

# The use of WAP Technology in Question Answering

Fernando Zacarías F.<sup>1</sup>, Alberto Tellez V.<sup>2</sup>, Marco Antonio Balderas<sup>3</sup>,  
Guillermo De Ita L., and Barbara Sánchez R.<sup>4</sup>

Benemérita Universidad Autónoma de Puebla,  
<sup>1,3,4,5</sup>Computer Science and <sup>2</sup> Collaborator - INAOE  
14 Sur y Av. San Claudio, Puebla, Pue.  
72000 México

<sup>1</sup>fzfflores@yahoo.com.mx, <sup>2</sup>albertotellezv@ccc.inaoep.mx  
<sup>3</sup>balderasespmarco@gmail.com, <sup>4</sup>brinza@hotmail.com

**Abstract.** The experience of Puebla Autonomous University on using WAP technology in the development of novel applications is deployed. The goal is to enhance question answering through innovative mobile applications providing new services and more efficiently. The architecture proposed based on WAP protocol, moves the issue of Question Answering to the context of mobility. This paradigm ensures that QA is seen as an activity that provides entertainment and excitement. This characteristic gives to Question Answering an added value. Furthermore, the method for answering definition questions is very precise. It could answer almost 95% of the questions; moreover, it never replies wrong or unsupported answers. Considering that the mobile-phone has had a boom in the last years and that a lot of people already have mobile telephones (approximately 3.5 billions), we propose a new application based on Wikipedia that makes Question Answering something natural and effective for work in all fields of development. This obeys to that the new mobile technology can help us to achieve our perspectives of growth. This system provides to user with a permanent service in anytime, anywhere and any device (PDA's, cell-phone, NDS, etc.). Furthermore, our application can be accessed via Web through iPhone and any device with internet access.

**Keywords:** Mobile devices, Question Answering, WAP, GPRS.

## 1 Introduction

Each generation of mobile communications has been based on a dominant technology, which has significantly improved spectrum capacity. Until the advent of IMT-2000, cellular networks had been developed under a number of proprietary, regional and national standards, creating a fragmented market.

- First Generation was characterized for Advanced Mobile Phone System (AMPS). It is an analog system based on FDMA (Frequency Division Multiple Access) technology. However, there were also a number of other proprietary systems, rarely sold outside the home country.

- Second Generation, it includes five types of cellular systems mainly:
  - Global System for Mobile Communications (GSM) was the first commercially operated digital cellular system.
  - GSM uses TDMA (Time Division Multiple Access) technology.
  - TDMA IS-136 is the digital enhancement of the analog AMPS technology. It was called D-AMPS when it was first introduced in late 1991 and its main objective was to protect the substantial investment that service providers had made in AMPS technology.
  - CDMA IS-95 increases capacity by using the entire radio band with each using a unique code (CDMA or Code Division Multiple Access)
  - Personal Digital Cellular (PDC) is the second largest digital mobile standard although it is exclusively used in Japan where it was introduced in 1994.
  - Personal Handyphone System (PHS) is a digital system used in Japan,
- Third Generation, better known as 3G or 3rd Generation, is a family of standards for wireless communications defined by the International Telecommunication Union, which includes GSM EDGE, UMTS, and CDMA2000 as well as DECT and WiMAX. Services include wide-area wireless voice telephone, video calls, and wireless data, all in a mobile environment. Thus, 3G networks enable network operators to offer users a wider range of more advanced services while achieving greater network capacity through improved spectral efficiency.

Currently, mobile devices are part of our everyday environment and consequently part of our daily landscape [5]. The current mobile trends in several application areas have demonstrated that training and learning no longer needs to be classroom. Current trends suggest that the following three areas are likely to lead the mobile movement: m-application, e-application and u-application. There are estimated to be 2.5 billion mobile phones in the world today. This means that this is more than four times the number of personal computers (PCs), and today's most sophisticated phones have the processing power of a mid-1990s PC. Even, in a special way, many companies, organizations, people and educators are already using iPhone, iPod, NDS, etc., in their tasks and curricula with great results. They are integrating audio and video content including speeches, interviews, artwork, music, and photos to bring lessons to life. Many current developments, just as ours [5, 3, 6], incorporate multimedia applications.

In the late 1980's, a researcher at Xerox PARC named Mark Weiser [4], coined the term "Ubiquitous Computing". It refers to the process of seamlessly integrating computers into the physical world. Ubiquitous computing includes computer technology found in microprocessors, mobile phones, digital cameras and other devices. All of which add new and exciting dimensions to applications.

As pragmatic uses grow for cellphones, mobile technology is also expanding into creative territory. New public space art projects are using cellphones and

other mobile devices to explore new ways of communicating while giving everyday people the chance to share some insights about real world locations.

While your cellphone now allows you to play games, check your e-mail, send text messages, take pictures, and oh, yeah, make phone calls, it can perhaps serve a more enriching purpose. Thus, we think that widespread internet access and collaboration technologies are allowing businesses of all sizes to mobilise their workforce. Such innovations provide additional flexibility without the need to invest in expensive and complex on-premise infrastructure requirements. Furthermore, it makes “eminent sense“ to fully utilise the web commuting options provided by mobile technology.

The problem of answering questions has been recognized and partially tackled since the 70’s for specific domains. However, with the advent of browsers working with billions of documents in internet, the need has newly emerged, having led to approaches for open-domain QA. Some examples of such approaches are emergent question answering engines such as *answers.com*, *ask.com*, or additional services in traditional browsers, such as *Yahoo*.

Recent research in QA has been mainly fostered by the TREC and CLEF conferences. The first one focus on English QA, whereas the second evaluates QA systems for most European languages except English. To do, both evaluation conferences have considered only a very restricted version of the general QA problem. They basically contemplate simple questions which assume a definite answer typified by a named entity or noun phrase, such as factoid questions (for instance, “How old is Cher?” or “Where is the Taj Mahal?”) or definition questions (“Who is Nelson Mandela?” or “What is the quinoa?”), and exclude complex questions such as procedural or epaculative ones.

Our paper is structured as follows: In section 2 we describe the state of the art about QA and similar works. Next, we present the method for question answering for definitions questions in section 3. After, in section 4 we present the WAP technology as support for our mobile application. Section 5 shows our application on the two variants, WiFi and WAP protocol. Section 6 describe our perspectives about our future work. Finally, the conclusions are drawn in section 7.

## 2 The state of the art

One of the oldest problems of human history is raising questions about several issues and conflicts that torments our existence. Since children this is the mechanism we use to understand and adapt to our environment. The counterpart to ask questions is to answer the questions that we do, an activity that also requires intelligence. This activity has a difficulty level that has tried to delegate to computers, almost since the emergence of these. The issue of question an-

swering for a computer has been recognized and tackled from the decade of the 70s century past for specific domains. In Mexico, have been obtained excellent results in this context, for this reason we propose to bring these same results with mobile technologies.

Recent research has focused on developing systems for question answering to open domain, ie systems that takes as their source of information a collection of texts on a variety of topics, and solve questions whose answers can be obtained from the collection of departure. From question answering systems developed so far, we can identify three main phases:

1. *Analysis of the question.* This first phase will identify the type of response expected from the given question, that is expected to be a question of "when" a kind of response time, or a question "where" will lead us to identify a place. Response rates are most commonly used personal name, name organization, number, date and place.
2. *Recovery of the document.* In the second stage performs a recovery process on the collection of documents using the question, which is to identify documents on the question that probably contain the kind of response expected. The result of this second stage is a reduced set of documents and preferably specific paragraphs.
3. *Extraction of the response.* The last phase uses the set of documents obtained in the previous phase and the expected type of response identified in the first phase, to locate the desired response.

Questions of definition require a more complex process in the third stage, since they must obtain additional information segments and at the same time are not repetitive. To achieve a good "definition" must often resort to various documents [1].

Currently the question answering on mobile devices for open domains is in a development stage. The project QALL-ME, is a project of 36 months, funded by the European Union and will be conducted by a consortium of seven institutions, including four academic and three industrial companies. The aim is to establish a shared infrastructure for developing a QA infrastructure via mobile phone for any tourist or citizen can instantly access to different information regarding the services sector, be it a movie in the cinema, a theater or restaurant of a certain type of food. All this in a multilingual and multimodal mode for mobile devices. The project will experiment with the potential of open domain QA and evaluation in the context of seeking information from mobile devices, a multimodal scenery which includes natural speech as input, and the integration of textual answers, maps, pictures and short videos as output.

The architecture proposed in the QALL-ME project is a distributed architecture in which all modules are implemented as Web services using standard language for defining services. In figure 1 shows the main modules of this architecture. The architecture of the QALL-ME described as follows:

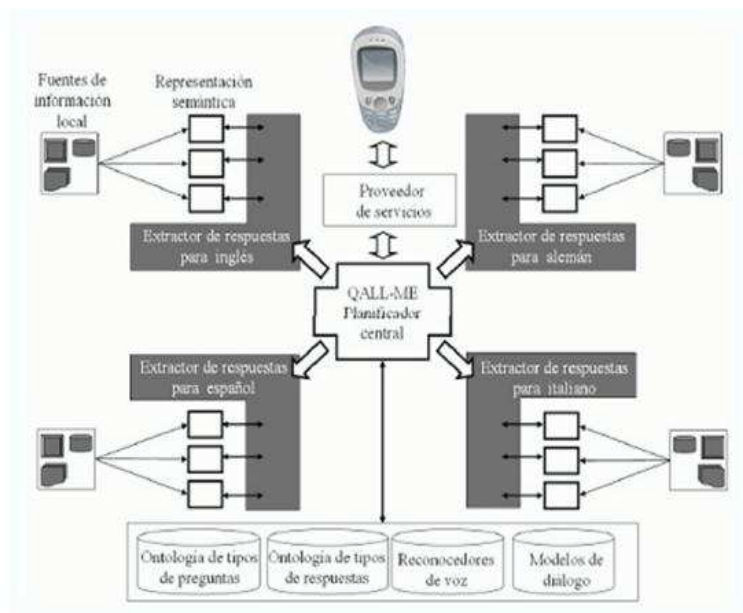


Fig. 1. Main QALL-ME Architecture [8]

“The central planner is responsible for interpreting multilingual queries. This module receives the query as input, processes the question in the language in which it develops and, according to the parameters of context, directs the search for required information. Extractor to a local response. The extraction of the response is made on different semantic representations of the information depends on the type of the original source data from which we get the answer (if the source is plain text, the semantic representation is an annotated XML document if the source is a website, the semantic representation is a database built by a wrapper). Finally, the responses are returned to the central planners to determine the best way to represent the requested information” [8].

### 3 Mobile Question Answering for Definitions Questions

The method for answering definition questions uses Wikipedia [10] as target document collection. It takes advantage of two known facts: [10] Wikipedia organizes information by topics, that is, each document concerns one single subject and, [11] the first paragraph of each document tend to contain a short description of the topic at hand. This way, it simply retrieves the document(s) describing the target term of the question and then returns some part of its initial paragraph as

answer. Figure 2 shows the general process for answering definition questions. It consists of three main modules: target term extraction, document retrieval and answer extraction.

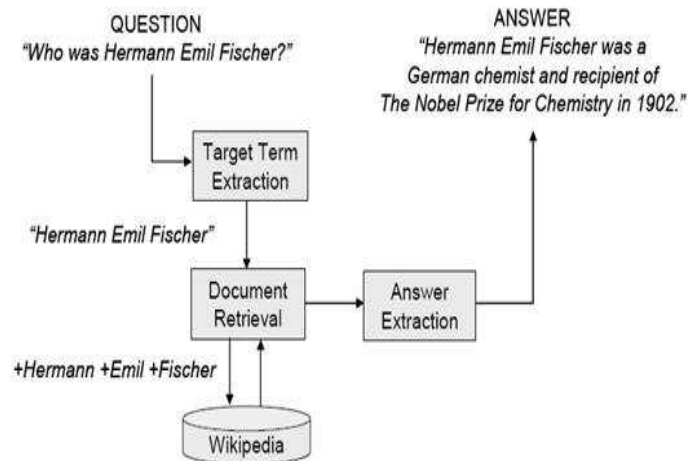


Fig. 2. Process for answer definition questions [7]

### 3.1 Finding Relevant Documents

In order to search in Wikipedia for the most relevant document to the given question, it is necessary to firstly recognize the target term. For this purpose the method uses a set of manually constructed regular expressions such as: “What—Which—Who—How” + “any form of verb to be” + <TARGET> + “?”, “What is a <TARGET> used for?”, “What is the purpose of <TARGET>?”, “What does <TARGET> do?”, etc. Then, the extracted target term is compared against all document names and the document having the greatest similarity is recovered and delivered to the answer extraction module. It is important to mention that, in order to favor the retrieval recall, we decided using the document names instead of the document titles since they also indicate their subject but normally they are more general (i.e., titles tend to be a subset of document names). In particular, the system uses the Lucene [11] information retrieval system for both indexing and searching.

### 3.2 Extracting the Target Definition

As we previously mentioned, most Wikipedia’s documents tend to contain a brief description of its topic in the first paragraph. Based on this fact, this method for answer extraction is defined as follows:

- Consider the first sentence of the retrieved document as the target definition (the answer).
- Eliminate all text between parenthesis (the goal is to eliminate comments and less important information).
- If the constructed answer is shorter than a given specified threshold<sub>2</sub>, then aggregate as many sentences of the first paragraph as necessary to obtain an answer of the desire size.

For instance, the answer for the question “Who was Hermann Emil Fischer?” (refer to Figure 2) was extracted from the first paragraph of the document “Hermann.Emil.Fischer”: “Hermann Emil Fischer (October 9, 1852 - July 15, 1919) was a German chemist and recipient of the Nobel Prize for Chemistry in 1902. Emil Fischer was born in Euskirchen, near Cologne, the son of a businessman. After graduating he wished to study natural sciences, but his father compelled him to work in the family business until determining that his son was unsuitable”.

### 3.3 Evaluation Results of our method

This section presents the experimental results about the participation [7] at the monolingual Spanish QA track at CLEF 2007. This evaluation exercise considers two basic types of questions, definition and factoid. However, this year there were also included some groups of related questions. From the given set of 200 test question, our QA system treated 34 as definition questions and 166 as factoid. Table 3.3 details our general accuracy results.

Table 1. System’s general evaluation

	Right	Wrong	Inexact	Unsupported	Accuracy
<b>Definition</b>	30	-	4	-	88.23%
<b>Factoid</b>	39	118	3	6	23.49%
<b>TOTAL</b>	69	118	7	6	34.50%

It is very interesting to notice that our method for answering definition questions is very precise. It could answer almost 90% of the questions; moreover, it never replies wrong or unsupported answers. This result evidenced that

Wikipedia has some inherent structure, and that our method could effectively take advantage of it. [7]

## 4 WAP technology in Question Answering

Wireless Application Protocol (WAP) is a secure specification that allows users to access information instantly via handheld wireless devices such as mobile phones, pagers, two-way radios, Smart phone and communicators.

WAP is designed to be user-friendly and innovative data applications for mobile phones easily. There are three types of terminals have been defined [12]:

- Feature phones, which offer high voice quality with the capability of text messaging and Internet browsing.
- Smart phones, with similar functionality but with larger display.
- The communicator, which is an advanced terminal designed with the mobile professional in mind, similar in size to a palm-top with a large display.

WAPs that use displays and access the Internet run what are called micro browsers; browsers with small file sizes that can accommodate the low memory constraints of handheld devices and the low-bandwidth constraints of a wireless-handheld network.

WAP uses Wireless Markup Language (WML), which includes the Handheld Device Markup Language (HDML) developed by Phone.com. WML can also trace its roots to eXtensible Markup Language (XML). A markup language is a way of adding information to your content that tells the device receiving the content and what to do with it. The best known markup language is Hypertext Markup Language (HTML). Unlike HTML, WML is considered a Meta language. Basically, this means that in addition to providing predefined tags, WML lets you design your own markup language components. WAP also allows the use of standard Internet protocols such as UDP, IP and XML.

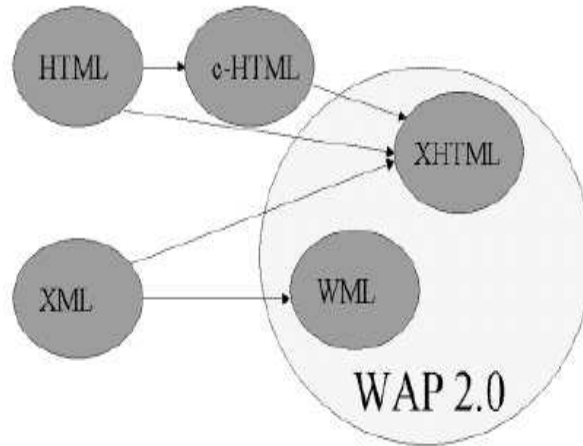
Although WAP supports HTML and XML, the WML language (an XML application) is specifically devised for small screens and one-hand navigation without a keyboard. WML is scalable from two-line text displays up through graphic screens found on items such as smart phones and communicators.

WAP also supports WML Script. It is similar to JavaScript, but makes minimal demands on memory and CPU power because it does not contain many of the unnecessary functions found in other scripting languages. Because WAP is fairly new, it is not a formal standard yet. It is still an initiative that was started by Unwired Planet, Motorola, Nokia, and Ericsson.

There are three main reasons why wireless Internet needs the Wireless Application Protocol:



## Markup language migration



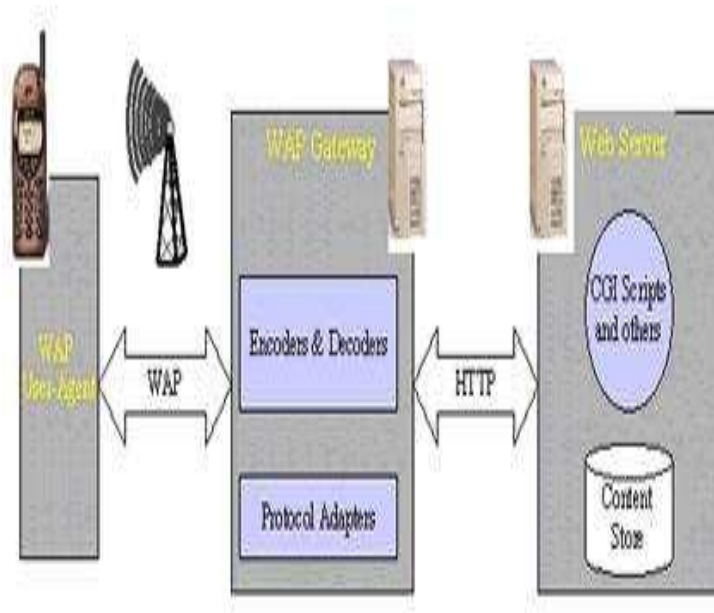
**Fig. 3.** Migration of Markup language

- Transfer speed: most cell phones and Web-enabled PDAs have data transfer rates of 14.4 Kbps or less. Compare this to a typical modem, a cable modem or a DSL connection. Most Web pages today are full of graphics that would take an unbearably long time to download at 14.4 Kbps. In order to minimize this problem, wireless Internet content is typically textbased in most cases.
- Size and readability: the relatively small size of the LCD on a cell phone or PDA presents another challenge. Most Web pages are designed for a resolution of 640x480 pixels, which is fine if you are reading on a desktop or a laptop. The page simply does not fit on a wireless device's display, which might be 150x150 pixels. Also, the majority of wireless devices use monochrome screens. Pages are harder to read when font and background colors become similar shades of gray.
- Navigation: navigation is another issue. You make your way through a Web page with points and clicks using a mouse; but if you are using a wireless device, you often use one hand to scroll keys.

WAP takes each of these limitations into account and provides a way to work with a typical wireless device.

Here's what happens when you access a Web site using a WAP-enabled device:

- You turn on the device and open the mini-browser.



**Fig. 4.** WAP Technology Infrastructure

- The device sends out a radio signal, searching for service.
- A connection is made with your service provider.
- You select a Web site that you wish to view.
- A request is sent to a gateway server using WAP.
- The gateway server retrieves the information via HTTP from the Web site.
- The gateway server encodes the HTTP data as WML.
- The WML-encoded data is sent to your device.
- You see the wireless Internet version of the Web page you selected.

Although WML is well suited to most mundane content delivery tasks, it falls short of being useful for database integration or extremely dynamic content. PHP fills this gap quite nicely-integrating into most databases and other Web structures and languages. It's possible to "crossbreed" mime types in Apache to enable PHP to deliver WML content. WML pages are often called "decks". A deck contains a set of cards. A card element can contain text, markup, links, input-fields, tasks, images and more. Cards can be related to each other with links.

When a WML page is accessed from a mobile phone, all the cards in the page are downloaded from the WAP server. Navigation between the cards is done by the phone computer (inside the phone) without any extra access communications to the server.

## 5 Application mobile

As we mentioned at the beginning, our proposal is the combination of mobile technologies and web technologies. First, we have developed a mobile application (as you can see in figure 5) based on WAP technology. This application allows users to use at anytime and anywhere at very low cost, 2 cents per search. Furthermore, this application is available for most types of mobile phones. The figure 5 shows the main interface, as well as the request and response from the user's search.

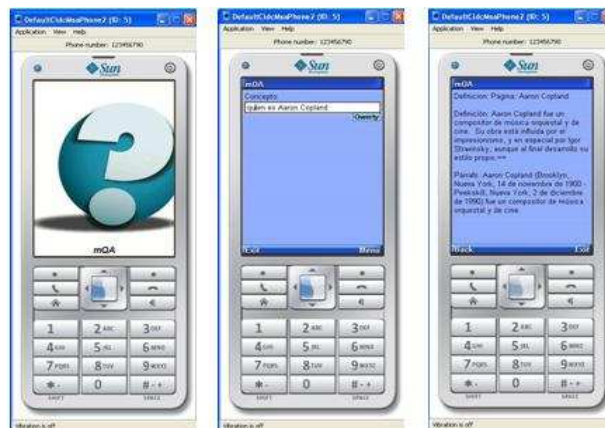


Fig. 5. Mobile application through WAP Technology

On the other hand, the figure 6 shows how our application mQAB can be accessed from web via iPhone through Wi-Fi. This is another channel of access to our application via wireless network. This feature allows our application covering all existing wireless and mobile devices.

## 6 Perspectives and Future work

People throughout the world are increasingly relying on cell phones and mobile devices to keep them plugged in. Obviously, search will play an ever increasing role in the evolution of mobile. When will mobile search surpass desktop search? We have been expecting better search capabilities from mobile devices for some time, and know that Asia is far ahead of North America in this respect at the current time. Today, experts discuss their views about the evolution of search in North America. And, what we are sure, is that we must continue working on this line. For this purpose, the next phase of development is the implementation of the Mobile Question Answering System for Spanish and English. Furthermore, we



**Fig. 6.** Mobile application through iPhone

seek the application of such search in some opportunity niches such as education.

To sum up the results expected from our architecture presented in this article are:

- Architecture presented here, unlike other proposals based on short text messages [2] is cheaper, such as was presented in section 4.
- Our proposal gives a better performance because the communication via WAP is much more reliable than that based on SMS. This is mainly due to SMS-based systems have a 80 percent certainty. While the WAP protocol provides a 100 percent reliability.
- Our proposal makes use of only a servlet on the server side and a simple midlet on the side of mobile device.
- Furthermore, our proposal will benefit from the availability of Spanish WIKIPEDIA.
- Finally, our proposal is based on Java Micro Edition, thus it will be independent of Operating Systems (OS).

## 7 Conclusions

A consortium of companies are pushing for products and services to be based on open, global standards, protocols and interfaces and are not locked to proprietary technologies. The architecture framework and service enablers will be independent of Operating Systems (OS). There will be support for interoperability of applications and platforms, seamless geographic and intergenerational roaming. Mobile architecture proposed in this paper has the advantage of being adaptable to any system and infrastructure, following the current trend that mobile technologies demand.

We believe the selection of topics covered in encyclopedias like WIKIPEDIA for a language is not universal, but reflects the salience attributed to themes in a particular culture that speaks the language. Our approach also would benefit from the availability of the Spanish WIKIPEDIA and the English WIKIPEDIA.

## 8 Acknowledgments

Thank you very much to the Autonomous University of Puebla for their financial support. This work was supported under project VIEP register number 15968. Also, we thank the support of the academic body: Sistemas de Informacin.

## References

1. A. Lopez. La busqueda de respuestas, un desafio computacional antiguo y vigente. La jornada de Oriente <http://ccc.inaoep.mx/cm50-ci10/columna/080721.pdf>, 1(1):1-2, July 2008.
2. L. Jochen, The Deployment of a mobile question answering system. Search Engine Meeting. Boston, Massachusetts, 1(1), April 2005.
3. F. Zacaras Flores, F. Lozano Torralba, R. Cuapa Canto, A. Vzquez Flores. English's Teaching Based On New Technologies. The International Journal of Technology, Knowledge & Society, Northeastern University in Boston, Massachusetts, USA. ISSN: 1832-3669, Common Ground Publishing, USA 2008.
4. Weiser, M. (1991). The computer for the twenty-first century. Scientific American, September, 94-104.
5. Zacarías F., Sánchez A., Zacarías D., Méndez A., Cuapa R. FINANCIAL MOBILE SYSTEM BASED ON INTELLIGENT AGENTS in the Austrian Computer Society book series, Austria, 2006.
6. F. Zacaras Flores, R. Cuapa Canto, F. Lozano Torralba, A. Vzquez Flores, D. Zacarias Flores. u-Teacher: Ubiquitous learning approach, pp. 9–20, june 2008.
7. Alberto Tellez, Antonio Juarez, Gustavo Hernandez, Claudia Denicia, Esau Villatoro, Manuel Montes, Luis Villasenor, INAOE's Participation at QA@CLEF 2007, Laboratorio de Tecnologas del Lenguaje, Instituto Nacional de Astrofisica, ptica y Electronica (INAOE), Mexico.
8. Izquierdo R., Ferrndez O., Ferrndez S., Toms D., Vicedo J.L., Martinez P. and Surez A. QALL-ME: Question Answering Learning technologies in a multiLingual and multiModal Envinroment, Departamento de Lenguajes y Sistemas Informticos, Universidad de Alicante.
9. <http://java.sun.com/developer/technicalArticles/javaserverpages/wap>
10. <http://ilps.science.uva.nl/WikiXML/database.php>
11. <http://lucene.apache.org/>
12. J. AlSadi, B. AbuShawar, *MLearning: The Usage of WAP Technology in E-Learning*, International Journal of Interactive Mobile Technologies/Vol. 3, (2009)