pregunta 9

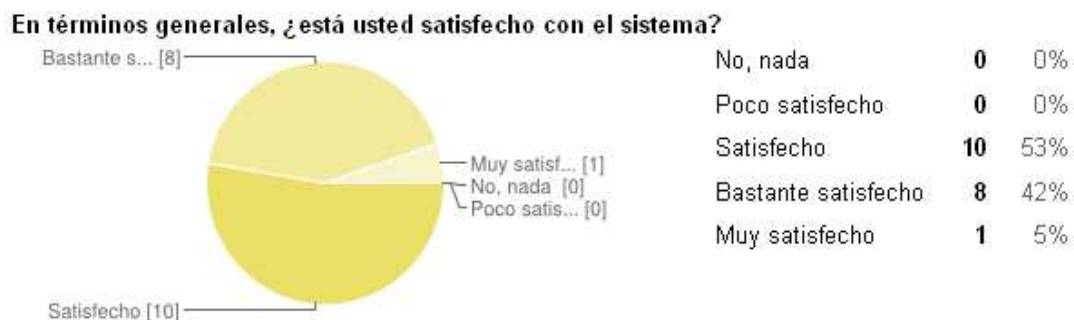


Figura 62 Resultados de la evaluación. Pregunta 10

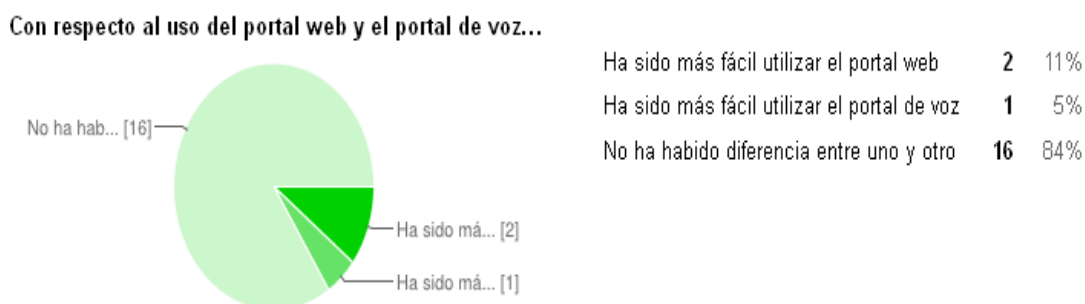


Figura 63 Resultados de la evaluación. Pregunta 11

Y por último, en la Figura 64 se pueden observar algunas de las respuestas proporcionadas por los usuarios en la última pregunta de la encuesta.

¿Qué contenidos le han parecido más interesantes en la aplicación? ¿Cuáles mejoraría?

Ampliaría la información acerca del alojamiento. La previsión meteorológica y añadir más información de los destinos. Nada. El contenido más interesante es poder acceder a la información de dos formas diferentes, ya que la hace muy flexible. Incluiría información sobre los horarios de los transportes. Lo mas interesante es la centralización de diferentes tipos de informacion en una misma aplicacion.

Figura 64 Resultados de la evaluación. Pregunta 12

Si se realiza un análisis de los resultados obtenidos se puede concluir en lo siguiente:

- El nivel de conocimiento de los usuarios a los que se les ha realizado la encuesta sobre nuevas tecnologías y uso de sistemas de diálogo es variado aunque se puede observar una tendencia a unos conocimientos medios.
- A la gran mayoría de usuarios les ha parecido adecuada la interacción con el sistema y también se valora como adecuado el ritmo de la misma.
- El sistema ha entendido bien los mensajes de los usuarios y los usuarios han entendido muy bien los mensajes generados por la aplicación.

5.3 Descripción de los módulos del acceso oral

- En cuanto a la pregunta de la similitud del comportamiento del sistema con un humano, la mayoría creen que el parecido se produce sólo a medias.
- En general, los usuarios obtienen de una forma fácil la información que buscan y están satisfechos o bastante satisfechos con el sistema.
- Con respecto a la pregunta acerca de las diferencias encontradas entre el uso de la aplicación web y la aplicación de voz, la mayoría de los usuarios afirman que no hay diferencia de uso entre ambas aplicaciones, y que por tanto, la facilidad de uso es igual en ambas.

Capítulo 7

Conclusiones y trabajo futuro

En este capítulo se presentan las principales conclusiones extraídas de la realización del presente proyecto de fin de grado. A partir de los resultados logrados se comprueba que los objetivos iniciales han sido cumplidos. Por último se hace un análisis de los posibles trabajos futuros que se podrían llevar a cabo en la aplicación para aumentar su funcionalidad y su eficiencia.

7.1 Conclusiones

En este proyecto de fin de grado se ha implementado una aplicación cuya interfaz es multimodal, es decir, el usuario puede interactuar con la aplicación de diferentes modos para proveer las entradas y obtener las salidas generadas por la aplicación.

Las modalidades de interacción que ofrece la aplicación desarrollada son la visual mediante una página web dinámica, y la oral mediante una aplicación de voz que ofrece la misma información que podría ser consultada visualmente.

La aplicación implementada consiste en un planificador de viajes que permite a los usuarios consultar información acerca de una ciudad determinada que deben proporcionar. La información que se puede consultar es la siguiente:

- Información acerca de cómo llegar al destino.
- Información sobre los medios de transporte más adecuados.
- Información acerca de las posibilidades de alojamiento.
- Información acerca de las atracciones turísticas de interés.
- Información meteorológica.

Estos contenidos se almacenan en una base de datos de forma que puedan ser actualizados y verídicos.

La base de datos se ha realizado mediante el sistema de gestión de bases de datos relacionales MySQL. Este gestor de bases de datos presenta numerosas ventajas como por ejemplo la garantía de integridad referencial o la evitación de duplicidad de registros. Gracias a estas ventajas junto con el diseño del modelo conceptual realizado, se ha conseguido que la base de datos esté adaptada a las necesidades de la aplicación y que las consultas sobre la misma sean rápidas y eficientes.

Con respecto a la página web, se ha utilizado todo el potencial del desarrollo web de forma que la página web creada es accesible y usable. La interfaz se ha diseñado de forma que se ajusta a los principios de usabilidad y accesibilidad, por lo que los menús y elementos principales se mantienen fijos a lo largo de todas las páginas con el objetivo de hacer más fácil al usuario la navegación en el sitio web.

Para el desarrollo del acceso oral se ha analizado el estado del arte de los sistemas de diálogo en la actualidad incluyendo su arquitectura, criterios para su clasificación y principales funcionalidades, así como del lenguaje utilizado para la implementación de la aplicación de voz, VoiceXML. Este estándar ha permitido construir diálogos hablados de forma sencilla y ha ofrecido numerosas funcionalidades como por ejemplo, reconocimiento de voz, reconocimiento de entrada DTMF, control del flujo del diálogo, etc. Otra característica fundamental de este estándar que ha facilitado el desarrollo del módulo oral es que VoiceXML separa la interacción del código de usuario de la lógica de servidor, siendo una tecnología independiente de la plataforma que permite la portabilidad y transferencia de datos entre aplicaciones heterogéneas.

La plataforma que ha proporcionado la infraestructura y componentes de reconocimiento y síntesis de voz necesarios para crear y utilizar la presente aplicación ha sido Voxeo. Esta plataforma presenta una gran facilidad para la creación y uso de aplicaciones de voz y ofrece múltiples herramientas que se han utilizado para el desarrollo y optimización, como son un depurador, soporte y la completa documentación.

El diseño de la aplicación de voz se ha llevado a cabo mediante distintos módulos implementados mediante la estructura de menús en los que el usuario puede interactuar con la aplicación mediante la voz o mediante señalización DTMF para conseguir que la aplicación de voz sea fácil de usar. Así pues, las gramáticas utilizadas se han diseñado para conseguir una mejor manejabilidad de los servicios por parte del usuario, de tal forma que en cada punto se sepa qué opción elegir para encontrar la información buscada o realizar la acción deseada.

La aplicación de voz permite a los usuarios acceder a toda la información disponible en el portal web, por tanto, se ha cumplido uno de los objetivos principales establecido al inicio de este proyecto de fin de grado.

Mediante el portal de voz se ha elaborado un sistema de acceso a la información sin la necesidad de disponer de un dispositivo que permita el acceso a un navegador web, ofreciendo los mismos contenidos que el navegador pero a través de otros dispositivos o medios alternativos como un teléfono, ya sea fijo o móvil, o incluso mediante Skype. De este modo, se proporciona movilidad, flexibilidad y facilidad de acceso al permitir el uso de la aplicación en entornos y situaciones en los que no podrían ser utilizados un ordenador con teclado y ratón o no existir acceso a Internet.

Como conclusión final, los objetivos establecidos al inicio de este proyecto de fin de grado se han cumplido. Se ha diseñado e implementado una aplicación cuya interfaz permite el acceso a la misma a distintos tipos de usuarios. Esto se ha conseguido mediante la implementación de la aplicación de voz, la cual permite que usuarios con impedimentos físicos o visuales que no puedan acceder a la información visualmente, así como usuarios que no dispongan de un dispositivo y/o conocimientos informáticos para acceder a la página web, puedan hacer uso de esta aplicación.

7.2 Trabajo futuro

A continuación se detallan los aspectos sobre los que se podría trabajar para obtener mayores prestaciones en la aplicación desarrollada. El trabajo futuro sobre la aplicación desarrollada se puede dividir en dos ámbitos principales:

- Contenidos y funcionalidades adicionales

Los módulos individuales de la aplicación podrían ampliar el contenido y funcionalidad de la misma con el objetivo de proporcionar más información útil a los usuarios.

- En el módulo ‘Cómo llegar’ se podría implementar la infraestructura necesaria, tanto en el módulo de acceso visual como en el oral, de forma que los usuarios pudieran reservar billetes de avión, tren, etc en el caso de que lo estimen oportuno.
- En el módulo ‘Transportes’ se podría facilitar el horario o rutas concretas de los transportes a partir de una fecha indicada.
- En el módulo ‘Alojamiento’ se podrían incluir críticas de los hoteles proporcionadas por usuarios que ya hayan visitado dichos hoteles con el objetivo de facilitar al usuario de la aplicación la elección del alojamiento. Además, se podría considerar la implementación de la plataforma necesaria para permitir a los usuarios realizar las reservas de hoteles a partir de la página web o de la aplicación de voz.

- En el módulo ‘Condición Meteorológica’ se podría incluir una previsión meteorológica de los próximos días en el destino de forma que el usuario se haga una idea del clima que tendrá cuando visite el citado destino.

Con respecto a la administración de la aplicación, se podría incluir una herramienta de gestión que facilitará la administración de la información referente a la página web y a la aplicación de voz, incluyendo el mantenimiento e inclusión de gramáticas de forma que el administrador de la aplicación pudiera actualizar o añadir más información o funcionalidad de forma más sencilla e intuitiva. Esta herramienta podría recoger información de los módulos más consultados de forma que se sepa que información consideran más importante los usuarios y proporcionar así un mejor servicio.

- Idiomas

Dado que la plataforma Voxeo permite crear aplicaciones y asignar números de teléfono válidos en más de 40 países, una línea de trabajo futuro de esta aplicación, sería su implementación en distintos idiomas, de forma que se pudiera extender su uso mundialmente al resto de países de habla no hispana. A su vez, la plataforma Voxeo, proporciona servicios de reconocimiento de voz en distintos idiomas, lo que facilitaría esta línea de trabajo.

De este modo, en función del idioma en el que el usuario desee utilizar la aplicación existiría un número de teléfono válido para ello, y del mismo modo, en el portal web, se daría la opción de acceder a la información en el idioma deseado. A partir de este punto se gestionaría el resto de la arquitectura de la aplicación según el idioma elegido.

Presupuesto

A continuación se presenta el presupuesto del presente proyecto de fin de grado. Se incluye un desglose de las fases del mismo, así como un desglose de costes de personal y material. Por último, se indica el coste total del proyecto.

- **Tareas**

- Fase de planificación

- Definición de objetivos. Duración → 31 días.
- Documentación. Duración → 20 días.
- Realización de ejemplos prácticos. → 10 días.

- Fase de desarrollo

- Análisis y diseño de la base de datos. Duración → 18 días.
- Análisis y diseño de la aplicación web. Duración → 12 días.
- Análisis y diseño de la aplicación de voz. Duración → 11 días.
- Desarrollo de la base de datos. Duración → 12 días.
- Desarrollo de la aplicación web. Duración → 22 días.
- Desarrollo de la aplicación de voz. Duración → 25 días.
- Pruebas unitarias. Duración → 17 días.

- Pruebas de integración. Duración → 7 días.
- Evaluación de la aplicación. Duración → 10 días.
- Fase de documentación
 - Redacción de la memoria. Duración → 31 días.
 - Realización de la presentación. Duración → 10 días.
- **Recursos**

Los recursos utilizados para la realización del presente proyecto son los que se detallan a continuación.

 - Recursos software
 - Editor de textos Notepad ++. Coste → 0 €
 - Plataforma Voxeo. Coste → 0 €
 - Skype. Coste → 0 €
 - Licencia de Microsoft Office 2010. Coste → 269 €
 - Recursos hardware
 - Ordenador Portátil. Coste → 950 €
 - Auriculares con micrófono. Coste → 90 €
 - Servidor x10Hosting. Coste → 0 €
 - Recursos humanos

En la realización de este proyecto han participado dos personas, el director de proyecto y el desarrollador. Las horas dedicadas al proyecto han sido de una media de 6 horas diarias, contemplando semanas de 7 días. El coste de un ingeniero es de 40 €/hora.

En la siguiente tabla se indica el desglose de costes en recursos humanos en cada fase del proyecto.

Fase	Días	Importe
Fase de planificación	61	14.640 €
Fase de desarrollo	134	32.160 €
Fase de documentación	41	9.840 €
Total	236	56.640 €

Tabla 1 Desglose de costes en recursos humanos

Por tanto, el **coste total** del proyecto se detalla en la Tabla 10.

Concepto	Importe
Recursos Software	269 €
Recursos Hardware	1.040 €
Recursos Humanos	56.640 €
Subtotal	57.949 €
(21% IVA)	12.169,29 €
TOTAL	70.118,29 €

Tabla 2 Coste total del proyecto

El presupuesto total de este proyecto asciende a la cantidad de **SETENTA MIL CIENTO DIECIOCHO EUROS Y VEINTINUEVE CENTIMOS.**

Madrid, a 11 de Febrero

Fdo. Jennifer Catalina Vaquero

Glosario

CSS	<i>Cascading Style Sheets</i>
DTMF	<i>Dual-Tone Multi-Frequency</i>
HTML	<i>HyperText Markup Language</i>
IVR	<i>Interactive Voice Response</i>
PHP	<i>Hypertext Pre-processor</i>
SLS	<i>Spoken Language Systems</i>
SQL	<i>Structured Query Language</i>
URI	<i>Uniform Resource Identifier</i>
VoiceXML	<i>Voice eXtensible Markup Language</i>
W3C	<i>World Wide Web Consortium</i>
WAI	<i>Web Accessibility Initiative</i>
UTF-8	<i>8-bit Unicode Transformation Format</i>
SMF	<i>Simple Machines Forum</i>
FTP	<i>File Transfer Protocol</i>

API	<i>Application programming interface</i>
CCXML	<i>Call Control Extensible Language</i>

Referencias

- [1.1] Incera, José A. Nuevas Interfaces y sus Aplicaciones en las Tecnologías de Información y Comunicaciones. Reporte técnico 1-1007. Octubre 2007. Disponible [Internet]: <http://allman.rhon.itam.mx/~jincera/EvolHCI.pdf> [17 de enero de 2013]
- [1.2] Bruno Dumas, Denis Lalannel, Sharon Oviatt. Multimodal Interfaces: A Survey of Principles, Models and Frameworks. 2009. Disponible [Internet]: http://diuf.unifr.ch/people/lalanned/Articles/mmi_chapter_final.pdf [17 de enero de 2013]
- [1.3] *HyperText Markup Language (HTML)* Disponible [Internet]: <http://www.w3.org/html/> [10 de febrero de 2013]
- [1.4] *Web Accessibility Initiative (WAI)* Disponible [Internet]: <http://www.w3.org/WAI/> [17 de enero de 2013]
- [1.5] *World Wide Web Consortium (W3C)* Disponible [Internet]: <http://www.w3c.es/> [17 de enero de 2013]
- [1.6] Carlos Zapata, Carlos M., Jhon E. Mesa. Los Modelos de Diálogo y sus Aplicaciones en Sistemas de Diálogo Hombre-Máquina. Diciembre 2009. Disponible [Internet]: <http://www.redalyc.org/redalyc/pdf/496/49612068021.pdf> [17 de enero de 2013]
- [1.7] Llisterri Joaquin. Los sistemas de diálogo. Disponible [Internet]: http://liceu.uab.es/~joaquin/speech_technology/tecnol_parla/dialogue/dialogo_general/sistemas_dialogo.html [17 de enero de 2013]
- [1.8] *Voice Extensible Markup Language (VoiceXML)* Disponible [Internet]: <http://www.w3.org/TR/voicexml20/> [17 de enero de 2013]

- [1.9] *Plataforma Voxeo* Disponible [Internet]: <https://evolution.voxeo.com/> [17 de Febrero de 2013]
- [1.10] *phpMyAdmin* Disponible [Internet]: http://www.phpmyadmin.net/home_page/index.php [17 de Febrero de 2013]
- [1.11] *Skype* Disponible [Internet]: <http://www.skype.com/es/> [17 de Febrero de 2013]
- [2.1] Pieraccini, Roberto, Rabiner Lawrence. *The Voice in the Machine: Building Computers That Understand Speech*. March 23, 2012. ISBN-10: 0262016850 ISBN-13: 978-0262016858
- [2.2] López Cozar Delgado, Ramón, Masahiro, Araki. *Spoken, Multilingual and Multimodal Dialogue Systems: Development and Assessment*. November 11, 2005. ISBN-10: 0470021551 ISBN-13: 978-0470021552
- [2.3] Llisterri, J. (2006) "Introducción a los sistemas de diálogo", in LLISTERRI, J.-MACHUCA, M. J. (Eds.) Los sistemas de diálogo. Bellaterra - Soria: Universitat Autònoma de Barcelona, Servei de Publicacions - Fundación Duques de Soria (Manuals de la Universitat Autònoma de Barcelona, Lingüística, 45), págs. 11-21.
- [2.4] David Griol, Zoraida Callejas, Ramón López-Cózar, Ana Gutiérrez. Utilización De los sistemas de diálogo hablado para el acceso a la información en diferentes dominios. II Conferencia Internacional sobre Brecha Digital e Inclusión Social. (Leganés, Madrid, del 28-30 de octubre de 2009).
- [2.5] Eduardo Sosa. Procesamiento del lenguaje natural: revisión del estado actual, bases teóricas y aplicaciones (Parte I). Sección científica de Biblioteconomía y Documentación. Universitat Pompeu Fabra. Barcelona.
- [2.6] Carlos García Ibáñez, Raquel Hervás Ballesteros, Pablo Gervás. *Una arquitectura software para el desarrollo de aplicaciones de generación de lenguaje natural*. Disponible [Internet]: <http://www.sepln.org/revistaSEPLN/revista/33/33-Pag111.pdf> [10 de Febrero de 2013]
- [2.7] Griol, David. Desarrollo y evaluación de Diferentes Metodologías para la Gestión Automática del Diálogo. Tesis Doctoral. Universidad Politécnica de Valencia, 2007.
- [2.8] *MASK (Multimodal-Multimedia Automated Service Kiosk)* Disponible [Internet]: <http://www.limsi.fr/tlp/kiosk-sncf.html> [19 de enero de 2013]
- [2.9] *August* Disponible [Internet]: <http://www.speech.kth.se/august/> [19 de enero de 2013]
- [2.10] *Mercury* Disponible [Internet]: <http://groups.csail.mit.edu/sls/research/mercury.shtml> [19 de enero de 2013]

- [2.11] *DARPA Communicator* Disponible [Internet]: <http://ssli.ee.washington.edu/ssli/projects/communicator.html> [19 de enero de 2013]
- [2.12] *LARRI (Language-based Agent for Retrieval of Repair Information)* Disponible [Internet]: <http://repository.cmu.edu/cgi/viewcontent.cgi?article=2346&context=compsci> [19 de enero de 2013]
- [2.13] *Jupiter* Disponible [Internet]: <http://groups.csail.mit.edu/sls/research/jupiter.shtml> [19 de enero de 2013]
- [2.14] *ITSPOKE* Disponible [Internet]: <http://acl.ldc.upenn.edu/N/N04/N04-3002.pdf> [19 de enero de 2013]
- [2.15] *Vocaliza* Disponible [Internet]: <http://www.vocaliza.es/> [19 de enero de 2013]
- [2.16] *VICO (Virtual Intelligent Co-Driver)* Disponible [Internet]: <http://www.cs.cmu.edu/~dgroup/papers/geutner02.pdf> [19 de enero de 2013]
- [2.17] *AthosMail* Disponible [Internet]: <http://www.connexor.com/nlplib/?q=node/178> [19 de enero de 2013]
- [2.18] *COGNIRON* Disponible [Internet]: <http://www.cogniron.org/final/Home.php> [19 de enero de 2013]
- [2.19] *LET'S GO!* Disponible [Internet]: <http://www.speech.cs.cmu.edu/letsgo/> [19 de enero de 2013]
- [2.20] *DisCoH* Disponible [Internet]: <http://research.microsoft.com/enus/um/people/dbohus/docs/conquest-paper-final.pdf> [19 de enero de 2013]
- [2.21] *Siri* Disponible [Internet]: <http://www.apple.com/ios/siri/> [19 de enero de 2013]
- [2.22] *Uniform Resource Identifiers (URI): Generic Syntax*, IETF RFC 2396, 1998. Disponible: <http://www.ietf.org/rfc/rfc2396.txt> [19 de enero de 2013]
- [2.23] *Voxeo*. Disponible [Internet]: <http://voxeo.com/> [10 de Febrero 2013]
- [2.24] *CCXML* Disponible [Internet] <http://www.w3.org/TR/ccxml/> [10 de Febrero de 2013]
- [2.25] *CallFire* Disponible [Internet]: <https://www.callfire.com/> [20 de enero de 2013]
- [2.26] *IVR Tech Group* Disponible [Internet]: <http://www.ivrtechgroup.com/products-services/custom-ivr> [20 de enero de 2013]

- [2.27] *SiatVox* Disponible [Internet]: <http://www.prosodie.es/index.php?id=1944> [20 de enero de 2013]
- [2.28] *Loquendo*. Disponible [Internet]: <http://www.loquendo.com/es/> [10 de Febrero de 2013]
- [3.1] *X10Hosting*. Disponible [Internet]: <http://x10hosting.com/> [10 de Febrero de 2013]

