

# Desarrollo de portales de voz municipales interactivos y adaptados al usuario\*

## *Development of Interactive and User-Centered Voice Portals to Provide Municipal Information*

David Griol, María García-Jiménez, José Manuel Molina, Araceli Sanchis

Universidad Carlos III de Madrid

Av. de la Universidad, 30

28911 - Leganés (Spain)

{david.griol, josemanuel.molina, araceli.sanchis}@uc3m.es, 100025080@alumnos.uc3m.es

**Resumen:** El objetivo principal de este artículo es mostrar la aplicación práctica de los sistemas de diálogo y del estándar VoiceXML para el desarrollo de portales de voz de información municipal. A través del portal desarrollado, los usuarios pueden interactuar telefónicamente para consultar información sobre el Ayuntamiento y la ciudad, realizar diversas gestiones y trámites, completar encuestas, acceder al buzón del ciudadano y ser transferido a la centralita del Ayuntamiento. Estas funcionalidades, conjuntamente con los mecanismos introducidos para su adaptación al usuario y a la evolución del diálogo, hacen que la aplicación desarrollada pueda considerarse como un avance considerable con respecto al desarrollo actual de los portales de voz en España. El artículo describe los servicios proporcionados por el portal, arquitectura y tecnologías utilizadas.

**Palabras clave:** Portal de Voz, Sistemas de Diálogo, VoiceXML, Adaptación al Usuario.

**Abstract:** In this paper, we describe a Voice Portal designed to provide municipal information by phone. It includes the set of modules required to automatically recognize users' utterances, understand their meaning, decide the following response and generate a speech response. The different functionalities include to consult information about the City Council, access city information, carry out several steps and procedures, complete surveys, access citizen's mailbox to leave messages for suggestions and complaints, and be transferred to the City Council to be attended by a teleoperator. The voice portal is, therefore, pioneer in offering an extensive and comprehensive range of user-centered services accessible through speech, creating a new communication channel which is useful, efficient, and easy to use. The paper describes the application software, architecture, and infrastructures required for its operation 24 hours a day.

**Keywords:** Voice Portals, Spoken Dialog Systems, VoiceXML, User Adaptation.

## 1 Introducción

Gracias a que la voz es un medio natural e intuitivo para interactuar y comunicarse, las aplicaciones basadas en sistemas de diálogo (Pieraccini, 2012) se han convertido en una de las opciones para facilitar la interacción con dispositivos electrónicos. Estos programas informáticos tienen como principal finalidad interactuar con los

usuarios oralmente o de forma multimodal para proporcionarles información o un determinado servicio de forma automática. El número de dominios de aplicación actuales de estos sistemas es enorme.

Para la implementación de los sistemas de diálogo que proporcionan información en la web, el World Wide Web Consortium (W3C) propone el uso del estándar VoiceXML (Will, 2012). Este lenguaje de programación permite la interacción persona-máquina integrando funcionalidades como la síntesis de texto-a-voz, reproducción de audio, reconocimiento del habla y de tonos DTMF

\* Trabajo parcialmente financiado por los proyectos TRA2011-29454-C03-03, MINECO TEC2012-37832-C02-01, CICYT TEC2011-28626-C02-02 y CAM CONTEXTS (S2009/TIC-1485).

(Dual-Tone Multi-Frequency), grabación de voz, control de flujo de diálogo y funciones de telefonía. VoiceXML es además una tecnología independiente de la plataforma que permite la portabilidad y transferencia de datos entre aplicaciones heterogéneas. Su utilización de manera conjunta con otros estándares y lenguajes de programación proporciona una base sólida para el desarrollo de sistemas de diálogo.

En este trabajo proponemos la utilización del estándar VoiceXML para el desarrollo de portales de voz que proporcionen información municipal. En este contexto, la Ley 11/2007 sobre el acceso electrónico de los ciudadanos a los Servicios Públicos incide en el fomento de múltiples canales de acceso a la información como una de las principales obligaciones de los Ayuntamientos, y reconoce explícitamente el derecho de los ciudadanos a relacionarse con las Administraciones Públicas por medios electrónicos.

No obstante el desarrollo de portales de voz municipales en España es actualmente muy limitado. De hecho, las pocas aplicaciones existentes actualmente ofrecen únicamente el acceso a una grabación para la elección de una determinada área o sección entre un conjunto reducido, a la que el usuario es redirigido para ser atendido personalmente por un operador del Ayuntamiento (por ejemplo, el portal desarrollado para el Cabildo de Gran Canaria).

Adicionalmente, existen otros tipos de acceso oral a la información municipal. Por ejemplo, en la página Web del Ayuntamiento de Santander se puede descargar una aplicación que permite navegar oralmente por las páginas que integran el portal del Ayuntamiento, además de poder escuchar el contenido de estas páginas web con una voz sintetizada. Otros ejemplos del uso de voz sintetizada son el portal web del Ayuntamiento de Zaragoza y del Ayuntamiento de Alicante.

De este modo, el portal de voz que se describe en este artículo es pionero en ofrecer una extensa y completa oferta de servicios municipales accesibles mediante el habla a través de un número de teléfono, así como en cuanto a las posibilidades ofrecidas para la adaptación del servicio proporcionado, tal y como se detalla en la siguiente sección.

## 2 Portal de voz desarrollado

El desarrollo de la aplicación se ha resuelto utilizando una arquitectura cliente-servidor cuyo esquema se puede observar en la Figura 1. La interacción con el sistema comienza cuando el usuario inicia una llamada, bien mediante la línea telefónica o bien mediante cualquier cliente de VoIP (por ejemplo, Skype). La interacción con el usuario y la provisión de las diferentes funcionalidades se lleva a cabo gracias a la disposición de un servidor VoiceXML y de servidores web.

En el servidor VoiceXML (para nuestro proyecto, la plataforma Voxeo Evolution<sup>1</sup>), 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. 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. Para el desarrollo del portal de voz se ha utilizado las tecnologías Prophecy Premium ASR y TTS<sup>2</sup>.

En cuanto al servidor web utilizado, éste alberga las diferentes páginas con información dinámica programadas utilizando el lenguaje PHP, así como las bases de datos MySQL que contienen la información estática de la aplicación. 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 la página web del Ayuntamiento de Alcorcón, y se almacena, perfectamente 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 un fichero VoiceXML. Ejemplos de este tipo de información son la historia de Alcorcón, los accesos a la ciudad, los datos de contacto de un hotel de la ciudad o de la oficina de empleo del municipio.

<sup>1</sup> evolution.voxeo.com

<sup>2</sup> help.voxeo.com/go/help/evolution.platforms.chooseplat.eu

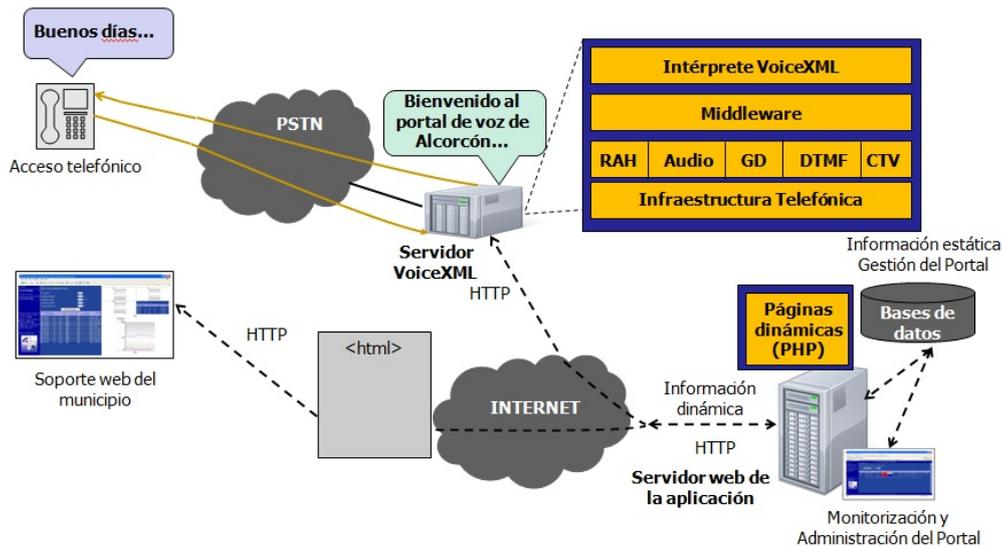


Figura 1: Arquitectura del portal de voz desarrollado

La información dinámica utilizada por la aplicación, en lugar de recargarse periódicamente en las bases de datos de la aplicación, se consulta directamente de portales web externos y se procesa sintácticamente a partir de páginas en PHP. Ejemplos de este tipo de información son las noticias y los eventos municipales, la información meteorológica, la encuesta municipal y la cartelera de los cines de Alorcón. Las gramáticas dinámicas utilizan información que varía con el tiempo y suelen tratar gran cantidad de datos. Si se quisieran incluir en los ficheros VoiceXML de forma manual habría que modificarlas cada poco tiempo, tarea que sería muy laboriosa debido al gran número de ítems de los que constan. Ejemplos de estas gramáticas son las diseñadas para consultar información sobre los comercios municipales, acceder a las páginas amarillas o rellenar las encuestas disponibles.

## 2.1 Funcionalidades de la aplicación

El módulo Inicio inicia la interacción con el usuario. Las opciones que éste puede elegir se reparten en 5 módulos bien diferenciados según el tipo de interacción y de datos que se van a proporcionar: información, gestiones y trámites, encuesta, buzón del ciudadano, y operador. Por tanto, es en este módulo inicio donde se bifurca el diálogo entre el resto de módulos. Las principales acciones que se llevan a cabo en él son reproducir un mensaje

de bienvenida al usuario, posibilitarle la selección de idioma y darle a elegir la acción que desea realizar. Una vez que el usuario elija una de estas opciones, se continúa la rutina en el fichero principal correspondiente a esta opción elegida, ya almacenado en el servidor web externo.

En el módulo Información el usuario puede acceder a toda la información del municipio de Alorcón. La información, según su tipo, se divide en 6 grandes submódulos bien diferenciados y clasificados de tal forma que facilitan el acceso del usuario a la información que esté buscando:

- **Ayuntamiento:** Proporciona toda información relacionada con el Equipo de Gobierno, los Órganos de Gobierno y las Áreas de Gobierno municipales.
- **Ciudad:** En este submódulo se accede a la información referente a Alorcón como ciudad. Se pueden consultar los datos de la ciudad, la historia, los accesos, y unas páginas amarillas compuestas por los bares, cafés, restaurantes, tiendas, hostales, hoteles y cines (incluida la información sobre la cartelera) del municipio.
- **Áreas temáticas:** Cualquier otro tipo de información que busque el usuario se ha repartido y clasificado en una de las 15 áreas que forman esta sección. Para cada una de estas áreas se proporciona información general, competencias y

datos de contacto. Simplemente con la incorporación de más información estática en la base de datos se podrían añadir más funcionalidades y áreas en este módulo de la aplicación.

- Noticias: Proporciona las noticias del municipio. Se facilita la fecha, título y subtítulo de cada noticia.
- Eventos: Reproduce el listado de eventos del municipio. Se proporciona el área temática, título, fecha, lugar y descripción de cada evento.
- Información meteorológica: El usuario puede obtener la información meteorológica actual del municipio y la previsión para los siguientes dos días.

Mediante el módulo Gestiones y Trámites el usuario puede verificar si está incluido en un listado, comprobar el estado de sus expedientes, reservar una instalación municipal o pedir cita para ser atendido en un servicio municipal.

Otra de las opciones del portal de voz es la realización de encuestas. De esta forma, se puede recoger de forma fácil y rápida la opinión de los ciudadanos sobre algún tema en particular que se plantee sobre el municipio. La encuesta se realiza de forma totalmente anónima, no solicitándose ningún tipo de identificación. Una vez que el usuario haya terminado de contestar, se le da la opción de escuchar los resultados parciales almacenados hasta el momento.

En el módulo Buzón del Ciudadano se implementa la funcionalidad de grabar un mensaje de voz por parte del usuario y que este mensaje sea clasificado y almacenado para su posterior tramitación. De esta forma, el ciudadano a cualquier hora y desde cualquier punto, puede hacer llegar al Ayuntamiento sus solicitudes, quejas, reclamaciones o comentarios. Además, en el caso en el que el ciudadano deje sus datos, ya sea teléfono fijo, móvil o correo electrónico, se puede contactar con él para dar una respuesta personalizada a su solicitud.

Finalmente, en el módulo Tele-Operador se transfiere la llamada del usuario al teléfono de la centralita del Ayuntamiento de Alorcón.

Uno de los aspectos fundamentales en el desarrollo del portal de voz ha consistido en la incorporación de diferentes funcionalidades

para adaptar el sistema al estado actual del diálogo y a características específicas de cada usuario. El primero de los puntos es relativo al tratamiento de los diferentes eventos proporcionados por VoiceXML. El portal almacena además en la base de datos de gestión de la aplicación los números de teléfono desde los que se accede al mismo, así como las diferentes funcionalidades de la aplicación que se han consultado en cada una de ellas. Esta información es utilizada por el sistema para asignar cuáles son las preferencias de los usuarios, en cuanto a consultas previas más frecuentes.

### 3 Conclusiones

En este artículo se ha descrito un portal de voz desarrollado mediante el estándar VoiceXML para proporcionar información municipal. Los servicios ofrecidos por el portal de voz municipal, distribuidos en diferentes módulos por los cuales se conduce al usuario en función de las decisiones que vaya tomando durante el diálogo, permiten que la aplicación del portal para proporcionar estos servicios en la ciudad de Alorcón pueda considerarse como un avance considerable con respecto al desarrollo actual de estos portales en España.

Las líneas de trabajo que se desarrollan actualmente están relacionadas con la incorporación de funcionalidades adicionales en los diferentes módulos descritos, así como la mejora e incorporación de nuevas técnicas que amplíen los mecanismos descritos para la adaptación del portal. Para ello, se desea incorporar perfiles de usuario que incluyan información más detallada que facilite el uso de la aplicación, sirva además para adaptar la información proporcionada en los mensajes del sistema y disminuya los errores que se pudieran producir durante la interacción.

### Bibliografía

- Pieraccini, R. 2012. *The Voice in the Machine: Building Computers That Understand Speech*. MIT Press.
- Will, T. 2012. *Creating a Dynamic Speech Dialogue: How to implement dialogue initiatives and question selection strategies with VoiceXML agents*. AV Akademikerverlag.