



Universidad
Carlos III de Madrid

ESCUELA POLITÉCNICA SUPERIOR

I.T.T. SISTEMAS DE TELECOMUNICACIÓN

Proyecto Fin de Carrera

**FACILEASYMO: DESARROLLO DE UN SISTEMA DE DIÁLOGO
MULTIMODAL PARA LA RESOLUCIÓN DE INCIDENCIAS
INFORMÁTICAS**

Autor: Ricardo Bárzano Santamaría

Tutor/Director: David Griol Barres

Título: FACILEASYMO: DESARROLLO DE UN SISTEMA DE DIÁLOGO MULTIMODAL PARA LA RESOLUCIÓN DE INCIDENCIAS INFORMÁTICAS.

Autor: Ricardo Bárzano Santamaría

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

Quiero agradecer en primer lugar a mis padres, Juan y Charo, por su constante sacrificio y apoyo, por no dejarme tirar la toalla y recordarme siempre cual era la meta.

A mi novia Laura por su apoyo incondicional y por ser mi refugio en los buenos y sobre todo en los malos momentos.

A los amigos y a la familia, por su ánimo y aliento que sin duda han hecho más llevadero el camino.

A Víctor, mi compañero de proyecto, de carrera, de trabajo, vecino y AMIGO, por las muchas vivencias que aún nos quedan por compartir.

Y sin olvidarme de nuestro tutor, David, por su gran ayuda y confiar en nosotros para llevar a cabo este proyecto.

Sin todos ellos hubiera sido imposible. GRACIAS por creer en mí.

Resumen del proyecto

Este Proyecto Fin de Carrera consiste en el desarrollo de una herramienta que ayude a resolver sencillas y rutinarias incidencias informáticas, tanto a nivel doméstico como profesional. Esto se conseguirá a través de la ayuda de un asistente que conseguirá identificar el motivo de la incidencia, y a continuación ofrecerá una serie de sencillos pasos para ayudar al interlocutor a solucionarlo.

Esta herramienta ofrece una comunicación multimodal, permitiendo al usuario interactuar con la herramienta y también a través del ordenador mediante una página web creada, en la cual se almacenaran los diferentes datos facilitados por cada uno de los usuarios y el historial de incidencias del que consta. En esta plataforma, podremos también acceder a distintos manuales de aparatos informáticos, pudiendo ayudar al usuario a entender un poco mejor su problema. Además en el momento que lo desee podrá modificar cualquiera de sus datos personales, permitiendo al portal estar constantemente actualizado.

Esta página también está adaptada para dispositivos móviles, de esta manera facilitamos el acceso desde móviles y tablets, ajustándose a cada uno de ellos.

La herramienta nos permite solucionar nosotros mismos incidencias desde el no funcionamiento del teclado, ratón, altavoces, etc., hasta poder averiguar por qué nuestra impresora no nos imprime o por qué el monitor no se enciende.

La aplicación se apoya además en otras tecnologías adicionales, como son la utilización de bases de datos (MySQL), servidores web y de VoiceXML (x10Hosting y Voxeo Evolution) y uso de diferentes lenguajes de programación (PHP, HTML, JavaScript), que la hacen más dinámica y flexible, aumentando su calidad y eficiencia.

El proyecto se complementa con un estudio detallado y análisis de los sistemas de diálogo y de la aplicación del estándar VoiceXML para su desarrollo, llevado a cabo con el objetivo de aplicar todo su potencial para completar con éxito este Proyecto Final de Carrera.

Palabras clave: sistemas de diálogo, interacción oral, portal de voz, VoiceXML, Voxeo, página web, dispositivos móviles, FACILEASYMO.

Abstract

This project consists in the development of a tool to help solve simple and routine informatics incidents, domestic and professional problems. This will be achieved through the help of an assistant to get identify the cause of the incident, and then offer a series of simple steps to solve the problem.

This tool offers a multimodal communication, allowing users to interact with the tool, also via computer by a website created, in which the different information provided by each of the users and the historical of the incidences were saved. On this platform, we can also access to different manuals informatics devices, being able to help the user to understand a little better your problem. Also in any moment the user can modify any of your personal data quickly, allowing the website to be constantly updated.

This website is also adapted for mobile devices, so we facilitate access from mobile and tablets, adjusting to each.

The tool allows us to resolve ourselves different incidents, for example the not operation to keyboard, mouse, speakers, etc., or even you can figure out why the printer does not print or the monitor does not turn on.

The application rests on other additional technologies, such as the use of databases (MySQL), web servers and VoiceXML (x10Hosting and Voxeo Evolution) and using different programming languages (PHP, HTML, JavaScript), which make it more dynamic and flexible, increasing their quality and efficiency.

The project is complemented by a detailed study and analysis of dialogue systems and the implementation of the VoiceXML standard for development, carried out in order to implement their potential to successfully complete this project.

Keywords: dialogue systems, oral interaction, voice portal, VoiceXML, Voxeo, website, mobile devices, FACILEASYMO.

Índice General

1	Introducción	13
1.1	Motivación.....	14
1.2	Objetivos.....	16
1.3	Fases del desarrollo	16
1.4	Medios y documentación utilizados.....	19
2	Estado del arte	20
2.1	Los sistemas de dialogo.....	20
2.1.1	Introducción a los sistemas de dialogo.....	20
2.1.2	Clasificación de los sistemas de diálogo	23
2.1.3	Ejemplos de aplicaciones.....	26
2.2	Voice Extensible Markup Language (VoiceXML)	28
2.2.1	Introducción al estándar VoiceXML.....	28
2.2.2	Control y gestión del flujo del diálogo	29
2.2.2.1	Etiquetas básicas sobre diálogos en VoiceXML.....	29
2.2.2.2	Etiquetas de control de diálogos.....	35
2.3	Plataforma Voxeo	37
2.3.1	¿Qué es Voxeo?	37
2.3.2	¿Cómo funciona Voxeo?	38
2.3.3	Desarrollo de una aplicación VoiceXML con la plataforma Voxeo	39
2.3.3.1	Primer contacto con Voxeo	39
2.3.3.2	Selección de la plataforma de la aplicación de voz.....	42
2.3.3.3	Manejo de ficheros	43
2.3.3.4	Creación de una aplicación y vinculación a un número de teléfono.....	44
2.3.3.5	Llamada a la aplicación	47
2.3.4	Otras implementaciones del estándar VoiceXML.....	47
3	Descripción general del sistema desarrollado	50
3.1	Presentación del sistema	50

3.2	Tecnologías utilizadas.....	52
3.2.1	Intérprete VoiceXML: Voxeo	52
3.2.2	Base de datos: MySQL.....	53
3.2.3	Administrador de base de datos: phpMyAdmin	54
3.2.4	Servidor web	55
3.3	Implementación de las operaciones generales	56
3.3.1	Tratamiento de la información.....	56
3.3.1.1	Información estática	57
3.3.1.2	Información dinámica	57
3.3.2	Gramáticas.....	58
3.3.2.1	Gramáticas estáticas.....	58
3.3.2.2	Gramáticas dinámicas.....	58
3.4	Gestión de bases de datos	59
3.4.1	Formato de almacenamiento	59
3.4.2	Codificación.....	60
3.4.3	Jerarquía de la base de datos del sistema	61
4	Descripción detallada de los módulos del sistema	64
4.1	Aplicación de voz	64
4.1.1	Diagrama general de ejecución	65
4.1.2	Index.....	66
4.1.3	Nueva incidencia	66
4.1.4	Gestor de incidencia.....	66
4.1.5	Encuesta	66
4.1.6	Exit	67
4.1.7	Ejemplo práctico de llamada a Facileasymo.....	67
4.2	Plataforma web.....	69
4.2.1	Vista principal	69
4.2.2	Registro y acceso	72
4.2.3	Cambio de contraseña olvidada	74
4.2.4	Inicio	76

4.2.5	Mi panel	77
4.2.6	Categoría	82
5	Evaluación de la aplicación	84
5.1	Metodología de evaluación.....	84
5.2	Resultados de la evaluación.....	85
6	Conclusiones y trabajo futuro	87
6.1	Conclusiones	87
6.2	Trabajo futuro	88

Índice de Figuras

Figura 1. Diagrama WBS representando las tareas definidas para el proyecto.....	18
Figura 2. Estructura de un sistema de dialogo	21
Figura 3. Visión general del código de una aplicación VoiceXML.....	29
Figura 4. Visión general del código de una aplicación VoiceXML.....	30
Figura 4. Interpretación de una página web	38
Figura 5. Interpretación de una página VoiceXML.....	38
Figura 7. Registro usuario nuevo en Voxeo.....	39
Figura 8. Acceso a la cuenta de Voxeo.....	39
Figura 9. Vista general cuenta de Voxeo	40
Figura 10. Directorio /root.....	43
Figura 11. Directorio /root/www	44
Figura 12. Application Manager	44
Figura 13. Create a new application.....	45
Figura 14. Tras especificar URL del código fuente, creada con éxito	45
Figura 15. "Contact Methods" de la aplicación 'PRUEBA'	46
Figura 16. Aplicaciones creadas en la plataforma.....	46
Figura 17. Llamada a la aplicación mediante Skype.....	47
Figura 18. Arquitectura de la aplicación	51
Figura 19. Página de inicio de PHPMyAdmin.....	55

Figura 20. Panel de control de usuario de x10hosting.....	56
Figura 21. Bases de datos utilizadas.....	61
Figura 22. Base de datos 'facileas_contenido'	61
Figura 23. Base de datos 'facileas_sesion'	62
Figura 24. Diagrama general de la aplicación de Voz	65
Figura 25. Pantalla de inicio de Facileasymo.x10.bz desde navegador de escritorio.....	69
Figura 26. Características del sistema de voz.....	70
Figura 27. Descripción de característica de responsive design.....	70
Figura 27. Vista principal de facileasymo.x10.bz desde dispositivo móvil	71
Figura 29. Ligthbox de registro usuario nuevo	72
Figura 30. Email de verificación de registro de usuario.....	73
Figura 31. Email de verificación de registro de usuario.....	73
Figura 32. Pantalla de acceso para usuario ya registrado	74
Figura 33. No recuerdo mi contraseña	75
Figura 34. Ejemplo de email "No recuerdo mi contraseña"	75
Figura 35. Cambio de contraseña correcto.....	76
Figura 36. Pantalla inicio navegador de escritorio.....	76
Figura 37. Pantalla inicio desde dispositivo móvil.....	77
Figura 38. Historial de llamadas desde navegador PC.....	78
Figura 39. Historial de llamadas desde dispositivo móvil.....	78
Figura 40. Datos personales desde navegador de escritorio	79

Figura 41. Datos personales desde dispositivo móvil.....	79
Figura 42. Cambio de contraseña.....	80
Figura 43. Cambio de correo electrónico.....	81
Figura 44. Email de confirmación de cambio correo electrónico	81
Figura 45. Darse de baja como usuario.....	82
Figura 46. Vista 'Categoría' desde navegador PC.....	82
Figura 47. Vista 'Categoría' desde dispositivo móvil.....	83
Figura 48. Encuesta desde navegador de escritorio	85
Figura 49. Resultados globales de la encuesta	85
Figura 50. Diagrama de Gantt del 22/08/2013 al 07/12/2013	91
Figura 51. Diagrama de Gantt del 07/12/2013 al 25/03/2014	91
Figura 52. Diagrama de Gantt del 25/03/2014 al 17/07/2014	92
Figura 53. Diagrama de ciclo de relación marca - consumidor potencial	96

Índice de Tablas

Tabla 1 – Ejemplo <form>	31
Tabla 2 – Ejemplo <field>.....	32
Tabla 3 – Ejemplo <filed>.....	33
Tabla 4 – Ejemplo <record>	34
Tabla 5 – Ejemplo <subdialog>.....	35
Tabla 6 – Ejemplo <block>	36
Tabla 7 – Ejemplo <initial>.....	37
Tabla 8 – Ejemplo gramática estática.	58
Tabla 9 – Ejemplo gramática dinámica. Captura de número de teléfono.....	59
Tabla 10 – Ejemplo gramática dinámica. Se guarda el teléfono en una variable	59
Tabla 11 – Ejemplo gramática dinámica. Coteja número de teléfono con base de datos.....	59
Tabla 12 – Ejemplo gramática dinámica. Bienvenida personalizada	59
Tabla 13 – Transcripción de una llamada de resolución de incidencia.	68
Tabla 14 – Valoraciones de los usuarios.....	86
Tabla 15 – Tareas de las fases del proyecto	90
Tabla 16 – Detalle de Costes de Recursos Humanos del Proyecto.....	95
Tabla 17 – Detalle de Coste Total del Proyecto.....	95

1 Introducción

En la sociedad actual, donde las tecnologías de la información juegan un papel fundamental, la interacción de los seres humanos con los sistemas informáticos, es cada vez más importante. Una de las muchas formas de interactuar con dichos sistemas, es mediante el uso de sistemas de diálogo, y en particular, como se verá en adelante, sistemas de diálogo hablado.

El ser humano, por su naturaleza, necesita de la comunicación con sus semejantes, y concretamente, mediante la voz, es capaz de hacerlo de forma natural y eficiente. El uso de la voz en las tecnologías de información, aporta dos ventajas:

Permite el uso de aplicaciones y sistemas en los que es imposible utilizar los interfaces tradicionales de entrada salida (ratón, teclado, pantallas, etc.), por ejemplo, en un automóvil.

Posibilita el acceso a dichas aplicaciones y sistemas a personas con cierto tipo de discapacidades (por ejemplo visuales o motoras).

Se puede definir un sistema de diálogo como un sistema informático destinado a interactuar con un ser humano de forma coherente, mediante el uso de voz, texto, etc.

Dichos sistemas pueden se pueden englobar según la siguiente clasificación:

- Según su modalidad: basados en texto, basados en diálogo hablado, basados en GUI's, multimodales, etc.
- Según el dispositivo usado: sistemas telefónicos, PDA's, dispositivos de automóvil, sistemas robóticos, etc.
- Según su estilo: basados en órdenes, basados en menús, basados en lenguaje natural, etc.
- Según la iniciativa del diálogo: iniciativa del sistema, iniciativa del usuario, iniciativa mixta.

- Según su funcionalidad: servicios de información, entretenimiento, educacionales, sistemas de avisos, asistencia sanitaria, etc.

Los sistemas de diálogo hablado, en los que se centra el presente trabajo, se componen principalmente de los siguientes elementos:

En los siguientes capítulos se introducirá un estado del arte más detallado, describiendo en más profundidad las temáticas principales relacionadas con el desarrollo de los sistemas de diálogo hablados.

Asimismo, a continuación se comentarán más extensamente los objetivos que pretende cubrir la realización del presente proyecto.

Se mostrará igualmente el trabajo realizado, así como una descripción de la aplicación, un diseño detallado de la misma, su arquitectura y los diferentes componentes que la integran.

Seguidamente se describirá el conjunto de pruebas funcionales que permiten comprobar el correcto funcionamiento de la aplicación, así como una serie de estadísticas de uso recogidas en un período de tiempo durante el cual el servicio de voz ha estado disponible a un conjunto de usuarios.

Finalmente, también se comentarán las conclusiones alcanzadas después de la realización del proyecto, y posibles futuras líneas de trabajo que pueden contribuir a la mejora del mismo. Como punto final, se recogen una serie de anexos con información extra complementaria a lo expuesto en el resto del documento.

1.1 Motivación

En el actual siglo XXI se está produciendo cada vez más un auge en el uso de las tecnologías, desde el punto de vista del desarrollo y puesta en escena y también el descubrimiento de muchas de ellas que permiten facilitarnos el día a día. Este avance tan rápido y positivo, también conlleva dificultades, es decir, al igual que las nuevas generaciones ven normal el uso de estas tecnologías a diario, a otras les cuesta un poco más integrarlas a su vida cotidiana.

Es así como nace este Proyecto Fin de Carrera, tras la experiencia que ambos tuvimos durante el periodo que compartimos beca en el departamento de soporte informático del Ayuntamiento de Leganés en el cual se recibían llamadas a diario de problemas en muchas ocasiones sencillos de resolver y que implicaban el desplazamiento del técnico a los distintos edificios repartidos por todo el municipio. De esta manera surge la idea de crear Facileasymo, con las ganas y el objetivo de poder ayudar a esas personas que se puedan encontrar con dificultades comunes, y que en ocasiones hacen que la persona pierda el interés y la motivación por su uso.

Este proyecto tiene dos pilares fundamentales, uno de ellos trata la comunicación multimodal, que brevemente podríamos resumir como la forma de comunicarse por parte del ser humano con una máquina utilizando diferentes modalidades de entrada y/o de salida. Para conseguir esta comunicación diseñamos la herramienta de manera que fuese posible la interacción oral con la misma, facilitando además su uso por parte de personas mayores o con discapacidades motoras o visuales.

El otro pilar del proyecto es la plataforma web que da soporte visual a la información en todo momento a todos los usuarios que deseen y puedan disponer de una conexión a Internet. A través de esta página podrán registrarse y acceder con su usuario para poder consultar sus últimas incidencias registradas y poder leer artículos de información relacionados con diferentes dispositivos electrónicos. Asimismo, podrá modificar cualquiera de los datos personales que utilizó a la hora del registro y de esta manera tener actualizado su usuario en todo momento.

A esta plataforma web podrá accederse desde cualquier dispositivo, ya que se adapta a todo tipo de pantalla. Así, se podrá dar uso de ella, desde un móvil, tablet e incluso una televisión.

De esta manera, pensamos que la aplicación es recomendable para todos los públicos ya que si alguno de los usuarios en el momento de la incidencia no dispone de conexión a internet, puede solucionarlo a través de una sencilla y rápida llamada telefónica.

Durante el transcurso de los 3 cursos académicos de la Ingeniería Técnica, al cursar asignaturas que se basaban en aprender y poner en práctica algunos conceptos importantes de un sitio web, como son la usabilidad, accesibilidad y funcionalidad, ha sido de gran ayuda para elegir este proyecto y para llevarlo a cabo de la mejor manera posible.

1.2 Objetivos

El objetivo fundamental de este Proyecto Fin de Carrera es el de realizar una herramienta que sirva como asistente para el usuario con la finalidad de que pueda servirle para poder solucionar pequeñas y cotidianas incidencias informáticas que pudieran ocurrirle en los distintos dispositivos electrónicos de los que dispusiera y también hacerlo de manera visual mediante la plataforma web, pudiendo el usuario interactuar con la aplicación por diferentes medios, ya sea modo convencional (ratón y teclado) o por voz.

Esta herramienta tiene que ser fácil de utilizar para el usuario y que no conlleve una gran dificultad para poder hacer uso de ello, también debe entenderse su funcionamiento de manera clara y concisa.

En base a este objetivo principal, se definieron los siguientes objetivos parciales:

- Interactuar con el usuario, a través de un asistente que ofrezca posibles soluciones para la incidencia correspondiente.
- Obtener información del usuario para una atención más personalizada, y que a su vez pueda trasladarse a la plataforma web.
- Obtener un sitio web accesible y navegable a través de cualquier dispositivo móvil.
- Ofrecer atención las 24 horas del día los 365 días del año.
- Minimizar las consecuencias que pudieran sufrir los usuarios que se encuentren con incidencias informáticas, permitiéndoles solucionarlas y seguir utilizando su dispositivo con normalidad.

1.3 Fases del desarrollo

Para establecer las fases del desarrollo, este proyecto se ha debido en 3 fases principales a destacar:

Fase de planificación:

- **Estudio de los sistemas de diálogo:** aproximación a los sistemas de diálogo y estado actual de la investigación en este campo.

- **Planificación y análisis de requisitos del portal de voz:** determinación y selección de las funcionalidades que proporcionan los Sistemas de Diálogo para realizar la aplicación del portal de voz.
- **Estudio de la plataforma web:** Relacionar la plataforma web con el sistema de dialogo, y determinación de los diferentes puntos y niveles de la página web. También la elección del contenido que contendrá la misma.
- **Estudio de las tecnologías necesarias:** estudio de la plataforma Voxeo^[1], del estándar VoiceXML, del lenguaje de programación PHP, del gestor de base de datos, lenguaje HTML^[2], JavaScript, así como implementación de ejemplos prácticos para la asimilación de estas tecnologías.

Fase de Desarrollo:

- **Análisis y diseño inicial:** división de las distintas funcionalidades en los diferentes módulos y submódulos del sistema.
- **Implementación del sistema:** Desarrollo del código, de todos los módulos y submódulos, así como de las páginas y enlaces necesarios para facilitar la interacción entre ellos. Destacado la implementación de la plataforma web a cualquier dispositivo móvil.
- **Pruebas unitarias:** estudio, configuración y realización de pruebas funcionales para cada módulo.
- **Pruebas de integración y de sistema:** proceso de pruebas del sistema completo hasta alcanzar una versión completamente estable. Prueba de la conexión entre los datos recogidos por sistema oral y mostrados correctamente en la plataforma web.
- **Evaluación de la aplicación:** estudio de los sistemas de evaluación para interfaces orales, realización de las preguntas de la encuesta de evaluación, implementación de la página HTML con el cuestionario, recogida y análisis de resultados. Estudio de la respuesta de la plataforma web, y la rapidez de carga de los distintos contenidos.

Fase de documentación:

- **Memoria del Proyecto Final de Carrera:** Redacción del presente documento de memoria del Proyecto de Final de Carrera.
- **Preparación de la presentación.**

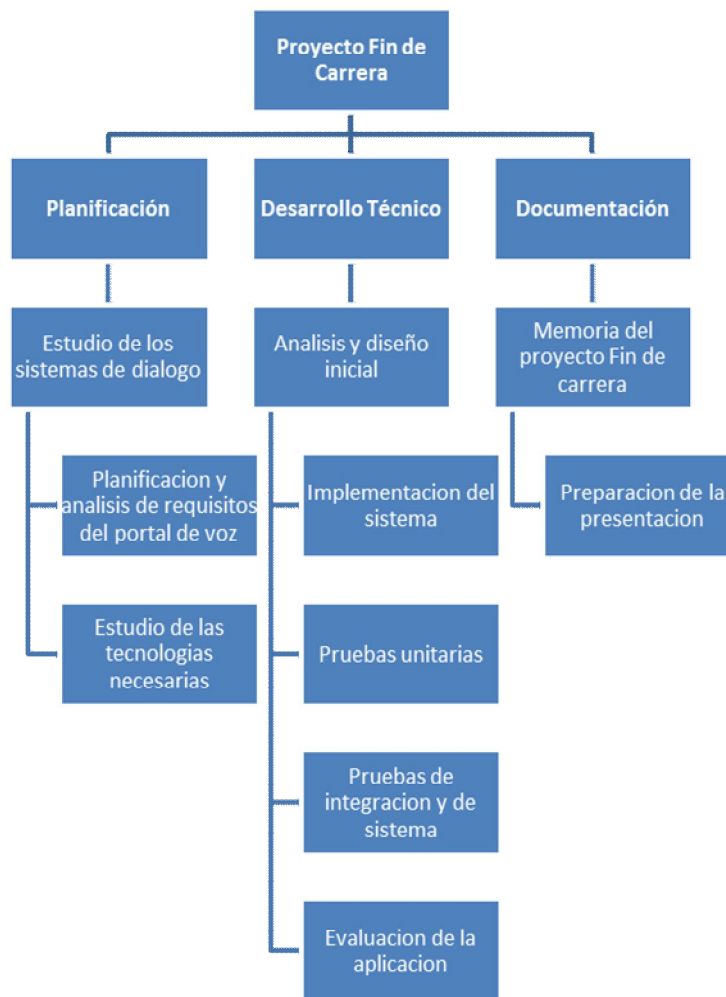


Figura 1. Diagrama WBS representando las tareas definidas para el proyecto

1.4 Medios y documentación utilizados

Los medios empleados para la realización del Proyecto Final de Carrera han sido los siguientes:

- Dispositivos Hardware:
 - Ordenador portátil.
 - Auriculares con micrófono.
 - Periféricos habituales (teclado, ratón).
 - Servidor web x10Hosting.

- Aplicaciones Software:
 - Plataforma Voxeo.
 - Skype.
 - Paquete Microsoft Office 2010.
 - Editor de programación Sublime Text, Notepad++.

En cuanto a la documentación utilizada han sido numerosos libros y distintos medios digitales, los cuales están recogidos en la correspondiente bibliografía.

2 Estado del arte

En este apartado se describirá un estado del arte relativo al campo de investigación relacionado con la aplicación desarrollada. En primer lugar se mostrará la evolución de los sistemas de diálogo, estándares relacionados, así como una serie de aplicaciones ya existentes desarrolladas con el estándar VoiceXML, y por último el funcionamiento y las múltiples opciones ofrecidas en la plataforma web.

La finalidad de estos sistemas es emular el comportamiento inteligente de un ser humano en un diálogo con otra persona, con el objetivo de que se realice una cierta tarea.

Un sistema de diálogo ideal reconocería el habla espontánea, comprendería enunciados sin restricciones de contenido, proporcionaría respuestas con sentido, gramaticalmente bien formadas y pragmáticamente adecuadas, respondería con voz completamente natural y sería multimodal.

2.1 Los sistemas de dialogo

2.1.1 Introducción a los sistemas de dialogo

Los sistemas de diálogo hablado (spoken dialogue systems) son sistemas informáticos que reciben como entrada frases del lenguaje natural expresadas de forma oral y generan como salida frases del lenguaje natural expresadas asimismo de forma oral. La finalidad de estos sistemas es emular el comportamiento inteligente de un ser humano que realiza una tarea concreta, utilizándose en la actualidad por distintas empresas para proporcionar información de forma automática, por ejemplo, petición de citas médicas, horarios de salida y llegada de una estación de tren o aeropuerto, movimiento bancarios, etc.

Un sistema de diálogo ideal reconocería el habla espontánea, comprendería enunciados sin restricciones de contenido, proporcionaría respuestas con sentido, gramaticalmente bien formadas y pragmáticamente adecuadas, respondería con voz completamente natural y sería multimodal.

Actualmente, los sistemas de diálogo están limitados debido a encontrarse sujetos a las restricciones del reconocimiento automático del habla, la comprensión y limitados a dominios específicos,

necesitan estrategias de verificación, y sin duda el problema del diálogo que surge de forma espontánea.

La siguiente figura 2 muestra el esquema general de un sistema de diálogo hablado. Los bloques amarillos se implementan mediante cinco tecnologías fundamentales en el diseño de los sistemas actuales: reconocimiento automático de habla, procesamiento del lenguaje natural, gestión del diálogo, generación de lenguaje natural y síntesis de habla. Los bloques azules representan componentes adicionales y el bloque verde representa la base de datos en la que el sistema obtiene la información que el usuario necesita conocer (horarios de salida de aviones, partes meteorológicos, etc.)^[3].

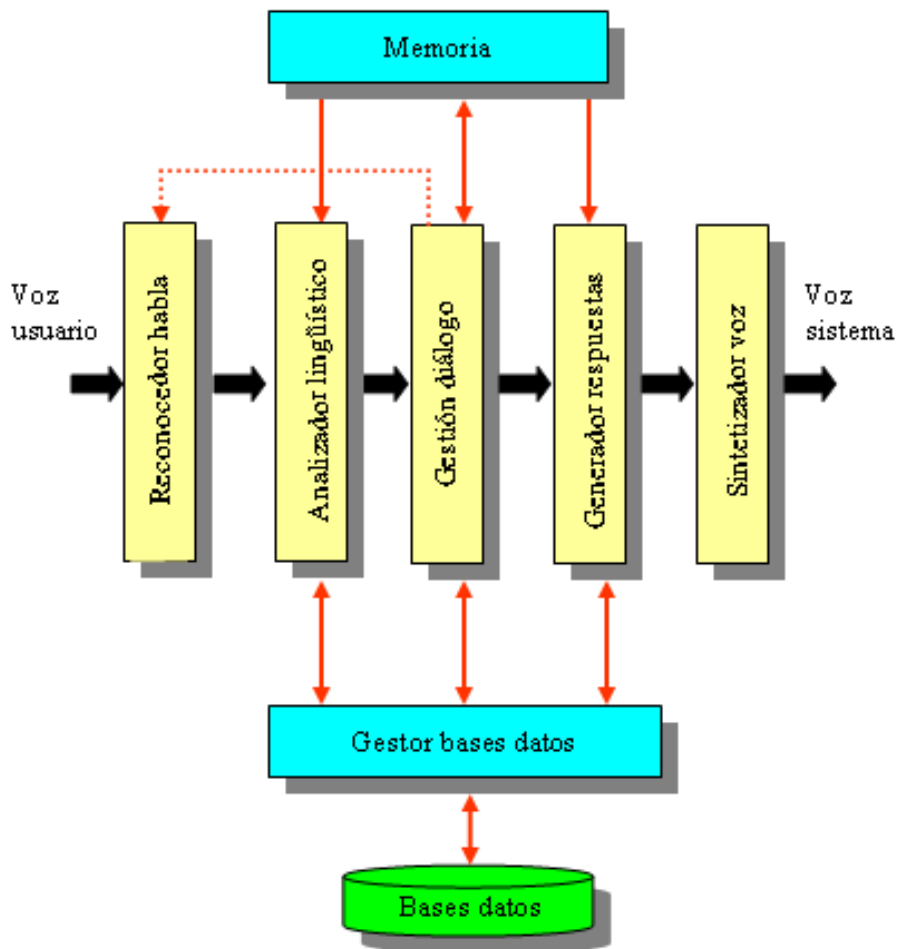


Figura 2. Estructura de un sistema de diálogo

La descripción de los módulos y sus principales características se explican a continuación:

- **Reconocimiento Automático del Habla (RAH):** Procesa la voz de usuario y la transforma en la secuencia de palabras reconocidas con mayor probabilidad.
- **Procesamiento del Lenguaje Natural (PLN):** Extrae el significado de las palabras reconocidas en el módulo anterior, obteniendo la representación semántica (significado) de la frase y expresándolo en términos de un lenguaje especificado para la tarea.
- **Gestor del diálogo (GD):** Decide qué paso debe dar el sistema tras cada intervención del usuario. Puede decirse que este es el módulo fundamental del sistema, pues su finalidad es lograr que la interacción con el usuario sea lo más cómoda e “inteligente” posible. Para ello, se basa en la interpretación semántica de la petición del usuario, la historia del proceso de diálogo, la información disponible en ese punto, el estado actual del sistema, la información obtenida de la base de datos, la estrategia definida, etc.
- **Generación del Lenguaje Natural (GLN):** Tiene como función la generación de una frase, gramaticalmente correcta y en un lenguaje lo más cercano posible al lenguaje natural, que transmita el mensaje generado por el gestor de diálogo.
- **Conversión Texto Habla (CTH):** Transforma la frase de respuesta en señal de audio.
- **Módulo de memoria:** Almacena las representaciones semánticas obtenidas a lo largo de la interacción así como las frases previamente generadas por el sistema, proporcionando esta información histórica a los módulos de procesamiento del lenguaje natural, gestor del diálogo y generación del lenguaje natural. De esta forma, el sistema puede resolver las referencias anafóricas existentes en las frases pronunciadas por los usuarios, puede conocer qué frases ha expresado el usuario previamente, y puede utilizar información contextual (mediante el uso de anáforas y elipsis) durante la generación de las frases.

- **Módulo gestor de consulta de información de la aplicación:** Se encarga de generar las consultas necesarias al repositorio de información, procesarlas y proporcionar la información obtenida al módulo gestor del diálogo.
- **Repositorio de información:** Se trata de bases de datos o páginas web donde se almacena la información necesaria para el sistema de diálogo y que será consultada por el gestor de consulta de información.

El abanico de posibilidades en cuanto a entorno y tareas en los que se pueden aplicar estos sistemas es muy amplio, ya que permiten el acceso a servicios y control de máquinas por vía telefónica. La posibilidad de convertir cualquier línea telefónica en un punto de acceso a los servicios les proporciona un gran valor económico y social.

Cabe destacar el complemento que hoy en día supone los portales de voz al acceso WAP, estando cada día más unidos Internet con la telefonía. De esta manera se facilita y posibilita aún más la comunicación en ciertos casos en los que las interfaces tradicionales como el teclado, o pantalla no serían posibles.

Muchos ejemplos los encontramos en personas con discapacidades visuales o motoras graves, aplicaciones accesibles desde distintos vehículos, en los cuales el conductor no puede distraer la atención y desviar la vista, utilización de teleservicios como la información del tiempo o información sobre accidentes de tráfico, etc.).

2.1.2 Clasificación de los sistemas de diálogo

Podemos clasificar los sistemas de diálogo de la siguiente manera:

- **Iniciativa del diálogo**

Sistemas de diálogo guiados por el sistema: La interacción se realiza mediante turnos concretos y fijados entre pregunta y respuesta. Las iniciativas del usuario se ven reducidas.

Sistemas de iniciativa mixta: 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.

Sistemas de diálogo guiados por el usuario: Es el usuario el que lleva la iniciativa de diálogo en cualquier instante de la interacción, permitiéndole al usuario que utilice lenguaje natural y no delimitando los mensajes del sistema (por ejemplo, ¿Puede indicarme cuál es su incidencia?).

- **Dependencia del hablante**

Dependiente del hablante: Un sistema dependiente del hablante está desarrollado para funcionar con un sólo hablante. Estos sistemas, normalmente, son más fáciles de implementar, más baratos y más precisos, pero se muestran menos flexibles que los sistemas adaptables al hablante o los sistemas independientes del hablante.

Independiente del hablante: Un sistema independiente del hablante está desarrollado para funcionar para cualquier hablante de un determinado tipo (por ejemplo, inglés americano). Estos sistemas son los más complicados de desarrollar, los más caros y la precisión es menor que la de los sistemas dependientes del hablante. Sin embargo, son más flexibles.

- **Tipo de comunicación**

Unimodal (oral): Sólo se utiliza el habla.

Multimodal: Se utilizan varios canales de comunicación. Por ejemplo, dispositivos de entrada como el habla, el teclado, el ratón, el micrófono, la cámara, una pantalla o un PDA y canales de salida para proporcionar información como voz, texto, gráficos o imágenes.

- **Tipo de idioma**

Monolingüe: Permiten interacción sólo mediante un idioma.

Multilingüe: Permiten interacción mediante varios idiomas. Por ejemplo, castellano e inglés.

- **Tipo de discurso**

Reconocimiento de palabras aisladas: En el reconocimiento de palabras aisladas el patrón acústico de cada palabra del vocabulario se almacena como una secuencia temporal de características derivadas,

usando LPC, análisis de banco de filtros o alguna otra técnica de análisis del lenguaje. El reconocimiento se lleva a cabo comparando el patrón acústico de la palabra a reconocer con los patrones almacenados, seleccionando la palabra que mejor encaje con la palabra desconocida.

Reconocimiento de palabras conectadas: En el reconocimiento de palabras conectadas, 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. Ejemplos de esto son las cadenas de dígitos conectados donde el vocabulario es un conjunto de 10 dígitos, o el reconocimiento de letras conectadas, donde el vocabulario es el conjunto formado por el abecedario. No hay que confundir esto con el reconocimiento continuo de voz, donde el reconocimiento se basa en unidades lingüísticas denominadas fonemas, sílabas, etc..., lo que supone separar la voz en estas unidades y etiquetarlas subsecuentemente.

Reconocimiento continuo: Un sistema de reconocimiento continuo funciona sobre un lenguaje en el que 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. Primero, es difícil encontrar el comienzo y el final de las palabras. Otro problema es la coarticulación. La generación de cada fonema se ve afectada por la generación de los fonemas adyacentes, y de modo parecido, el comienzo y final de las palabras se ven afectados por las palabras que les preceden y suceden. El reconocimiento del lenguaje continuo también se ve condicionado por la frecuencia de habla (un discurso rápido suele ser más difícil).

Reconocimiento discreto: Un sistema de reconocimiento discreto funciona con palabras simples, necesitando de una pausa entre la dicción de cada palabra. Esta es la forma más sencilla de reconocimiento para llevar a cabo ya que los puntos de finalización son más fáciles de encontrar y la pronunciación de una palabra no afecta a las demás. De este modo, y ya que las ocurrencias de las palabras son más consistentes, es más fácil reconocerlas.

- **Adaptación**

Sistemas adaptativos: El sistema es capaz de aprender nuevas estrategias comunicativas en función del comportamiento del usuario.

Una línea de investigación actual consiste en la incorporación de las emociones a los sistemas de diálogo. Para ello, el análisis del comportamiento del usuario no debe incluir únicamente las emociones básicas sino también todos sus estados emocionales y cambios en el comportamiento. Por ejemplo, si se detecta que el usuario puede estar enfadado, se inician estrategias de recuperación o se transfiere la llamada a un operador humano, o si se detecta algo similar a la alegría, se aprovecha para ofrecer un nuevo producto.

2.1.3 Ejemplos de aplicaciones

La interacción oral puede ser la única manera de acceder a la información en muchos casos, como cuando la pantalla disponible es demasiado pequeña para mostrar la información (dispositivos de bolsillo) o cuando los ojos del usuario están ocupados en otras tareas (conducción), por ello en este apartado se van a destacar algunas de las muchas que hoy en día existen^{[4][5]}.

Centralita telefónica

Un ejemplo sencillo de aplicación de voz es una centralita telefónica cuyo único cometido es transferir automáticamente la llamada a otro número distinto.

. Atos^[6] - Centralita telefónica. España

Consulta de información

Los sistemas de diálogo que se utilizan para obtener información son aplicaciones en las que los datos que se consultan se encuentran en un repositorio de información, como puede ser una base de datos o página web a los que accederá el sistema de diálogo.

. Larri^[7] - Sistema de diálogo multimodal que asiste y guía a personal de aviones de combate durante tareas de mantenimiento. EE.UU

. Dihana^[8] - Información de viajes en tren. España.

. Let's Go^[9] - Información acerca de los autobuses, orientado principalmente a extranjeros y ancianos. EE.UU.

- . Conquest^[10] - Información de horarios en congresos. EE.UU.

Reservas

Existen otros sistemas que además de proporcionar información realizan tareas algo más complejas, como son reservas a través de la aplicación de voz:

- . DARPA Communicator^[11] - Interfaz conversacional inteligente para la reserva de billetes de avión. EE.UU.
- . Mercury^[12] - Interfaz conversacional que proporciona información sobre vuelos además de permitir hacer reservas. EE.UU.

Transacciones

También los sistemas conversacionales pueden llevar a cabo transacciones. Son las aplicaciones relacionadas con el comercio electrónico, como la venta de entradas o de billetes de tren y de avión, y los relacionados con la banca electrónica:

- . Cineentradas^[13] - Venta telefónica de entradas de cine. España.
- . CTT-Bank^[14] - Sistema que permite solicitar transferencias o domiciliaciones, vender o comprar valores. Se comprueba la identidad del cliente mediante una clave personal de acceso. España.

Control remoto

La interacción oral también es útil para el control de dispositivos, especialmente en entornos inteligentes. Un ejemplo de este tipo de sistemas es:

- . MIMUS^[15] - Sistema de diálogo multimodal para controlar un hogar inteligente. El sistema MIMUS incluye un componente específico para la gestión de conocimiento que integra el núcleo de gestión de diálogo con las ontologías del dominio. España.

Compañeros virtuales

Los agentes virtuales y compañeros constituyen la aplicación más compleja de los sistemas de diálogo:

. Ikea^[16] - Pregunta a Anna. Compañera relacionada con la tienda IKEA. España.

2.2 Voice Extensible Markup Language (VoiceXML)

2.2.1 Introducción al estándar VoiceXML

El lenguaje VoiceXML (Voice Extensible Markup Language)^[17] tuvo sus orígenes en 1995. Fue diseñado como un lenguaje de diálogo basado en XML (Extensible Markup Language)^[18], que pretendía simplificar el proceso de desarrollo de aplicaciones de reconocimiento de voz dentro de un proyecto de AT&T llamado PML (Phone Markup Language).

En 1998, el W3C (Consortio World Wide Web) organizó una conferencia sobre navegadores por voz. En este momento, AT&T y Lucent tenían diferentes variantes de su PML original, mientras que Motorola e IBM habían desarrollado sus propios lenguajes VoiceXML y SpeechML respectivamente. Muchos otros asistentes a la conferencia también estaban diseñando lenguajes parecidos, por ejemplo, TalkML de HP y VoiceHTML de PipeBeach.

Debido a esta situación, AT&T, IBM, Lucent y Motorola decidieron formar el Foro de VoiceXML para aunar sus esfuerzos. La misión del Foro VoiceXML era definir un lenguaje estándar que los desarrolladores pudieran utilizar para crear aplicaciones de voz. Eligieron XML como base para ello.

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

Finalmente, VoiceXML 2.0 es el resultado de ese trabajo basado en la colaboración conjunta de las empresas miembros de W3C [W3C], otros grupos de trabajo del W3C y el público. La versión 2.1 aparece como recomendación del W3C en junio de 2007. El working draft de VoiceXML 3.0 fue aprobado en diciembre de 2010.

VoiceXML es un lenguaje diseñado para crear sistemas de diálogo mediante voz que incluye reconocimiento de entradas de voz y de entradas DTMF, funciones de telefonía, salida de voz sintetizada y ficheros de audio, grabación de diálogos y control del flujo del diálogo.

Su objetivo es aprovechar las ventajas de las tecnologías Web para el desarrollo de aplicaciones controladas mediante la voz.

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

Figura 3. Visión general del código de una aplicación VoiceXML.

2.2.2 Control y gestión del flujo del diálogo

Según la RAE, el dialogo es la “plática o conversación entre dos o más personas, que alternativamente manifiestan sus ideas o afectos”. En esta sección se trata de estilizar nuestro código de los documentos VoiceXML, relacionar nuestro documento con un tipo de dialogo que se pueda dar entre el ordenador y el usuario, puede haber varios sub-diálogos dentro de un diálogo, esta construcción del diálogo ayudará y guiará la interacción entre ordenador y usuario.

2.2.2.1 Etiquetas básicas sobre diálogos en VoiceXML

VoiceXML ofrece 2 tipos de diálogos “form” y “menú”, mientras que el tipo de diálogo menú es sólo para tipos especiales de aplicaciones VoiceXML, es mucho más corriente el uso del dialogo “form” para aplicaciones más sencillas. Este estudio se basa en el tipo de diálogo “form”, se implementa con la etiqueta “<form>” que tendrá en su interior etiquetas como “<field>”, “<record>” y “<subdialog>” que a continuación se puede ver con más detalle^[19].

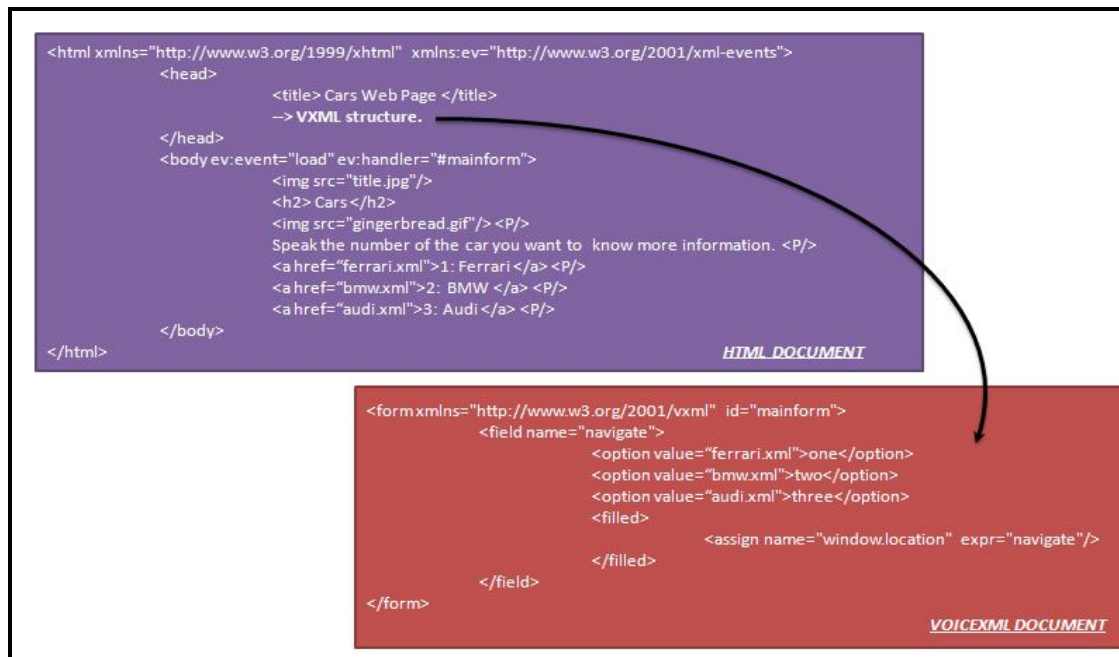


Figura 4. Visión general del código de una aplicación VoiceXML.

<FORM>

El uso de esta etiqueta está reservado para la construcción de diálogos sencillos, que se pueden incluir en páginas web sencillas, este elemento es el elemento raíz del que luego dependerá otros para el control y seguimiento del diálogo.

Los atributos de los que dispone son los siguientes:

- **Id:** que indicará la cadena de caracteres con la que se podrá identificar a este diálogo "form".
- **Scope:** indicará el ámbito del diálogo, puede tomar los valores dialogo (lo más común) o documento que indicará que es un formulario de gramáticas.

```
<?xml version="1.0" ?>
<vxml version="2.0" xmlns="http://www.w3.org/2001/vxml" xml:lang="es-ES">
  <property name="inputmodes" value="dtmf" />
  <form>
    <field name="num" type="digits?minlength=3;maxlength=8">
      <prompt>Introduzca su codigo PIN.</prompt>
      <filled>
        <if cond="num.length != 4">
          <prompt>No se han introducido exactamente 4 digitos.</prompt>
          <clear/>
          <reprompt/>
        <else/>
          <prompt>Ha introducido el codigo <value expr="num"/></prompt>
        </if>
      </filled>
    </field>
  </form>
</vxml>
```

Tabla 1 – Ejemplo <form>

<FIELD>

Trabaja como subetiqueta del diálogo “form”, este elemento declara una entrada en el formulario, con la entrada que da el usuario, el sistema intenta ver si corresponde con alguna de las opciones de la gramática que se ha definido previamente.

El valor de la entrada se guarda en una variable que puede ser usada dentro del documento. Algunos de sus atributos más importantes son:

- **Name:** atributo obligatorio usado para definir el nombre de la variable. Éste atributo debe ser único y no puede ser repetido en el mismo “<form>”.
- **Type:** atributo obligatorio, usado para definir el tipo de datos que contendrá este “<field>”. Algunos de los valores que puede tomar son: “boolean” (true o false), “date” (tiempo), “currency” (moneda), “phone” (teléfono)...

- **Expr:** atributo opcional, usado para definir el valor inicial de la variable. A continuación se pueden ver unos ejemplos de esta etiqueta, en este primer ejemplo se puede apreciar cómo se le pide al usuario que diga un número y posteriormente la aplicación se lo reproduce.

```
<?xml version="1.0" ?>
<vxml version="2.0" xmlns="http://www.w3.org/2001/vxml" xml:lang="es-ES">
  <property name="inputmodes" value="dtmf" />
  <form>
    <field name="num" type="digits?minlength=3;maxlength=8">
      <prompt>Introduce código PIN.</prompt>
      <filled>
        <if cond="num.length != 4">
          <prompt>No ha introducido 4 dígitos.</prompt>
          <clear/>
          <reprompt/>
        </if>
        <prompt>Introduzca el código <value expr="num"/></prompt>
      </filled>
    </field>
  </form>
</vxml>
```

Tabla 2 – Ejemplo <field>

<FILLED>

La etiqueta "<filled>" es la encargada de permitir al desarrollador especificar las acciones a tomar cuando se produce una coincidencia entre el "input" recibido por la gramática y lo que hay dentro de la misma.

Es decir, depende donde se emplace esta etiqueta se puede tener un flujo del diálogo u otro, lo que mejor le convenga para el desarrollador.


```
<?xml version="1.0" ?>
<vxml version="2.0" xmlns="http://www.w3.org/2001/vxml" xml:lang="es-ES">
  <property name="inputmodes" value="dtmf" />
  <form>
    <field name="num" type="digits?minlength=3;maxlength=8">
      <prompt>Introduce codigo PIN.</prompt>
      <filled>
        <if cond="num.length != 4">
          <prompt>No ha introducido 4 digitos.</prompt>
          <clear/>
          <reprompt/>
        <else/>
          <prompt>Introduzca el codigo <value expr="num"/></prompt>
        </if>
      </filled>
    </field>
  </form>
</vxml>
```

Tabla 3 – Ejemplo <filed>

<RECORD>

Esta etiqueta se usa para recoger datos sobre del usuario mediante un archivo de audio, y todo lo grabado es graba en una variable. Es posible utilizar la grabación en posteriores secciones del documento VoiceXML ya que se graba en el sistema. Algunos de los atributos de esta etiqueta son:

- **Name:** atributo requerido usado para definir el nombre de la variable. El Name debe ser único y no puede ser repetido en el mismo "<form>".
- **Expr:** atributo opcional, usado para definir el valor inicial de la variable.
- **Beep:** atributo opcional, si el valor del atributo es true, un tono se inserta antes de que empiece la grabación, el valor por defecto es false.
- **Maxtime:** es opcional, define el tiempo máximo de grabación.

A continuación se puede apreciar un ejemplo, en el cual se utiliza la etiqueta "<record>".

```
<?xml version="1.0" ?>
<vxml version="2.0" xmlns="http://www.w3.org/2001/vxml" xml:lang="es-ES">
  <form>
    <record name="mensaje" beep="true" maxtime="10s" dest="nombre_archivo" finalsilence="4s"
tmfterm="true">
      <prompt>
        Registro tras el pitido.
      </prompt>
    </record>
  </form>
</vxml>
```

Tabla 4 – Ejemplo <record>

<SUBDIALOG>

Esta etiqueta se usa para invocar a un diálogo que previamente ya se ha escrito en el mismo documento o que está en otro fichero distinto, cuando se llama a esta etiqueta se ejecutara un nuevo proceso que contiene la información del diálogo. Algunos de sus atributos más importantes son:

- **Name:** es imprescindible, sirve para definir el nombre de la variable que guarda el resultado del diálogo raíz.
- **Namelist:** atributo obligatorio, mantiene una lista de variables de la etiqueta.
- **Src:** atributo obligatorio, indica la URI del subdiálogo, puede ser dentro del documento o fuera del mismo.
- **Method:** atributo obligatorio, indica el método que se usa para la respuesta del subdiálogo, puede tomar los valores [get](#) y [post](#).

A continuación se puede ver más claro, un ejemplo sobre esta etiqueta.

```

<?xml version="1.0" ?>
<vxml version="2.0" xmlns="http://www.w3.org/2001/vxml" xml:lang="es-ES" >
  <form id="main">
    <block>Vamos a llamar a un sub-dialogo</block>
    <subdialog name="result" src="#subbie">
      <param name="hello" value="goodbye"/>
      <param name="goodbye" value="goodbye"/>
    </subdialog>
    <block>
      Volvemos del sub-dialogo.
      Resultado hello tiene valor <value expr="result.hello"/>,
      Resultado goodbye tiene valor <value expr="result.goodbye"/>
    </block>
  </form>
  <form id="subbie"> <!-- Este es el sub-dialogo -->
    <!-- las variables se envían como parámetros -->
    <var name="hello" expr="hello"/> <!-- la inicialización no se toma en cuenta -->
    <var name="goodbye"/>
    <block>
      Esto es el sub-dialogo.
      hello tiene valor <value expr="hello"/>
      goodbye tiene valor <value expr="goodbye"/>,
      <return namelist="hello goodbye" />
    </block>
  </form>
</vxml>

```

Tabla 5 – Ejemplo <subdialog>

2.2.2.2 Etiquetas de control de diálogos

Existen 2 elementos básicos para el control del diálogo dentro de la etiqueta `form`: “<block>” y “<initial>”, se puede ver estas dos etiquetas con más profundidad.

<BLOCK>

Se utiliza para construir un bloque de contenidos ejecutables que contienen el texto, éste será reproducido íntegramente al usuario por parte de la aplicación. Dispone de atributos, el más usual es “name” que se utiliza para asignar un nombre de variable a ese bloque para que pueda ser utilizado en otros lugares del documento. También dispone del elemento “expr” el cual indica el valor por defecto

que tendrá ese bloque. En el siguiente ejemplo se puede ver que se crea un bloque el cual tiene una variable Equipo_Futbol y con un valor por defecto "Real Madrid", en la siguiente frase se utiliza la etiqueta `value` con un el valor referente a la variable "Equipo_Futbol". El Resultado que se le ofrecería al usuario sería: "El mejor equipo de fútbol es el Real Madrid".

```
<form id="bienvenida">
  <block>
    <prompt >Hola. Bienvenido a facilisimo. ¿Vuelves a tener problemas con tu
dispositivo?</prompt>
  </block>
</form>
```

Tabla 6 – Ejemplo <block>

<INITIAL>

Esta etiqueta se utiliza para los diálogos de iniciativa mixta (donde las expresiones del primer usuario determinan el flujo de la aplicación), permitiendo al usuario completar toda la información del formulario y según la información dar una respuesta distinta al usuario.

Algunos de sus atributos más importantes son:

- **Name:** es imprescindible, sirve para definir el nombre de la variable que guarda la información que contiene la etiqueta "<initial>".
- **Cond:** atributo opcional, se usa para saber si este elemento ha sido utilizado por el documento, tiene como valores true o false.

En el siguiente ejemplo se puede ver como se utilizan 2 tipos distintos de obtener información uno de la manera fácil con el uso de 2 "<field>" uno para la ciudad de salida y otro para la ciudad de destino y la otra manera es el uso de la etiqueta "<initial>" la cual pide al usuario a la vez tanto la ciudad de origen como la ciudad de destino. Para la búsqueda de gasolineras en su trayecto.

```
<form id="ObtenerInformacion">
  <grammar src=http://gramaticaejemplo.com/gramatica1.vxml type="application/grammar+xml"/>
</block>
  Bienvenido al Localizador de Gasolineras PETROLSEARCH
</block>
  <initialName="Inicial">
    <prompt>Por favor indique la ciudad de salida y de llegada de su viaje para poder buscar todas las
gasolineras que habrá en su trayecto.</prompt>
    <nomatchcount='1'> Por ejemplo "De Madrid a Valencia" </nomatch>
  </initial>

  <field name="Ciudad_Salida">
<grammar src="gramatica2.vxml" type="application/grammar+xml"/>
<prompt>Por favor indique la ciudad de salida.</prompt>
  </field>

  <field name="Ciudad_Llegada">
<grammar src="gramatica3.vxml" type="application/grammar+xml"/>
<prompt>Por favor indique la ciudad de llegada.</prompt>
  </field>
</form>
```

Tabla 7 – Ejemplo <initial>

2.3 Plataforma Voxeo

2.3.1 ¿Qué es Voxeo?

Voxeo Corporation (www.voxeo.com) es una compañía que ha desarrollado la plataforma VoiceCenter, la cual permite crear, implementar y probar aplicaciones de voz desarrolladas en VoiceXML de forma telefónica y completamente gratuita.

Además, Voxeo incluye una web estática local para el alojamiento de las aplicaciones de voz, un número de teléfono dedicado para su uso durante la prueba de las aplicaciones, registro y depuración en tiempo real de la aplicación, conectividad de voz sobre IP, foros de soporte donde se puede interactuar con la Corporación Voxeo y con otros miembros de la comunidad y soporte extremo 24 horas los 7 días de la semana.

2.3.2 ¿Cómo funciona Voxeo?

Para explicar el funcionamiento de la plataforma de voz Voxeo se va a comparar el modo de operación de un servidor Web con el de la red Voxeo.

En la Web, los exploradores realizan solicitudes a los servidores Web por medio de protocolos de Internet como HTTP^[20]. Los servidores Web alojan páginas estáticas y aplicaciones dinámicas HTML que se devuelven al explorador. El resultado es una página visual que se ve en un navegador Web. Este mecanismo lo podemos observar en la Figura 4.

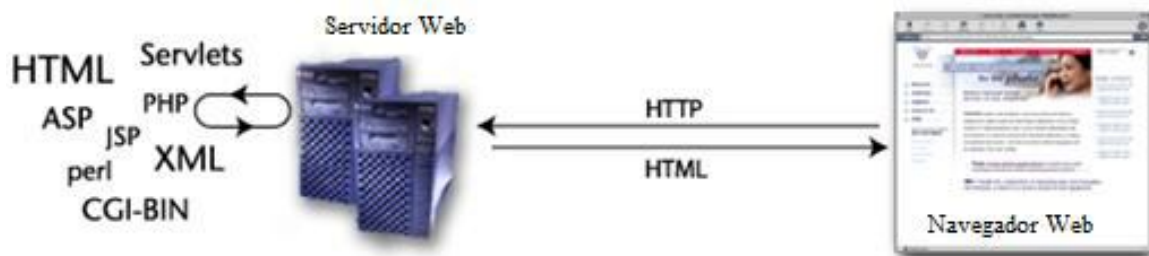


Figura 5. Interpretación de una página web

En la red de Voxeo, las aplicaciones de voz siguen el mismo modelo de navegador que las páginas Web estándares. La diferencia es que, en vez de HTML, se utilizan lenguajes de marcado para voz (VoiceXML, CCXML^[21] o CalXML^[22]) para escribir la página. Cuando estas páginas son procesadas por la red de Voxeo, lo que se devuelve es un audio. La Figura 5 muestra este procedimiento.



Figura 6. Interpretación de una página VoiceXML

2.3.3 Desarrollo de una aplicación VoiceXML con la plataforma Voxeo

En este apartado se explica cómo crear una cuenta de Voxeo, qué herramientas ofrece esta plataforma y cómo configurar de forma detallada una aplicación VoiceXML en ella.

2.3.3.1 Primer contacto con Voxeo

La cuenta de Voxeo incluye todo lo necesario para desarrollar aplicaciones de voz. Para crear una cuenta hay que registrarse en www.evolution.voxeo.com según muestra la Figura 7.

ACCOUNT DOCUMENTATION TOOLS & DOWNLOADS PAID SERVICES EXTREME SUPPORT

u: p:

Account > Registration

REGISTRATION: TERMS OF SERVICE

Membership is free, however you must agree to the following terms of service:

CONTAINED ON THE SITE, ANY QUESTIONS REGARDING THE INFORMATION SHOULD BE DIRECTED TO THE PROVIDERS OF SUCH INFORMATION.

IN NO EVENT WILL VOXEO BE LIABLE TO ANY PARTY FOR ANY DIRECT, INDIRECT, SPECIAL, PUNITIVE OR OTHER CONSEQUENTIAL DAMAGES FOR ANY USE OF THIS WEB SITE, OR ON ANY OTHER HYPERLINKED WEB SITE, INCLUDING, WITHOUT LIMITATION, ANY LOST PROFITS, BUSINESS INTERRUPTION, LOSS OF PROGRAMS, OTHER DATA ON YOUR INFORMATION HANDLING SYSTEM, OR OTHER ECONOMIC ADVANTAGE, EVEN IF WE ARE EXPRESSLY ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. YOU HAVE SOLE RESPONSIBILITY FOR ADEQUATE PROTECTION AND BACKUP OF DATA AND/OR EQUIPMENT USED IN CONNECTION WITH THE SITE AND WILL NOT MAKE A CLAIM AGAINST VOXEO FOR LOST DATA, RUN TIME, INACCURATE OUTPUT, WORK DELAYS OR LOST PROFITS RESULTING FROM THE USE OF THE SITE OR INFORMATION.

YOU AGREE TO INDEMNIFY AND HOLD VOXEO HARMLESS FROM, AND YOU COVENANT NOT TO SUE VOXEO FOR, ANY CLAIMS THAT MAY BE ASSERTED, ARISING OUT OF, RELATING TO, OR BASED ON USING THE SITE OR ANY OF THE SERVICES PROVIDED ON OR MADE ACCESSIBLE BY THE SITE.

Phone numbers provisioned to this account for development purposes may be changed or revoked at any time without warning.

You acknowledge and agree that the Site and information are subject to the U.S. Export Administration Laws and Regulations. Diversion of such information contrary to U.S. law is prohibited. You agree to comply strictly with all U.S. export laws and assume sole responsibility for obtaining licenses to export or re-export as may be required.

Confirmation of Terms and Conditions

I agree to Terms of Service shown above

I agree that my applications are not used in any scenario that is a matter of life and death, such as 911 or emergency healthcare

Account Introduction

As a member of our community, you'll have access to everything you need to create and test your own voice and messaging applications, including phone numbers on our free test network, file hosting, and 24x7 customer support.

And the best part... it's free!

So what are you waiting for, go ahead and create your new account.

Figura 7. Registro usuario nuevo en Voxeo

Una vez se haya realizado el registro, puede accederse a la plataforma introduciendo el username y la password anteriormente elegidos.

ACCOUNT DOCUMENTATION TOOLS & DOWNLOADS PAID SERVICES EXTREME SUPPORT

u: p:

Account > Login

username: password:

Membership is required to deploy and manage voice applications on our network. As a registered member, you'll get direct local phone access and toll-free access for testing and tuning your voice application. You will also have free access to our industry-leading Extreme Support team.

Figura 8. Acceso a la cuenta de Voxeo

La vista general una vez accedido se muestra en la Figura 9.

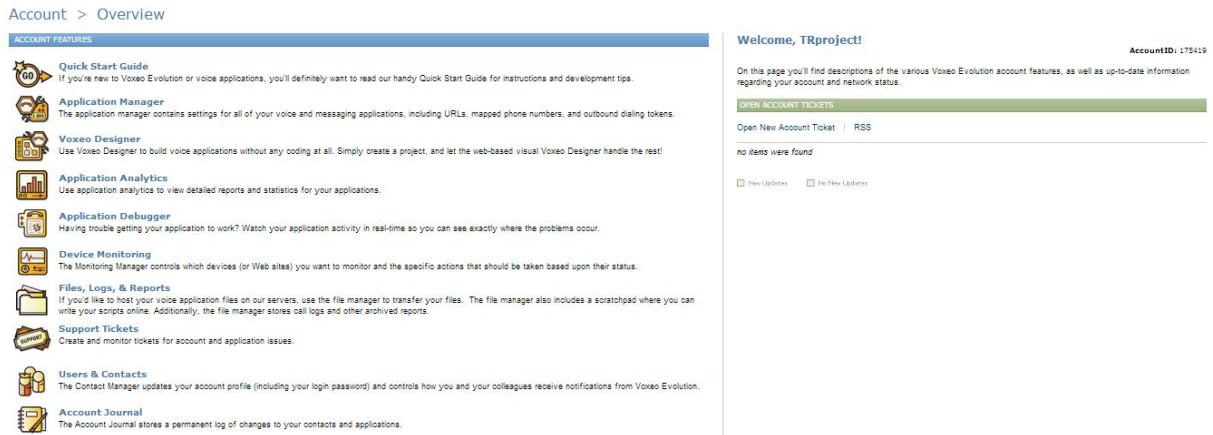


Figura 9. Vista general cuenta de Voxeo

A continuación se muestran las distintas opciones que ofrece Voxeo:

- **Quick Start Guide**

Guía rápida de iniciación de la plataforma Voxeo para aquellos usuarios nuevos en Voxeo Evolution o en las aplicaciones de voz.

En esta guía se detallan las herramientas de desarrollo disponibles y las instrucciones a seguir para desarrollar aplicaciones.

- **Application Manager**

La herramienta Application Manager contiene la configuración de las aplicaciones de voz y mensajería, incluyendo URLs, números de teléfono asignados a las aplicaciones, y tokens de marcado.

- **Voxeo Designer**

Se utiliza Voxeo Designer para crear aplicaciones de voz sin necesidad de codificación. Simplemente, se crea un proyecto de forma gráfica y Voxeo Designer lo gestiona.

- **Prophecy Log Search**

Herramienta para gestionar los registros de la llamada realizada. Además de facilitar la búsqueda de errores, se pueden ver las estadísticas de la llamada y generar una amplia variedad de informes de esta llamada.

- **Application Analytics**

Aplicación para ver informes detallados y estadísticas de las aplicaciones implementadas.

- **Application Debugger**

Si se producen problemas para conseguir que la aplicación funcione, con Application Debugger se puede ver la actividad de las aplicaciones en tiempo real para comprobar exactamente dónde ocurren estos errores.

- **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.

- **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. Este administrador también incluye un scratchpad donde se pueden escribir scripts y además, almacena los registros de llamadas y otros informes de archivado.

- **Support Tickets**

Con la herramienta Support Tickets se crean y controlan los tickets para cuestiones de cuenta y de aplicación.

- **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.

- **Account Journal**

Herramienta que almacena un registro permanente de los cambios de los contactos y aplicaciones de un usuario.

2.3.3.2 Selección de la plataforma de la aplicación de voz

Voxeo ofrece tres plataformas de aplicaciones de voz: CallXML, CCXML y VoiceXML. Las características de cada una de ellas se explican a continuación.

- **Prophecy - CallXML 3.0**

La plataforma CallXML está disponible exclusivamente en Voxeo. CallXML 3.0 agrega la capacidad de utilizar menús de reconocimiento de voz simples dentro de sus aplicaciones, y emplea las operaciones de lógica condicional integrada para extender aún más sus capacidades.

CallXML se adapta a las necesidades de la mayoría de las aplicaciones de telefonía que utilizan entrada DTMF. Si en la aplicación a desarrollar no es necesario el reconocimiento de voz complejo, CallXML puede ser la opción correcta como plataforma.

CallXML 3.0 está disponible para su uso alojada en la red Voxeo, o como instalación de prueba libre para un equipo local.

- **Prophecy - CCXML W3C 1.0**

Voxeo ofrece la primera plataforma compatible con CCXML 1.0. Habiendo distribuido más de 1.000 millones de llamadas desde su introducción comercial en el año 2002, es una buena elección para implementar aplicaciones de conferencia de nueva generación y de enrutamiento de llamadas y garantizar a su vez que permanezcan en la vanguardia de la tecnología.

CCXML 1.0 está disponible para su uso en el servidor de Voxeo o como una instalación de prueba gratis en un equipo local.

- **Prophecy - VoiceXML 2.1**

VoiceXML 2.1 está disponible como una solución alojada en el propio servidor de Voxeo o en un equipo local. Incluye el reconocedor de voz Voxeo y es el primero y único en el mundo compatible al 100% con la certificación VoiceXML 2.0.

2.3.3.3 Manejo de ficheros

VoiceXML 2.1 permite desarrollar aplicaciones que no es necesario modificar si se migran a otra plataforma compatible. Además, soporta todas las ampliaciones y mejoras de VoiceXML 2.1, así como los estándares de formato de gramáticas SISR/SRGS. También incluye soporte para los formatos de gramática antiguos GSL y JSGF, ofreciendo una enorme flexibilidad y variedad de opciones para las implementaciones IVR (Interactive Voice Response) activadas por voz.

Las figuras 10 y 11 muestran gráficamente los directorios donde se deben alojar los ficheros según se ha explicado anteriormente, y se puede observar cómo se ha cargado con éxito el fichero *holamundo.vxml*.

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		3/12/2013	
recordings		3/12/2013	
reports		3/12/2013	
www		3/12/2013	

ADD NEW FILE OR DIRECTORY

Upload File:
 Ningún archivo seleccionado

New Directory:

New File:

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

Introduction

Each Voxeo Evolution account automatically receives free hosting for static files.

This hosting space provides a convenient location to host and develop your static voice application files.

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

Editing & naming files: Click the pencil icon to edit files with the following extensions: .dot, .grammar, or .vxml (any extension ending in .xml)

File preview: Click on a file name to open it in your Web browser.

RECORD AUDIO

Call the following number to record audio to your [www/audio](#) directory: (678) 275-5252

Account ID: 175419
 PIN: 4200

Figura 10. Directorio /root

The screenshot shows the VoXeo Evolution File Manager interface. At the top, there is a navigation bar with links: ACCOUNT, DOCUMENTATION, TOOLS & DOWNLOADS, PAID SERVICES, and EXTREME SUPPORT. The user is logged in as TProject - 175419. The main content area is titled "Account > Files, Logs, Reports" and shows the "EXISTING FILES" section for the path "/root/www". A table lists a file named "holamundo.vxml" with a size of 125 bytes and a date modified of 3/14/2014 9:11 AM (EDT). Below the table, there are sections for "ADD NEW FILE OR DIRECTORY", "Upload File:" (with a "Seleccionar archivo" button and "Upload" button), "New Directory:" (with a "Create" button), and "New File:" (with a "Create" button). A note states: "New file names must have one of the following extensions: .txt, .grammar, or *.xml (any extension ending in xml)". On the right side, there is an "OPERATION STATUS" section indicating "Successfully renamed: index.vxml to holamundo.vxml" and an "Introduction" section with a folder icon and text explaining the hosting space and file naming conventions.

Figura 11. Directorio /root/www

2.3.3.4 Creación de una aplicación y vinculación a un número de teléfono

Para vincular la aplicación con un número de teléfono, se debe seleccionar "Application Manager" en el menú "Account". Dentro de este gestor de aplicaciones, se hace clic en el botón "Add application" como muestra la Figura 12.

The screenshot shows the VoXeo Evolution Application Manager interface. At the top, there is a navigation bar with links: ACCOUNT, DOCUMENTATION, TOOLS & DOWNLOADS, PAID SERVICES, and EXTREME SUPPORT. The user is logged in as TProject - 175419. The main content area is titled "Account > Applications" and shows the "EXISTING APPLICATIONS - APPLICATION VIEW" section. A table lists an application named "FACILEASYMO" with an application type of "Prophecy VoiceXML, Premium ASR/TTS" and a deployment type of "Staging". Below the table, there is an "Add Application" button. On the right side, there is an "Introduction" section with a gear icon and text explaining the Application Manager's function. Below that, there is an "APPLICATION SUMMARY" section with the following statistics: Application Count: 1, Phone# Count: 2, and Token Count: 0.

Figura 12. Application Manager

En la siguiente pantalla ("Create application"), se asigna un nombre a la aplicación, se selecciona el tipo de plataforma de las explicadas en el apartado 2.3.3.2 y se escribe la URL de acceso público para el código de la aplicación de voz en el campo "Voice URL". Si se utiliza el espacio de hospedaje gratuito de VoXeo, sólo hay que hacer click en el vínculo del "File Manager" para asignar el archivo que se ha cargado según el apartado anterior. En el caso de este proyecto se ha utilizado un servidor web llamado

X10hosting que se explicará más adelante, y para ello copiamos en el “Voice URL” la ruta en la que tenemos alojado el archivo que queremos utilizar de dicho servidor.

Finalmente, hay que completar todos los campos necesarios y hacer click en el botón “Create application”.

En las figura 13 y 14 se presentan gráficamente los pasos a seguir para la creación de la aplicación ‘PRUEBA’.

Account > Applications

CREATE A NEW APPLICATION

* Application Name: PRUEBA

* 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	Europe	CXXML	DTMF-Only	Prophecy VoiceXM
	USA	CXP (VoiceObjects)	Nuance 9	
		CallXML	Premium ASR	
		Designer	Premium ASR/TTS	
		VoiceXML		

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

* Voice URL: file manager
 http://facileasymo.x10.bz/vxml/index.php

+ Add a failover URL

Phone Number: Spain (+34) Madrid (91) - Voice Only

Create Application

Instructions

This form will collect all of the information you need to connect your voice or messaging application to our network.

You'll begin by giving your application a name and telling us how you'd like to communicate with your application. After that, you'll specify an application type and tell us where your application is located. That's it!

If you need help, our [Choosing a Platform](#) guide will help you pick the best platform for your voice application.

TEXT MESSAGING

Our SMS HTTP POST interface is an API that allows you to send and receive text messages using an HTTP POST request. Please visit our [text messaging documentation](#) for more information and sample code.

Your account does not have access to send outbound SMS messages. Please [open a ticket](#) to request access.

EXTREME SUPPORT

If you have any questions or problems, contact our 24/7 Extreme Support team at support@vox2o.com

Figura 13. Create a new application

Account > Applications > PRUEBA

APPLICATION SETTINGS

Successfully created the application!

Application Settings | Contact Methods

* Application Name: PRUEBA

* 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	CXXML	DTMF-Only	Prophecy VoiceXML
		CXP (VoiceObjects)	Nuance 9	
		CallXML	Premium ASR	
		Designer	Premium ASR/TTS	
		VoiceXML		

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

* Voice URL: file manager | view file
 http://facileasymo.x10.bz/vxml/index.php

+ Add a failover URL

Update Application | Delete Application

Introduction

Use this page to manage your application settings, including its URL's, phone numbers, instant messaging networks, and outbound dialing tokens.


TEXT MESSAGING

Our SMS HTTP POST interface is an API that allows you to send and receive text messages using an HTTP POST request. Please visit our [text messaging documentation](#) for more information and sample code.

Your account does not have access to send outbound SMS messages. Please [open a ticket](#) to request access.

VOICE COMMUNICATION OPTIONS

In addition to calling your application with any standard phone or SIP VoIP phone, you can use the following methods to call your application:

 Use the [Skype client](#) to dial: +990009369991487899

 Use an [inum client or compatible telephone network](#) to dial: +883510012351212

VOXEO HOSTED FILES

Use the [file manager](#) links located above each Start URL box, to map your application to your Vooxo-hosted files.

Figura 14. Tras especificar URL del código fuente, creada con éxito

Tras seguir los pasos anteriores, la aplicación nos muestra una nueva pestaña "Contact Methods". En ella se pueden consultar los teléfonos asignados para la aplicación que se acaba de crear. A continuación se muestra los datos de la aplicación 'PRUEBA' recién creada.

Account > Applications > PRUEBA

APPLICATION SETTINGS

Successfully created the application!

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 1980703
<input type="checkbox"/> iNum Number (Voice Only)	+883510012351212
<input type="checkbox"/> USA Toll Free PIN Access (Voice Only)	(800) 289-5570 then PIN: 9991487899
<input type="checkbox"/> USA Domestic PIN Access (Voice Only)	(407) 386-2174 then PIN: 9991487899
<input type="checkbox"/> Skype VoIP	+990009389991487899
<input type="checkbox"/> SIP VoIP	sip:9991487899@sip.voxeo.net
<input type="checkbox"/> Phono Number	sip:9991487899@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

Introduction

Use this page to manage your application settings, including its URL's, phone numbers, instant messaging networks, and outbound dialing tokens.

TEXT MESSAGING

Our SMS HTTP POST Interface is an API that allows you to send and receive text messages using an HTTP POST request. Please visit our [text messaging documentation](#) for more information and sample code.

Your account does not have access to send outbound SMS messages. Please [open a ticket](#) to request access.

VOICE COMMUNICATION OPTIONS

In addition to calling your application with any standard phone or SIP VoIP phone, you can use the following methods to call your application:

skype Use the [Skype client](#) to dial: +990009389991487899

inum Use an [iNum client or compatible telephone network](#) to dial: +883510012351212

Figura 15. "Contact Methods" de la aplicación 'PRUEBA'

Una vez se ha creado la aplicación que se quiere, es posible acceder en todo momento a ella para modificarla o directamente eliminarla, a través de la pestaña 'Application manager' seleccionamos la opción 'Existing Applications'.

ACCOUNT DOCUMENTATION TOOLS & DOWNLOADS PAID SERVICES EXTREME SUPPORT

TRproject - 175419 | logout

Account > Applications

EXISTING APPLICATIONS - APPLICATION VIEW

Application View | Number View

Application Name	Application Type	Deployment Type
FACILEASYMO	Prophecy VoiceXML, Premium ASR/TTS	Staging
PRUEBA	Prophecy VoiceXML, Premium ASR/TTS	Staging

Add Application

Introduction

The Application Manager maps your voice and messaging applications to phone numbers, instant messaging networks, and tokens.

Click on an application name to view full details, including linked phone numbers.

APPLICATION SUMMARY

Application Count: 2

Phone# Count: 4

Token Count: 0

CALLING FROM SKYPE

Figura 16. Aplicaciones creadas en la plataforma

2.3.3.5 Llamada a la aplicación

La aplicación de voz se puede probar llamando, por ejemplo, mediante Skype según se muestra en la Figura 17.

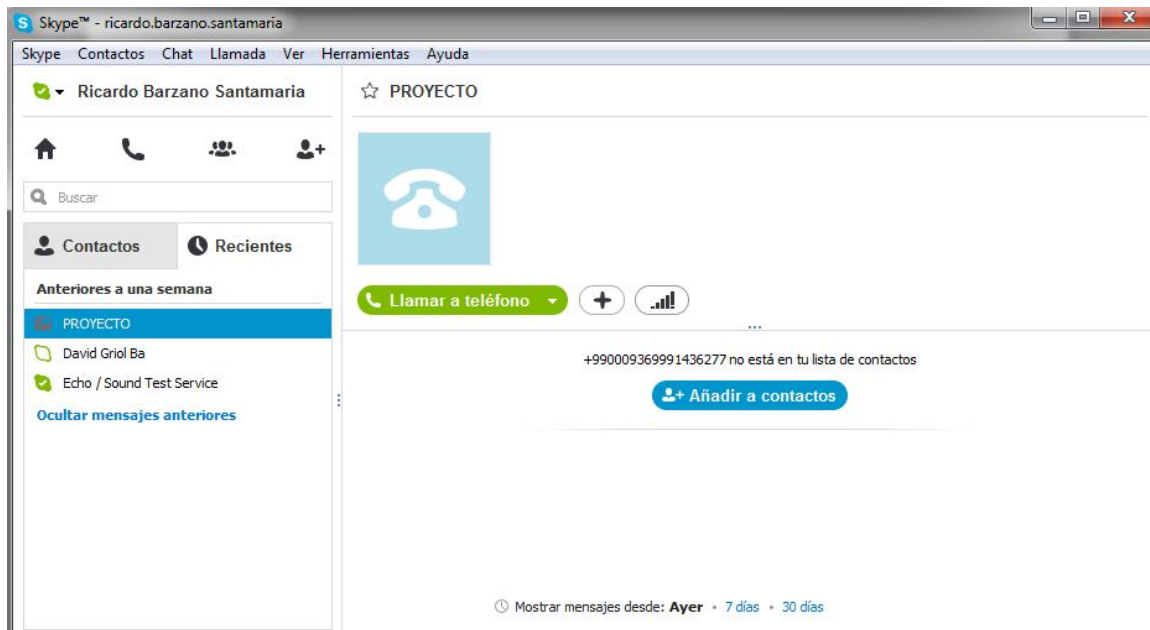


Figura 17. Llamada a la aplicación mediante Skype

2.3.4 Otras implementaciones del estándar VoiceXML

Además de la plataforma ofrecida por Voxeo, existen diversas implementaciones de las especificaciones de interfaces vocales del W3C para VoiceXML 2.0 y 2.1. Las más relevantes se describen brevemente a continuación.

- **Bevocal Café**^[23] es un entorno de desarrollo que proporciona un intérprete de VoiceXML 2.0 y las herramientas necesarias para depurar y probar aplicaciones. El intérprete de VoiceXML soporta las últimas recomendaciones del W3C, incluyendo mejoras como la verificación del usuario o la utilización de datos en formato XML.
- **Conita**^[24] ha implementado un intérprete VoiceXML basado en Open VXI.

- **Eloquent**^[25] ha implementado un intérprete de VoiceXML 2.0. Este intérprete soporta las especificaciones de interpretación semántica definidas por el W3C para el reconocimiento de la voz y ofrece hospedaje para aplicaciones vocales.
- **HP ofrece la plataforma OpenCall**^[26] que soporta VoiceXML 2.0 en entornos HP-Unix y Linux para protocolos de telefonía ISUP/SIP/RDSI, software de reconocimiento y síntesis MRCP (por ejemplo, Nuance y SpeechWorks) y CCXML 1.0. También está disponible un escritorio VoiceXML SDK libre para el desarrollo de aplicaciones y el entorno OpenCall Speech Web para una gestión eficiente e implementación de portales de voz a gran escala.
- **IBM**^[27] tiene una variedad de productos y herramientas para el desarrollo, depuración e implementación de aplicaciones VoiceXML. Los productos de aplicaciones VoiceXML de IBM incorporan la mayoría de las especificaciones VBWG, incluyendo VoiceXML 2.0 y SRGS 1.0.
- **Omvia Media Server**^[28] desarrollado por Intervoice incluye un navegador de VoiceXML 2.0. Intervoice también ofrece InVision Studio, una herramienta para el diseño, desarrollo y depuración de aplicaciones de VoiceXML. El Omvia Media Server soporta múltiples reconocedores y sintetizadores de voz.
- **JVoice XML**^[29] es un intérprete de VoiceXML de código libre desarrollado en JAVA que soporta APIs como JSAPI y JTAPI. Su objetivo principal es tener una plataforma independiente de la implementación que se pueda utilizar de forma gratuita.
- **Loquendo**^[30] ha desarrollado un intérprete de VoiceXML que administra documentos tanto VoiceXML 2.0 como 2.1. El intérprete está actualmente integrado en la plataforma Loquendo VoxNauta y se utiliza en una amplia gama de servicios de voz, incluso a gran escala. Por otra parte, LoquendoCafé proporciona a los desarrolladores recursos y herramientas para aprender a crear aplicaciones de voz y les permite ejecutar su aplicación de VoiceXML en una plataforma y escuchar el servicio que han creado por teléfono en diversos idiomas.
- **Lucent Technologies**^[31] ofrece la pasarela MiLife VoiceXML, que proporciona acceso telefónico a servicios de voz en entornos web.

- **Motorola licencia el sistema VoxGategay**^[32], ha sido integrado en gran número de plataformas de voz.
- **Nuance**^[33] ofrece una plataforma VoiceXML, herramientas gráficas para el desarrollo de aplicaciones, una plataforma para el desarrollo y prueba de prototipos y un navegador de voz basado en VoiceXML sin coste para los desarrolladores.
- **OpenVXI**^[34] es un intérprete de VoiceXML de código libre y portable desarrollado por la Universidad Carnegie Mellon y SpeechWorks. Puede ser usado de forma gratuita en aplicaciones comerciales y permite añadir módulos propietarios. Cumple la mayoría de las especificaciones definidas para VoiceXML 2.0.
- **OptimSys ofrece OptimTalk**^[35] una plataforma VoiceXML modular, flexible y escalable. Incluye intérpretes compatibles con VoiceXML y CCXML, y un SDK para agregar soporte a las soluciones VoiceXML y CCXML existentes. OptimTalk permite la elección de las tecnologías de reconocimiento y síntesis de voz y del hardware de telefonía.
- **Tellme**^[36] es una plataforma que responde a más de un millón de llamadas de teléfono VoiceXML 2.0 cada día. Los desarrolladores VoiceXML pueden acceder a Tellme Studio para construir gratis sus propias aplicaciones.
- **VoiceGenie**^[37] proporciona una plataforma que soporta los protocolos PSTN y SIP simultáneamente, 4 tipos de reconocedores de voz (Speech Works, BBN, AT&T Watson y Nuance) y 7 sintetizadores (incluyendo entre ellos SpeechWorks, AT&T Natural Voices, Rhetorical y SVOX). Además, dispone de un portal web donde las aplicaciones pueden desarrollarse y probarse de forma gratuita. VoiceGenie también ofrece la plataforma VoiceGenieCCXML que implementa las especificaciones CCXML del W3C y es capaz de interactuar con la plataforma VoiceXML.
- **Voxpilot**^[38] ofrece un entorno para el desarrollo e implementación online de aplicaciones multilingües VoiceXML. Incluye números de acceso locales en la mayoría de países europeos, proporcionando una robusta infraestructura VoIP.

3 Descripción general del sistema desarrollado

En este capítulo se describen las características fundamentales del conjunto de módulos que conforman el sistema desarrollado: funcionalidad, arquitectura y esquema general de bloques. A continuación se analizan las distintas tecnologías utilizadas para su desarrollo. Finalmente, se incluyen las operaciones más generales y codificaciones especiales que se han tenido en cuenta a lo largo de la implementación.

3.1 Presentación del sistema

El portal de voz Facileasymo desarrollado para este Proyecto Final de Carrera ofrece a personas de todas las edades acceso telefónico a las distintas soluciones informáticas. También a través de la plataforma web el usuario puede consultar en todo momento distintos manuales que le servirán para lograr solucionar cualquier incidencia informática que se le resista.

- Consultar información relacionada con un equipo informático.
- Realizar encuestas.
- Realizar registro de usuario para atención más personalizada.
- Acceso a la plataforma desde cualquier dispositivo móvil.

El desarrollo de la aplicación se ha hecho utilizando una arquitectura cliente-servidor cuyo esquema se puede observar en la Figura 18.

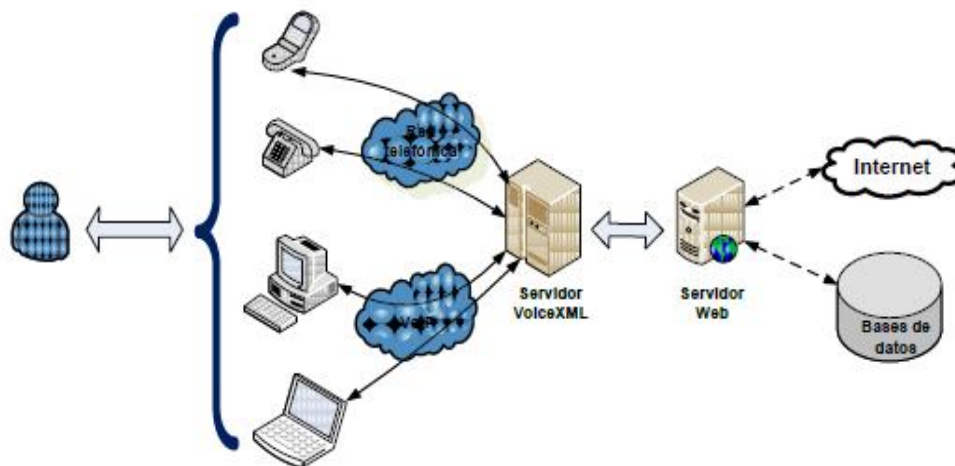


Figura 18. Arquitectura de la aplicación

La interacción con el sistema comienza cuando el usuario inicia una llamada, utilizando la línea telefónica o mediante cualquier cliente de VoIP (por ejemplo, Skype).

En el servidor VoiceXML (para nuestro proyecto, la plataforma Voxeo Evolution), el intérprete de VoiceXML se encarga por un lado de responder las llamadas de los usuarios y, por otro, de interpretar los documentos VoiceXML para ofrecer el servicio al usuario.

También es el encargado de solicitar los recursos necesarios para la ejecución de la aplicación, de seguir la lógica del servicio y de mantener el estado de sesión de los usuarios actuando en consecuencia. Para llevar a cabo estas acciones el intérprete VoiceXML debe disponer de una serie de elementos:

- Un sistema que se encargue de atender las llamadas de los usuarios. Para ello debe tener conexión al canal de acceso, ya sea por red telefónica o por voz sobre IP.
- Un sistema que gestione la comunicación con los servidores y vaya solicitando en cada momento los recursos necesarios identificados por sus URLs mediante, por ejemplo, los protocolos HTTP y FTP.
- Un sistema que reproduzca los ficheros de audio.

- Un sistema que genere audio a partir de texto: conversión de texto a voz, TTS (Text-to-Speech).
- Un sistema que se encargue de la recogida de datos del usuario, ya sea mediante mecanismos de reconocimiento de voz, como de reconocimiento de tonos DTMF.
- Un sistema que posibilite la grabación de voz.
- Un sistema que se encargue de la ejecución de la lógica del servicio, del mantenimiento de las sesiones y de la gestión de eventos.

En el servidor web se almacenan los ficheros vxml y php que implementan cada uno de los servicios prestados por el portal de voz, además de las bases de datos que contienen la información utilizada. Todos los ficheros se interrelacionan de tal forma que el usuario en su llamada puede navegar por cualquiera de los servicios ofrecidos, pudiendo realizar más de una acción en cada llamada.

3.2 Tecnologías utilizadas

3.2.1 Intérprete VoiceXML: Voxeo

Actualmente existen multitud de intérpretes del lenguaje VoiceXML, tal y como se ha explicado en el apartado 2.3. Uno de ellos es Voxeo, que proporciona la infraestructura y componentes necesarios para poder crear y usar aplicaciones de voz. Dado que en el apartado 2.3 se analizó en detalle esta herramienta, en las siguientes líneas se razonarán los motivos por los que ha sido elegida para la implementación y alojamiento de la aplicación desarrollada en este Proyecto Final de Carrera.

En primer lugar, Voxeo permite crear una aplicación VoiceXML y acceder a ella a través de distintos medios, ya que proporciona un número de teléfono local (según país y provincia en el caso de España) y un número Skype para llamadas desde la aplicación. Voxeo permite además hacer un seguimiento en tiempo real de las llamadas realizadas al servicio, así como visualizar y almacenar los ficheros de logs de las mismas, muy útil a la hora de depurar y optimizar la aplicación, estudiar el uso que se hace de ella e incluso obtener estadísticas. 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. Cabe añadir que la plataforma posee un

sistema de soporte rápido y eficaz, que incluye foros, tickets de soporte y una documentación muy completa. Además, todo ello de forma gratuita.

Por último, Voxeo proporciona el intérprete VoiceXML y los componentes ASR y TTS necesarios para la ejecución de la aplicación. En el presente Proyecto se ha utilizado la implementación Prophecy 9 - Multi-Language VXML proporcionada por Voxeo, que ha permitido desarrollar la aplicación para su uso en idioma castellano.

3.2.2 Base de datos: MySQL

MySQL es un sistema de gestión de bases de datos relacional, multihilo y multiusuario. Relacional debido a que se sustenta en un modelo de datos basado en la lógica de predicados y en la teoría de conjuntos. Multihilo, ya que permite el acceso concurrente a la información de forma ordenada y segura. Finalmente multiusuario, porque permite proveer servicio a múltiples usuarios simultáneamente.

MySQL AB (desde enero de 2008 una subsidiaria de Sun Microsystems y ésta a su vez de Oracle Corporation desde abril de 2009) desarrolla MySQL como software libre en un esquema de licencia dual.

Por un lado se ofrece bajo la GNU GPL para cualquier uso compatible con esta licencia, pero para aquellas empresas que quieran incorporarlo en productos comerciales deben comprar a la empresa una licencia específica que les permita este uso. Además de la venta de estas licencias privativas, la compañía ofrece soporte y servicios.

Se trata además del sistema de base de datos con uso más generalizado en aplicaciones web, como Drupal o phpBB, en plataformas (Linux/Windows-Apache-MySQL-PHP/Perl/Python), y por herramientas de seguimiento de errores como Bugzilla. Su popularidad como aplicación web está muy ligada a PHP, que a menudo aparece en combinación con MySQL.

Varios ejemplos de sitios web importantes y populares que hoy en día utilizan MySQL son Wikipedia, Google (aunque no para búsquedas), Facebook, Twitter, Flickr, y YouTube.

En el presente proyecto se optó por MySQL debido a su facilidad de manejo y a que aporta las siguientes ventajas:

- Escalabilidad: es posible manipular bases de datos enormes, del orden de seis mil tablas y alrededor de cincuenta millones de registros, y hasta 32 índices por tabla.
- MySQL está escrito en C y C++ y probado con multitud de compiladores y dispone de APIs para muchas plataformas diferentes.
- Conectividad: permite conexiones entre diferentes máquinas con distintos sistemas operativos. Es corriente que servidores Linux o Unix, usando MySQL, sirvan datos para ordenadores con Windows, Linux, Solaris, etc. Para ello se usa TCP/IP, tuberías, o sockets Unix.
- Tal y como se ha comentado, es multihilo, con lo que puede beneficiarse de sistemas multiprocesador.
- Permite manejar multitud de tipos para columnas.
- Permite manejar registros de longitud fija o variable.
- Uso gratuito tanto del sistema como de diferentes administradores de bases de datos MySQL, como el que se presenta en la siguiente sección.

3.2.3 Administrador de base de datos: phpMyAdmin

PhpMyAdmin es una herramienta escrita en PHP con la intención de manejar la administración de MySQL a través de páginas web, utilizando Internet. Actualmente puede crear y eliminar Bases de Datos, crear, eliminar y alterar tablas, borrar, editar y añadir campos, ejecutar cualquier sentencia SQL, administrar claves en campos, administrar privilegios, exportar datos en varios formatos y está disponible en 62 idiomas. Se encuentra disponible bajo la licencia GPL.

Las principales características de phpMyAdmin son:

- Multiplataforma.
- Multilenguaje (más de 50).
- Licencia GPL.
- Escrito en PHP.

Por todas las anteriores propiedades se ha decidido hacer uso de phpMyAdmin versión 3.3.9.2 para administrar y gestionar las bases de datos utilizadas en la aplicación del portal de voz. Además,

posee una interfaz gráfica muy intuitiva que facilita su uso, la cual se puede observar en la captura de pantalla de la Figura 19.

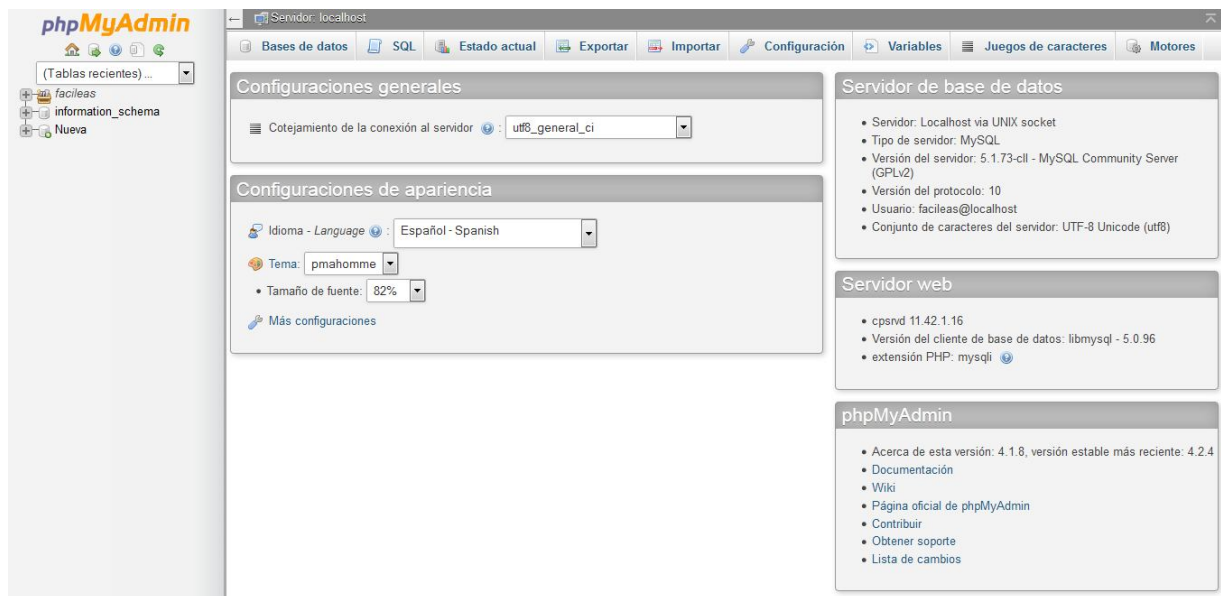


Figura 19. Página de inicio de PHPMyAdmin

3.2.4 Servidor web

Servidor web para el almacenamiento de los ficheros PHP y de las bases de datos de la aplicación se ha utilizado x10hosting. Este servidor tiene las siguientes características para su plan gratuito:

- cPanel 11+.
- 300 MB de almacenamiento.
- 10 Gb de banda ancha.
- Cuentas FTP.
- Posibilidad de agregar dominios o subdominios.
- Servidor Apache 2.2.
- PHP 5 y 3 bases de datos MySQL.

- CGI, PHP mail, POP3, IMAP, Sendmail y cuentas de e-mail.
- Auto-instalador de scripts.
- 99.9 % online y soporte técnico 24/7.

Otras ventajas de x10hosting es que permite subir archivos comprimidos, descomprimir en el servidor y subir archivos al directorio no web (directorio de acceso no público).

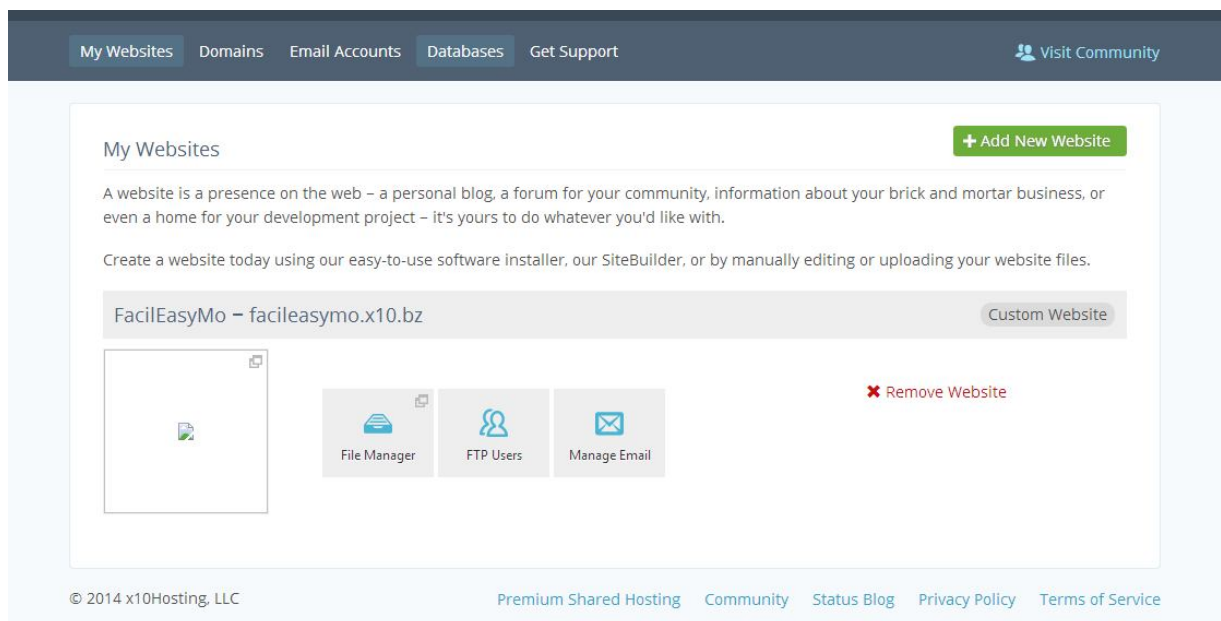


Figura 20. Panel de control de usuario de x10hosting

3.3 Implementación de las operaciones generales

En este apartado se explican en detalle las operaciones comunes al conjunto de módulos que ha sido necesario tratar e implementar a lo largo del desarrollo del portal de voz.

3.3.1 Tratamiento de la información

En la aplicación desarrollada se trabaja con dos tipos de información: información estática e información dinámica. Cada una de ellas se trata de diferente forma a lo largo de la aplicación, tal y como se explica a continuación.

3.3.1.1 Información estática

Se ha considerado información estática a aquella que no cambia con el paso del tiempo, o al menos no lo hace en un tiempo considerable. Este tipo de información se recopila de páginas web, principalmente de páginas relacionadas con problemas e incidencias informáticas y todas aquellas con información de distintos dispositivos electrónicos. Toda esa información se almacena clasificada, en la base de datos de la aplicación. Cuando el usuario solicita este tipo de información, el sistema accede a ella en la base de datos y la devuelve encapsulada en VoiceXML.

Ejemplos de este tipo de información son los distintos pasos de solución que se ofrece a través de la aplicación de voz como los diferentes manuales de los dispositivos electrónicos que encontramos en el portal web.

3.3.1.2 Información dinámica

La información dinámica es aquella que cambia a lo largo del tiempo. Esta información se tiene que recargar en la base de datos cada cierto tiempo, ya que si no se actualizara podría llegar un momento en que la base de datos contuviera datos obsoletos o erróneos. Ejemplos de este tipo de información son el historial de incidencias de cada usuario, disponible en todo momento en el portal web y del cual se vale la plataforma de voz para poder ofrecer soluciones más personalizadas al usuario.

Una vez se llama a la plataforma de voz Facileasymo, ésta realiza una lectura de la base de datos para poder consultar las últimas incidencias registradas por el nombre de usuario. Mediante esta tarea se ahorra tiempo al usuario y se consigue una atención más ágil y personalizada.

Ejemplos de este tipo de información es el número de teléfono con el que el usuario realiza la llamada, al recibir la llamada se compara con los distintos números almacenados en la base de datos y si existe coincidencia se ofrece un saludo personalizado y posibilidad de acceder a la incidencia más frecuente en él.

3.3.2 Gramáticas

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. Además, 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, a continuación se indicaran para qué casos han sido elegidas cada una de ellas.

3.3.2.1 Gramáticas estáticas

Las gramáticas estáticas tratan información que no varía a lo largo del tiempo, además de tener un número pequeño de opciones a elegir. Se codifican de forma fija en el mismo fichero donde se utilizan.

Un ejemplo de este tipo de gramática es la despedida que nos hace la aplicación de voz una vez hemos finalizado el dialogo, y que contiene siempre el mismo mensaje. El en caso del portal web, otro ejemplo que encontraríamos sería las preguntas que se realizan en la encuesta al salir de la plataforma y que se explicará con más detalle más adelante.

```
<filled>  
    <audio src="despedida.wav"></audio>  
</filled>
```

Tabla 8 – Ejemplo gramática estática.

3.3.2.2 Gramáticas dinámicas

Las gramáticas dinámicas utilizan información que varía con el tiempo y suelen tratar gran cantidad de datos. Si se quisieran incluir en los ficheros vxml de forma manual habría que modificarlas cada poco tiempo, tarea que sería muy laboriosa debido al gran número de ítems de los que constan.

Un ejemplo de esta gramática es el tratamiento personalizado que recibe el usuario al llamar a la aplicación de voz, y que en caso de reconocer el número de teléfono ofrece una solución a su incidencia

más frecuente, por otro lado en la plataforma web y acorde al ejemplo que se puso anteriormente, los resultados de la encuesta que se muestran por pantalla una vez se ha finalizado la misma.

```
<var name="idllamante" expr="session.callerid"/>
<var name="idaplicacion" expr="session.calledid"/>
```

Tabla 9 – Ejemplo gramática dinámica. Captura de número de teléfono

```
$_SESSION['telefono'] = $telefono;
```

Tabla 10 – Ejemplo gramática dinámica. Se guarda el teléfono en una variable

```
$database = 'facileas_sesion';
require('../php/s_base_de_datos.php');
$sql = "SELECT * FROM usuarios WHERE telefono='$telefono'";
```

Tabla 11 – Ejemplo gramática dinámica. Coteja número de teléfono con base de datos

```
<form id="bienvenida">
<block>
<prompt bargein="false">Hola <?php echo $nombre; ?>. Bienvenido a facilisimo. ¿Vuelves a tener problemas con
tu <?php echo $dispositivo ?>?</prompt>
</block>
```

Tabla 12 – Ejemplo gramática dinámica. Bienvenida personalizada

3.4 Gestión de bases de datos

3.4.1 Formato de almacenamiento

MySQL soporta distintas tecnologías de almacenamiento de datos, entre ellas destacan MyISAM e InnoDB:

- InnoDB es una tecnología de almacenamiento de datos de fuente abierta para MySQL, incluido como formato de tabla estándar en todas las distribuciones de MySQL AB a partir de las

versiones 4.0. Su característica principal es que soporta transacciones de tipo ACID (Atomicity, Consistency, Isolation and Durability), bloqueo de registros e integridad referencial.

- MyISAM es la tecnología de almacenamiento de datos usada por defecto por el sistema administrador de bases de datos relacionales MySQL. Este tipo de tablas están basadas en el formato ISAM, incluyendo además nuevas extensiones.

Como el mejor rendimiento de uno u otro formato depende de la aplicación específica, en el desarrollo del proyecto se ha elegido en cada caso concreto de tabla el formato que proporciona la mejor relación de calidad.

De este modo, cuando la aplicación hace un uso elevado de INSERT y UPDATE se ha escogido InnoDB, ya que proporciona un aumento de rendimiento en cuanto a fiabilidad y consistencia con respecto a MyISAM. En aquellos casos en los que predominan las consultas SELECT a la base de datos, se utiliza MyISAM, ya que se obtiene mayor velocidad en general a la hora de recuperar datos.

3.4.2 Codificación

El Estándar Unicode es un estándar de codificación de caracteres diseñado para facilitar el tratamiento informático, transmisión y visualización de textos de múltiples lenguajes y disciplinas técnicas. El término Unicode proviene de los tres objetivos perseguidos: universalidad, uniformidad y unicidad.

Uno de los formatos Unicode es UTF-8 (8-bit Unicode Transformation Format). Se trata de un formato de codificación de caracteres que utiliza símbolos de longitud variable. Está definido como estándar por la RFC 3629 de la Internet Engineering Task Force (IETF). Actualmente es una de las tres posibilidades de codificación reconocidas por Unicode y lenguajes web, y de las cuatro definidas en ISO 10646.

MySQL tiene dos conjuntos de caracteres Unicode: ucs2 (UCS-2 Unicode) y utf8 (UTF-8 Unicode). Usando estos conjuntos de caracteres se puede almacenar texto en unos 650 idiomas.

Las colaciones utf8_spanish_ci y utf8_spanish2_ci se corresponden con español moderno y español tradicional respectivamente. En ambas colaciones, 'ñ' es una letra independiente, entre 'n' y 'o'. Además, para español tradicional 'ch' es una letra, ordenada entre 'c' y 'd', y 'll' es una letra que se coloca entre 'l' y 'm'.

Para poder utilizar el idioma español moderno en la codificación de todos los datos que se usan en la aplicación del portal de voz, se almacena la información en tablas con cotejamiento utf8_spanish_ci.

3.4.3 Jerarquía de la base de datos del sistema

Para el almacenamiento de toda la información necesaria en la centralita de voz y la plataforma web, se utilizan dos bases de datos: 'facileas_contenido' y 'facileas_sesion'.

Base de datos	Cotejamiento
facileas_contenido	utf8_general_ci
facileas_sesion	utf8_general_ci

Figura 21. Bases de datos utilizadas

- **Base de datos 'facileas_contenido'**

En esta base de datos se almacena toda la información de contenido estático tanto de la página web como del sistema de voz. Consta de 3 tablas que se explican a continuación.

Tabla	Acción	Filas	Tipo	Cotejamiento	Tamaño	Residuo a depurar
articulos	Examinar Estructura Buscar Insertar Vaciar Eliminar	10	MyISAM	utf8_spanish_ci	73.9 KB	-
encuesta_opciones	Examinar Estructura Buscar Insertar Vaciar Eliminar	~28	InnoDB	utf8_spanish_ci	32 KB	-
encuesta_preguntas	Examinar Estructura Buscar Insertar Vaciar Eliminar	~10	InnoDB	utf8_spanish_ci	32 KB	-
soluciones_incidencias	Examinar Estructura Buscar Insertar Vaciar Eliminar	1	MyISAM	utf8_spanish_ci	2.3 KB	-
4 tablas	Número de filas	49	MyISAM	latin1_swedish_ci	140.2 KB	0 B

Figura 22. Base de datos 'facileas_contenido'

Tabla 'articulos': contiene los artículos de información correspondiente a cada dispositivo electrónico mostrados en la plataforma web.

Tabla 'encuesta_opciones': almacena las distintas opciones que podrá elegir el usuario para contestar a las preguntas de la encuesta. Además se utiliza para realizar las estadísticas de resultados de la encuesta.

Tabla 'encuesta_preguntas': incluye las preguntas de la encuesta ofrecida tanto por la aplicación de voz como por la plataforma web.

Tabla 'soluciones incidencias': contiene todos los registros de incidencias con sus correspondientes soluciones a los cuales accederá el gestor de incidencias de la aplicación de VOZ.

- **Base de datos 'facileas_sesion'**

En esta base de datos se almacena todos lo necesario para la gestión de la información de los usuarios registrados, desde los datos de sus conexiones a la web, información personal de contacto, historial de llamadas, validación y recuperación de cuentas, etc. Toda esta información es dinámica ya que está abierta a posibles cambios por parte del usuario. Para ello consta de 7 tablas que se explican a continuación

Tabla	Acción	Filas	Tipo	Cotejamiento	Tamaño
<input type="checkbox"/> cambio_correo	Examinar Estructura Buscar Insertar Vaciar Eliminar	~0	InnoDB	utf8_spanish_ci	16 KB
<input type="checkbox"/> control_acceso	Examinar Estructura Buscar Insertar Vaciar Eliminar	~119	InnoDB	utf8_spanish_ci	48 KB
<input type="checkbox"/> historial	Examinar Estructura Buscar Insertar Vaciar Eliminar	6	MyISAM	utf8_spanish_ci	2.8 KB
<input type="checkbox"/> historial_sin_registro	Examinar Estructura Buscar Insertar Vaciar Eliminar	0	MyISAM	utf8_spanish_ci	1 KB
<input type="checkbox"/> recuperar	Examinar Estructura Buscar Insertar Vaciar Eliminar	~0	InnoDB	utf8_spanish_ci	16 KB
<input type="checkbox"/> usuarios	Examinar Estructura Buscar Insertar Vaciar Eliminar	~2	InnoDB	utf8_spanish_ci	16 KB
<input type="checkbox"/> validacion	Examinar Estructura Buscar Insertar Vaciar Eliminar	~2	InnoDB	utf8_spanish_ci	16 KB
7 tablas	Número de filas	129	MyISAM	utf8_general_ci	115.8 KB

Figura 23. Base de datos 'facileas_sesion'

Tabla 'cambio_correo': incluye claves aleatorias enviadas al usuario para validar el cambio de su correo electrónico.

Tabla 'control_acceso': almacena los diferentes registros de los usuarios que inician sesion en la plataforma web. Información tal como la versión del navegador y del sistema operativo, la IP desde donde accede, hora de acceso, etc.

Tabla 'historial': contiene el historial de anteriores incidencias para usuarios registrados en la web.

Tabla 'historial_sin_registro': en esta tabla se encuentra el historial de anteriores incidencias para usuarios aún no registrados. Cuando el usuario se registre en la plataforma, los datos de estas tablas asociados al número de teléfono registrado se moverán a la tabla de historial de

usuarios registrados. De este modo, podrá tener acceso a incidencias anteriores al registro, y evitaremos que se pierda información de utilidad.

Tabla 'recuperar': almacena claves aleatorias enviadas al usuario para cambiar la contraseña cuando ésta se ha olvidado.

Tabla 'usuarios': incluye datos personales y de inicio de sesión del usuario registrado. En esta tabla se asociarán los números de teléfono para identificar al usuario cuando llama por teléfono.

Tabla 'validación': en esta tabla se guardan las claves aleatorias enviadas por correo electrónico al usuario para hacer la validación de su cuenta en el sistema.

4 Descripción detallada de los módulos del sistema

Se va a proceder a analizar en detalle cada uno de los módulos de la aplicación desarrollada: descripción de las funcionalidades que proporcionan, comentarios sobre la arquitectura, visualización gráfica de la información ofrecida y, finalmente, la exposición de sus escenarios de mayor utilización.

4.1 Aplicación de voz

La aplicación de voz tratará de dar soporte a todo aquel usuario que decida llamar por una incidencia informática. A través de una llamada telefónica, la aplicación se ocupa de reconocer al usuario y se darán dos tipos de casuísticas a tratar. Por un lado si el usuario fuera nuevo la aplicación recopilará sus datos para dejarlos almacenados en la base de datos y mediante unas sencillas preguntas se encargará de identificar nuestra incidencia, y ofrecernos unos sencillos pasos para lograr solucionarla. En todo momento el usuario puede retroceder o repetir cualquiera de estos pasos en caso de no haberlo escuchado bien o por la razón que quisiese. El otro camino que podría coger la llamada es que el usuario ya haya sido registrado a través del portal web, por lo que la aplicación le dará la bienvenida personalizada y le preguntará si el motivo de su llamada está relacionado con la incidencia que por frecuencia más veces haya sido registrada en su caso, de ser así se le ofrecerá directamente la solución, si no fuera ese el motivo seguirá adelante con la llamada como en el anterior caso hasta poder ofrecerle una solución acorde a la incidencia.

Toda la información que recopile la aplicación durante la llamada será guardada en la base de datos y dejará un registro vinculado al usuario, de manera que siempre este actualizada y haya una sincronización con la información de la plataforma web. En caso de que el usuario no estuviera aun registrado esos datos se guardarían una tabla auxiliar de historial para usuarios aun no registrados en la web.

A continuación se explican en mayor detalle cada uno de los distintos módulos que se han diseñado para la aplicación de voz, y las fases por las que pasará la llamada.

4.1.1 Diagrama general de ejecución

El diagrama de flujo general de ejecución del sistema de resolución de incidencias por voz es el mostrado a continuación.

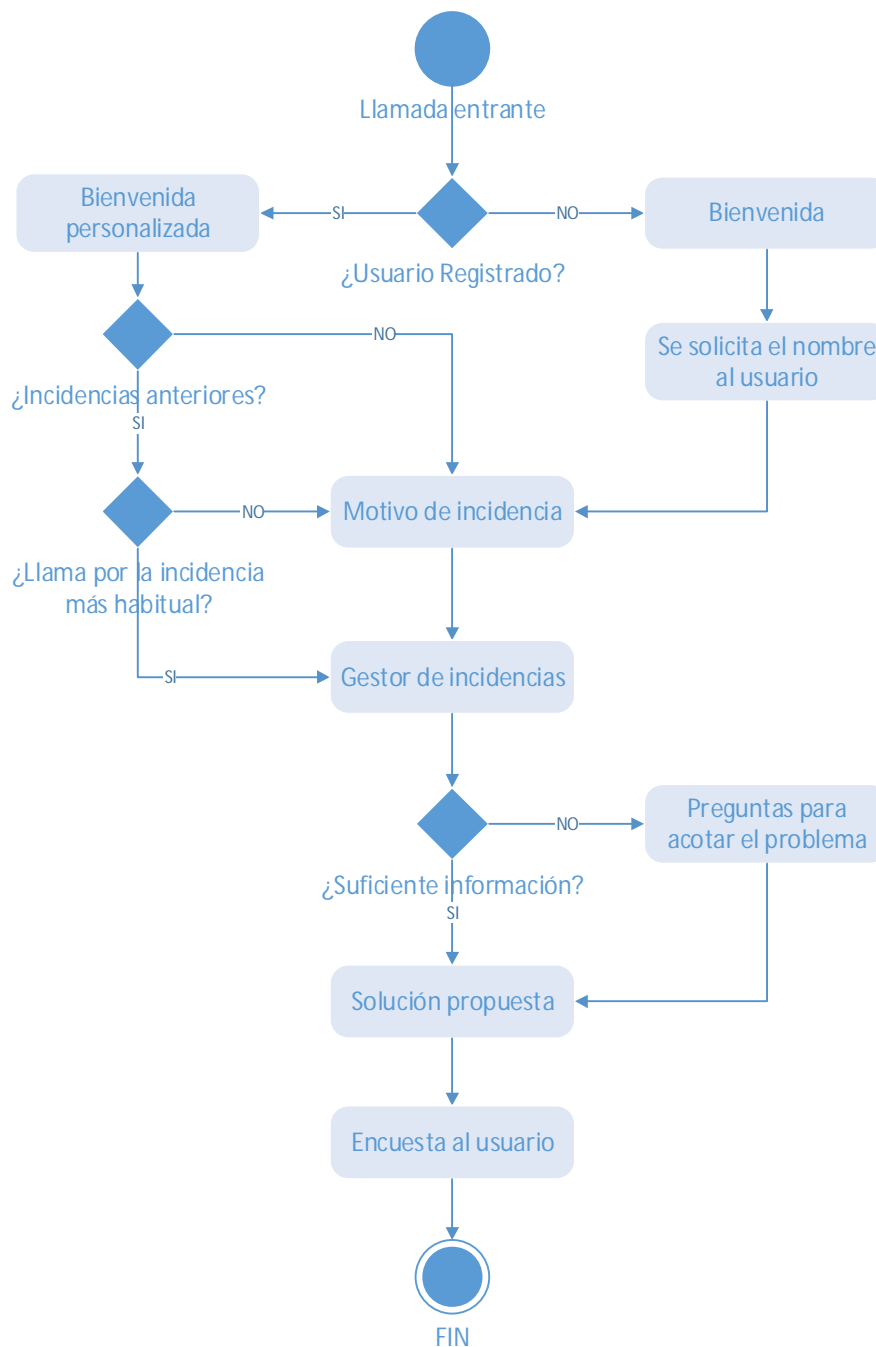


Figura 24. Diagrama general de la aplicación de Voz

4.1.2 Index

El módulo **index** implementa el primer diálogo que se le proporciona al usuario al llamar al número de teléfono del portal de voz. Si el usuario ya ha sido registrado previamente recibirá un saludo personalizado y se le ofrecerá la solución a la incidencia que estadísticamente ha consultado con mayor frecuencia. Si el usuario no hubiera sido registrado y fuera la primera vez que utiliza la plataforma, se le pedirá su nombre para registrarlo automáticamente y quede vinculado al teléfono desde el cual ha realizado la llamada, y a continuación pasará al módulo "**nueva incidencia**" donde se le pedirá que aporte detalles sobre su incidencia.

4.1.3 Nueva incidencia

La aplicación llega a este módulo para captar la incidencia por la cual el usuario ha llamado, para ello pregunta al usuario el motivo de su llamada y recoge el mensaje que este le diga y que tratará en el siguiente módulo "**Gestor de incidencia**".

4.1.4 Gestor de incidencia

La aplicación llega a este módulo una vez ha captado la incidencia del usuario, y mediante las palabras clave almacenadas en la base de datos es capaz de cotejar el mensaje del usuario y redirigirlo a la solución planteada para dicha incidencia. Una vez conoce el motivo del problema se le ofrecerá una serie de sencillos pasos a seguir para lograr solucionarlo. Si el usuario una vez escuchado y probada la solución propuesta continuará teniendo el problema se le ofrecerá la posibilidad de volver a indicar la incidencia por si hubiera cambiado en el transcurso de la llamada o bien no se hubiera entendido correctamente, en caso contrario y de resultar correcta la solución la aplicación pasaría al siguiente módulo "**Encuesta**". En todo momento el usuario podrá repetir un paso de la solución o volver a escuchar los anteriores, simplemente diciendo repetir, o anterior y siguiente para continuar.

4.1.5 Encuesta

Tras dar por concluida la incidencia y antes de finalizar la llamada mediante el módulo "**Exit**", se le ofrecerá al usuario realizar una sencilla y rápida encuesta que nos permitirá recopilar información de

primera mano y en la que se podrá valorar tanto la atención recibida como aspectos a mejorar, de esta manera se podrá evolucionar acorde a las necesidades prioritarias de los usuarios.

4.1.6 Exit

Tras haber solucionado la incidencia con la que se encontraba el usuario antes de realizar la llamada y haber respondido las preguntas de la encuesta, la aplicación volcará toda la información en la base de datos para que a su vez sea mostrada en el portal web, de esta forma si el usuario quisiera acceder tras colgar el teléfono ya vería reflejada su llamada en el historial personal.

4.1.7 Ejemplo práctico de llamada a Facileasymo

Se muestra un ejemplo de una llamada realizada por un usuario no registrado aún, y que por lo tanto no dispone de historial de incidencias.

Maquina: Hola. Bienvenido al sistema de incidencias de facilísimo. ¿Cuál es tu nombre?

Usuario: Ricardo.

Maquina: Te he entendido: "Ricardo". ¿Es correcto?

Usuario: Sí.

Maquina: Describe brevemente qué problema tienes.

Usuario: Mi impresora no funciona.

Maquina: ¿Has dicho: "Mi impresora no funciona"?

Usuario: Sí.

Maquina: Un momento por favor, estamos buscando una solución.

A continuación te daré los pasos necesarios para solucionar tu problema. Di siguiente para continuar; repetir para volver a escuchar; o anterior para volver atrás. Comenzamos.

Paso 1. Apague y desconecte su impresora.

Usuario: Siguiente.

Maquina: Paso 2. Extraiga la bandeja tirando suavemente de ella.

Usuario: Siguiente.

Maquina: Paso 3. Retire el papel atascado en los rodillos.

Usuario: Siguiente.

Maquina: Paso 4. Vuelva a colocar la bandeja en su sitio.

Usuario: Siguiente.

Maquina: Último paso. Conecte y encienda de nuevo su impresora.

Usuario: Siguiente.

Maquina: ¿Se ha solucionado tu problema?

Usuario: Si.

-----ENCUESTA-----

Maquina: Vale Ricardo. Antes de terminar quiero hacerte un cuestionario muy rápido. Me va servir para continuar mejorando. Por favor, responde con la opción que te parezca más adecuada.

¿El sistema de voz te ha ayudado a solucionar tu problema?

1. Si. 2. No

Usuario: Si.

Maquina: Establece el nivel de dificultad del sistema de voz.

1. Fácil. 2. Normal. 3. Difícil.

Usuario: Uno.

Maquina: ¿Qué tal te ha entendido el sistema de voz?

1. Bien. 2. Regular.

Usuario: Bien.

Maquina: La interacción con el sistema de voz fue...

1. Rápida. 2. Adecuada. 3. Lenta.

Usuario: Rápida.

Maquina: Ya hemos acabado Ricardo. Muchas gracias por utilizar los servicios de facilísimo. Hasta pronto.

-----FIN DE LA LLAMADA-----

Tabla 13 – Transcripción de una llamada de resolución de incidencia.

4.2 Plataforma web

A continuación se van a detallar las distintas funcionalidades de las que se ha dotado a la plataforma web, así como el funcionamiento de cada una de ellas. Para hacerlo más ilustrativo se van a acompañar con varias imágenes que servirán para mostrar el aspecto final que podrá encontrarse el usuario.

4.2.1 Vista principal

Una vez visitamos Facileasymo (<http://www.facileasymo.x10.bz/>), veremos la primera vista general que se encuentra antes de proceder al registro y acceso a la zona personal. En este lugar podremos ver las características principales y la propuesta de valor de la plataforma.

Uno de los puntos fuertes es que se ha apostado por utilizar **AJAX para cargar los contenidos de la plataforma**. De este modo se recargará solo la parte que se necesite y no todo la página. Con esta tecnología se consiguen grandes resultados de velocidad y fluidez en la navegación, además de reducir los datos enviados y recibidos entre el cliente y el servidor.

En las figuras 25, 26 y 27 se detallan algunas de las funcionalidades del sistema completo. Todas ellas tienen apariencia adaptadas para un navegador de escritorio.

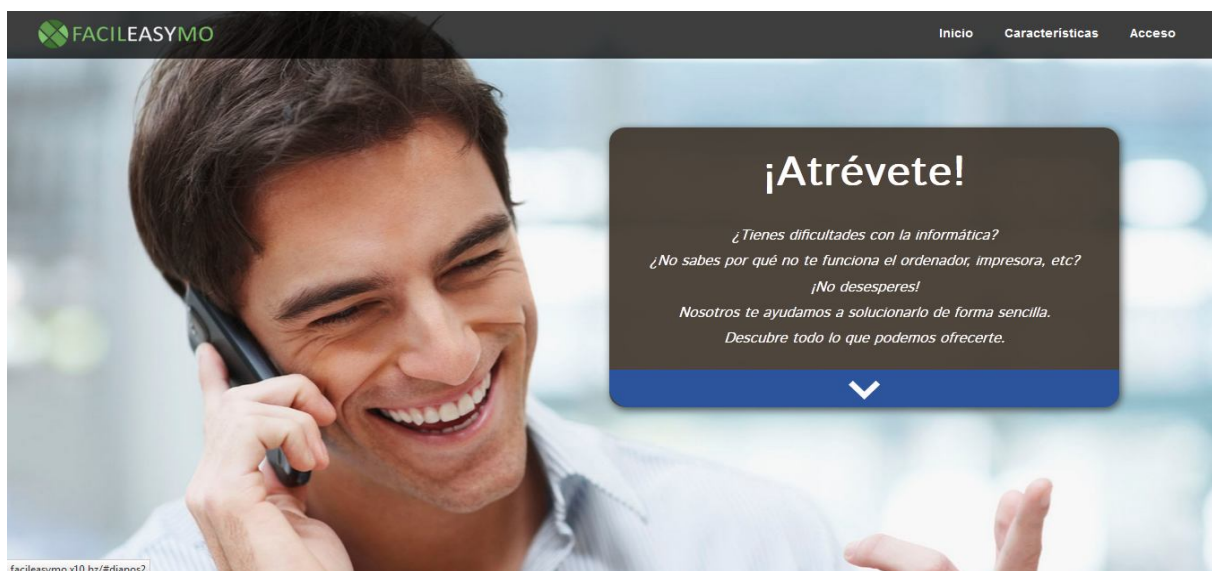


Figura 25. Pantalla de inicio de Facileasymo.x10.bz desde navegador de escritorio

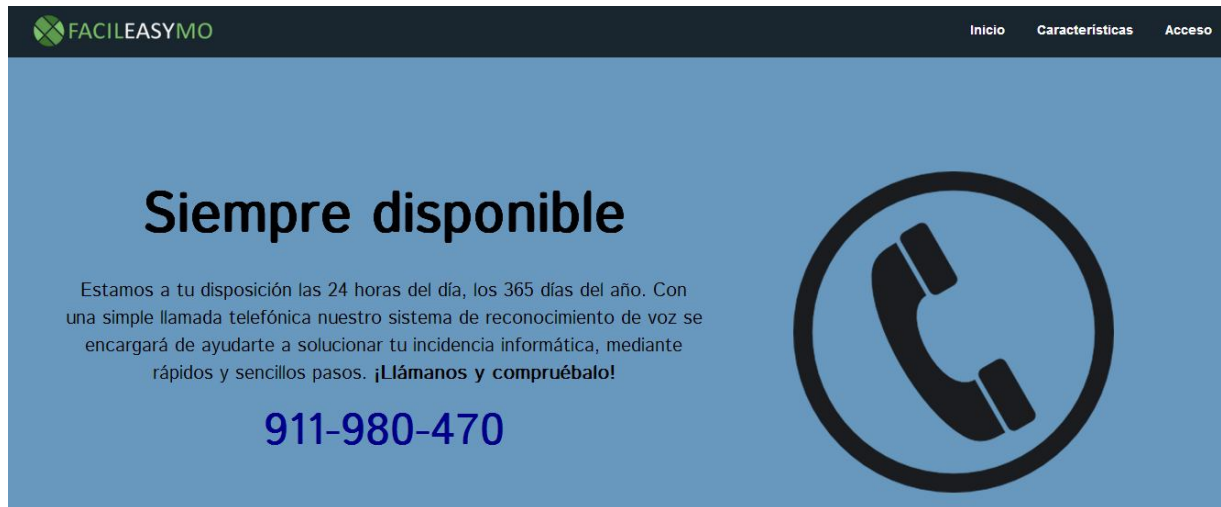


Figura 26. Características del sistema de voz



Figura 27. Descripción de característica de responsive design

El objetivo principal es invitar al usuario desde el primer momento a atreverse a solucionar el mismo sus propias incidencias, mostrándoselo de manera dinámica y didáctica, para aprender desde el momento en el que comience a utilizar Facileasymo.

Como se puede ver en la figura 27, resaltamos la característica de poder acceder desde cualquier sitio, es decir, la plataforma se adapta a cualquier tamaño de pantalla pudiendo acceder desde distintos dispositivos como tablets y smartphones, incluso televisores, abarcando la gran variedad de tamaños existentes en el mercado. Para ello, se ha adaptado un sencillo menú que nos ofrece las mismas opciones que en el navegador de escritorio, pero de una forma más dinámica.

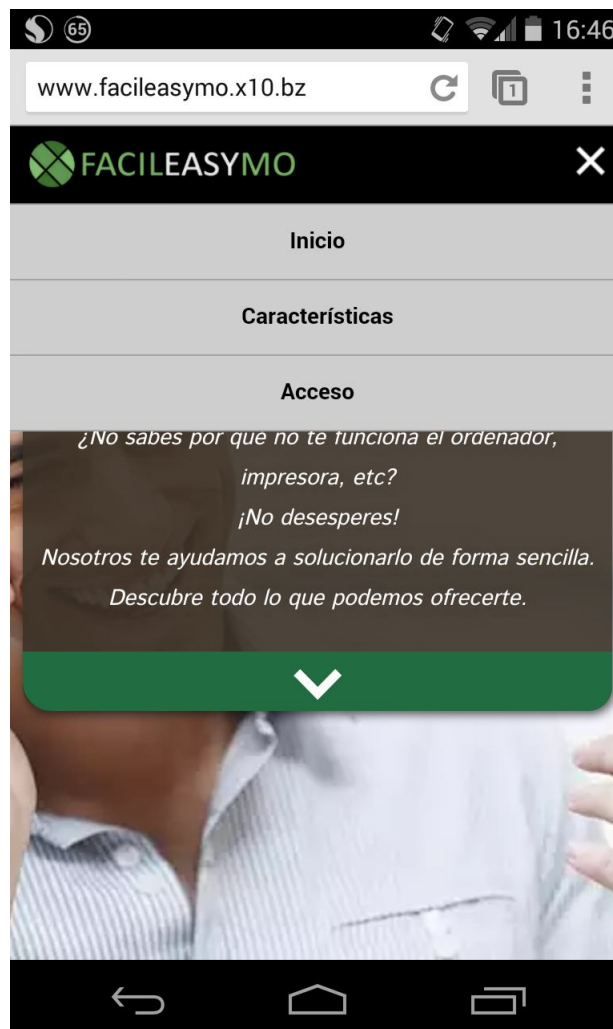


Figura 28. Vista principal de facileasymo.x10.bz desde dispositivo móvil

4.2.2 Registro y acceso

En el proceso de registro únicamente se pedirán los datos estrictamente necesarios para el uso de la herramienta, y que podrán ser modificados más adelante desde la zona personal del usuario. Para verificar la cuenta del usuario, se le envía un correo electrónico a la dirección introducida anteriormente, en el que se le pide que haga click en un enlace que sirve para verificar sus datos y su correo, y de este modo activar el acceso a ese usuario a la plataforma. En el caso de que no haya activado su cuenta y quiera acceder a la plataforma, **se le notificará en el momento del inicio de sesión con un mensaje de advertencia pidiendo que complete la verificación de su cuenta**. La figura 29 muestra la apariencia del lightbox para el registro de usuario.

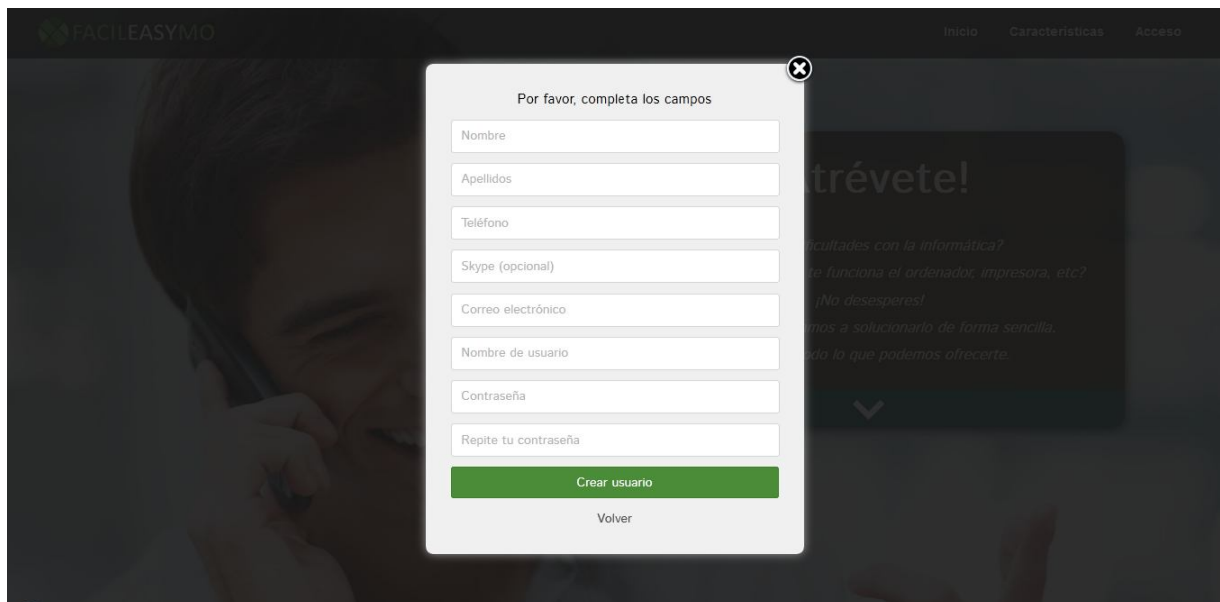


Figura 29. Ligthbox de registro usuario nuevo

La figura 30 es un ejemplo del correo de verificación de cuenta que podrá recibir el usuario. Como se explicaba anteriormente vendrá con el enlace preparado para activar la cuenta, además de recordar los datos de acceso.



Figura 30. Email de verificación de registro de usuario

Cuando el usuario pinche en el enlace se abrirá una ventana nueva en la que se le informará si se ha podido validar su usuario correctamente o por el contrario ha habido algún problema con el tratamiento de sus datos.

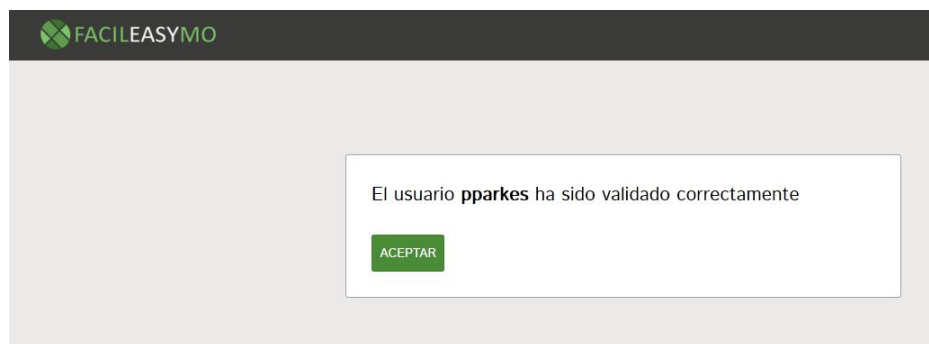


Figura 31. Email de verificación de registro de usuario

A partir de aquí el usuario se encuentra en condiciones de poder acceder a la plataforma web y de este modo ver todo lo el contenido de la misma.

En la figura 32 se muestra el lightbox de inicio de sesión. Se realizará una comprobación de si los datos de acceso introducidos por el usuario cumplen con los requisitos del sistema (longitud del nombre de usuario superior a cuatro caracteres, longitud de contraseña mayor de seis caracteres...). Estas comprobaciones se harán por JavaScript, es decir, antes de enviar los datos al servidor.

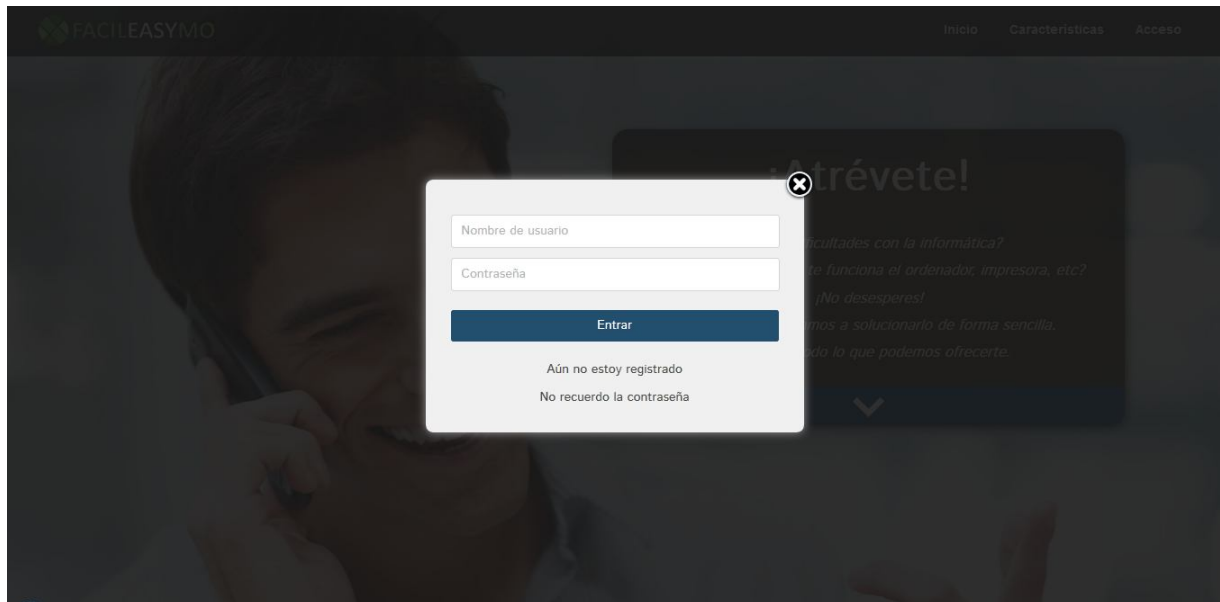


Figura 32. Pantalla de acceso para usuario ya registrado

4.2.3 Cambio de contraseña olvidada

El sistema también cuenta con la posibilidad de realizar un cambio de contraseña cuando el usuario no la recuerda. El procedimiento a seguir es, acceder al menú de acceso y hacer click en el enlace de "No recuerdo la contraseña". A continuación te solicita el nombre de usuario y el correo electrónico con el que esté vinculado este mismo. La figura 33 ilustra la apariencia del lightbox destinado a esta funcionalidad.

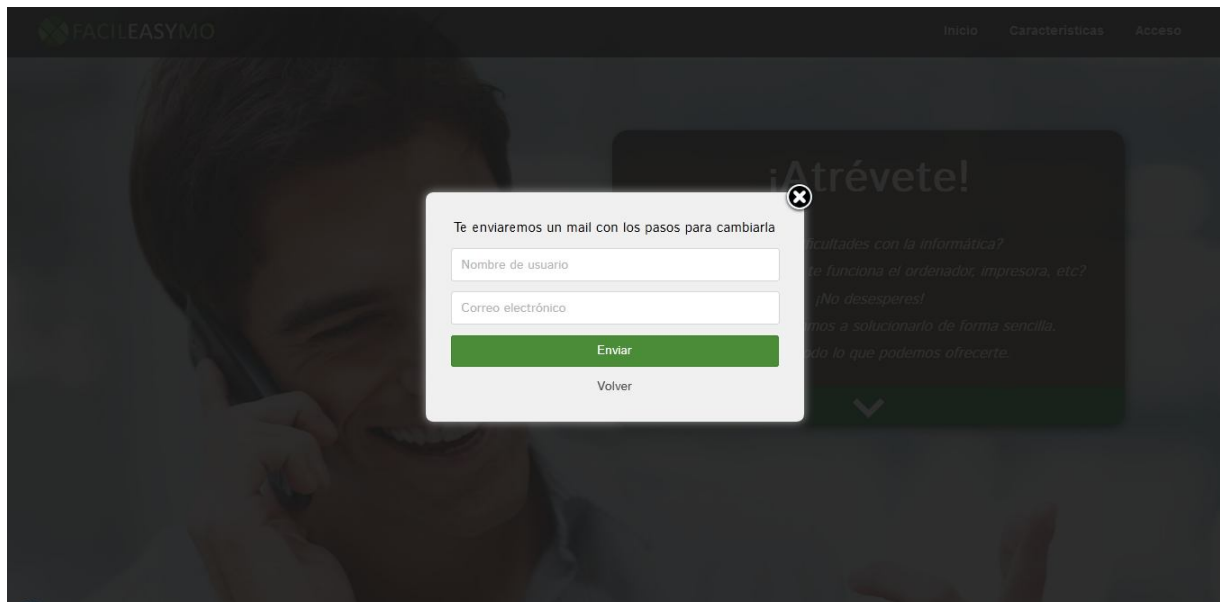


Figura 33. No recuerdo mi contraseña

Seguidamente, el usuario recibirá un correo electrónico con un enlace que redirigirá a la pantalla que le permitirá sustituir su contraseña anterior por una nueva. La figura 34 muestra un ejemplo de correo electrónico que recibirá el usuario.

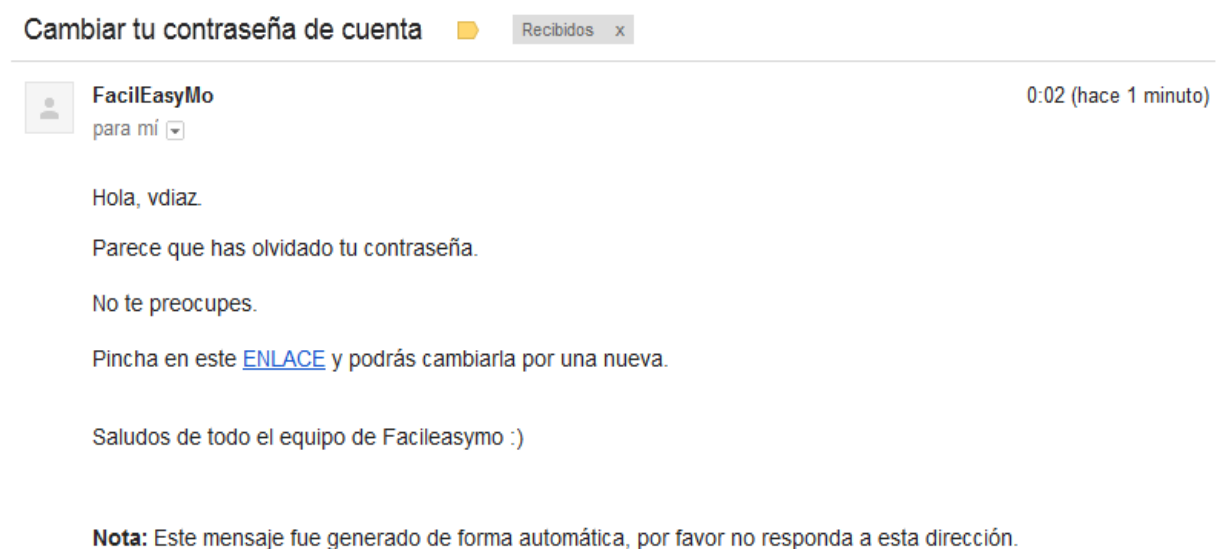


Figura 34. Ejemplo de email "No recuerdo mi contraseña"

Una vez cambiada la contraseña se informa con el siguiente mensaje (figura 35).

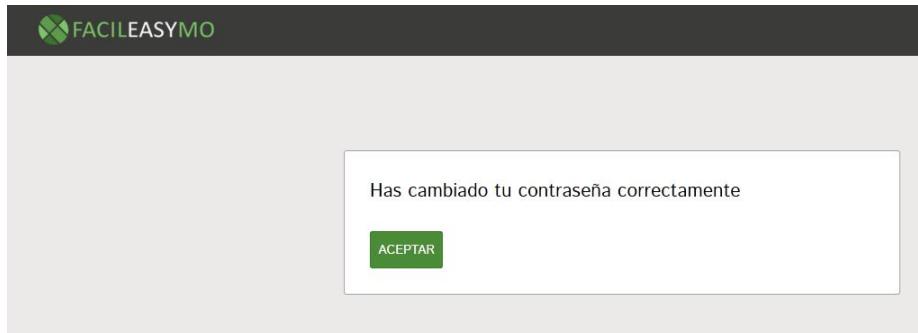


Figura 35. Cambio de contraseña correcto

4.2.4 Inicio

Una vez registrado y se haya iniciado sesión en la plataforma web, el usuario visualizará los distintos módulos de información.

El primer módulo, que será el que por defecto se muestre al acceder es la sección de inicio, mostrará los últimos artículos (figura 36).



Figura 36. Pantalla inicio navegador de escritorio

Si se está accediendo desde un dispositivo móvil la apariencia variará ligeramente para favorecer su visualización y usabilidad (figura 37). Para mostrar el menú se puede hacer click sobre el botón indicado para ello o se da la posibilidad de captar el gesto del usuario de desplazar el dedo hacia la derecha.

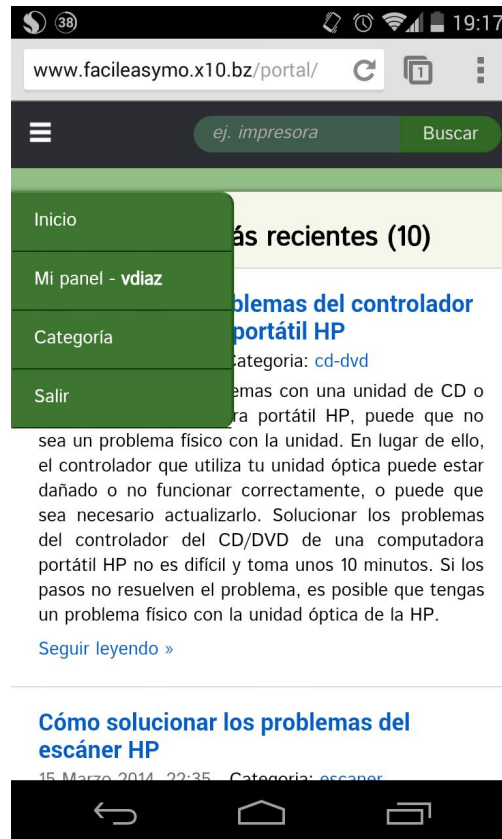


Figura 37. Pantalla inicio desde dispositivo móvil

4.2.5 Mi panel

El siguiente módulo (figuras 38 y 39) de la barra que permanece fija a la derecha de la pantalla es “Mi panel – nombre_de_usuario”, el cual nos será muy útil para, como se comentó anteriormente, modificar cualquiera de los datos personales facilitados, realizar consultas a nuestro historial de incidencias e incluso poder dar de baja la cuenta.

Fecha de llamada	Incidencia	Categoría	Manual
2014-07-12 11:41:52	no tengo acceso a internet	red	No disponible
2014-07-08 23:18:48	mi impresora esta atascada	impresora	Ver articulo
2014-07-08 07:52:40	mi impresora esta atascada	impresora	Ver articulo
2014-07-07 23:59:31	mi impresora esta atascada	impresora	Ver articulo
2014-07-05 14:06:49	impresora está atascada	impresora	Ver articulo

Figura 38. Historial de llamadas desde navegador PC

Fecha de llamada	Incidencia	Categoría	Manual
2014-07-12 11:41:52	no tengo acceso a internet	red	No disponible
2014-07-08 23:18:48	mi impresora esta atascada	impresora	Ver articulo

Figura 39. Historial de llamadas desde dispositivo móvil

Como se puede apreciar en la imagen del historial, se ofrece un enlace que nos lleva directamente a artículos relacionados con el dispositivo con el que tuvimos la incidencia, pudiendo servirle al usuario para obtener un mayor conocimiento y poder enfrentarse a posibles futuras incidencias con una mayor soltura.

Dentro del Panel de Control el usuario podrá acceder y realizar modificaciones de sus datos personales. En las figuras 40 y 41 vemos como el panel muestra al usuario su nombre, apellido, teléfono, datos de login (nombre de usuario y correo electrónico), como datos de registro (tanto la fecha de alta como la de último acceso).

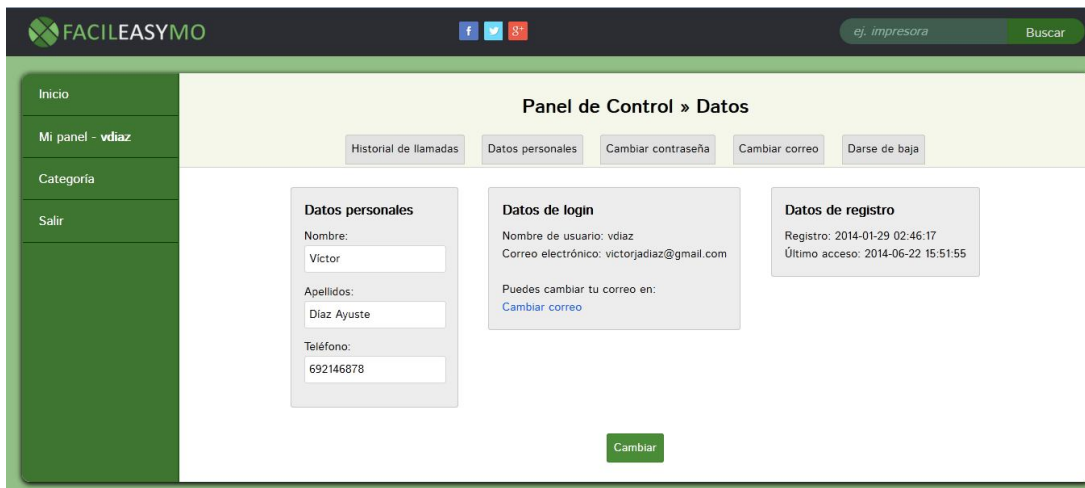


Figura 40. Datos personales desde navegador de escritorio



Figura 41. Datos personales desde dispositivo móvil

El usuario puede realizar otras modificaciones como el cambio de contraseña. Para realizar este cambio tendrá que introducir la contraseña actual y la contraseña nueva dos veces, verificando así que es correcta (figura 42).



The screenshot displays the 'Panel de Control » Contraseña' interface. On the left is a green sidebar with navigation options: 'Inicio', 'Mi panel - wdiaz', 'Categoría', and 'Salir'. The main content area has a title 'Panel de Control » Contraseña' and a sub-menu with 'Historial de llamadas', 'Datos personales', 'Cambiar contraseña', 'Cambiar correo', and 'Darse de baja'. The 'Cambiar contraseña' section contains three input fields: 'Contraseña actual:' (filled with 6 dots), 'Contraseña nueva:' (filled with 8 dots), and 'Repite la contraseña:' (filled with 8 dots). Below the fields is a green confirmation message 'Contraseña cambiada correctamente' and a green 'Cambiar' button. The top navigation bar includes the Facileasymo logo, social media icons, a search bar with the text 'ej. impresora', and a 'Buscar' button.

Figura 42. Cambio de contraseña

Las contraseñas en Facileasymo deben cumplir siempre una condición: la longitud debe ser superior a seis caracteres.

También podrá realizar un cambio de correo electrónico. Para ello, el usuario acudirá a la pantalla que se muestra en la figura 43 e introducirá el nuevo correo electrónico. El sistema le mandará un mensaje a la dirección especificada anteriormente con un código de cuatro caracteres (figura 44). Este código debe ser introducido en Facileasymo para que el cambio de correo sea válido.

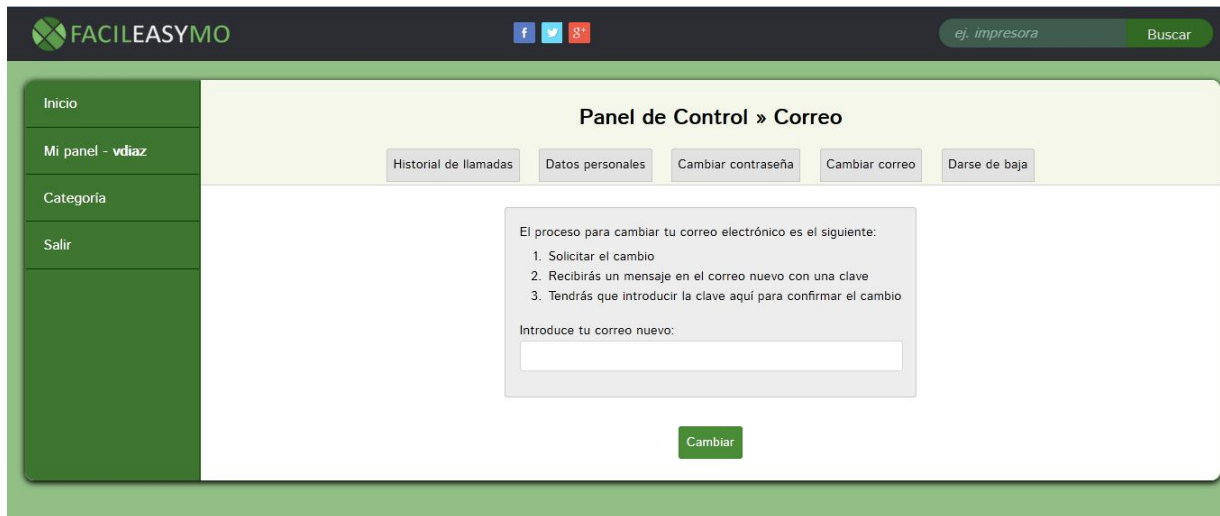


Figura 43. Cambio de correo electrónico



Figura 44. Email de confirmación de cambio correo electrónico

Por último existe la posibilidad de dar de baja el usuario (figura 45). En este caso se borrarán todos los datos salvo el nombre de usuario, IP, junto con otros datos de conexión, que se almacenarán por motivos de seguridad. Este proceso es transparente para el usuario final.

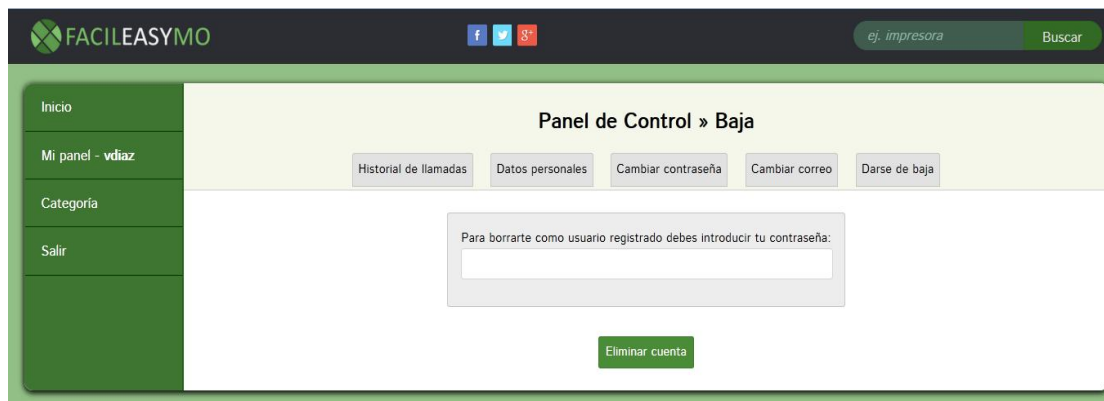


Figura 45. Darse de baja como usuario

4.2.6 Categoría

El siguiente y último módulo, "Categoría", contiene tutoriales para la resolución de incidencias e información para profundizar sobre los dispositivos, tanto de uso en el ámbito laboral como doméstico. Este contenido está catalogado por la tipología de dispositivos más comunes que tratamos en la plataforma (figuras 46 y 47).



Figura 46. Vista 'Categoría' desde navegador PC



Figura 47. Vista 'Categoria' desde dispositivo móvil

Para finalizar, y una vez realizadas todas las tareas oportunas, saldremos de la plataforma web y esta nos ofrecerá realizar una sencilla y rápida encuesta que nos ayudará a mejorar el sistema. En el caso de la plataforma de voz, una vez finalicemos la llamada, tendremos una encuesta relacionada principalmente con el sistema de diálogo.

5 Evaluación de la aplicación

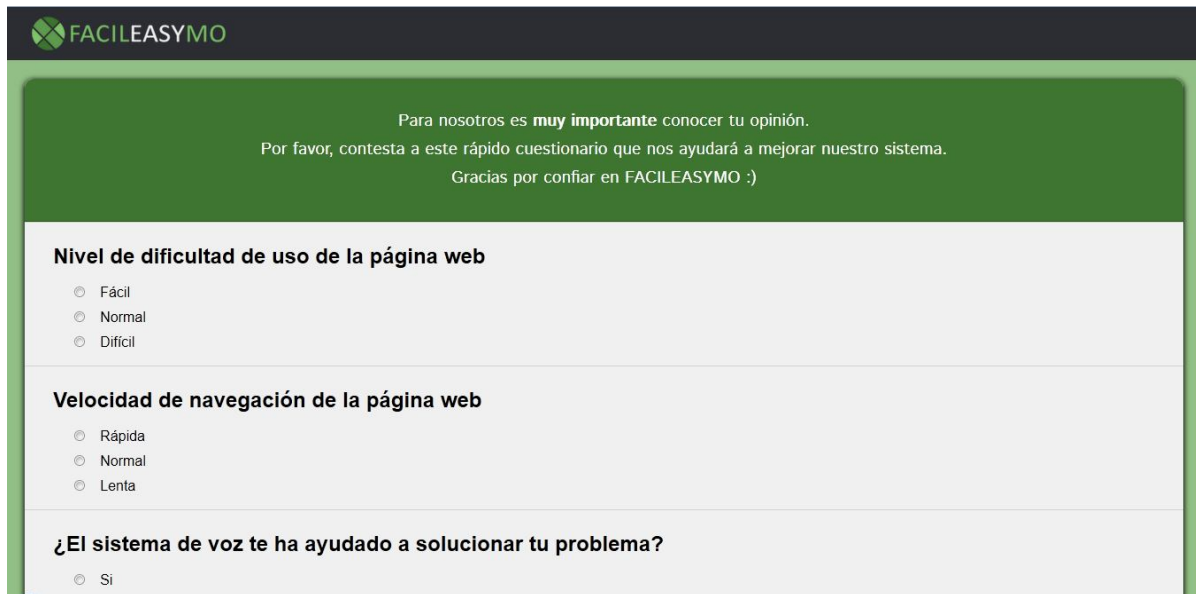
En este capítulo se describe la evaluación preliminar que se ha llevado a cabo del portal de voz y la plataforma web desarrollados para el Proyecto Final de Carrera. Para ello, se ha incluido una sección en la página web con un cuestionario que recoge la valoración subjetiva de los usuarios que han utilizado el portal de voz y la web. Los resultados se almacenan en una base de datos de la aplicación y posteriormente se han analizado para obtener, de esta forma, la completa evaluación de la aplicación (voz y web).

5.1 Metodología de evaluación

La evaluación, el estudio del rendimiento y el análisis de la usabilidad de los sistemas de diálogo son procedimientos necesarios para minimizar costes y optimizar resultados en las aplicaciones que hagan uso de ellos.

La evaluación de la presente aplicación se ha llevado a cabo a través de valoraciones de calidad. Para ello, se ha realizado un cuestionario 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.

Los aspectos que se han querido analizar son: el grado en el cual el usuario valora que es entendido por el sistema y entiende los mensajes del mismo, la velocidad percibida de la interacción, la presencia de errores, la 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 analizará los diferentes aspectos de la página web, como su grado de intuición, su rapidez al seleccionar las distintas secciones, como también el aspecto visual y los diferentes puntos incluidos en ella. Finalmente recopilamos información adicional de los usuarios sobre su grado de conocimiento acerca de nuevas tecnologías y uso de sistemas de diálogo para que sirva para hacer una idea del perfil de estos usuarios.



Para nosotros es **muy importante** conocer tu opinión.
Por favor, contesta a este rápido cuestionario que nos ayudará a mejorar nuestro sistema.
Gracias por confiar en FACILEASYMO :)

Nivel de dificultad de uso de la página web

- Fácil
- Normal
- Dificil

Velocidad de navegación de la página web

- Rápida
- Normal
- Lenta

¿El sistema de voz te ha ayudado a solucionar tu problema?

- Si

Figura 48. Encuesta desde navegador de escritorio

5.2 Resultados de la evaluación

El usuario después de realizar la encuesta, puede visualizar los resultados estadísticos de las votaciones de todos los usuarios que haya completado la encuesta. De este modo, puede ver la valoración general de Facileasymo. La figura 49 muestra los gráficos de la encuesta.

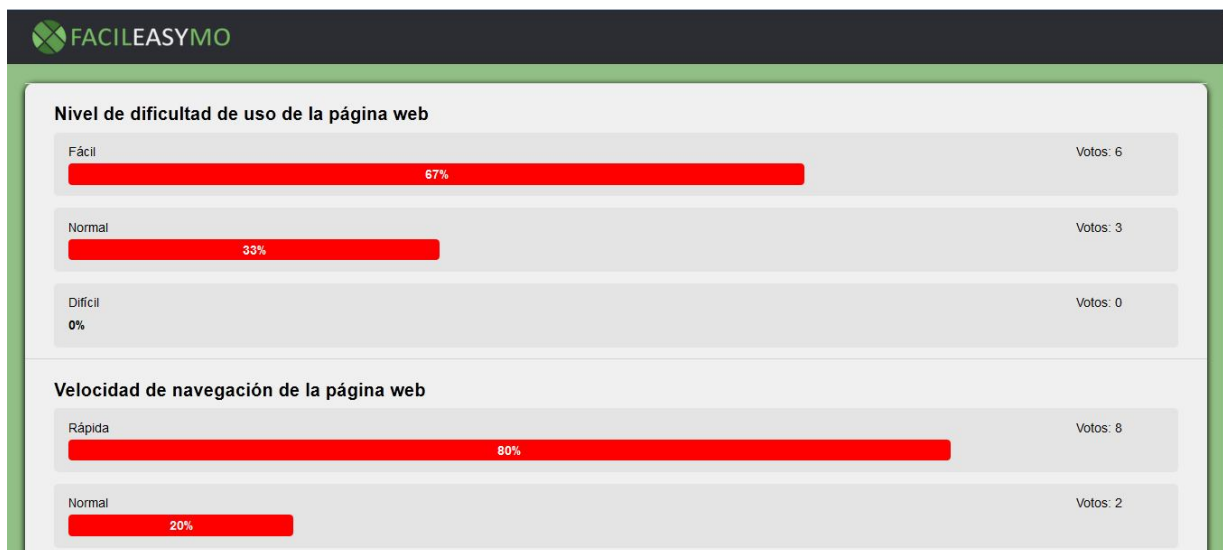


Figura 49. Resultados globales de la encuesta

En la tabla 14 se muestran los resultados de la valoración del sistema según la opinión de los usuarios.

Pregunta	Opción	Votos
Nivel de dificultad de uso de la página web	Fácil	6
	Normal	3
	Difícil	0
Velocidad de navegación de la página web	Rápida	8
	Normal	2
	Lenta	0
¿Los artículos y manuales disponibles te han ayudado a resolver tus dudas?	Si	10
	No	0
Si accedes desde un dispositivo móvil (tablet, smartphone, tv), ¿qué te parece la navegación?	Buena	6
	Normal	3
	Mala	0
El nivel de categorías de incidencias de la página web te parece...	Suficiente	5
	Normal	4
	Escaso	1
¿El sistema de voz te ha ayudado a solucionar tu problema?	Si	17
	No	2
Establece el nivel de dificultad del sistema de voz	Fácil	12
	Normal	8
	Difícil	0
¿Qué tal te ha entendido el sistema de voz?	Bien	14
	Regular	6
	Mal	0
La interacción con el sistema de voz fue...	Rápida	12
	Adecuada	7
	Lenta	1
En términos generales, ¿estás satisfecho con el sistema?	Si, bastante	9
	Satisfecho	1
	No, nada	0

Tabla 14 – Valoraciones de los usuarios

De los datos obtenidos se puede sacar como conclusión; como **puntos fuertes** en la plataforma web tenemos la **velocidad de navegación** y la **calidad de los artículos y manuales**. Sin embargo como punto a mejorar puede ser el nivel de categorías de la web. Respecto al sistema de voz, los usuarios han valorado positivamente el **reconocimiento de voz** y que les ha servido para **solucionar su incidencia**.

6 Conclusiones y trabajo futuro

En este capítulo se hace un balance del trabajo realizado en el presente Proyecto Final de Carrera. Partiendo de la revisión de los resultados logrados, se comprueba que los objetivos fijados inicialmente han sido cumplidos y se exponen las conclusiones obtenidas. Para finalizar, se presentan los posibles trabajos futuros que se podrían desarrollar sobre la aplicación, de forma que aumentara su eficiencia y el número de funcionalidades.

6.1 Conclusiones

En este Proyecto Final de Carrera se ha desarrollado una aplicación basada en el estándar VoiceXML, apoyada por una plataforma web. Se trata de una centralita automática de resolución de incidencias informáticas, en la cual los usuarios pueden resolver distintos problemas con los que a diario nos encontramos tanto en el mundo profesional como en el resto de la jornada cotidiana.

El usuario mediante una simple llamada telefónica puede identificar de manera fácil y rápida el motivo de la incidencia y recibir sencillos consejos para poder solucionar la misma.

La plataforma web también permite al usuario en cualquier momento acceder a consultar su historial de incidencias, y acceder a distintos manuales de diferentes dispositivos electrónicos, de manera que no solo consiga solucionar sus incidencias informáticas si no adquirir mayores conocimientos de distintos equipos. También se ofrece realizar una rápida encuesta a través de la web, consiguiendo así una comunicación constante con el usuario y permitiéndonos mejorar y actualizarnos día a día.

Los servicios ofrecidos por el portal de voz Facileasymo se reparten en módulos por los cuales se conduce al usuario en función de las decisiones que vaya tomando en cada diálogo.

Como conclusión final, los objetivos planteados en un principio se han cumplido. Con la realización de este Proyecto Final de Carrera se ha llevado a cabo un estudio completo de los Sistemas de Diálogo, y del lenguaje estándar VoiceXML para poder así aplicar todas las posibilidades de las interfaces orales y las ventajas de VoiceXML en el desarrollo de la nueva aplicación, e implementar las funcionalidades de un portal de voz con el fin de lograr una comunicación lo más efectiva y natural

posible. También se ha conseguido implementar una plataforma web, dinámica, de fácil manejo y sobre todo intuitiva de manera que cualquier usuario pueda acceder a cada una de las distintas opciones que se encontrará y poder consultar las distintas soluciones ofrecidas a los problemas más comunes.

Mediante el portal de voz se ha elaborado un sistema de acceso a la información y obtención de unos conocimientos suficientes para solucionar distintas incidencias informáticas por parte de los usuarios sin la necesidad de disponer de un navegador, 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 posible acceso a Internet. Estos entornos podrían ser un automóvil, un lugar al aire libre, etc.

Otro de los objetivos más importantes cumplidos es el de facilitar la accesibilidad a los servicios proporcionados por esta aplicación a usuarios que desconocen el uso de un ordenador o de Internet, así como a personas con discapacidades motoras o visuales. Este portal de voz puede promover un mayor conocimiento, acercamiento e integración de estos colectivos dentro del universo de las nuevas tecnologías. También para aquellos que ya se encuentran más acostumbrados al uso de las tecnologías móviles, podrán acceder sin problemas desde cualquier dispositivo móvil como tablets o smartphones al portal web *facilesasymo*, ya que debido a su responsive design se adapta a cualquier tamaño de pantalla facilitando su uso y visibilidad.

Finalmente, uno de los objetivos logrados y que más destaca, es el de la unión conseguida entre la página web y la plataforma de voz, consiguiendo con esto una atención más personalizada y haciendo sentir al usuario más cómodo y más atraído a utilizarla. De esta manera, se obtiene una actualización constante de los datos almacenados del usuario, permitiendo llevar un historial de las distintas incidencias con las que se ha encontrado en los últimos meses.

6.2 Trabajo futuro

Tras la finalización de este proyecto, se detallan a continuación los puntos adicionales sobre los que se podría trabajar para obtener mayores prestaciones en la aplicación desarrollada.

En primer lugar consideramos futuras mejoras como ampliar el número de incidencias contempladas, y la cantidad de información para solucionar las mismas, de manera que se pudieran recoger numerosos dispositivos electrónicos, de los cuales cada vez dependemos más en nuestro día a día. Otro punto a tener en cuenta sería el de poder obtener más información del usuario, como aficiones o intereses para poder realizar sugerencias más personales, es decir, poder incluir mayor información relacionada con el mundo tecnológico, de esta manera a través del portal podríamos hacerle sugerencias personalizadas acorde a sus intereses y gustos.

Por otra parte, se podrían introducir una serie de funcionalidades globales para el sistema que permitirían conseguir una mayor globalización de la plataforma y que ampliaría muchísimo su repercusión y uso, como podría ser ofrecer al usuario la posibilidad de seleccionar el idioma a la hora de realizar la llamada y también de visualizar el portal. De esta manera la plataforma podría extenderse a múltiples países y dar cobertura a un número superior de usuarios, y en nuestro propio país poder acercarse a la población extranjera.

El requisito fundamental para añadir este servicio sería que la aplicación estuviera alojada en una plataforma que proporcionara un reconocedor de voz y un sintetizador de texto en voz que soportara los idiomas añadidos. En este caso, Voxeo con su Prophecy 9 Multi-Language VXML permite la utilización de los principales idiomas. En cuanto a la implementación, en el formulario inicial se daría la opción al usuario de elegir el idioma. A partir de este punto se gestionaría el resto de la arquitectura de la aplicación según el idioma elegido.

Por último, y debido a la cada vez mayor frecuencia en la que la gente accede a estudios a distancia, y consigue formarse de manera telemática, se podrían incluir una especie de tutorías a través del portal, en las cuales a determinados horarios y días pudiera ofrecerse una sesión con un técnico especialista, que pudiera contestar al usuario a cualquier tipo de pregunta relacionada con las incidencias informáticas o tecnológicas, pudiendo este remitirle en ese mismo momento a páginas web o manuales concretos en los cuales pudiera encontrar un aporte de información que le permitiera conocer aún más su dispositivo electrónico, o incluso poder asesorarle a la hora de adquirir un nuevo equipo.

Planificación

Esta es la planificación inicial de todas las fases principales de las que consta el Proyecto Fin de Carrera. En la tabla 15, se muestra el inicio del proyecto desde los primeros pasos de planificación y análisis, hasta el final del desarrollo de la memoria y su posterior presentación.

Nombre de tarea	Duración	Comienzo	Fin
Plan de Proyecto	236 días	jue 22/08/13	jue 17/07/14
Análisis	40 días	mar 17/09/13	lun 11/11/13
Estudiar arquitecturas	20 días	mar 17/09/13	lun 14/10/13
Estudio tecnología VoiceXML	13 días	mar 17/09/13	jue 03/10/13
Estudio tecnologías Web	7 días	jue 03/10/13	vie 11/10/13
Análisis de requisitos	10 días	lun 14/10/13	vie 25/10/13
Elegir motor de Base de Datos	5 días	vie 25/10/13	jue 31/10/13
Estudio lenguajes de programación	5 días	jue 31/10/13	mié 06/11/13
Reunión seguimiento análisis	2 días	mié 06/11/13	jue 07/11/13
Fase de diseño	30 días	jue 07/11/13	mié 18/12/13
Diseño Web	15 días	jue 07/11/13	mié 27/11/13
Diseño estructura de voz	13 días	mié 27/11/13	vie 13/12/13
Reunión seguimiento fase de diseño	2 días	vie 13/12/13	lun 16/12/13
Implementación	65 días	lun 16/12/13	vie 14/03/14
Desarrollo plataforma web	40 días	lun 16/12/13	vie 07/02/14
Desarrollo aplicación de voz	40 días	lun 16/12/13	vie 07/02/14
Sincronización mediante base de datos	21 días	vie 07/02/14	vie 07/03/14
Reunión seguimiento implementación	4 días	vie 07/03/14	mié 12/03/14
Pruebas	37 días	mié 12/03/14	jue 01/05/14
Pruebas comportamiento Plataforma Web	15 días	jue 01/05/14	mié 21/05/14
Responsive design: adaptación a distintos dispositivos	5 días	mié 21/05/14	mar 27/05/14
Llamadas de prueba a centralita telefónica	10 días	mar 27/05/14	lun 09/06/14
Resolución diferentes incidencias por teléfono	5 días	lun 09/06/14	vie 13/06/14
Encuestas por teléfono y web	2 días	vie 13/06/14	lun 16/06/14
Documentación	220 días	vie 06/09/13	jue 10/07/14
Documentación y memoria del proyecto	220 días	vie 06/09/13	jue 10/07/14
Presentación del proyecto	10 días	dom 06/07/14	jue 17/07/14

Tabla 15 – Tareas de las fases del proyecto

En las figuras 50, 51 y 52, se puede visualizar el diagrama de Gantt de las diferentes etapas del proyecto. Muchas de estas tareas al ser dos personas se han podido realizar en paralelo y así avanzar más rápido, teniendo reuniones periódicas de seguimiento que han permitido a ambos conocer el estado del proyecto en cada momento. Por ello el número de días de la planificación es menor que el utilizado para calcular los costes totales del proyecto.

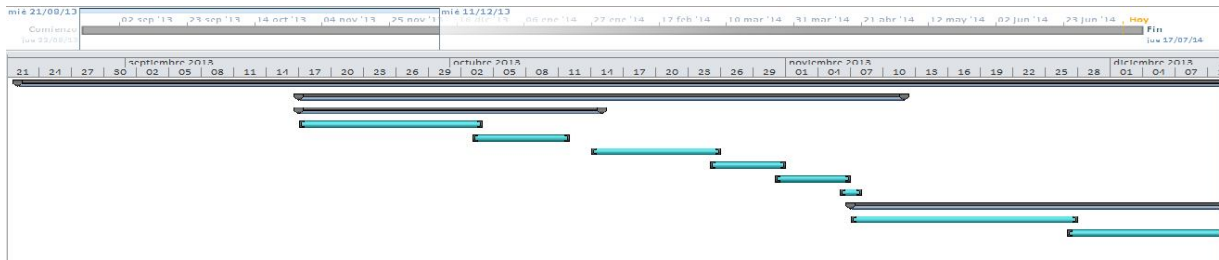


Figura 50. Diagrama de Gantt del 22/08/2013 al 07/12/2013

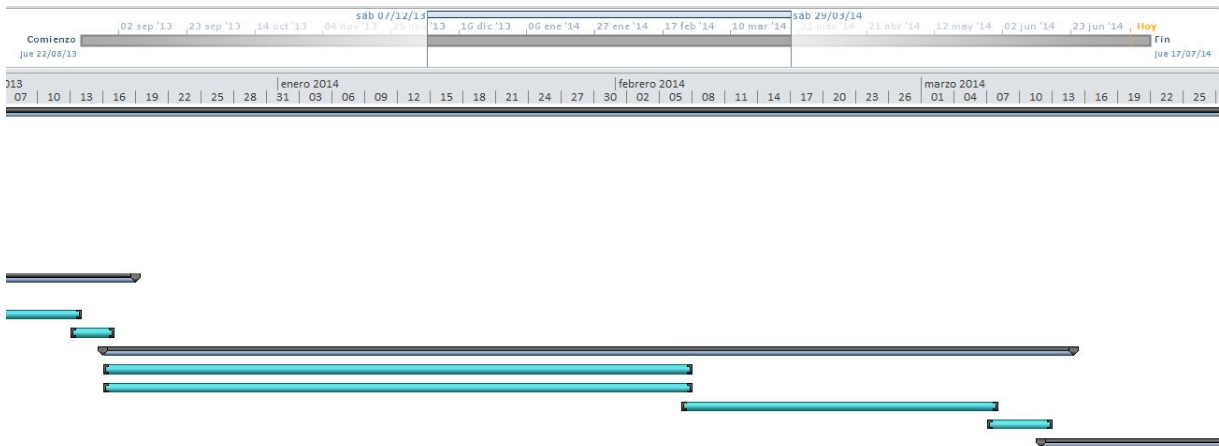


Figura 51. Diagrama de Gantt del 07/12/2013 al 25/03/2014

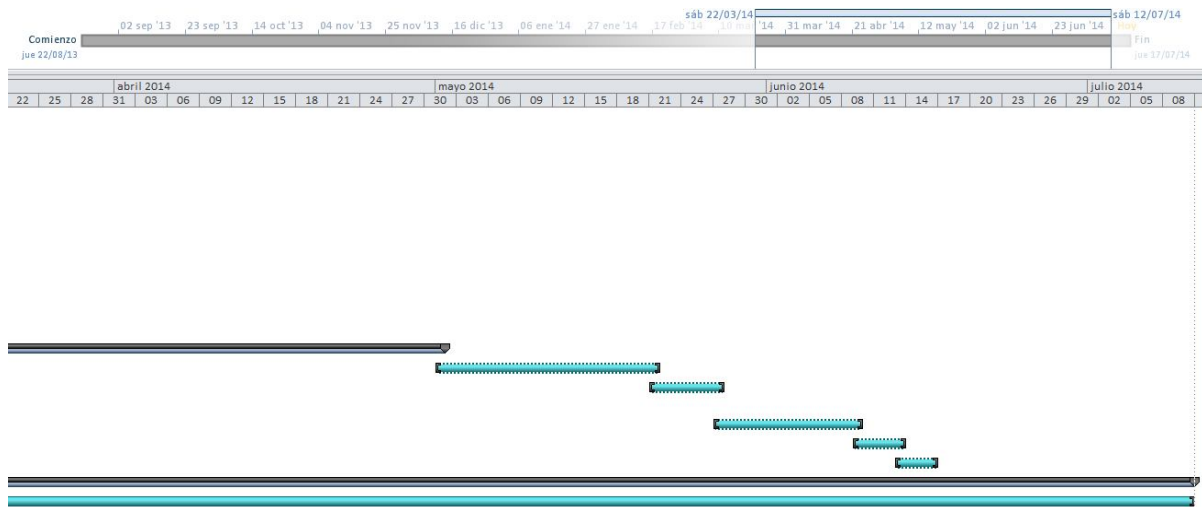


Figura 52. Diagrama de Gantt del 25/03/2014 al 17/07/2014

Presupuesto

En esta sección se presenta el presupuesto del proyecto. En él se contempla la duración de las distintas fases y tareas, y se incluye un desglose de costes de personal, costes de material y costes totales.

1. Tareas

Fase 1: Planificación.

- Estudio de los sistemas de diálogo.
 - Duración: 30 días.
- Planificación y análisis de requisitos del portal de voz.
 - Duración: 21 días.
- Estudio de las plataformas web.
 - Duración: 35 días.
- Planificación y análisis de la plataforma Web.
 - Duración: 20 días.
- Estudio de las tecnologías necesarias.
 - Duración: 30 días.

Fase 2: Desarrollo.

- Análisis y diseño inicial.
 - Duración: 19 días.
- Implementación del sistema.
 - Duración: 145 días.
- Pruebas unitarias.
 - Duración: 35 días.

Fase 3: Documentación.

- Memoria del Proyecto Final de Carrera.
 - Duración: 49 días.
- Preparación de la presentación.
 - Duración: 15 días.

2. Recursos

Para la realización del presente proyecto se han utilizado los siguientes recursos:

- Recursos software:
 - Plataforma Voxeo: 0 €
 - Skype: 0 €
 - Licencia Microsoft Office 2010: 350 €
 - Licencia Microsoft Office Project 2010: 280 €
 - Editor de programación Sublime Text: 0 €
 - Editor de programación Notepad ++: 0 €
 - Banco de imágenes de Shutterstock: 39 €

- Recursos hardware:
 - Ordenador portátil: 900 €
 - Auriculares con micrófono: 100€
 - Servidor X10Hosting: 0€

- Recursos humanos:
 - En la realización de este proyecto han participado dos personas, dos desarrolladores, uno de los cuales asume las funciones de director de Proyecto.
 - Las horas dedicadas al proyecto han sido de una media de 2 horas en días laborales y hasta 8 en festivos y fines de semana, por lo que la media diaria sería de 4 horas.
 - Coste de un Ingeniero: 30 €/hora.

3. Resumen de costes

El coste de recursos humanos de la aplicación se resume en la siguiente tabla 16.

TAREA	DIAS	IMPORTE
FASE 1	136	16320
FASE 2	199	23880
FASE 3	64	0
Subtotal	399	47880

Tabla 16 – Detalle de Costes de Recursos Humanos del Proyecto

El coste total del sistema viene detallado en la tabla 17.

CONCEPTO	IMPORTE
Recursos humanos	47880
Recursos software	23880
Recursos hardware	1000
Subtotal	72760
(21% IVA)	15279
TOTAL	88039

Tabla 17 – Detalle de Coste Total del Proyecto

El presupuesto total de este proyecto asciende a la cantidad de **OCHENTA Y OCHO MIL TREINTA Y NUEVE EUROS**.

Madrid a 11 de Julio de 2014

Fdo. Ricardo Bárzano Santamaría

Fdo. Víctor Javier Díaz Ayuste

Plan de lanzamiento

Desde el punto de vista del marketing definimos el ciclo de relación de la marca con el consumidor potencial para fijar, de esta forma, los objetivos principales.

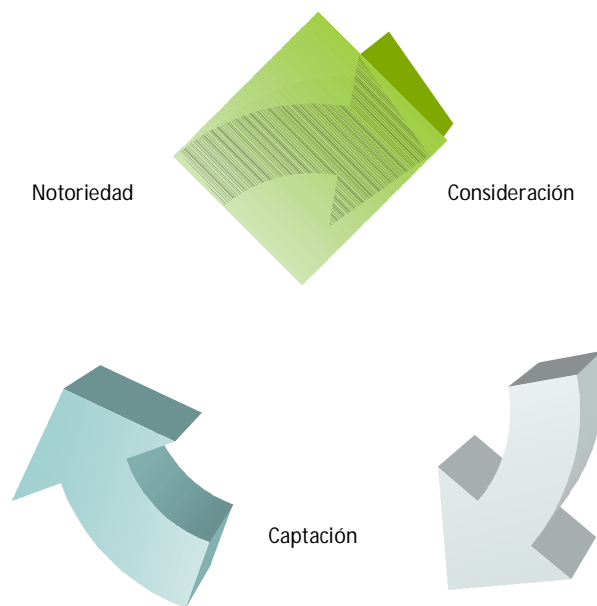


Figura 53. Diagrama de ciclo de relación marca - consumidor potencial

- **Notoriedad:** que el usuario nos conozca, que encuentre nuestro servicio. Esto nos da la oportunidad de poner en valor el producto creado de cara al registro del usuario. En el momento de lanzamiento debemos centrarnos en este punto ya que si no adquirimos suficiente notoriedad para los usuarios, no podremos trabajar el resto de objetivos.
- **Consideración:** que el usuario nos tenga en cuenta. Para que este objetivo sea exitoso, debemos tener muy clara la propuesta de valor de nuestro servicio y hacérselo llegar.

- **Conversión:** que el usuario se registre. Este objetivo es relevante de cara a un futuro modelo de monetización del proyecto (bien sea mediante banners en la web o acciones conjuntas con otras marcas del sector).

Para cada uno de los objetivos definimos una serie de acciones que nos ayuden en su consecución.

Notoriedad - Plan de captación

De cara a conseguir nuevos usuarios planteamos las siguientes acciones:

- **Mejora del posicionamiento en buscadores:** desde el punto de vista del SEO realizaremos una consultoría previa de la web y una estrategia de contenido. Con la estrategia del contenido buscamos generar artículos en un blog que encajen con el tipo de búsquedas que realizan nuestros clientes potenciales como mejoras y cuidado en sus terminales o soluciones de problemas básicos.
- **Acuerdos con empresas/ Marketing directo:** buscar empresas que puedan ofrecer a sus empleados, iniciados en el uso de terminales electrónicos pero acostumbrados a tener un sistema de soporte en sus empresas, nuestro servicio.

Consideración - Propuesta de valor

Como hemos dicho anteriormente, hacer llegar la propuesta de valor a los usuarios potenciales, será clave de cara a la conversión.

- **Mostrar los servicios de la empresa en la web:** ser claros en la definición para que los usuarios potenciales entiendan el sistema. Para esta fase crearemos materiales y vídeos con formato *screencast* o recorrido virtual por la plataforma desde los casos de uso más interesantes.
- **Generar contenidos propios:** posicionarnos como expertos a la hora de detectar y solucionar las incidencias que un usuario medio pueda tener. Si somos reconocidos de esta forma, conseguiremos que nuestra reputación afecte en la conversión.

Conversión – Modelo de informe

Las acciones que nos harán mejorar en la conversión se basan en el análisis de nuestra web y del servicio.

- **Análisis UX:** definir un *sistema de reporting* de analíticas web donde podamos ver qué pasa en la web a nivel de usabilidad.
- **Análisis de las encuestas:** realizar un análisis del sistema de encuestas de cara a la mejora de los resultados.

Glosario

- **CGI** (Common Gateway Interface): es una importante tecnología de la World Wide Web que permite a un cliente (navegador web) solicitar datos de un programa ejecutado en un servidor web. CGI especifica un estándar para transferir datos entre el cliente y el programa. Es un mecanismo de comunicación entre el servidor web y una aplicación externa cuyo resultado final de la ejecución son objetos MIME. Las aplicaciones que se ejecutan en el servidor reciben el nombre de CGIs.
- **CALLXML**: es un lenguaje basado en etiquetas XML creado por Voxeo Community, que permite implementar aplicaciones telefónicas basadas en voz convencional o voz sobre IP (VoIP) que interactúan con servidores CallXML, los cuales usan el código de la aplicación para controlar las llamadas. CallXML ha sido diseñado para facilitar la implementación de aplicaciones web que puedan interactuar con cualquier tipo de llamada telefónica, soportando la iniciación de llamadas salientes, conferencias, interacción entre múltiples llamadas, funciones encuéntreme-sígame, etc.
- **CCXML** (Call Control Extensible Language): es un lenguaje de marcación basado en XML diseñado para añadir capacidades de control de llamada a sistemas de telefonía de forma estándar. CCXML ha sido diseñado para implementar aplicaciones telefónicas complejas, con posibilidad de lanzar nuevas aplicaciones IVR (Interactive Voice Response) en llamadas subsiguientes. CCXML está impulsado por el W3C.
- **DTMF** (Dual-Tone Multi-Frequency signaling): es un sistema de marcación por tonos, también llamado sistema multi-frecuencial. En este sistema cuando el usuario pulsa en el teclado de su teléfono la tecla correspondiente al dígito que quiere marcar, se envían dos tonos, de distinta frecuencia, uno por columna y otro por fila en la que esté la tecla, que la central descodifica a través de filtros especiales, detectando instantáneamente que dígito se marcó.
- **ECMAScript**: es una especificación de lenguaje de programación publicada por ECMA International. El desarrollo empezó en 1996 y estuvo basado en el popular lenguaje JavaScript propuesto como estándar por Netscape Communications Corporation. Actualmente está aceptado como el estándar ISO 16262. ECMAScript define un lenguaje de tipos dinámicos ligeramente inspirado en Java y otros lenguajes del estilo de C. Soporta algunas características de la programación orientada a objetos

mediante objetos basados en prototipos y pseudoclasas. La mayoría de navegadores de Internet incluyen una implementación del estándar ECMAScript.

- **IMAP** (Internet Message Access Protocol): es un protocolo de red de acceso a mensajes electrónicos almacenados en un servidor. Mediante IMAP se puede tener acceso al correo electrónico desde cualquier equipo que tenga una conexión a Internet. IMAP tiene varias ventajas sobre POP, que es el otro protocolo empleado para obtener correo desde un servidor. Por ejemplo, es posible especificar en IMAP carpetas del lado servidor. Por otro lado, es más complejo que POP ya que permite visualizar los mensajes de manera remota y no descargando los mensajes como lo hace POP.
- **IVR** (Interactive Voice Response): es un sistema telefónico que capaz de recibir una llamada e interactuar con el humano a través de grabaciones de voz y el reconocimiento de respuestas simples, como "sí", "no" u otras. Se trata de un sistema automatizado de respuesta interactiva, orientado a entregar y/o capturar información a través del teléfono, permitiendo el acceso a servicios de información u otras operaciones.
- **MySQL**: es un sistema de gestión de bases de datos relacionales multihilo, multiusuario y robusto. El software MySQL tiene licencia dual, pudiéndose usar de forma gratuita bajo licencia GNU o bien adquiriendo licencias comerciales de MySQL AB en el caso de no desear estar sujeto a los términos de la licencia GPL. MySQL es una marca registrada de MySQL AB.
- **PHP** (Hypertext Pre-processor): es un lenguaje de programación interpretado, diseñado originalmente para la creación de páginas web dinámicas. Es usado principalmente para la interpretación del lado del servidor (server-side scripting) pero actualmente puede ser utilizado desde una interfaz de línea de comandos o en la creación de otros tipos de programas incluyendo aplicaciones con interfaz gráfica.
- **POP3** (Post Office Protocol): En informática se utiliza el POP3 en clientes locales de correo para obtener los mensajes de correo electrónico almacenados en un servidor remoto. Es un protocolo de nivel de aplicación en el Modelo OSI.
- **PTSN** (Public Switched Telephone Network): es una red con conmutación de circuitos tradicional optimizada para comunicaciones de voz en tiempo real. Cuando se llama a alguien, se

cierra un conmutador al marcar y se establece así un circuito con el receptor de la llamada. PSTN garantiza la calidad del servicio (QoS) al dedicar el circuito a la llamada hasta que se cuelga el teléfono.

- **Script:** es un programa usualmente simple, que por lo regular se almacena en un archivo de texto plano. Los script son casi siempre interpretados, pero no todo programa interpretado es considerado un script. El uso habitual de los scripts es realizar diversas tareas como combinar componentes, interactuar con el sistema operativo o con el usuario.
- **Sendmail:** es un popular "agente de transporte de correo" (MTA - Mail Transport Agent) en Internet, cuya tarea consiste en "encaminar" los mensajes de forma que estos lleguen a su destino. Se afirma que es el más popular MTA, compatible con sistemas Unix y el responsable de la mayoría de envíos del correo de Internet, aunque se critica su alto número de alertas de seguridad además de no ser sencillo de configurar.
- **SIP** (Session Initiation Protocol): es un protocolo desarrollado por el grupo de trabajo MMUSIC del IETF con la intención de ser el estándar para la iniciación, modificación y finalización de sesiones interactivas de usuario donde intervienen elementos multimedia como el video, voz, mensajería instantánea, juegos en línea y realidad virtual.
- **Sistema de diálogo:** son programas informáticos cuya finalidad es interactuar con los usuarios oralmente o de forma multimodal para proporcionar diversos servicios.
- **URI** (Uniform Resource Identifier): es una cadena corta de caracteres que identifica inequívocamente un recurso (servicio, página, documento, dirección de correo electrónico, enciclopedia, etc.). Normalmente estos recursos son accesibles en una red o sistema. Algunos URI pueden ser URL, URN o ambos.
- **VoiceXML** (Voice eXtensible Markup Language): es un formato XML estándar del W3C para la especificación de diálogos interactivos entre un humano y una máquina.
- **Voxeo:** es una plataforma de implementación del estándar VoiceXML que cumple por completo con las especificaciones del W3C para VoiceXML 2.0 y 2.1.

- **W3C** (Consortio World Wide Web): es una asociación internacional formada por organizaciones miembros del consorcio, personal y el público en general, que trabajan conjuntamente para desarrollar estándares web. W3C pretende guiar la web hacia su máximo potencial a través del desarrollo de protocolos y pautas que aseguren el crecimiento futuro de la web.
- **XML** (eXtensible Markup Language): es un metalenguaje extensible de etiquetas desarrollado por el World Wide Web Consortium (W3C). Es una simplificación y adaptación del SGML y permite definir la gramática de lenguajes específicos (de la misma manera que HTML es a su vez un lenguaje definido por SGML). Por lo tanto XML no es realmente un lenguaje en particular, sino una manera de definir lenguajes para diferentes necesidades. Algunos de estos lenguajes que usan XML para su definición son XHTML, SVG, MathML.
- **RWD** (Responsive Web Design): es una filosofía de diseño y desarrollo cuyo objetivo es adaptar la apariencia de las páginas web al dispositivo que se esté utilizando para visualizarla. Esta tecnología pretende que con un sólo diseño web tengamos una visualización adecuada en cualquier dispositivo.
- **CC** (Creative Commons): es una organización sin ánimo de lucro, cuya oficina central está ubicada en la ciudad de Mountain View, en el estado de California, en los Estados Unidos de América, que permite usar y compartir tanto la creatividad como el conocimiento a través de una serie de instrumentos jurídicos de carácter gratuito.

Bibliografía

A continuación se listan las diferentes fuentes de información a las que se ha consultado durante la elaboración de este Proyecto Fin de Carrera.

^[1] **Voxeo**

<http://evolution.voxeo.com/> - 05 de enero de 2014.

^[2] **HTML**

<http://www.w3.org/TR/1999/REC-html401-19991224> - 05 de enero de 2014.

^[3] **Multimodal**

http://www.ugr.es/~rlopezc/sistemas_dialogo.htm - 05 de enero de 2014.

^[4] **Sistemas de dialogo**

Griol, David. Desarrollo y evaluación de Diferentes Metodologías para la Gestión Automática del Diálogo. Tesis Doctoral. Universidad Politécnica de Valencia, 2007.

^[5] **Aplicaciones**

Pérez, Guillermo. Contribuciones a los sistemas del diálogo multimodales. Universidad Politécnica de Madrid, 2009.

^[6] **Atos**

<http://www.es.atosorigin.com/es-es/servicios/soluciones/default.htm> - 13 de febrero de 2014.

^[7] **Larri**

Callejas, Zoraida. Desarrollo de Sistemas de Diálogo Oral Adaptativos y Portables: Reconocimiento de Emociones, Adaptación al idioma y Evaluación de Campo. Tesis doctoral. Universidad de Granada, 2008.

^[8] **Dihana**

David Griol, Lluís F. Hurtado, Encarna Segarra, Emilio Sanchis: A statistical approach to spoken dialog systems design and evaluation. Speech Communication 50(8-9): 666-682 (2008).

[9] Let's Go

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).

[10] Conquest

Dan Bohus, Sergio Grau Puerto, David Huggins-Daines, Venkatesh Keri, Gopala Krishna, Rohit Kumar, Antoine Raux, Stefanie Tomko: Conquest – an Open-Source Dialog System for Conferences.

[11] DARPA Communicator

Ing. Stanislav Ondáš, supervisor: doc. Ing. Jozef Juhár, CSc: VOICEXML-BASED SPOKEN LANGUAGE INTERACTIVE SYSTEM.

[12] Mercury

<http://groups.csail.mit.edu/sls/research/mercury.shtml> - 25 de enero de 2014.

[13] Cineentradas

<http://www.ydilo.com/es/> - 13 de febrero de 2014.

[14] CTT-Bank

<http://www.ctt.pt/fectt/wcmservlet/ctt/en/landingpage.html> - 13 de febrero de 2014.

[15] MIMUS

Pérez, G., Amores, G., Manchón, P., 2006. A multimodal architecture for home control by disabled users. In: Proc. of IEE/ACL Workshop on Spoken Language Technology (SLT). Palm Beach, Aruba, págs. 134-137.

[16] Ikea

http://www.ikea.com/ms/es_ES/customer_service/index_new.html - 25 de enero de 2014.

[17] VoiceXML

<http://www.w3.org/TR/voicexml20/> - 25 de enero de 2014.

[18] XML

<http://www.w3.org/TR/2000/REC-xml-20001006> - 03 de diciembre de 2013.

[19] **Etiquetas**

<http://guias.ovh.es/printPage/TelFaqSVIVXMLejemplo#field> - 03 de diciembre de 2013.

[20] **Protocolo HTTP**

[Hypertext Transfer Protocol -- HTTP/1.1](http://www.ietf.org/rfc/rfc2616.txt), IETF RFC 2616, 1999. 05 de enero de 2014.

[21] **CCXML**

<http://www.w3.org/TR/ccxml> - 03 de diciembre de 2013

[22] **CallXML**

<http://docs.voxeo.com/callxml/3.0/home.htm> - 03 de diciembre de 2013.

[23] **Bevocal Café**

<http://cafe.bevocal.com/> [09 de febrero de 2011] - 03 de febrero de 2014.

[24] **Conita**

Griol, David. Desarrollo y evaluación de Diferentes Metodologías para la Gestión Automática del Diálogo. Tesis Doctoral. Universidad Politécnica de Valencia, 2007.

[25] **Eloquent**

<http://www.eloquant.com/index.php> - 12 de febrero de 2014.

[26] **OpenCall**

<http://www.hp.com> - 03 de febrero de 2014.

[27] **IBM**

<http://www.ibm.com> - 03 de febrero de 2014.

[28] **Omvia Media Server**

<http://www.convergys.com/> - 12 de febrero de 2014.

[29] **JVoice XML**

<http://www.jvoicexml.sourceforge.net> - 12 de diciembre de 2013.

[30] **Loquendo**

<http://www.loquendo.com> - 12 de enero de 2014.

[31] **Lucent Technologies**

<http://www.alcatel-lucent.com> - 12 de diciembre de 2013.

[32] **VoxGategay**

<http://www.motorola.com> - 12 de enero de 2014.

[33] **Nuance**

<http://www.nuance.com> - 12 de diciembre de 2013.

[34] **OpenVXI**

<http://www.speech.cs.cmu.edu/openvxi> - 26 de diciembre de 2013]

[35] **OptimSys ofrece OptimTalk**

<http://www.optimsys.cz> - 22 de diciembre de 2013.

[36] **Tellme**

<https://studio.tellme.com> - [12 de diciembre de 2013.

[37] **VoiceGenie**

<http://www.voicexml.org/solution/voicegenie-71> - 12 de diciembre de 2013.

[38] **Voxpilot**

<http://www.vocabase.com/en/content/view/113/258/> - 22 de diciembre de 2013.