



Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

GAT: Platform for automatic context-aware mobile services for m-tourism

M.C. Rodriguez-Sanchez^{a,*}, J. Martinez-Romo^b, S. Borrromeo^a, J.A. Hernandez-Tamames^a^a Department of Electronics, Universidad Rey Juan Carlos, Madrid, Spain^b NLP & IR Group, Departamento de Lenguajes y Sistemas Informáticos, UNED, Madrid, Spain

ARTICLE INFO

Keywords:

Web technologies
 Tourist information system
 Ubiquitous computing
 Mobile computing
 Collaborative process integration
 Semantic web services
 Emerging technologies

ABSTRACT

Despite the recent advances in mobile tourism systems, most of the wayfinding applications have still to deal with some problems: a huge amount of tourist information to manage, guidance for indoor and outdoor environments, and the need of users to have programming knowledge about many mobile phone platforms. In this study, we propose the GAT platform to overcome these problems. In GAT, users are able to generate wayfinding applications for indoor and outdoor environments through a web form without the need for programming skills, assisted by a system of automatic generation and update of points of interest.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

The widespread availability and portability of mobile phones has led them to become the de facto platform for ubiquitous computing. The emergence of new services is a consequence of the mobile devices ubiquity. The use of mobile technologies in some scenarios, such as tourism, has increased in the last couple of years. Tourism is an activity of great importance. As mobile battery life and capabilities of phones continue to grow, there are supporting increasingly complex tourist applications that leverage information about wayfinding functionalities. This kind of service should be adapted to the environment with high levels of heterogeneity (network topology, physical connections, objects, devices, and user preferences). This heterogeneity should be constantly recomputed. For this purpose, ubiquitous computing has emerged, which according to Mark Weiser, can be described as “by making many computers available through the physical environment, while making them effectively invisible to the user” (Lee, Lim, & Kim, 2009; Weiser, 1993).

There has also been work that has studied tourism and drawn implications specifically for the design of these mobile technologies. New technologies for mobile tourism (m-tourism) systems have emerged. The term m-tourism is related to the tourism industry and new technologies to obtain tourist information.

In this study we present GAT platform, a solution for fast building and automatic knowledge management of context-aware services for tourism in a novelty way, for indoor and outdoor environments. Our proposal consists of four key issues: generation and updating wayfinding mobile applications, automatic knowledge management, multiplatform architecture available as native

application, indoor and outdoor location and technologies applied to wayfinding functionalities (GPS, aGPS, 3G, Wi-Fi, Bluetooth and Qr-code compatibility). We explain how Information and Communication Technologies (ICTs) are being applied into the GAT platform to manage tourism applications more effectively. Our design goals are: multiplatform support, extensibility of sensor and information functionalities, data accessibility for m-tourism, user experience and not knowledge experience in Smartphones.

There are many proposals that included the use of one or several functionalities described above. However, there are currently no systems that encompass all these functionality in one system like GAT.

1.1. State of art

Some projects related to context-aware and wayfinding mobile applications use multimedia technologies in the fields of cultural organizations for interactive tourism (Lehn & Heath, 2003; Wilson, 2004; Woodruff, Aoki, Hurst, & Szymanski, 2004). Mobidenk (Baldzer et al., 2004), Lola (Michlmayr, 2002), CRUMPET (Poslad et al., 2001), Gullivers Genie (Hristova, O'Hare, & Lowen, 2003) and GUIDE (Cheverst, Davies, Mitchell, Friday, & Efstratiou, 2002) provide information delivery services for a far more heterogeneous tourist population. They use location, device, network and user context properties. The most of them use GPS and web services for outdoor location. The most of projects described before collect information manually. In this sense, the tourist information management is another issue to improve tourist services efficiently. Another researches work in automatic knowledge management for tourism.

Some projects such as (Rayman-Bacchus & Molina, 2001) have been proposed to offer web service technologies to integrate disparate Internet tourist resources. However, they did not focus

* Corresponding author. Tel.: +34 914888238.

E-mail address: cristina.rodriguez.sanchez@urjc.es (M.C. Rodriguez-Sanchez).

on tourist requirements or a software development perspective contrast to (Chiu, Yueh, Leung, & Hung, 2001; García-Crespo, López-Cuadrado, Colomo-Palacios, González-Carrasco, & Ruiz-Mezcua, 2011). Besides intelligent software agents can run on these devices and can provide personalized assistance to tourists during their trip. Together with traditional information agents such as hotel broker agents, tour planning agents, and other disparate tourist resources, they form a Multi-Agent Information System (MAIS) for collaborative and intelligent assistance to tourists (Chiu, Cheung, & Leung, 2005). Although there are a great amount of tourist information and service resources available on the Internet, it is still difficult to choose the most suitable resource. MAIS uses an agent cluster comprises several types of agents to achieve the goals of the major tasks of a tourist's trip, such as, information gathering, preference matchmaking, planning and mobile services.

Another projects using crawling methods in order to improve this knowledge management (Herrero et al., 2012; Kim, Kim, Gautam, & Lee, 2005). The application of technologies based on the semantic web using crawling can improve access to tourist information integrated from multiple sources of information scattered. Contur platform (Herrero et al., 2012) and Tour Guide System (Kim et al., 2005) use crawling to obtain tourist information from Internet. ConTur makes recommendations about a tourist destination. Tour Guide System is a web application that it makes crawling on demand to obtain context-aware information. They only offer wayfinding applications for outdoor environments. However, their solutions are not applicable for indoor environments. Besides, both solutions require a permanent network connection to receive information, thus taking a longer response time.

The most previous solutions require rich multimedia content in highly interactive interface using web services. To minimize user hesitance and enable affordable usage, the access to content and services should not require constant network connection. The solution for these problems could be presented as a wayfinding application with static information (Cheverst, Davies, Mitchell, Friday, & Efstratiou, 2000; Han, Twidale, Gutierrez, & Farivar, 2011; Lamsfus, Alzua-Sorzabal, Martin, & Lopez de Ipiña, 2010; Micha & Economou, 2005; Pielot & Boll, 2010; Pielot, Henze, & Boll, 2009), which can have the most relevant tourist information about a location, context, and user preferences.

Another important issue is to provide techniques to facilitate the development of these mobile applications. The deployment of wayfinding applications for different cities and environments could be more flexible. However, building mobile applications is a complex and time-consuming task. Some frameworks can provide supports to simplify the development tasks and some aspects of native application functionality such as (Biegel & Cahill, 2004; Economou Gavalas, Kenteris, & Tsekouras, 2008). Context Toolkit offers one solution for supporting context-aware application prototyping (Dey, Abowd, & Salber, 2001) for a specific target platform. Betelgeuse (Kukkonen & et al., 2009) is a data collection platform for mobile devices that uses a microkernel Java ME architecture with native programming languages to collect sensor data. The data can be used by local mobile applications. The Context-Phone (Raento, Oulasvirta, et al., 2005) platform follows a similar idea, yet it also supports web protocols for publishing information to SNS only for Nokia. MultiMAD (2012) is a visual tool for the generation of multi-modal applications for mobile devices. However, this tool generates code only for a specific platform: one for J2ME MIDP and another for WML. XMobile (Viana & Andrade, 2008) consists of a framework and a generation tool that facilitates the fast prototyping for pervasive applications just in JavaME mobile devices. Although these proposals include interesting features like rapid development and reusability, there are important needs of knowledge programming to use them.

Most of these proposals are oriented to make native applications for a specific Operating System (OS). In most cases it is necessary to develop a native application for each specific platform (iOS, Android, Blackberry, WindowsPhone, Symbian, etc.). Nevertheless, currently there are a big variety of mobile phone models with different specifications. One of the drawbacks in the development of mobile applications is that developers face the restrictions in the mobile phone hardware and the device's specifications (screen size, memory, processing capacity and operating system) (Chaisatien & Tokuda, 2009). They also present differences in the display properties and how users interact with them. In addition, the support of programming platforms changes from one device to another, which makes it difficult to adapt a single platform for the application development process. Thus, programmers need to have knowledge about different programming languages, since the Software Development Kits (SDKs) released by the platform creators are usually linked to a specific language. This heterogeneity between mobile devices causes that the development of a valid application for different platforms are very expensive for developers; this requires the development in different programming languages and using the APIs and platform for specific technologies.

In this way, there are several frameworks that provide technologies to develop mobile applications that can be exported to different platforms and can run on several different types of devices. To reduce the overhead caused by having to develop and maintain the same application on different mobile platforms, many developers have decided to use context-aware web services in order to build web applications (Kapitsaki, Kateros, & Venieris, 2008) or native applications. This measure not only save development costs but also ensure that their applications can be used by as many users as possible since most of the mobile phones of today include a web browser (Hernandez, 2009).

The architecture of COMPASS (Setten, Pokraev, & Koolwaaij, 2004), PinPoint (Roth, 2002) and Sightseeing4U (Baldzer et al., 2004) provides context-aware recommendations and services using web services for tourism. COMPASS builds on the open Web Architectures for Services Platforms (WASP). This kind of platform requires a permanent network connection. In this sense, SEF-AGI (Chaari & Laforest, 2005) and iConAwa (Pascual, González, Sanjuán, Pelayo, & Cueva, 2012) are tools with a web service based architecture to automatically generate mobile applications. Sense-Sation system (Shirazi, Winkler, & Schmidt, 2010) facilitates the development of web applications based on community of mobile phones. The system consists of a web based application platform. However, an additional runtime environment is installed on the phone. Nevertheless, there are many technical differences for using hardware API between native applications and web applications. One example is the management of device hardware components such as sensors, GPS, accelerometer, camera, Bluetooth and son on (Gossweiler, McDonough, Lin, & Want, 2001). Wayfinding applications could use the hardware of the Smartphones to communicate with smart objects or other Internet of thing systems (Kortuem, Kawsar, Fitton, & Sundramoorthy, 2010).

PhoneGap (Allen, Graupera, & Lundrigan, 2010) allows applications using web technologies and specific libraries that access the device hardware API. PhoneGap is an open source framework for building native mobile applications using HTML, CSS, and JavaScript for iPhone, Android, BlackBerry, Palm webOS, and Symbian WRT (Nokia). Web developers need programming knowledge in HTML5, Java and Javascripts to use PhoneGap.

Location in indoor environments is another feature that should be addressed in the guidance systems. Some mobile applications have been designed for indoor environments (Lehn & Heath, 2003; Micha & Economou, 2005; Wilson, 2004). Thus, a solution where a user can manage the tourist services for indoor and

outdoor environments is an important issue to have a complete guidance.

Finally, in many cases developers need previous experience in Smartphone technologies to use their frameworks. Some tourism entities need users with computer science skills to manage contents, marketing and wayfinding applications. However, some of these entities do not have enough economical resources to contract technical experts.

1.2. Our proposal

Considering the state of the art, there exist some restrictions in the present solutions for the generation of wayfinding mobile applications. First, mobile guide applications require specialist programmers for all platforms. That is, developers require mobile programming knowledge to extend the platforms described before. Second, in some cases the same tourist contents need to be generated or copied again for new applications. When changing some contents, the programmer should build new versions of the same application. Therefore, they need to design, implement, and validate the application once again. This requires a great deal of time, money and maintenance. The most applications require permanent network connection to access information. Lastly, the vast tourist information available on the web could be used to automatically generate contents.

To overcome these problems, in this study, we propose a new approach that allows the management of contents and tourist wayfinding applications without the need of programming. GAT is a global solution to improve the technologies for tourism combines different functionalities. This platform provides means to reduce development time and maintenance of wayfinding applications. GAT collects or manages information from Internet and generates information manually or from opinions of tourist users. Moreover, the most system uses default routes to guide users. In GAT, the user can use automatic routes or manage routes for a customized service. The guide services can be for indoor and outdoor environments. Besides, the final wayfinding applications can take the advantage of using dynamic or static information in any place. The location and guide functionalities use attributes from sensors, indoor and outdoor capabilities using the input and output capabilities of the phone. Finally, the final user can make wayfinding applications for any Smartphone in any place using a laptop, computer or SmartPhone.

Our design goals are: multiplatform support, extensibility of sensor and information functionalities, data accessibility for tourism, user experience and not knowledge experience in Smartphones.

2. Architecture

Fig. 1 illustrates the architecture of the GAT platform. On the one hand, this system can acquire autonomously tourist information from some websites (Wikipedia, DBpedia, etc.). On the other hand, GAT allows to generate native wayfinding applications using a “Manager Website”.

The generated wayfinding application by GAT provides static and dynamic tourist information related to different points of interest (POIs). A place could be different POIs. For example, a city has different POIs such as museums, theaters, restaurants, hotels, etc. Besides, in contrast to other solutions, the user does not need network connection to access this information in real time. Moreover, the generated application offers guidance in outdoor and indoor environments. For this purpose, the application interacts with the environment combining different technologies (GPS, aGPS, 3G,

Wi-Fi, Camera, sensors of the Phone, Bluetooth communication or bi-dimensional code readers).

There are two types of users for GAT. The first one is called the “Manager”. The manager can use a “Manager Website” to generate wayfinding applications in anyplace without knowledge programming for different Smartphones platforms (iOS, Android, Blackberry and Nokia). The second one is the “Tourist” who will use these generated applications for Smartphones.

The architecture of GAT (Fig. 1) consists mainly of a Multiplatform Application Generator (MAG) and an Information Management Unit (IMU). The MAG provides the functionality to generate the wayfinding applications from a “Manager Website”. The manager only needs a browser to use it. The main advantage of MAG is that does not require the user to be a specialist in programming to make wayfinding applications for indoor and outdoor environments. The manager can choose the Smartphone platforms for the application, information of several POIs and manages routes for indoor and outdoor environments.

The IMU manages a Universal Tourist Database (UTDb). A web crawling process is used to collect tourist information from the Internet, and generate points of interest using this information. This tourist information is stored and organized into an ontology following a standardized structure for this kind of information.

3. Multiplatform application generator

The Multiplatform Application Generator (MAG) allows making guided applications using a Manager Website. The main function is to create mobile applications for different targets.

3.1. Backend server – Manager Website

The manager only needs to make five steps to create a wayfinding application (see Fig. 2): introduce the location of guidance, select the points of interest related to this location, configuration for an indoor/outdoor route guidance or both, choose the mobile platform, and press the button to generate and to distribute the mobile application. The GUI of the website is mostly implemented with web technologies to make it platform independent so that it can be used across different PC desktops, PC browsers, and different mobile operating systems. This module consists of a “Manager Website” and a “Generator Module”.

The system needs to make some tasks to process the manager request to generate a wayfinding application. In the first step, the manager requests the information to the server selecting a location (city, town, country, etc.). Then, the IMU carries out a search in the UTDb in order to find the most appropriate points of interest (POI) for the selected location.

In the second step, the system shows different POIs with information (monuments, hotels, museums, bus stops, subway, etc.). The manager can choose the POIs (some or all). For this step, the system use a database agent retrieves non-spatial information from the database UTDb. A geo-spatial agent retrieves spatial information from the database about waypoints and performs a range of geo-spatial computations on that information.

In the third step, the user can select to include tracks in the wayfinding application. This application can locate user and can provide guidance to user in different environment. A route agent computes and manages routes (tracks) and their segments automatically for indoor and outdoor environments. Besides, the manager can modify these routes or add new routes. In this case, the manager could choose some waypoints in a route to indicate the path between origin and destiny, then the system generates automatically more waypoints to offer more references in the route. A map agent in the mobile applications generates and handles maps

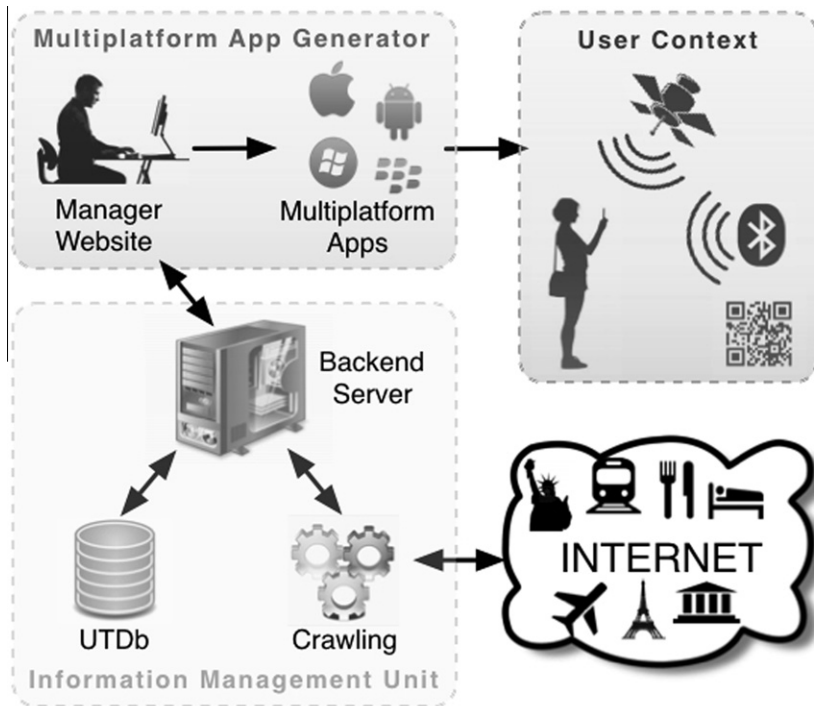


Fig. 1. Architecture of GAT: Multiplatform Application Generator and Information Management Unit.

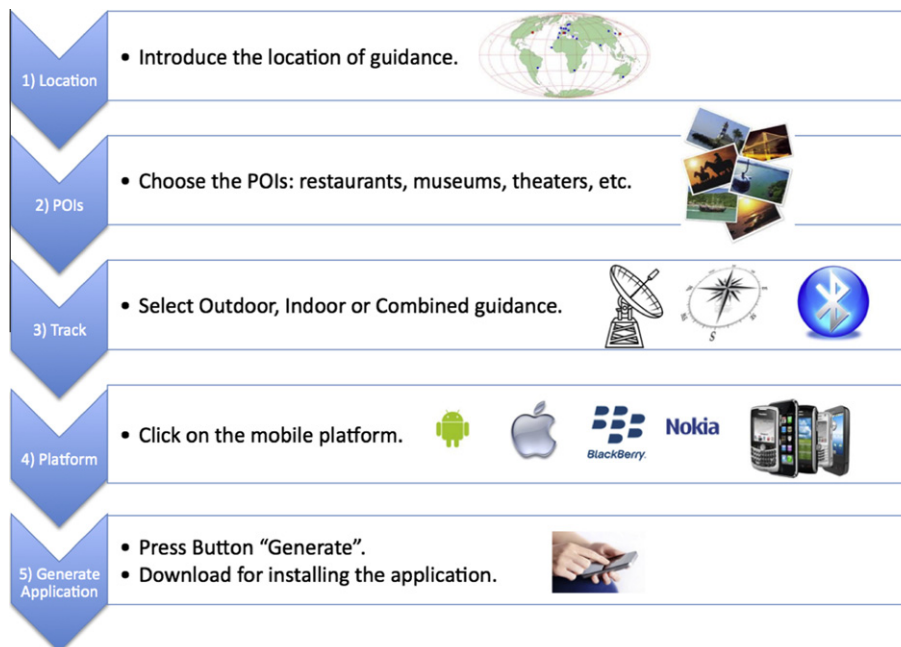


Fig. 2. Steps to generate a wayfinding application from the Manager Website.

and their visualization. GAT allows to use different types of technologies to provide location and guidance in the wayfinding application: GPS, AGPS, the camera to read bar codes or augmented reality, sensors and Bluetooth technologies. All of these functionalities combine with voice and tactile commands to provide information in the better way for each tourist and for each environment. These technologies have been chosen to offer a track more accurately, step by step, with orientation. The precision could be 1–3 m for an outdoor environment and one meter for an indoor environment. Due to GPS is unavailable for indoors, we use Bluetooth

access point (Bluegiga Technologies, 2012) and QR-codes. In this case, the managers can choose and configure Bluetooth Access Point from the Manager Website. The manager only needs to put the Media Access Control (MAC) of the Bluetooth Access Point. These devices are beacons with Bluetooth technologies, which can be installed in any place to provide indoor location. These Bluetooth devices can provide contextual information to the tourist by proximity. Moreover, the system has a QR-code for each waypoint which has been automatically generated when the manager or the crawler insert a new waypoint. The tourist only

needs the application to read the QR-code installed at a POI to receive context information. Therefore, the mobile application can receive in real time information from waypoints, location and guidance using Bluetooth, QR-code or both.

In the fourth step, the manager needs to indicate the mobile platform/s. The final application could be generated for different platforms: Android, Blackberry, Nokia or iOS. The system creates temporary database with all selected information described above. Using this temporary database, the manager can edit and modify information, pictures, and suggestions about the points of interest. These requirements are stored in a XML file.

In the five step, the user clicks on the “Generate” button. Then, the system runs the “Generator Module”. This process may take a few seconds. For example, for an application of 70 POIs could be 15 s.

3.2. The Generator Module

This module generates the wayfinding applications for different Smartphones to be sent to the tourists. This module consists of different types of template generators for Symbian, Android, Blackberry, and iOS.

As template generator consists of wayfinding files and a special script for each platform. The wayfinding files contents several generic classes and libraries for a wayfinding application. The script has the set instructions to generate the final mobile application in real time using the template files for each target. XML file with tourist information generated previously has the instructions about the functionalities for the final application. The wayfinding application uses this file in the presentation layer to process the context information and guidance.

When a GAT manager sets the generator for a platform, the script starts. For Blackberry, Android and Nokia, the Java code will be compiled and transformed into a final application on the server side using a Java compiler. The script follows instructions from the XML file to include functionalities in the final application. For iOS, we compile using an Objective-C compiler. The images and multimedia contents are embedded in it. Then, the application is pushed and dynamically distributed for downloading. The system will include the information automatically in these versions. Finally, the generated applications are uploaded to the “Application Repository”. This repository stores all the instances of the same application.

This generated wayfinding application could be modified anywhere and anytime using the Manager Website: add, update, modify or delete items, contents or functionalities. Furthermore, the XML file has the same structure for all platforms. Therefore, the migration to new mobile platforms would be automatic. We only need the template files of the new target.

4. Information Management Unit

The Information Management Unit (IMU) (Fig. 1) is responsible for managing the information system to generate mobile applications later. On the one hand, the IMU has the function to collect tourist information from the Internet, and compile and generate points of interest using this information. On the other hand, this module stores the POIs into the system. For that, a universal database has been built to store the information from the POIs. Touristic contents obtained from the Internet are organized into an ontology following a standardized structure for this kind of information.

4.1. Universal Tourist Database

In the state of the art for tourism systems, there is mainly a defect in which we want to deepen and propose a solution. There is

no standard information to represent the information from points of interest in a universal way. Thus we propose the use of an ontology that could be used for different types of applications and services for tourism. In GAT this ontology is managed and stored permanently by a module described below.

Fig. 3 shows the Universal Tourist Database (UTDb) that is the module responsible for managing an universal ontology and store new POIs such as hotels, museums, monuments, journeys, etc. For this universal access to the information, managers have to be registered in GAT previously. Then, they can access the ontology and use these contents for building any kind of wayfinding application.

All points of interest in the UTDb have a set of common attributes described in the ontology. GAT defines three types of mandatory attributes: descriptive attributes such as name, type (museum, church, hotel, public transport, etc.), description, date, and location (city, country, etc.); multimedia attributes such as images; and location attributes, such as GPS coordinates.

In addition, different points of interest (types) may have a specific set of attributes; e.g., hotels can have an attribute for room price.

4.2. UTDb input methods

The UTDb mainly has two input methods: a manual insertion of POIs carried out for a manager with tourist knowledge, and an automatic generation of POIs using a web tourist crawler that is responsible for searching relevant documents on the Internet and extract the information needed to complete the ontology represented in the UTDb. The UTDb constantly monitors the changes in its sources (webs of tourism) and updates its information accordingly.

4.2.1. Points of interest generation

There exists a web form in GAT that can be used by the users by adding new POIs manually. For each type of interest point, a web form will be presented to the manager with the corresponding fields in the ontology. In addition, when a new POI is added to the UTDb, the IMU automatically generates a QR-code associated to that POI, which will allow the interaction between the mobile application, user, and context. Once the user has finished the edition of a POI successfully, this information is stored permanently in the UTDb. This new POI, only can be modified for the manager who created it, but any manager will be able to use that information in the future.

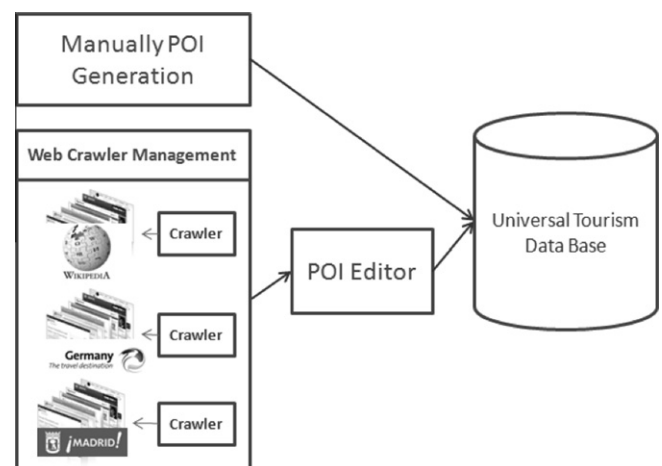


Fig. 3. Components of the Universal Tourism Database.

Manual editing of POIs is a possibility offered to the user of GAT, however managers can find and use a lot of tourist information on-line (Wikipedia, DBpedia, web of cities, etc.). We realized that it is necessary to automate these searching processes and provide to the user this information in a easy and automatic way. Thus, we propose the use of a “web tourist crawler” to integrate this new functionality and to automate the process. We believe this is one of the main contributions of our work, because even though there are many works in the literature that have used crawlers, GAT propose the combined use of a universal ontology, a central and permanent database, and a simple usage for users, fully integrated into the system.

4.2.2. Automatic approach for inserting information

The task of generating information from one interest point is, in most of the cases redundant. The number of public or private entities that generate tourist information for visitors as complete and simple as possible is increasing every day. Thus, resources keep multiplying on the Internet.

In this study, we propose a new method for automatic insertion of information in the UTDb with a triple purpose: Automation of the process of generation of POIs to avoid the presence of a person engaged in this task; assistance if manual insertion is required, accelerating the process of elaboration of the information; and finally, automatic update of the POIs (e.g., schedule of a museum or a city population). This last proposal uses a scheduler of web tourist crawlers, that is, a software agent that manages the periodic execution of different crawlers optimizing their use and integrating the updated information according to the data in the UTDb.

4.2.3. Tourist web crawling

A web crawler (Olston & Najork, 2010) is a software agent that starts with a URL or a list of URLs to visit, called seeds. When the crawler visits these URLs, it identifies all the hyperlinks in the page, adding them to the list of the URLs to be visited. These URLs are recursively visited according to a set of policies.

Web crawlers have been used for many purposes. One of them has been to analyze the structure and properties of the web; Kumar, Raghavan, Rajagopalan, and Tomkins (1999) suggested that the distribution of in-degrees (the number of links to a page) follows a power law. Chen-Chuan Chang, He, Li, Patel, and Zhang (2004) studied the “deep web,” i.e., websites whose content is not reachable via hyperlinks, but, instead, can only be retrieved by submitting HTML forms.

However, the most widespread use of web crawlers is being one of the main components of web search engines, assembling a corpus of web pages and allowing users to achieve queries against the index and find the web pages that match the queries. In modern web retrieval systems, crawlers continuously run and download pages from the web, updating the content of the document cache incrementally (Castillo, 2005).

In our approach, we have integrated the web crawler in a tourism scenario to extract tourist information from the web documents. In Fig. 4, it is shown a block diagram with the links between the modules that make up the tourist web crawler used in GAT is shown. In this figure, besides the basic modules consisting of all crawlers (downloader, parser, scheduler, and URLs queue) (Olston & Najork, 2010), the information extraction module, created explicitly for this study, is presented. Therefore, the input of the crawler is one or more web pages and the output is a POI compiled by the information extraction module and dumped in the UTDb. This web crawler works as a user acceding to the target web sites interactively. The crawler browses the website pages looking for new points of interest. During this task, it stores the unvisited hyperlinks to visit them further recursively. In the case of finding a POI not previously stored, the crawler sends the web

page content to the information extraction module and this module compiles a POI with the obtained data to be dumped in the UTDb.

4.2.4. Information extraction techniques

Considering the general structure of the UTDb, each POI must contain a number of mandatory attributes. Thus, the module responsible for the extraction of information must extract the necessary information for each POI to fulfill each of the required attributes. Three types of attributes are extracted: descriptive attributes such as name, type (museum, church, hotel, public transport, etc.), description, and location (city, country, etc.); multimedia attributes such as images; and location attributes, such as GPS coordinates, used in the guide for outdoor environments. Regarding location attributes, the information extraction module generates a QR-code unique for each POI to be used in the process of indoor guidance.

Information extraction (IE) is the name given to any process that selectively structures and combines data that are found, explicitly stated, or implied, in one or more texts. Formally, an IE task is defined by its input and extraction target. The input can be unstructured documents, such as free text that are written in natural language, or semi-structured documents that are pervasive on the web, such as tables, itemized and enumerated lists. The extraction of structure from noisy or unstructured sources is a challenging task, which has engaged a veritable community of researchers for over two decades till now. In the past few years, many approaches to IE systems, including machine learning (Shen, Doan, Naughton, & Ramakrishnan, 2007) and pattern mining techniques (Choi, Cardie, Riloff, & Patwardhan, 2005; Feldman, Rosenfeld, & Fresko, 2006), have been proposed, with various degrees of automation.

Section 5.1 describes how the different information extraction techniques can be used in a sample of tourist web crawler.

5. Evaluation

This section presents the GAT system evaluation from a qualitative and quantitatively point of view. On the one hand, the most important advantages of our system are shown by analyzing the two main modules, the module responsible for the acquisition and management of information, and the module that provides generation of mobile applications multiplatform. On the other hand, we will analyze the state of the art to compare the key features that define the context-aware mobile systems for tourism. In Section 6, we will analyze the state of the art to compare the key features that define the context-aware mobile systems for tourism.

5.1. Use case: Wikipedia web crawler

For the present study, we have developed a web crawler that deals with the retrieval of all points of interest from all the existing provinces for the Spanish version of Wikipedia. It implements the three information extraction techniques defined in Section 4.2. The web crawler browses the pages of Wikipedia looking for new points of interest, such as museums, monuments, churches, palaces, castles, sculptures, etc. Fig. 5 shows the typical structure of a sample of a Wikipedia page (“Prado Museum” in Madrid, Spain). Also, it is possible to distinguish the different types of attributes, such as name, description, and a menu on the right side called Infobox. Precisely, these three types of attributes reflect the different information retrieval tasks undertaken in this project.

In the first case, given that Wikipedia pages have a common structure, the name of the POI is always in the same place, sur-

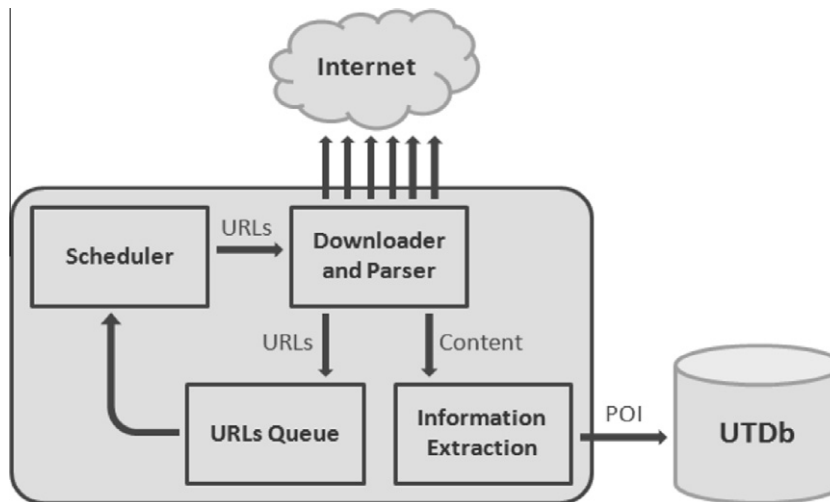


Fig. 4. Block diagram illustrating the connection of different modules that make up the crawler used in GAT 2.1.

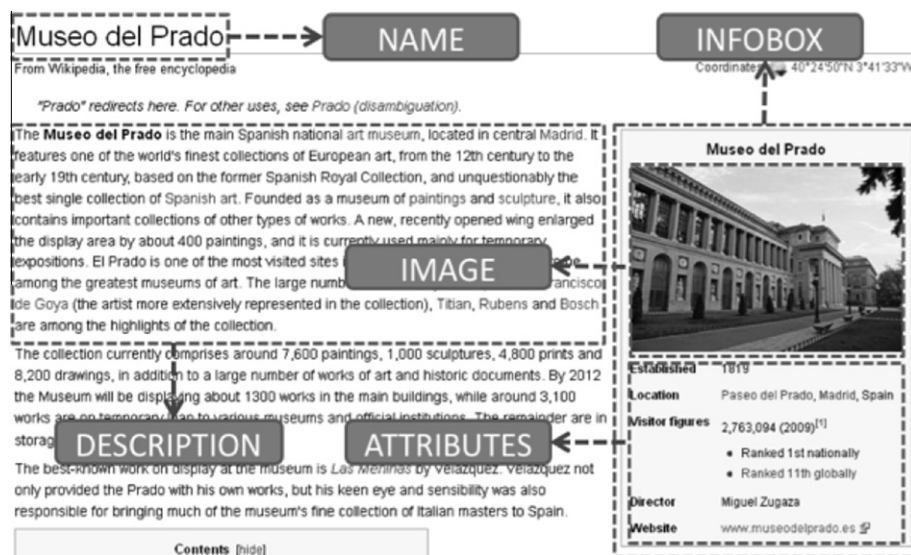


Fig. 5. "Prado Museum" page on Wikipedia.

rounded by a defined tag. Thus, this first example is the application of a structured extraction technique. Second, the description of each POI contains, in most of the cases, very useful information. However, these are fragments of text in which the information is not structured, and therefore, it is necessary to apply techniques of natural language processing such as named entity recognition (NER) to extract any information. Through the application of these techniques, trying to recognize named entities (persons, organizations or places) we have obtained features such as the number of paintings (7600) or the top artists (Velazquez, Goya, Rubens...), along with works in the museum. This information cannot be found in the Infobox because the templates used in the Wikipedia Infoboxes depend on the category of POI, and therefore, should have a fixed structure.

An Infobox, according to the definition by the Wikipedia, "is a fixed-format table designed to be added to the top right-hand corner of articles to consistently present a summary of some unifying aspect that the articles share and sometimes to improve navigation to other interrelated articles". The generalized Infobox feature

grew out of the original Taxoboxes (taxonomy Infoboxes) that editors developed to visually express the scientific classification of organisms. In this third type of source of information, it is therefore necessary to apply techniques of semi-structured information extraction to obtain attributes, such as locality, establishment, images, and GPS coordinates. This information is relatively easy to obtain because labels that define the type of information is known and therefore its extraction can be automated. These obtained features are used to help the tourist with a better context information.

On the other hand, the points of interest defined in the Wikipedia can belong to one or more categories, and these categories can be subcategories of a higher category at the same time. Therefore, one of the main objectives of the tourist web crawler is to visit and store all POIs included in a category and subcategories recursively. For instance, the Prado Museum belongs to the Buildings and structures in Madrid category, and, at the same time, this category has several sub-categories: Churches in Madrid, Palaces in Madrid, Skyscrapers in Madrid, etc.

5.2. MAG validation

This section presents the validation of the generation process of multiplatform tourist applications in GAT, focusing on the MAG unit. To this end, we will show some statistics about the GAT behavior in the last year. GAT has been tested in different activities by users outside the development team.

GAT developed 50 applications per month in 12 months. For a sample application of 25 POIs, different managers took 4 min and 30 s on average, whereas a typical wayfinding application can take 1–2 h as minimum (manually collect information, write information, design the application, and generate code). In GAT, a crawling process on a medium-sized city with about 50 POIs may take a minute at most. However, a user with wide tourist knowledge not takes less than half an hour to compile the necessary information to 50 waypoints and write them into a structured format. In addition, this user should update the information regularly, something that GAT automatically does. It is also important to note that this person or another person should have programming skills in different mobile platforms. GAT therefore not only allows a faster time for the generation of tourist applications, but also reduces the cost of generating and maintaining the application.

We used different Smartphones to validate the generated applications: Samsung Galaxy S, SII, SIII, Ace, HTC Desire and Motorola DEFY. Furthermore, we used Nokia devices, Blackberry, iPhone 3G, 4G, 4GS. Besides, the proposal is available and validated for iPhone 5.

This wayfinding application offers tourist information, location, multimedia contents, RSS news from the context, and guidance functionality for indoor and outdoor environment. The wayfinding application combines different technologies to guide the tourist in indoor and outdoor environments based on the tourist context. Thanks to this combination of technologies and the advantages of native applications to access hardware of the Phone, GAT achieves an accuracy of 1–3 m for both environments. The application guided user step by step providing static and dynamic information.

6. Discussion

Table 1 shows an in-depth comparative analysis of different current solutions to solve the problem described in this study. The main advantages of our system can be summarized as follows.

- *Architecture*: The architecture of the most solutions requires a permanent network connection to access tourist information. GAT can access dynamic and static information without permanent network connection.
- *Technologies and knowledge programming*: One of the main advantages of our system is that the manager does not need to have programming knowledge. Furthermore, our solution can be used for Blackberry, JavaME, Android, and iOS. Some projects for native applications just in offer solutions for one platform, and usually defined systems as multi-platform are really based in a Web application.
- *Crawling*: The information is stored in databases or tourism website where the information could be stored manually. Application technologies based on the semantic web using crawling can improve access and manage of tourist information (Contur and GAT). They integrated from multiple sources of information scattered.
- *Location technology*: GAT is the unique system able to generate wayfinding applications for indoor/outdoor with the next features of technologies in the phones: 3G, GPS, Camera, Sensors,

Bluetooth and QR-code acquisition. The wayfinding application uses these hardware functionalities to improve the location and guide. The accuracy is about 1–3 m for indoor/outdoor.

- *Time*: In the traditional procedure of developing mobile guide applications, the manager follows the same steps for each new guide application: design, deployment, information insertion about points of interest, validation, generation, and finally, application distribution. There are many tasks that increase the time to generate and distribute the mobile application for the managers. However, our system allows reducing this process to only two steps to generate applications: choose point of interest and choose application distribution from the GAT. The projects such as ConTur or Tourist Guide System require more time to access information from the final mobile application.

6.1. User feedback

In Table 2 is presented a statistical study about the opinion of 40 users that have used GAT. These users used the “Manager Website” to generate different wayfinding applications. Then, they tested the wayfinding applications and provide us the feedback.

The most users were able to generate guidance applications without a previous experience in programming applications. Many users were familiarized with typical wayfinding applications for outdoor environments. However, the most users had not used applications for an indoor environment. They recommended this kind of system to improve tourist services because it facilitates the access to tourist information and services. They also agreed that this kind of proposal improves quality of tourist information management.

To sum up, the obtained feedback shows that GAT has a positive evaluation regarding the process of generating and using of wayfinding applications in a novel way.

7. Conclusions and future research

Tourism presents considerable potential for the use of new mobile technologies. However, most of the wayfinding systems require constant maintenance and monitoring because tourist contents and users constantly change. There are three main issues for the new tourism technologies with which any system must deal: a huge amount of touristic information to manage, guidance for indoor and outdoor environments, and several mobile platforms with different requirements. There are many proposals to improve the mobile tourism described at this paper. However, there are currently no systems that solve all these issues in just a system like the proposal solution at this paper: GAT platform.

In Internet there are many structured and not structured data that could be used to obtain greater tourist information of better quality. However, most wayfinding systems generate and update this information manually. GAT facilitates the management of points of interests for wayfinding functionalities using a tourist web crawler. This crawler searches tourist spots automatically and by applying information extraction techniques, generates points of interest with descriptive information, multimedia resources and location data.

A complete wayfinding system should provide information in any environment to provide a complete guidance. In this regard, GAT provides guidance for indoor and outdoor environments, while related systems typically work on just one of the two environments. In order to the outdoor guidance, our system interacts with the environment combining different technologies (GPS, aGPS, 3G, Wi-Fi, and camera and sensors of the phone). In the case of indoor environments, mobile devices use technologies such as

Table 1
Comparative of the current works.

Project	Type of applications	Architecture	Technology	Location technology	Crawling	Knowledge programming	Time of generation
SEFAGI	Online application based on web services	J2SE and JavaME	Platform static adaptation and device dynamic adaptation	n/a	No	Yes	n/a
iConAwa	Online application based on web services	Web-based	HTML, XML	GPS, sensors	No	Yes	n/a
MultiMAD	Online/offline applications	WML 1.0/2.0, JavaME	Platform static adaption	n/a	No	Yes	n/a
Xmobile	Online/offline applications	Superwaba, JavaME	Static and dynamic adaptation.	n/a	No	Yes	n/a
COMPASS	Context-aware guide application	Web-based	Web service, technology, XML, OWL, P3P	GPS, 3G	No	Yes	n/a
CRUMPET	Location-aware guide application	Web-based	Agents	GPS	No	Yes	n/a
GUIDE	Location-aware guide application	Web-based	HTML	Cell-based wireless	No	Yes	n/a
Guilliver's genie	Location-aware guide application	Web-Based	Agents, Java	GPS	No	Yes	n/a
LoL@	Location-aware Guide application	Web-based	Java, applets, XML, XSL	GPS, 3G	No	Yes	n/a
MobiDENK	Location-aware guide application	Web-based	Java	GPS	No	Yes	n/a
PinPoint	Web recommendation system	Web-based	Net framework, C#	GPS	No	Yes	n/a
Sightseeing4U	Multimedia guide system	Web-based	Java	GPS	No	Yes	n/a
ConTur	Tourist website with recommendations	Web-based	Java, HTML, XML, SOAP	GPS	Yes	Yes	n/a
Tour Guide System	Guide application – recommendation system	Web-based	HTML, XML, SOAP	GPS	Yes – on demand from the mobile application	Yes	Many time to get information of waypoints
GAT	Indoor/outdoor wayfinding applications Manager	Native application – static and dynamic information	Web service, Technology, XML, OWL, P3P	GPS, aGPS, 3G, Wi-Fi, Bluetooth, Qr-Code, sensors	Yes – continuously update	No	Only waste time to choose waypoints.

Table 2
User's Feedback – a statistical study about the opinion of 40 users that we have tested GAT.

Questions	Answers
Age of the user	17–35
Do you use a computer normally?	Yes: 87.5%, No: 0%, Maybe: 12.5% (17–25 years old) Yes: 87.5%, No: 7.5%, Maybe: 5% (26–35 years old)
Do you have a smartphone?	Yes: 97.5%, No: 2.5% (17–25 years old) Yes: 75%, No: 25% (26–35 years old)
Have you ever use a wayfinding application for outdoor environment?	Yes: 50%, No: 50% (17–25 years old) Yes: 50%, No: 50% (26–35 years old)
Have you ever developed an application?	Yes: 37.5%, No: 62.5% (17–25 years old) Yes: 25%, No: 75%. (26–35 years old)
Have you ever developed a guide application?	Yes: 25%, No: 75% (17–25 years old) Yes: 0%, No: 100%. (26–35 years old)
How easy will it be for people to learn how to use the web to generate applications?	Easy: 75%, Very easy: 25%, Difficult: 15% (17–25 years old) Easy: 75%, Very easy: 25%, Difficult: 15% (26–35 years old)
How easy will it be for people to learn how to use the generated wayfinding application?	Easy: 50%, Very easy: 50%, Difficult: 0% (17–25 years old) Easy: 75%, Very easy: 25%, Difficult: 0% (26–35 years old)
How useful will people find the web?	Easy: 25%, Very easy: 75%, Difficult: 0% (17–25 years old) Easy: 87.5%, Very easy: 7.5%, Difficult: 5% (26–35 years old)
How useful will people find the wayfinding application?	Easy: 70%, Very easy: 30%, Difficult: 0% (17–25 years old) Easy: 87.5%, Very easy: 7.5%, Difficult: 5% (26–35 years old)
Will they recommend it to tourist entities?	Yes: 87.5%, No: 0%, Maybe: 12.5%. (17–25 years old) Yes: 87.5%, No: 0%, Maybe: 12.5%. (26–35 years old)

Bluetooth, bi-dimensional codes and its sensors to obtain its location. The accuracy can be 1–3 m for both environments.

Regarding the third issue introduced above, users that generate the wayfinding applications (managers) need to have program-

ming knowledge about all mobile phone platforms (Android, iOS, Symbian, Nokia, etc.). One of the main advantages of GAT is that managers use a Manager Website to generate wayfinding applications for indoor and outdoor environments. Thus, managers can

create any kind of tourism application in a few minutes making use of all mobile hardware capabilities without the need of having programming skills.

We consider the following research directions as future work. On the one hand, we would like to apply GAT in an accessibility project to generate wayfinding applications oriented to disability people. On the other hand, this proposal is aimed to primarily wayfinding and tourism applications for mobile phones devices, but it could also be applied to make applications for other devices such as personal computers and embedded systems. For this purpose, we are working on including PhoneGAP (Allen et al., 2010) support in GAT.

References

- Allen, S., Graupera, V., & Lundrigan, L. (2010). *Lee. Book: Pro smartphone cross-platform development Springer professional and applied computing*. Berlin: Springer. 131–152.
- Baldzer, J., Boll, S., Klante, P., Krösche, J., Meyer, J., & Rump, N., et al. (2004). Location-aware mobile multimedia applications on the Niccimon platform. In *2. Braunschweiger symposium – informationssysteme für mobile anwendungen*.
- Baldzer, J., Boll, S., Klante, P., Krösche, J., Meyer, J., & Rump, N., et al. (2004). Location-aware mobile multimedia applications on the Niccimon platform. In *Braunschweiger symposium – informationssysteme für mobile anwendungen IMA'04*. Brunswick, Germany.
- Biegel, G., & Cahill, V. (2004). A framework for developing mobile, context-aware applications. In *Paper presented at the proceedings of the second IEEE annual conference on pervasive computing and communications*. IEEE.
- Bluegiga Technologies. Device: Bluegiga access point. Last visited in 29/11/2012: <http://www.bluegiga.com/bluetooth_access_point_3201-1>.
- Castillo, C. (2005). Effective web crawling. *SIGIR Forum*, 39, 55–56.
- Chaari, T., & Laforest, F. (2005). SEFAGI: Simple environment for adaptable graphical interfaces – generating user interfaces for different kinds of terminals. In *Annals of international conference on enterprise information systems-ICEIS*. Cyprus.
- Chaisatien, P., & Tokuda, T. (2009). A web-based mashup tool for information integration and delivery to mobile devices. In *Proceedings of the 9th international conference on web engineering: Vol. 5648. LNCS* (pp. 489–492). New York: Springer.
- Chen-Chuan Chang, K., He, B., Li, C., Patel, M., & Zhang, Z. (2004). Structured databases on the web: Observations and implications. *SIGMOD Record*, 3(3), 61–70.
- Cheverst, K., Davies, N., Mitchell, K., Friday, A., & Efstratiou, C. (2002). Developing a context-aware electronic tourist guide: Some issues and experiences. In *Proceedings of the SIGCHI conference on human factors in computing systems – CHI '00*. ACM, Netherland: ACM Press: ACM.
- Cheverst, K., Davies, N., Mitchell, K., Friday, A., & Efstratiou, C. (2000). Developing a context-aware electronic tourist guide: some issues and experiences. In *Proceedings of the conference on human factors in computing systems (CHI'00)* (pp. 17–24).
- Chiu, D. K. W., Cheung, S. C., & Leung, H. F. (2005). A multi-agent infrastructure for mobile workforce management in a service oriented enterprise. Big Island, Hawaii: IEEE Computer Society press :IEEE.
- Chiu, D. K. W., Yueh, T. F. Y., Leung, H. F., & Hung, P. C. K. (2001). Towards ubiquitous tourist service coordination and process integration: A collaborative travel agent system architecture with semantic web services. *Information Systems Frontiers*, 11(3), 241–256.
- Choi, Y., Cardie, C., Riloff, E., & Patwardhan, S. (2005). Identifying sources of opinions with conditional random fields and extraction patterns. In *Proceedings of the conference on human language technology and empirical methods in natural language processing, HLT '05, (association for computational linguistics, Stroudsburg, 2005)* (pp. 355–362). Stroudsburg, PA, USA.
- Dey, A. K., Abowd, G. D., & Salber, D. (2001). A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. *Human-Computer Interaction*, 16(2), 97–166.
- Economou Gavalas, D., Kenteris, M., & Tsekouras, G. E. (2008). Cultural applications for mobile device: Issues and requirements for authoring tools and development platforms. *ACM SIGMOBILE Mobile Computer Communication Review*, 12(3), 18–22.
- Feldman, R., Rosenfeld, B., & Fresko, M. (2006). Teg – A hybrid approach to information extraction. *Knowledge and Information Systems*, 9, 1–18.
- García-Crespo, A., López-Cuadrado, J. L., Colomo-Palacios, R., González-Carrasco, I., & Ruiz-Mezcua, B. (2011). Sem-Fit: A semantic based expert system to provide recommendations in the tourism domain. *Expert Systems with Applications*, 38, 13310–13319.
- Gossweiler, R., McDonough, C., Lin, J., Want, R. (2011). Argos: Building a webcentric application platform on top of android. In *Pervasive computing* (Vol. 10, issue no. 4, pp. 10–14). IEEE.
- Han, J., Twidale, M., Gutierrez, A., & Farivar, R. (2011). Methods for applied mobile digital library research: A framework for extensible wayfinding systems. *The Reference Librarian*, 52, 106–116.
- Hernandez, E. (2009). A war of the mobile browsers. *IEEE Pervasive Computing*, 8(1), 82–85.
- Herrero, G., Campo, A., Gil, M., Garcia, A., Martin, D., & Ion Zugasti, et al. (2012). ConTur: An intelligent content management system for the tourism sector. In *19th eTourism community conference. Tourism present and future services and applications* (pp. 24–27). Helsingborg, Sweden.
- Hristova, N., O'Hare, G. M. P., & Lowen, T. (2003). Agent-based ubiquitous systems: 9 lessons learnt. In *Workshop on system support for ubiquitous computing (UbiSys'03), 5th international conference on ubiquitous computing*. Seattle, WA, USA.
- Kapitsaki, G. M., Kateros, D. A., & Venieris, I. S. (2008). Architecture for provision of context-aware web applications based on web services. In *Paper presented at IEEE 19th international symposium on the personal, indoor and mobile radio communications, PIMRC*.
- Kim, J. W., Kim, C. S., Gautam, A., & Lee, Y. (2005). Location-based tour guide system using mobile GIS and Web crawling. In *Web and wireless geographical information systems. Lecture notes in computer science* (Vol. 3428, pp. 51–63). Springer-Verlag.
- Kortuem, G., Kawsar, F., Fitton, D., & Sundramoorthy, V. (2010). Smart objects as building blocks for the internet of things. In *Internet computing* (Vol. 14, issue no. 1, pp. 44–51). IEEE.
- Kukkonen, J. et al. (2009). BeTelGeuse: A platform for gathering and processing situational data. In *IEEE Pervasive Computing*. IEEE.
- Kumar, R., Raghavan, P., Rajagopalan, S., & Tomkins, A. (1999). Trawling the web for emerging cyber-communities. *Computer Networks*, 31(11–16), 1481–1493.
- Lamsfus, C., Alzua-Sorzabal, A., Martin, D., & Lopez de Ipiña, D. (2010). Digital broadcasting for context-aware services in tourism. In *Paper presented at second IEEE region 8 conference on the history of the telecommunications conference (HISTELCON, 2010)*.
- Lee, H. N., Lim, S. H., & Kim, J. H. (2009). UMONS: Ubiquitous monitoring system in smart space. *IEEE Transaction on Consumer*, 55(3).
- Lehn, D., & Heath, C. (2003). Displacing the object: Mobile technologies and interpretive resources. In *Proceedings of the ICHIM 03. Archives and museum informatics Europe*. Last visited in 05/01/2012: <<https://www.amico.org/publishing/ichim03/088C.pdf>>.
- Micha, K., & Economou, D. (2005). Using personal digital assistants (PDAs) and to enhance the museum visit experience. In *Proceedings of the 10th panhellenic conference on informatics: Vol. 3746. LNCS* (pp. 188–198).
- Michlmayr, E. (2002). Flexible content management for the LoL@UMTS Application. Master Thesis. Vienna: University of Technology Vienna.
- MultiMAD – multimodal mobile application development tool. Available at: <<http://homepages.dcc.ufmg.br/~thiagohp/multimad/>>. Last access in January 2012.
- Olston, C., & Najork, M. (2010). Web crawling. *Foundations and Trends in Information Retrieval*, 4, 75–246.
- Pascual, J., González, R., Sanjuán, O., Pelayo, B. C., & Cueva, J. M. (2012). Extensible architecture for context-aware mobile web applications. *Expert Systems with Applications*, 39(10), 9686–9694.
- Pielot, M., & Boll, S. (2010). Tactile Wayfinder: comparison of tactile waypoint navigation with commercial pedestrian navigation systems. In *International conference on pervasive computing*. Helsinki, Finland.
- Pielot, M., Henze, N., & Boll, S. (2009). Supporting paper map-based navigation with tactile cues. In *MobileHCI '09: Human-computer interaction with mobile devices and services*.
- Poslad, S., Laamanen, H., Malaka, R., Nick, A., Buckle, P., & Zipf, A. (2001). CRUMPET: Creation of user-friendly mobile services personalised for tourism. In: *Proceedings of: 3G 2001 – second international conference on 3G mobile communication technologies*. London, United Kingdom.
- Raento, M., Oulasvirta, A., et al. (2005). ContextPhone: A prototyping platform for context-aware mobile applications. *IEEE Pervasive Computing*, 4(2), 51–59.
- Rayman-Bachus, L., & Molina, A. (2001). Internet-based tourism services: Business issues and trends. *Futures*, 33(7), 589–605.
- Roth, J. (2002). Context-aware web applications using the PinPoint. In *IADIS international conference www/internet 2002* (pp. 3–10). Lissabon, Portugal: IADIS Press.
- Setten, M. V., Pokraev, S., & Koolwaaij, J. (2004). Context-aware recommendations in the mobile tourist application COMPASS. In W. Nejdl, & P. De Bra (Eds.), *AH 2004: Vol. 3137. LNCS*: Springer-Verlag.
- Shen, W., Doan, A., Naughton, J. F., & Ramakrishnan, R. (2007). Declarative information extraction using datalog with embedded extraction predicates. In *Proceedings of the 33rd international conference on very large databases, VLDB '07* (pp. 1033–1044). VLDB, Endowment.
- Shirazi, A. S., Winkler, C., & Schmidt, A. (2010). SENSE-SATION: An extensible platform for integration of phones into the web. *Internet of Things IOT*, 1–8.
- Viana, W., & Andrade, R. M. C. (2008). XMobile: A MB-UID environment for semi-automatic generation of adaptive applications for mobile devices. *Journal of Systems and Software*, 81(3), 382–394.
- Weiser, M. (1993). Some computer science issues in ubiquitous computing. *Communications ACM*, 36(7), 74–84. ACM.
- Wilson, G. (2004). Multimedia tour programme at tate modern. In D. Bearman, & J. Trant (Eds.), *Proceedings of the international conference of museums and the web 2004. Archives and museum informatics*. Last visited in 01/12/2011: <<http://www.museumsandtheweb.com/mw2004/papers/wilson/wilson.html>>.
- Woodruff, A., Aoki, P. M., Hurst, A., & Szymanski, T. M. H. (2004). Electronic guidebooks and visitor attention. In *Proceedings of the 6th international cultural heritage informatics meeting* (pp. 437–454). Last visited in 12/01/2008: <<http://www2.parc.com/csl/members/woodruff/publications/2001-Woodruff-ICHIM2001-VisitorAttention.pdf>>.