VoiceCampus: An Automated Interactive Voice Response System for Students

Andreas Velonis, Theofilos Milonas, George Papazidis, George Dimitriou
Cardisoft S.A., 60 Monastiriou street, GR-54627, Thessaloniki, Greece
info@cardisoft.gr, www.cardisoft.eu

Keywords
Voice Campus, Voice Portal, Interactive Voice Response System, Text-to-Speech (TTS)

1. EXECUTIVE SUMMARY

VoiceCampus is an automated multi-lingual interactive voice response (IVR) telephone system utilizing high-end speech recognition and text to speech (TTS) technologies. It is especially designed to help students obtain the information they seek in a timely manner and is reliable, scalable, intelligent and user-friendly operating on a 24X7 basis. With VoiceCampus, students can obtain real-time information using just their telephone set or mobile phone via a simple voice call, even in situations where no Internet connection or PC is available to get this information.

VoiceCampus is based on a dialogue between the student and the system. The system, without the need of an agent, recognizes the words or phrases “spoken” by the student and immediately serves their requests without having them wait physically in line for hours at the administration offices or trying to find an Internet access point. In addition, routine information calls are handled by the system, saving valuable time for the administrative staff.

1.1. Background

VoiceCampus system is an application which belongs to the Voice Portal family of products that Cardisoft has deployed in numerous situations within the past 10 years. Over these years, Cardisoft has completed more than 100 successful Voice Portal installations with VoiceCampus being in the top of the list regarding education vertical market. VoiceCampus is utilizing Automatic Speech Recognition ASR) and Text To Speech (TTS) engines by the pioneer companies in the field and a dynamic dialog structure based on the VoiceXML and CCXML standards as defined by the W3 Consortium. The system can connect to any popular TDM or IP PBX\ACD system through analog, ISDN or VoIP line interface.

1.2. Services

VoiceCampus can provide automated voice services over the phone to the students such as:

- Information about their grades
- Teaching and Examination schedule
- Class Registration
- Certificate requests
- Library services
- General Information:
- Directory Services

In addition, VoiceCampus can provide several Unified Communication services to the students:

- SMS messaging,
- SMS Alerts & notifications
- Voicemail
VoiceCampus system

1.3. Conclusions

VoiceCampus is an easy-to-use voice application from anywhere, anytime. By making just a simple telephone call, the students can access personalized invaluable information which otherwise could find either by physically appear in the institute or using the internet. This is a system that can be extremely helpful to students from other countries or in holidays, where access is limited. Some of the benefits VoiceCampus introduces to an institute are:

- Innovative electronic voice services for the students
- Effective and efficient access to data by students over phone
- Less or even zero routine call-handling by the administrative staff
- Minimization of waiting queues in department’s secretariat and information desks
- Minimization of call waiting.
- 24X7 basis accommodation, on an always friendly and gentle communication approach

2. TECHNICAL OVERVIEW

In the following paragraphs, we will analyze the architecture, technology and interfaces that Voicecampus system utilizes.

2.1. Architecture

The main architecture of the VoiceCampus system consists of the following entities:

- Telephony Entity
- Speech Entity
- Administrative Entity

Telephony and Speech Entity is scalable according to the sizing and the capacity of expected simultaneous incoming calls. Later on, we will describe the Erlang analysis which describes the system’s capacity depending on various parameter inputs.

2.1.1. Telephony Entity

The Telephony subsystem is the entity responsible for the communication between the PBX\ACD infrastructure and the VoiceCampus system. Also, the telephony subsystem is the runtime environment of the dialog flow that is developed with the telephony software licenses depending on the capacity of the VoiceCampus system (simultaneous calls served). It consists of a hardware server which we describe as the Telephony Server and depending on the PBX\ACD infrastructure:

- TDM PBX requires a telephony hardware voice board with analog, ISDN or digital channels that connect to the PBX
- IP PBX requires no extra telephony hardware but SIP or H.323 trunks configured from the PBX’s VoIP gateway and an IP application middleware to translate SIP and H.323 events to the VoiceCampus system.

Depending on the capacity of the VoiceCampus system and the required protocol the telephony hardware voice board is selected among various models of the worldwide leading manufacturer in the field. Accordingly, VoiceCampus system uses an IP middleware application protocol in the case of VoIP implementations.
In addition, in this subsystem the telephony administration console is installed. This software application allows the administrator to monitor & manage:

- Telephony Resources (number of telephony channels)
- Telephony protocol and signaling
- Voice Resources
- Speech Resources (automatic speech recognition and text to speech licenses)
- Database connections (between the dialog application and back office systems)
- IVR scripts
- Log of calls
- Event viewer
- Alarm and notification in case of exceptions

2.1.2. Speech Entity

The Speech subsystem is hosting the Automatic Speech Recognition engines and Text To Speech engines that the VoiceCampus system is using. These engines offer to the VoiceCampus system the necessary resources depending on the overall capacity and dialog usage of speech recognition and text to speech within the dialog itself.

The CPU load in real-time calls which cannot accept delays in the dialog flow for optimal performance, requires a dedicated Speech Server where these engines and the software licenses of the engines are installed.

This subsystem covers all speech and text to speech sensitive parameters of the VoiceCampus application such as:

- Barge-In and Selective Barge-in
- Voice Activity Detection
- Background noise cancellation
- Acoustic models
- Speech grammars, synonyms and lexicons
- Linguistics and phonetics
- Text To Speech Language identifier
- Language Identifier
2.1.2.1. Multilingual Support

The speech entity of the VoiceCampus system offers a *multilingual* environment in both Speech Recognition and Text To Speech Engines. VoiceCampus currently supports the following languages as shown below:

<table>
<thead>
<tr>
<th>Language</th>
<th>Dialect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cantonese</td>
<td>Hong Kong</td>
</tr>
<tr>
<td>Catalan</td>
<td>Spain</td>
</tr>
<tr>
<td>Czech</td>
<td>Czech Republic</td>
</tr>
<tr>
<td>Danish</td>
<td>Denmark</td>
</tr>
<tr>
<td>Dutch</td>
<td>Netherlands</td>
</tr>
<tr>
<td>English</td>
<td>Australia, Great Britain, India, Singapore, USA</td>
</tr>
<tr>
<td>Euskera</td>
<td>Spain</td>
</tr>
<tr>
<td>Finnish</td>
<td>Finland</td>
</tr>
<tr>
<td>Flemish</td>
<td>Belgium</td>
</tr>
<tr>
<td>French</td>
<td>Belgium, Canada, France, Luxembourg, Switzerland</td>
</tr>
<tr>
<td>German</td>
<td>Austria, Germany, Luxembourg, Switzerland</td>
</tr>
<tr>
<td>Greek</td>
<td>Greece</td>
</tr>
<tr>
<td>Hebrew</td>
<td>Israel</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Language</th>
<th>Dialect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hungarian</td>
<td>Hungary</td>
</tr>
<tr>
<td>Italian</td>
<td>Italy, Switzerland</td>
</tr>
<tr>
<td>Japanese</td>
<td>Japan</td>
</tr>
<tr>
<td>Korean</td>
<td>Korea</td>
</tr>
<tr>
<td>Mandarin</td>
<td>PRC, Taiwan</td>
</tr>
<tr>
<td>Norwegian</td>
<td>Norway</td>
</tr>
<tr>
<td>Polish</td>
<td>Poland</td>
</tr>
<tr>
<td>Portuguese</td>
<td>Brazil, Portugal</td>
</tr>
<tr>
<td>Russian</td>
<td>Russia</td>
</tr>
<tr>
<td>Slovak</td>
<td>Slovakia</td>
</tr>
<tr>
<td>Slovene</td>
<td>Slovenia</td>
</tr>
<tr>
<td>Spanish</td>
<td>Argentina, Colombia, Spain, USA</td>
</tr>
<tr>
<td>Swedish</td>
<td>Finland, Sweden</td>
</tr>
<tr>
<td>Turkish</td>
<td>Turkey</td>
</tr>
<tr>
<td>Welsh</td>
<td>Great Britain</td>
</tr>
</tbody>
</table>

Table 1
2.1.2.2. Speech Recognition Accuracy

Speech Recognition Accuracy of the VoiceCampus system tested on various conditions and thousand of calls is shown in the table below:

<table>
<thead>
<tr>
<th>Task</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes/No</td>
<td>&gt; 99%</td>
</tr>
<tr>
<td>Zip code</td>
<td>&gt; 97%</td>
</tr>
<tr>
<td>Phone number</td>
<td>&gt; 95%</td>
</tr>
<tr>
<td>Date</td>
<td>&gt; 95%</td>
</tr>
<tr>
<td>Small Item List</td>
<td>&gt; 98%</td>
</tr>
<tr>
<td>4000 Company Names</td>
<td>&gt; 95%</td>
</tr>
</tbody>
</table>

Table 2

2.1.3. Administrative Entity

The Administrative subsystem is offering the Web management tool for the administrator. This software application allows the administrator to manage the dialog operations and get statistic reports of the usage of the VoiceCampus system. The administrative subsystem supports:

- Overall call list of incoming calls
- Specific reports per day, time or service used etc.
- Historical reports
- Dialog content management tool in order to edit the prompt “spoken” by the system and change it
- Service management tool in order to add, change or delete services and menus
- Dialog synthesis tool in order to change how a phrase is synthesized and spoken in the dialog, using simple Excel functions

The Administrative subsystem is a Web site utilizing Microsoft Share Point Services and can be installed in either the Telephony or Speech Server or anywhere in the network.
2.2. Topology
In the following figures you can see the topology of VoiceCampus in the network.

**TDM Implementation**

![TDM Implementation Diagram]

**VoIP Implementation**

![VoIP Implementation Diagram]
2.3. Interfaces & Technology

Having described the entities of the VoiceCampus system and the topology, we can analyze the various interfaces that are used in the interoperability of the overall system. In the following figure you can see the interfaces and technology that the VoiceCampus system is using:

![Figure 4](image.png)

The whole VoiceCampus approach is to use edge-technology, open interfaces and web tools in order to provide a dynamic dialog structure easily managed by the administrators of the system. Strict to this approach Cardisoft is using open voice and telephony standards like VoiceXML, CCXML, XML Web Services, SMTP, SOAP with external systems.

2.4. Redundancy

The VoiceCampus system is a redundant system taking advantage of the architecture described in the paragraphs above. The system is designed to offer redundancy by using either the Telephony or Speech Server as backup one to the other. This way, if for example the Telephony Server fails, calls are automatically routed and served by the Speech Server. The only limitation is that in the case of a TDM deployment where a telephony hardware board is required, a second stand-by telephony board, should be installed also in the Speech Server.
2.5. Sizing Considerations

VoiceCampus system as well as any other Voice Portal System is sized according to the simultaneous incoming calls it is expected to serve. All the software licenses (telephony, asr, tts) as well as the telephony hardware board is depended on the capacity of the VoiceCampus system meaning simultaneous calls served.

But how can we ensure that the sizing analysis will not drive us to blocked calls by using less channels than needed or exceeded cost by using more?

A very precise tool used adopted by the VoiceCampus sizing consideration is the Erlang model (http://www.erlang.com/calculator/erlb/). By using the Erlang model we can estimate safely the capacity of the VoiceCampus system.

The Erlang B traffic model is used by telephone system designers to estimate the number of lines required. The three variables involved are Busy Hour Traffic (BHT), Blocking and Lines:

- **Busy Hour Traffic (in Erlangs)** is the number of hours of call traffic there are during the busiest hour of operation of a telephone system.
- **Blocking** is the failure of calls due to an insufficient number of lines being available. E.g. 0.03 mean 3 calls blocked per 100 calls attempted.
- **Lines** is the number of lines in a trunk group.

Let’s suppose that the busiest hour within the day, based on historical data, we receive 400 calls with average duration of each call 3 minutes and that we do not want our institute to miss more than 10 calls, which means 2% blocked calls. In this example, the BHT is calculated as following: BHT=500*3=1500 minutes of calls=25 hours within the busiest hour and the Erlang model is shown in the figure below:

![Erlang B Calculator](image)

Which dictates us that our VoiceCampus system needs 34 lines.

2.6. Scalability

After having decided how many lines we will need for our VoiceCampus system, it is very important to know what happens in the situations that we have more calls and the system does not suffice to accommodate? VoiceCampus is 100% scalable in case of a VoIP implementation, by importing on-the-fly software license files increasing the capacity of the system. In the case of a TDM implementation the telephony board should also support the additional voice channels.

2.7. Reliability

VoiceCampus operates on a 24X7 basis. As a Voice Portal system it cannot afford downtimes which may occur the most inconvenient times or days. As a Voice Portal system a simple server restart can result into numerous lost calls, since the users of the system are outside callers and not employees. For that reason, Cardisoft is using the most reliable telephony hardware and speech engines to minimize any downtime possibility.
3. **VoiceCampus Features & Services**

VoiceCampus system provides a portfolio of automated voice services for the students. The students call to the institute, enter their student personal code and pin through the digits of their telephone set and have access to personalized information retrieved from the back-office systems of the institute. Furthermore, other services are “open” to the callers, without the need of logging to the system. Below we have the two main categories of services:

**Personalized Information**

- *Grades Information*: The student states the class and is informed of the grade achieved during the last examination period as well as other historical data in previous periods for this class.
- *Teaching and Examination schedule*: The student by stating his class to the system he can be informed of the weekly teaching schedule as well as the exam date for this class.
- *Class registration*: Students can apply and register to classes for the semester by phone, according to criteria defined by the institution rules and limitations.
- *Certificate requests*: Students can request certificate documents from the department such as grade analysis etc.
- *Library services*: Students can apply to borrow a specific title in the department’s library, learn the availability of specific documents etc.

**General Information**

- *General Information*: Callers can learn about facilities, registration procedure, access, general operation rules and any other information that the institute would like to publish over the phone.
- *Directory Services of the Campus*: Callers can be informed of contact information to specific departments of the Campus and may even use the system to connect to the direct extension of these departments.

**Unified Communication**

VoiceCampus can provide a wide range of additional services to the callers over the phone, utilizing Unified Communication technologies and the relevant module of the system:

- *SMS messaging*: This service can be offered to the students if they leave their mobile phone number in order to receive a SMS message
- *SMS Alerts & notifications*: Extending the previous service, students can request an alert or notification for specific services such as examination date, class registration deadline, availability of a library document etc.
- *Voicemail*: Callers are able to leave voice messages for specific employees or departments within the institute. Voice messages are stored in voice mailboxes and are sent as .wav attachments in the email of the corresponding voice mailbox.
- *Fax On Demand*: VoiceCampus system can send faxes of documents requested to the given fax number. Callers must enter their fax number.
- *Email On Demand*: In the case the system knows the email address of the caller who logged in, VoiceCampus can send email with requested information to this address.

**Additional Services**

As described in the paragraph § 2.1.3 Administrative Entity, administrators of the VoiceCampus system can introduce additional services to the system. An institute may need to offer further
services to the callers over phone which are not mentioned above. Administrators have the web tool to enrich the content of the VoiceCampus system with additional services, menus and messages.

4. Case Study - University of Crete VoiceCampus

One of the main strategies of the University of Crete (www.uoc.gr) was to adjust to international trends and practices in order to provide digital services for the university community members (students, teaching and administrative personnel, and citizens in general).

The projects carried out incorporated modern technology for an electronic University and created an infrastructure capable of providing high quality services within the educational and administrative departments.

These applications are dynamic and are capable of following the changes and needs of the University in its daily routines as well as with respect to its interactions with other University.

As part of these projects, Cardisoft S.A. was awarded and implemented VoiceCampus IVR system in order to provide services to students via telephone on a 24hour basis using Speech Recognition and Text to Speech technologies.

Students can call the automated telephone system which is connected to the Cardisoft e-University Administrative Suite and provides students with services such as access to grades, exam timetable, weekly deliverables programs, and requests for official documents etc.

The aim of the project was to minimize the total time spent by students waiting to be served by their departments as well as to minimize the number of routine calls that could be served by the IVR system.

The main element of the University’s Voice Portal is that the existing VoIP technology based on Cisco Systems technology and equipment was utilized and expanded to accommodate Voice Recognition and Text to Speech technologies.

Today, after operating for nearly 3 years, VoiceCampus in University of Crete serves more than 1,500 students daily and is acknowledged as the first Voice Portal system in Greece to introduce Speech Recognition and Text To Speech Technology in an IP PBX environment.

Services and Technologies

The Voice Portal developed for the University of Crete utilizes cutting edge technology with open architecture for the communication with the Cardisoft e-University Suite of Applications that was already in operation. The system uses an intelligent and friendly algorithm for the authentication of each student to utilize the personalized services via his/her Student ID Number and a PIN number issued. The services provided are:

- Grades
  The student states the subject and is informed of the grade achieved during the last examination period as well historical data other grades received for this subject.

- Teaching and Examination program.
  Students have access information regarding due dates and examination dates based on the subject.

- Curriculum
  Students are able to select the subjects to be studied for the semester. Criteria restrictions apply as in the case that the student applies in person.

- Requests for certificates
  Students may request various types of documents to be issued by their department such as analytical grades etc.

- Services for Postgraduate students
- **Services for first year students**
  Such as information regarding enrollment, necessary documents etc.
- **Information regarding the Library**

**Technology Utilized**
- Text to Speech technology (T.T.S. - Nuance Realspeak Telecom 4.0 engine) in order to process the large volume of subjects for all the schools and departments which change often and the use of prerecorded messages would render the system less flexible to changes.
- Speech Recognition technology (A.S.R. - Nuance OpenSpeech Recognizer 3.x engine) in order to ensure intelligent and friendly dialogs with the students.
- The system’s architecture is based on VoIP technology using Cisco Call Manager and Voice Gateways (www.cisco.com) that were already in operation. The technology was used for the first time in Greece with Speech Recognition and Text to Speech and interoperability with the Cardisoft e-University Platform is excellent since all systems are based on open standards and architecture.

5. **REFERENCES**


