

# A Multimodal Guide for the Augmented Campus

Salvatore Sorce<sup>1</sup> Agnese Augello<sup>1</sup> Antonella Santangelo<sup>1</sup> Giovanni Pilato<sup>2</sup>  
Antonio Gentile<sup>1</sup> Alessandro Genco<sup>1</sup> Salvatore Gaglio<sup>1,2</sup>

<sup>1</sup> DINFO - Dipartimento di ingegneria INFormatica, University of Palermo  
Viale delle Scienze, edificio 6 - 90128 Palermo – Italy  
Tel.: +39 091 7028526 – Fax.: +39 091 6598043

<sup>2</sup> ICAR - Istituto di CALcolo e Reti ad alte prestazioni, Italian National Research Council  
Viale delle Scienze, edificio 11 - 90128 Palermo – Italy  
Tel.: +39 091 238111 – Fax: +39 091 6529124

{sorce|augello|santangelo}@csai.unipa.it; g.pilato@icar.cnr.it; {gentile|genco|gaglio}@unipa.it

## ABSTRACT

The use of Personal Digital Assistants (PDAs) with ad-hoc built-in information retrieval and auto-localization functionalities can help people navigating an environment in a more natural manner compared to traditional audio/visual pre-recorded guides. In this work we propose and discuss a user-friendly, multi-modal guide system for pervasive context-aware service provision within augmented environments. The proposed system is adaptable to the user needs of mobility within a given environment; it is usable on different mobile devices and in particular on PDAs, which are used as advanced adaptive HEI (human-environment interaction) interfaces. An information retrieval service is provided that is easily accessible through spoken language interaction in cooperation with an auto-localization service. The interaction is enabled by speech recognition and synthesis technologies, and by a ChatBot system, endowed with common sense reasoning capabilities to properly interpret user speech and provide him with the requested information. This interaction mode turns to be more natural, and users are required to have only basic skills on the use of PDAs. The auto-localization service relies on a RFID-based framework, which resides partly in the mobile side of the entire system (PDAs), and partly in the environment side. In particular, RFID technology allows the system to provide users with context-related information. An implemented case study is showed that illustrates service provision in an augmented environment within university campus settings (termed "Augmented Campus"). Lastly, a discussion about user experiences while using trial services within the Augmented Campus is given.

## Categories and Subject Descriptors

H.5.2 [User Interfaces]: Natural language, Voice I/O, Input devices and strategies, Prototyping.

**General Terms:** Management, Documentation, Design, Experimentation, Human Factors.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

*SIGUCCS'07*, October 7–10, 2007, Orlando, Florida, USA.  
Copyright 2007 ACM 978-1-59593-634-9/07/0010...\$5.00.

## Keywords

Multimodal guides, pervasive services, mobile devices, human-environment interfaces

## 1. INTRODUCTION

In the ubiquitous computing vision, technology is seamlessly integrated into the environment and is used to provide humans with useful services in their everyday lives [27]. Context factors, such as who, when, where, are used to develop pervasive and context-aware systems with the aim to provide people with useful information in relation with the user profile and with the environment within the user is in. The advent of small and smart personal mobile devices, equipped with third-generation wireless communication technologies, can significantly contribute to the realization of such a vision, and they can be exploited to reach the user needs of mobility and focused information access.

Mobile devices are more and more used in experiments with the aim to make them suitable to pervasively access services in different application fields. On the other hand, new ways of service provision are studied (and brand-new services too) according to the features and capabilities of mobile devices [12] [6]. This common interest is justified by the wide diffusion of such devices, that are almost in everyone's pocket and that can be used almost any when and anywhere.

There is a large variety of application fields where services can be pervasively accessed by mobile devices, such as context-aware information provision within university campuses [13], interactive user profile-based guides in cultural heritage sites [22] [24], augmented reality objects assembly in mobility [15], healthcare systems [23].

The provision of services in a pervasive fashion within university campuses is one of the most used application field by several research groups, since it represents a very significant test-bed. In such contexts there is a large variety of users, with different ages, different skills and different needs, who can be provided with a huge range of possible services. Previous research results within actual campuses showed that both teachers, administrative employees and students, like to experience new ways to access well-known services, such as adaptive e-mail and calendar services [7], or flexible network storage [8]. Actually, young users better appreciate the use of a mobile phone as an access media for some pervasive service ubiquitously available [14], thus leading to a new way to conceive the access to existing services [19] in order to make them

less boring and to encourage students to use new services [10]. Last but not least, university campuses or school buildings often have an existing available framework (such as a web-based services network) which can be used as a basis for further developments.

In this paper we present a pervasive, multimodal virtual guide which can be accessed by means of smart personal mobile devices. The guide integrates different technologies in order to offer a more natural interaction between users and information provision services compared to traditional pre-recorded, audio-visual guides. The proposed system is a context-sensitive multimodal information service provider, whose aim is to provide people with information related to their position within an environment, using a PDA equipped with a Radio Frequency Identification (RFID) compact flash reader module.

In order to achieve a more natural interaction and to enable information access also by inexpert and/or disabled users, a conversational agent is designed based on the ChatBot software technology [1], integrated with a Cyc-based reasoning module [21][17], speech recognition/synthesis engine, and RFID-based auto-localization capability, which gives the system the context sensitiveness [20].

The use of a conversational agent as a user interface allows users to easily interact with the system, avoiding ad-hoc styli/touch pens and excluding requirements of uncomfortable interaction [3]. Besides, traditional use of mobile devices such as PDAs or smartphones is not a simple task mainly because of the limited dimensions of their screens. This drawback is emphasized if users are disabled or not technology-skilled. Vocal interaction is a more direct approach but it comes with a different set of constraints, such as needed processing power, speakers adaptation, narrow dictionaries and bounding grammars.

In the following we describe the system architecture, and an implemented prototype tested at the campus of the Palermo University, and in particular at the Dipartimento di Ingegneria Informatica (Computer Engineering Department, DINFO). The system prototype has been implemented using Asus A730W PDAs equipped with CF-RFID readers.

## 2. SYSTEM OVERVIEW

The proposed system is an intelligent campus guide which is accessible using a PDA equipped with an auto-localization module and provides users a natural language speech interface for ChatBot interaction. An overview of the overall system is shown in Figure 1.

The system is composed by two main areas: the Interactive Area and the Rational Area. The first one provides a multimodal access to the system services. This area embeds the vocal interaction, the traditional point-and-click interface, and the automatic self-location. It allows to acquire user inputs in multimodal form and it feeds the Rational Area with simple questions in textual form. After the Information Retrieval (IR) process, the Rational Area returns text, images and links to the Interactive Area for the output process.

Users interact with the Rational Area by means of the Interactive Area. The Rational Area is made of a Cyc ontology and a set of categories described with Artificial Intelligence Mark-up Language (AIML), both concerning the specific domain.

Users can request and access information through spoken commands instead of traditional visual and keyboard (or stylus) based commands by means of Automatic Speech Recognition (ASR) and Text-To-Speech (TTS) technologies embedded in a multimodal browser.

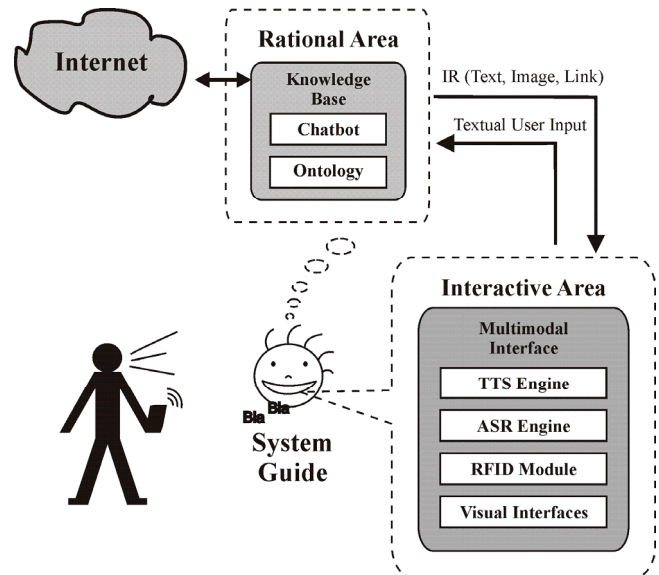


Figure 1. Architecture and system components

The ChatBot processes user questions and gives consistent replies using its KB enriched with ontology conceptualization and internet services too. For example, the ontology can be explored by the ChatBot in order to obtain information about the department, such as its organizational structure, the research groups working inside of it, information about the courses, and so on.

Smart interaction is therefore enabled by making deductions about the vocal information requests. An OpenCyc [21] ontology has therefore been created for the specific domain, and it has been combined with the agent dialogue capabilities. The system is equipped with a multimodal interface using the HTML+Voice language [18]. The vocal technologies embedded in a multimodal browser allow the information access through spoken commands as well as traditional visual and keyboard (or stylus) based commands. RFID module in the Interactive Area allows the system to sense changes in the environment in which the user is located and to automatically adapt itself and act accordingly to these changes based on user needs. This technology can be exploited to improve the users experience within augmented environments, thus allowing the access to relevant resources in a context-dependent fashion. Figure 2 shows a screenshot of the multimodal browser running on the PDA.

## 3. CONVERGING TECHNOLOGIES

### 3.1 A.L.I.C.E.

The ChatBot has been implemented by means of A.L.I.C.E. (Artificial Linguistic Internet Computer Entity) technology [1]. The dialogue mechanism of A.L.I.C.E. is based on algorithms for automatic detection of patterns in the dialogue data (*pattern matching*). Every time a user asks the ChatBot for a question, the ChatBot searches the best pattern matching the query in its knowledge base (KB). The ChatBot KB is composed by a set of files written in AIML (Artificial Intelligence Mark-up Language) [1], which layout is XML-like. Each file contains one or more question-answer modules called *categories*, which are described by specific AIML tags. In particular, the *pattern* tag describes the possible user expressions, while the *template* tag contains the corresponding ChatBot answers. The template can also include other AIML tags, thus allowing the ChatBot to save and get the

values of variables (*set* and *get* tags), or to start other programs (*system* tag), or to recursively call other categories (*srai* tag) and remember the context of the conversation (*that* tag). Figure 3 shows the case of an AIML file containing one category. The *pattern* and the *template* tags represent an example of a user's query and the corresponding ChatBot answer.

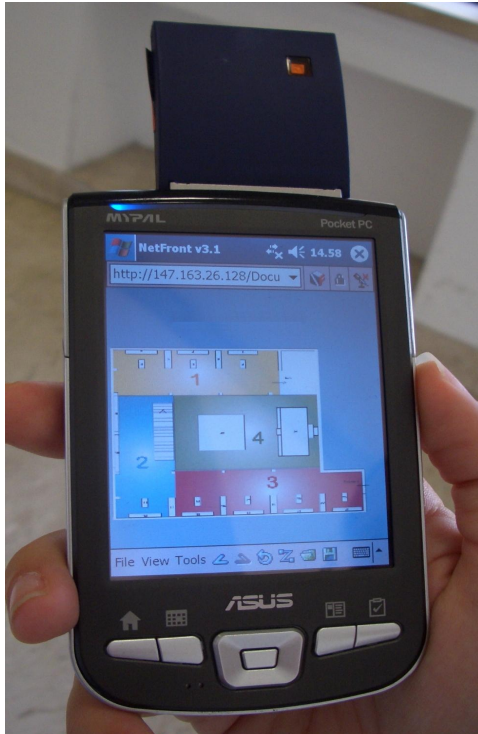


Figure 2. Multimodal browser screenshot

```
<?xml version="1.0">
<aiml>
<category>
<pattern>
  WHERE IS LOCATED THE CLASSROOM A320
</pattern>
<template>
  The Classroom A320 is located at the
  third floor of the building number 6.
</template>
</category>
```

Figure 3. A simple AIML single category

### 3.2 OpenCyc

OpenCyc is the open source version of the Cyc KB and reasoning engine [21]. The Cyc KB is a formalized representation of a vast quantity of human knowledge: facts, rules and heuristics for reasoning about concepts and events of everyday life. Cyc associates assertions that share a common set of assumptions into inherited clusters called microtheories. The KB is structured in four overlapped layers; at the lowest level there are the facts that are specific assertions for the involved domain and at the higher level there are domain-specific theories, core theories, and the upper ontology. The Cyc inference engine allows to obtain a lot of deductions regarding the KB. It is based on a series of heuristic related to a reasoning techniques about microtheories. CycL, a

formal language from Lisp language, describes the Cyc KB and it is possible to interact with Cyc also with a set of Java APIs.

### 3.3 Multimodal Technology

Multimodality allows user to move between different modes of interaction, from visual to voice, according to both changes in environment and user preferences. There is no complete standard based language for multimodal interfaces development. We focus our attention on XHTML+Voice [4], X+V for short, because it has been submitted to W3C for standardization.

X+V is a markup language for web application that implements the vocal part of the multimodal interface with a subset of VoiceXML and the visual part with XHTML. The VoiceXML is a W3C standard XML-based language, for pure vocal interface development. XHTML is HTML with a syntax that conforms to the XML one.

The multimodal interface has been developed with the IBM Multimodal Tools 4.1.2.2 [16] for WebSphere Studio V5.1.2. The Multimodal Tools include the IBM WebSphere Multimodal Toolkit and IBM Multimodal Browser, which are the developing environment and the technology for the application execution. This solution uses a traditional client/server architecture [5], as shown in Figure 4.

In this architecture, Automatic Speech Recognition (ASR) and Text-To-Speech (TTS) technologies are embedded in a multimodal browser.

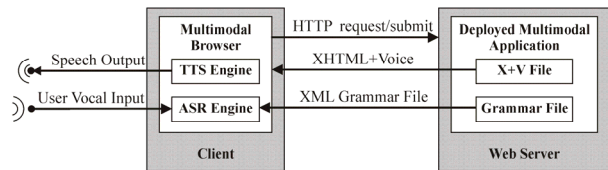


Figure 4. Multimodal interaction architecture

VoiceXML allows the access to web content through speech command providing a tag set for the vocal conversation flow controls. The Automatic Speech Recognition (ASR) process is achieved using grammar files loaded from the application at runtime. These grammar files set the dialogue context defining legal utterances user can say. X+V can support the XML Form and the ABNF Form of the W3C Speech Recognition Grammar Specification [2]. For the application presented in this paper we have chosen the XML Form. Synchronization between vocal and visual parts is made through XML Events, so that voice handlers can be invoked through a standard EventListener interface.

### 3.4 RFID Technology

A RFID tag is an electronic device that holds data. Typically these tags are attached to an item and contain a serial number or other data associated with that item.

RFID tags are categorized as either active or passive. Active RFID tags are powered by an internal battery, passive RFID tags operate without a separate external power source and obtain operating power generated from the reader. We used passive tags each holding an unique ID, and each placed at a point of interest. We also used a PDA equipped with a RFID compact flash reader module.

The RFID technology is used to detect changes occurring within the context the user is part of. This information is very useful for automatically adapting the behavior of systems in order to effectively satisfy the user needs. RFID technology can significantly

contribute to the realization of context-aware systems within pervasive augmented environments, as demonstrated by many researchers over the years [11]. Some examples are the Magic Medicine Cabinet [25], the augmentation of desktop items [26] and smart shelves [9]. These prototypes show that RFID technology has many benefits over other identification technologies because it does not require line-of-sight alignment, multiple tags can be identified almost simultaneously, and the tags can be suitably placed and hidden almost anywhere.

## 4. CASE STUDY

We implemented our proposed system within the campus of the University of Palermo, and in particular we carried out our experiments within the “Dipartimento di Ingegneria Informatica” (Computer Engineering Department, DINFO) domain. We built a domain-specific Cyc microtheory in order to represent knowledge about the DINFO. This can be used by the Chatbot in order to give appropriate answers to the user. The choice of using Cyc allows the system developers to exploit the big set of data already organized and described into this ontology. This makes the system more adaptable to domain changes, because it is not necessary to write the all set of knowledge every time, but only the most specific one.

### 4.1 Rational Area

#### 4.1.1 Common sense ontology about the “Augmented Campus”

We created the “DINFOMt” Cyc microtheory, which is connected to some pre-defined Cyc microtheories, and a hierarchy among them is established. In particular, “DINFOMt” is connected to the *AcademicOrganizationMt* Cyc microtheory, which describes the academic structure and the relations among several scholastic institutions by means of a specific set of collections and predicates. Most of used concepts are inherited from this microtheory. We created further concepts, predicates and inference rules in order to describe the specific domain and in particular the Italian academic course of study.

A small portion of the created ontology is illustrated in Figure 5, where light ovals represent Cyc Collections, and dark ovals represent instances of the collection.

In the following, a description of the main concepts is reported for clarity.

*University* is the collection of educational organizations at which university-level teaching and research takes place. *Unipa* is an instance of this collection representing the University of Palermo.

*AcademicDepartment* represents the collection of Academic sub-organizations. Each instance of *AcademicDepartment* is a major subdivision of an educational institution, with the primary purpose of research and/or teaching related to a particular academic discipline or subject-matter.

*ComputerScienceDepartment* is a specialization of *AcademicDepartment*. The instances of this collection are departments which conduct teaching and research in ComputerScience. An instance of *ComputerScienceDepartment* is *DINFO*, which represents the Department of Computer Science of Palermo.

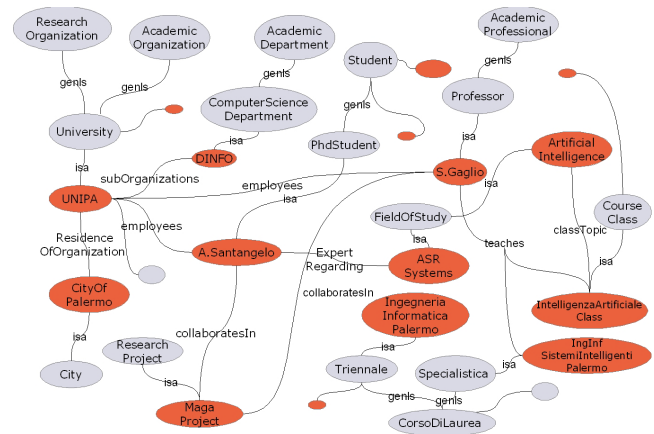


Figure 5. A detail of the DINFOMt ontology

*CorsoDiLaurea* represents the collection of the curricula available at the Italian universities which allow to achieve an academic degree. The Italian academic courses are mainly divided in two groups. The first group is the set of first level triennial curricula described by the collection *Triennale*; an instance of this collection is the bachelor course in “Ingegneria Informatica” at the University of Palermo (*IngegneriaInformaticaPalermo*). The second group is the set of second level biennial curricula, represented by the collection *Specialistica*; an instance of this collection is the Master of Science course in “Ingegneria Informatica per i Sistemi Intelligenti” at the University of Palermo (*IngInfSistemiIntelligentiPalermo*). Curricula of both groups are classified according to specific research and application fields.

*AcademicClass* represents the collection of the classes a student has to attend within a curriculum.

We also modeled relations among the ontology concepts by introducing specific predicates. For example we can represent that a *Professor* teaches in a *AcademicClass* belonging to a specific *CorsoDiLaurea*. We also represent people working at the University by specifying their main skills and tasks. For example *Antonella Santangelo* is a *PhdStudent* who is employed at the University of Palermo *UNIPA*, and who is an expert about *ASRsystems* (Automatic Speech Recognition systems).

#### 4.1.2 Interaction between the ChatBot and the ontology

The ChatBot is able to interact with the OpenCyc ontology by means of a module which integrates the AIML interpreter and the OpenCyc inference engine, and which is a Java porting of the CyN project [17]. This module allows to convert user requests in CycL/SubL expressions which can be interpreted by Cyc. The Cyc-enhanced template of the ChatBot can be seen as a meta-answer that must be processed to compose the appropriate response (Figure 6). The Cyc responses are made of automatically composed natural language sentences according to the presence of ad-hoc tags into the template.

The ontology plays a significant role in the dialogue with the user, because it enables the composition of answers that are not present in the traditional AIML KB.



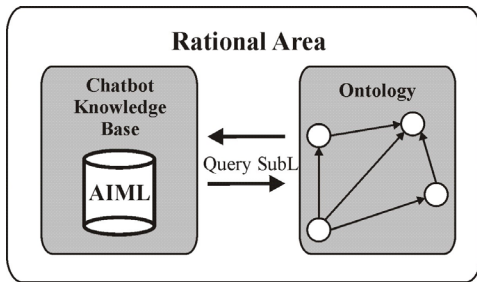


Figure 6. Interaction between the ChatBot and the ontology

### 4.1.3 Dialogue example

In the following a few examples of interaction are reported.

*User:* I need information about the Artificial Intelligence class.

*ChatBot:* The class is delivered by professor Salvatore Gaglio and provides the students with 12 CFU. If you want information about the program, you can visit the website: [www.csai.unipa.it/...](http://www.csai.unipa.it/)

*User:* Do you know some research project in which Professor Salvatore Gaglio is involved?

*ChatBot:* He is involved in the Maga Project. This project concerns the development of a multimodal interface for museum site tours.

*User:* Is someone in Maga project working on ASR systems?

*ChatBot:* The answer is Antonella Santangelo.

*User:* Which are the research activities at DINFO?

*ChatBot:* There are four main research groups, concerning Innovative Digital Computer Architecture, Image processing, Artificial Intelligence and Robotics.

Figure 7 shows the AIML category concerning a fragment of the previous dialogue. The user asks for people involved in a project and also dealing with a specific research topic. The *pattern* contains two wildcards "\*" that will be associated to the name of the project and to the research topic respectively. The tags belonging to the *template* allow the system to compose a SubL query with the values corresponding to the wildcards. As a consequence it is possible to retrieve the Cyc information satisfying the user request.

```

<category>
  <pattern> IS SOMEONE IN * PROJECT WORKING ON *
  </pattern>
  <template>
The answer is
<cycsystem>
  (cyc-query'
  (#$and (#$expertRegarding ?X
  <cycterm><star index='1' /> </cycterm>)
  (#$collaboratesIn ?X
  <cycterm><star index='2' /></cycterm>
  )))</cycsystem>
  </template>
</category>

```

Figure 7. An AIML category with wildcards

## 4.2 Interactive Area

The Interactive Area manages synchronization between different kind of user inputs and provides IR results to the user.

The system is accessible through a web page in a multimodal browser from the handheld device (Figure 8). The interaction occurs

by means of the loading of X+V pages, that can be triggered by user vocal and visual command or RFID detection.

Whenever the user submits his request visually, vocally or through RFID detection, the multimodal browser loads a new page built with the Rational Area reply.

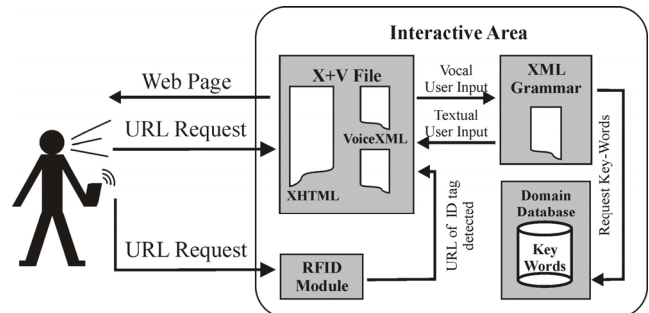


Figure 8. Detail of Interactive Area

Vocal inputs are obtained by means of the speech recognition process which is carried out through an ad-hoc built-in speech grammar.

The grammar has been built using XML Form from a set of keywords about the domain. These keywords are related to the ontology categories and are stored in a database. Besides, a particular domain-restricted clever grammar increases the user freedom of expression, enhancing the naturalness of the dialogue.

Keywords are extracted from the database and put into the grammar to build spoken user utterances. When the speech matching occurs, the application submits the recognized query to the ChatBot and waits for the answer. In the same way the application submits the query associated to the element selected with the touch-pen on the PDA screen.

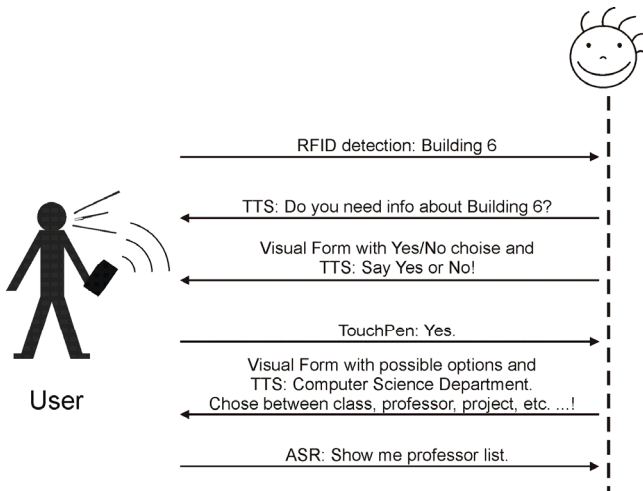
The interaction between the application running on the PDA and the system is also triggered by the detection of a RFID tag, which is used to estimate the PDA position within the environment. We have passive tags each one holding an unique ID, and each one attached to a point of interest. Once the RFID reader on the PDA detects a tag, its unique ID is passed to the server application over the wireless network. The link between the information sent to the PDA and the user's position is given by the detected ID.

This allows the ChatBot to start the position-related interaction with the user, who can go on asking the ChatBot with further questions or discard the information.

### 4.2.1 Multimodal Interaction

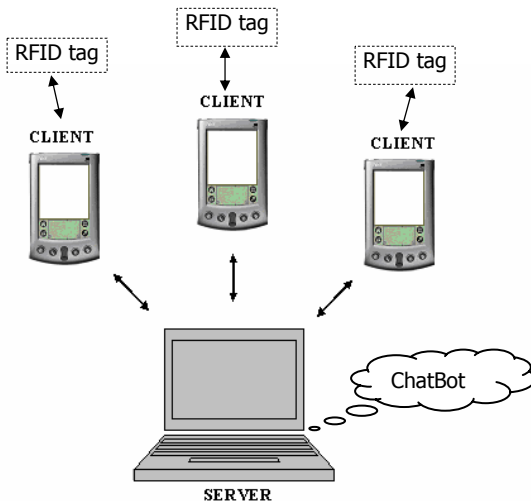
The Interaction Area synchronizes multimodal inputs during the user navigation. Figure 9 shows an example of an interaction between the user and the proposed system, in which all three available access modalities are involved.

The first step is to tag any point of interest within a given environment. People walk across such an environment with a PDA running our client application (Figure 10). When an RFID tag is detected by the reader module on the PDA, data read from the tag is passed to the PDA to be processed. Then the client application running on the PDA sends the detected tag ID to the server via the wireless connection. This allows the ChatBot to start the interaction with the PDA by providing it with some basic information related to its current position (Figure 9).



**Figure 9. Example of user-system interaction**

When a tag is detected, the RFID Module sends a query to the ChatBot about the selected item specifying that the input channel is the RFID. The ChatBot does not reply with information about the item but allows the user to accept interaction on the detected item or continue his previous dialogue. In the interaction shown in Figure 9 the user accepts the system proposal using traditional point-and-click visual interface. Hence, the ChatBot starts interaction about the detected item which in the specific example is the Building #6 inside the Campus. The system provides a short description of the building and a list of available information. The interaction goes on via the voice channel by which the user asks the ChatBot for more information about professors list. After every user request, the system builds web pages with contents retrieved by the Rational Area. Such contents can be visually and vocally navigated according to the user choices.



**Figure 10. RFID-based query architecture**

## 5. DISCUSSION AND FUTURE WORK

A multimodal pervasive user-friendly system has been presented to assist students to obtain information about the “Dipartimento di Ingegneria Informatica” (Computer Engineering Department, DINFO) domain. The system takes the advantages of ChatBot,

speech recognition/synthesis and RFID technologies embedded in a Personal Digital Assistant, which can be easily carried during the site tour.

The implemented information provision service has been tested by 115 students within our department, during a test period that lasted 30 days.

The Human-System interaction takes place by means of speech-to-text recognition (ASR) and text-to-speech synthesis (TTS). Research results in TTS are more ahead than in ASR due to asymmetries which are inherent the speech production and interpretation. Our proposed system makes use of a speaker independent voice recognition engine. At the present time, adaptive processing of multimodal interfaces is still in its infancy.

Informal studies on the feedbacks gave us by trial users of our multimodal guide, shows that 67% of them had successful interaction experience. In addition, observations revealed that the system behaved according to its specifications in 89% of the cases, with a speech recognition success at first time of 81%, over a total of about 1700 vocal questions.

Performance limitations, associated with slower processing speeds and memory decline, are mostly referred to users that speaks a bad English. However, it has to be noticed that users with bad English pronunciation, after a short practice with the guide by means of the on-line vocal help, improve their pronunciation and achieve better results. During our tests, we saw that users never changed the subject of the conversation before obtaining the needed information, and the maximum number of trials with the same question was five. This shows that the spoken interaction establishes a sort of “feeling” between the human users and the system, thus making users more patient and favorably disposed towards the system speech recognition failures.

In these preliminary studies we have also tested the system in noisy outdoor/indoor environments. When the noise source is not related with the human speech, the system performances decrease almost linearly with the drop in S/N ratio. We had different results when the noise – intended as non-wanted signal – is given by human talks around the microphone. With a crowded situation of two or three people talking each other within the microphone range, the system performance decays dramatically. We expected this result since the system is not tuned to a given voice, so it tries to recognize all words contained in the signal read by the microphone. In this cases, users spontaneously and instinctively talked closer to the microphone, thus exploiting the Automatic Gain Control (AGC) feature present on almost all audio acquisition devices.

We received positive feedback from students who found this way to access information more natural than traditional ones, such as info-points on fixed kiosks in which information are requested and presented by means of a Web browser. They reported that they felt the same experience as if they talked with human employee, with the advantage that they have not to reach the information desk, and they have not to wait any line to be served.

Confirming this informal result, the system log files showed that we had an average of 79 accesses per day. The test period lasted 30 days, and we saw a rapid growth from nine accesses the first day up to a peak of 148 accesses in 120 minutes on the fourth day. From the third day on, we had a minimum of 23 accesses per day. These results are remarkable if compared with the complete novelty of the system within our department. Furthermore, half of students who tested our system never used the corresponding information provision service given by human employees, for the reasons above described.

The proposed system is a first prototype intended as a smart replacement of the outdated traditional audio/visual pre-recorded guides. Future work will regard the enhancement of the system improving the interaction naturalness.

## 6. ACKNOWLEDGMENTS

This work has been partially funded by the Italian “Programma di Rilevante Interesse Nazionale” (PRIN) 2005, contract no. 2005103830\_002, entitled: "Artificial Intelligence Techniques for Processing Analysis, Preservation and Retrieval of Spoken Natural Language Archives".

## 7. REFERENCES

- [1] Alice: <http://www.alicebot.org>
- [2] Andrew, H. and McGlashan, S.: Speech Recognition Grammar Specification Version 1.0, W3C Recommendation, (2004)
- [3] Aridor, Y., Carmel, D., Maarek, Y. S., Soffer, A., and Lempel, R. 2002. “Knowledge encapsulation for focused search from pervasive devices.” *ACM Trans. Inf. Syst.* 20, 1 (Jan. 2002), 25-46.
- [4] Axelsson, J., Cross, C., Ferrans, J., McCobb, G., Raman, T. and Wilson, L.: XHTML+Voice profile 1.2, W3C Note. (2004)
- [5] Barnett, J., Bodell, M., Raggett, D. and Wahbe, A.: Multimodal Architecture and Interfaces, W3C Working Draft. (2006)
- [6] Berhe G., Brunie L., Pierson J.M., “Modeling Service-Based Multimedia Content Adaptation”, Proceedings of the 1st ACM Conference on Computing Frontiers CF '04, Ischia, Italy 2004, pp. 60-69
- [7] Black B., E. A. Larsson, “Beyond simple e-mail: upgrading an entire campus to enterprise e-mail and calendaring with groupwise”, in Proceedings of the 34th annual ACM conference on User services SIGUCCS '06, November 2006, Edmonton, Alberta, Canada, Pages: 24 – 27
- [8] Chen A., “Network Based Storage: Getting Students to Actually Use It”, in Proceedings of the 34th annual ACM conference on User services SIGUCCS '06, November 2006, Edmonton, Alberta, Canada, pages: 55-59.
- [9] Decker C., Kubach U., and Beigl. M., “Revealing the retail black box by interaction sensing”, in Proceedings of the 23rd Distributed Computing Systems Workshops (ICDCS 2003), 19-22 May 2003 Page(s): 328 – 333
- [10] Fujimura N., D. Masairo, “Collecting Students’ Degree of Comprehension with Mobile Phones”, in Proceedings of the 34th annual ACM conference on User services SIGUCCS '06, November 2006, Edmonton, Alberta, Canada, pages: 123-127
- [11] Genco, A., Sorce, S. and Varrica, E.: “Combined People Position and Compass Detection for Context-aware Service Provision”, proc. of IEEE 3rd International Workshop on Embedded Computing IWEC-06, Columbus, Ohio, USA, 14 Aug 2006, pp 368-375
- [12] Genco A., Sorce S., Reina G., Santoro G., “An Agent-Based Service Network for Personal Mobile Devices”, IEEE Pervasive Computing, vol. 5, no. 2, Apr-Jun 2006, pp. 54-61
- [13] Genco A., S. Sorce, G. Reina, G. Santoro, R. Messineo, R. Raccuglia, L. Lo vecchio, G. Di Stefano, “An Augmented Campus Design for Context-aware Service Provision”, Proceedings of the 33rd annual ACM SIGUCCS Conference on User Services, Monterey, CA, Nov. 6-11 2005, pp. 92-97
- [14] Grinter, R. E., and Eldridge, M., “Y do tngrs luv 2 txt msg?”, in Proceedings of the Seventh European Conference on Computer-Supported Cooperative Work (ECSCW'01), Bonn, Germany, 16-20 September, 2001, Kluwer Academic Publishers, Dordrecht, Netherlands, 219-238.
- [15] Henrysson A., M. Ollila, M. Billinghurst, “Mobile phone based AR scene assembly”, Proceedings of the 4th International Conference on Mobile and Ubiquitous Multimedia, Christchurch, New Zealand 2005, pp. 95-102
- [16] IBM Multimodal Tools: <http://www-306.ibm.com/software/pervasive/multimodal/>
- [17] Kino Coursey, Daxtron Laboratories, Inc., Living in CYN: mating AIML and CYC together with Program N. 2004.
- [18] McGlashan, S., Burnett, D. C., Carter, J., Danielsen, P., Ferrans, J., Hunt, A., Lucas, B., Porter, B., Rehor, K. and Tryphonas, S.: Voice Extensible Markup Language (VoiceXML) Version 2.0, W3C Recommendation, (2004)
- [19] Murnan, C.A., “Expanding Communication Mechanisms: They’re Not Just E-Mailing Anymore”, in Proceedings of the 34th annual ACM conference on User services SIGUCCS '06, November 2006, Edmonton, Alberta, Canada, pages: 267-272
- [20] Ni, L.M.; Yunhao Liu; Yiu Cho Lau; Patil, A.P.: LANDMARC: indoor location sensing using active RFID, Proceedings of the First IEEE International Conference on Pervasive Computing and Communications (PerCom 2003), 23-26 March 2003 Page(s): 407-415
- [21] OpenCyc: <http://www.opencyc.org>
- [22] Pilato, G., Augello, A., Santangelo, A., Gentile, A. and Gaglio, S.: An Intelligent Multimodal Site-guide for the “Parco Archeologico della Valle dei Templi in Agrigento, Proc. of First European Workshop on Intelligent Technologies for Cultural Heritage Exploitation at The 17th European Conference on Artificial Intelligence, Riva del Garda, Italy, August, 45-49 (2006)
- [23] Price, S.; Summers, R., “Mobile Healthcare in the Home Environment”, Proceedings of 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBS '06, New York, NY, Aug. 2006, pp. 6446-6448
- [24] Raptis D., Tselios N., Avouris N., “Context-based design of mobile applications for museums: a survey of existing practices”, Proceedings of the 7th ACM International Conference on Human-Computer Interaction with Mobile Devices & Services, Salzburg, Austria 2005, pp: 153-160
- [25] Wan. D., “Magic medicine cabinet: A situated portal for consumer healthcare.”, In Proceedings of the International Symposium on Handheld and Ubiquitous Computing, 1999, Lecture Notes in Computer Science, vol. 1707, pag. 352
- [26] Want R., Fishkin K.O., Gujar A., and Harrison B.L., “Bridging physical and virtual worlds with electronic tags.”, In Proc. ACM CHI '99, Pittsburgh, PA, May 15-20, 1999, pp. 370–377
- [27] Weiser M.: The Computer for the Twenty-First Century, Scientific American, pp. 94-100, (1991).