QALL-ME
Question Answering Learning technologies in a multiLingual and Multimodal Environment

Information Society Technologies

Sixth Framework Programme

FP6 IST-033860

http://qallme.itc.it

First showcase

Authors: Giorgio Sangalli; Nicolò Bargna

Affiliation: COM

Abstract

This document presents describes the first showcase of the QALL-ME project. The showcase includes two different demonstrators: the mobile version and the desktop version. The first one is supposed to help the user to access information about cinemas in Trento through a mobile phone, which interacts with the QALL-ME system. The desktop version of the QALL-ME showcase is represented by a web application that can be found at the web address http://qallme.itc.it/server/demo. The application takes as input a question in natural language, and outputs the answer formatted as an HTML page.

Date: 15/10/2007

Document Number: QALL-ME_D8.1_R_20071015

Status/Version: Final

Distribution Level: Restricted to a group specified by the consortium (including the Commission Services)

Project Reference	FP6 IST-033860
Project Acronym	QALL-ME
Project Full Title	Question Answering Learning technologies in a multiLingual and Multimodal Environment
Distribution Level	Restricted to a group specified by the consortium (including the Commission Services)
Contractual Date of Delivery	October '07
Actual Date of Delivery	October '07
Document Number	QALL-ME_D8.1_R_20071015
Type	Report
Status & Version	Final
Number of Pages	8
WP Contributing to the Deliverable	WP8, WP7
WP Task responsible	COM
Authors	Giorgio Sangalli (COM); Nicolò Bargna (COM)
Other Contributors	Arnaldo Claudiani (UBI), Enrico Craveia (WAY)
Reviewer	(WLV)
EC Project Officer	Xavier Gros
Keywords: First showcase, mobile and desktop version	

Abstract: This document presents describes the first showcase of the QALL-ME project. The showcase includes two different demonstrators: the mobile version and the desktop version. The first one is supposed to help the user to access information about cinemas in Trento through a mobile phone, which interacts with the QALL-ME system. The desktop version of the QALL-ME showcase is represented by a web application that can be found at the web address http://qallme.itc.it/server/demo. The application takes as input a question in natural language, and outputs the answer formatted as an HTML page.

Summary

1. Introduction

This document presents the description of the first prototype of the QALL-ME project. The showcase for the first development cycle of the project includes two different demos: a mobile version and a desktop version. The first one is supposed to help the user to access information about cinemas scheduled in the city of Trento through a mobile phone (Nokia N95, Symbian based is used), which interacts with the QALL-ME system. The input is in natural language, both speech (this is covered only for Italian at this stage) and as written messages (English, Italian, German and Spanish); answers consist of a text to speech data in four languages (Italian, Spanish, German and English), geo-referenced maps and route descriptions. The answer presentation over mobile devices can be multimedial. For the first prototype we only consider text to speech answers, maps, images and videos that support the answers.

The desktop version of the QALL-ME demo is represented by a web application that can be found at the web address http://qallme.itc.it/server/demo, showing similar functionalities as the mobile version. The application takes as input a question in natural language, and outputs the answer formatted as an HTML page.

For both demos the usage scenario for this first prototype is that of a tourist looking for information about cinemas and pharmacies. The data providers selected by each partner of the QALL-ME consortium are:

- LaNetro (http://www.lanetro.com/) for Spanish
- Zelluloid.de (http://www.zelluloid.de/) for German
- Trentino Cultura (http://www.trentinocultura.net); T-Net Consulting (http://www.tnetconsulting.com) for Italian
- Movie Tickets (http://www.movietickets.com) for English

The rest of the document will give detailed information about the two versions of the demonstrator.

2. Mobile Demo

For the first year project ShowCase, the choice is oriented towards the use of a mobile device (**Nokia N95**) with an installed client, appropriately configured to answer questions relative to the chosen scenario and categories. Within the tourism domain we have chosen to test the first QALL-ME application prototype in the town of Trento for the following categories and corresponding information:

Cinemas	Pharmacies
List of cinemas, movies,	List of pharmacis
Localizations (as maps)	Localizations (as maps)
Programming time tables	Opening times
Movie Plots	Open pharmacies directory

For the first ShowCase we have chosen to use a **Nokia N95** with an OS Symbian Series 60 3rd Edition equipped with:

- Internal Bluetooth connected GPS antenna;
- TCP connection on public APN option;
- Software module, for vocal analysis and separation in phonemes to be sent to the Automatic Speech Recognition (ASR) server;
- Text-To-Speech (TTS) module to vocally synthesize the texts received by the ASR server;
- Push-To-Talk mode to activate the reception of the user's voice to send to the ASR server;
- Integrated speakerphone;

We expect to use the mobile device on field and the potentialities of the prototype wil be showed during the Trento's review meeting.

We have selected a subset of the concepts and relations defined in the QALL-ME ontology as Expected Answer Types (i.e. the categories of the answers) activated for the first showcase:

- Cinema-Name
- Cinema-Phone Number
- Cinema-Address
- Pharmacies-Name
- Pharmacies-Phone Number
- Pharmacies-Address
- Pharmacies-Opening Hour
- Movie-Date
- Movie-Starting Hour
- Movie-Genre
- Movie-Title
- Movie-Director
- Movie-Actor
- Movie-Description

For the first Mobile Demo, questions answered by the system are still at a controlled level of complexity (i.e. nested questions are not allowed) and there are no dialogue capabilities. The following Italian patterns represent some of the questions allowed:

- "Qual è la farmacia più vicina?"
- "Cosa danno al cinema <nome cinema> ?"
- "Dove posso vedere il film <attore>/<regista> ?"
- "A che ora inizia il film <nome proiezione> ?"
- "Dove si trova il cinema <nome cinema> ?"

During the mobile demo, the user will activate the client using a Push-to-Talk modality and will ask, in a natural language, the information according to the issues available for this demo.

The client will manage the requests by dispatching them to the QALL-ME modules. First of all, it will make a call to the ASR module, which will return a written translation of the question made in natural language. This text will then be sent to a module which will recognize its semantic structure and will insert the related geographical coordinates. This module will apply its algorithms in order to transform the question in a query, which will translate the content in a correct

answer, according to its semantic structure and its meaning. The answer will be sent to the client, which will ask to the TTS module to transform the written text in spoken language (Italian, Spanish, German and English). Each data as output of the answer have a couple of coordinates, which could be used to ask for a circumscribed map or the way to get to a destination. In this case, it is possible for the user to see in the screen of the mobile phone a map and, possibly, the route to the destination choice.

In the following we present some images of the mobile client:

1 Splash page – Intro
"Starting page with the QALL-ME Logo"

2 Input question phase
"Where is the nearest pharmacy?"

3 Output answer phase
"TTS output with the list of the nearest pharmacy"

4 Multimedia content output – Map
"View of the pharmacy in the Trento's map"

5 Multimedia content output – Navigation "Route to the nearest pharmacy"

3. Mobile Demo components

QALL-ME Mobile Client with UbiNav technology inside

For the *QALL-ME Mobile Client* has been adopted an UbiEst engine for the cartographic and navigation for Symbian phones: *UbiNav*. In the QALL-ME project the entire mobile user interface and the communication interface to the QALL-ME server to submitting the natural language queries has been developed.

UbiNav is more than an off-board navigation service for mobile phones. *UbiNav* provides for users' needs with on line navigation and real-time content, bringing location and door-to-door guidance into daily activity helping plan movements and finding businesses.

UbiNav is a server based navigation: data is stored on a central server, without using the mobile device's memory space. *UbiNav* actually connects to a server only when the user requests searches, views maps or calculate routes. This ensures always up-to-date street database (NAVTEQ).

UbiNav navigator offers extra route features to plan personal multi-stops journeys, selecting along the route waypoints, moving them up or down in the itinerary list, checking out locations when you reach them and stopping/resuming itinerary while travelling.

UbiNav comes as a customizable, ready-to-use, off board navigation service, providing an environment for the development of Business and Consumer Mobile Location Based Services:

- Street navigation
- City guides, and other location aware contents

On-line Navigation

Select Navigation from *QALL-ME Mobile Client* main menu:

1. Start position is automatically set from GPS signal, select Start to change it.

2. Select Destination to enter your desired destination (e.g. manually by address search, from your Favourites).

Optionally you can select Settings from Options Menu to pick special route options, such as the fastest route, the shortest route or a route without highways.

Select Calculate route.

The application connects to the UbiNav server to calculate and download the route. As the route is being downloaded, you will begin to get vocal travel instructions and your immediate position will be displayed on screen.

With or without GPS Navigation

UbiNav engine allows you to use navigation features even if a GPS receiver is not available: manually set Start and Destination and select Start Navigation. In the navigation mode, click the right upper button under the screen to open the options menu and select Start demo to view a no GPS navigation movie.

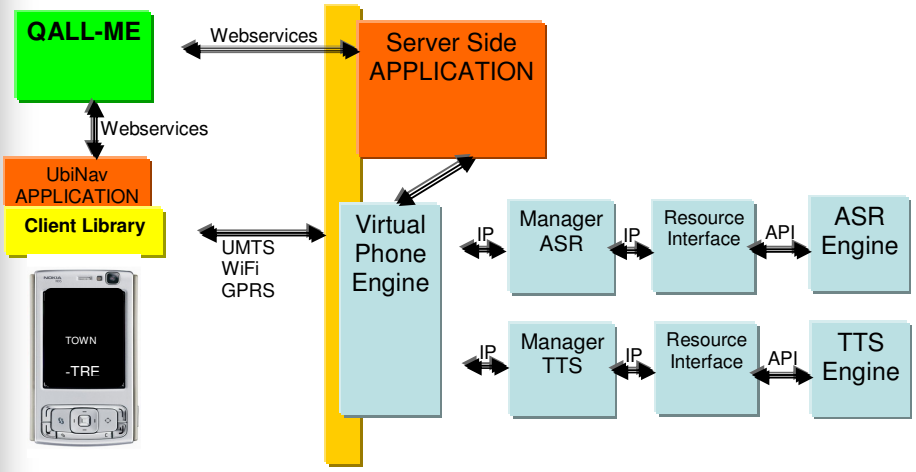
Speech Recognition Module

The Speech Recognition module consists of an integrated software library, available in the *QALL-ME Mobile Client* application. The module provides *QALL-ME Mobile Client* with specific functions to perform the following facilities:

- TCP/IP connection to the Server where ASR and TTS modules are installed.
- The library method "Call" starts the application flow on the server. The first action of the flow could be to send the client a vocal message which asks the user to ask a question, for example "Welcome to the QALL-ME service, please ask a question".
- Pressing the "Push-To-Talk" button the microphone begins to record and meanwhile the device sends the voice to the server (this modality is called "streaming"). At the end of the question the user releases the Push-To-Talk button.
- Through the ASRresult event, which is part of the library, the *QALL-ME Mobile Client* application receives a character string containing the word recognized by the ASR module.

- The library method "SendData" enables the *QALL-ME Mobile Client* to send the data collected by the device, for example the user GPS location. This information can be used by the server supposing it can manage the request to the QALL-ME "Web-Services".
- To make further questions based on the answers we obtained**,** we start a "Computer Initiative" dialogue flow**.** Using the library**, the** *QALL-ME Mobile Client* application keeps up the dialogue with ASR and TTS modules until the "Disconnect" method is activated.

The connection management between the client device described before (*QALL-ME Mobile Client*) and ASR and TTS modules is accomplished by a service called "Virtual Phone Engine".



ASR Translation module

The speech recognition system supplies its applicative modules with the capability of obtaining a character string from a sequence of vocal samples or acoustic features.

The speech recognition system used in this prototype is speaker independent, i.e. it can recognize words spoken by any speaker. This system is more complex than speaker dependent recognizers which have to be trained and can recognize only "known" voices.

The speaker independent modality is managed by considering a limited vocabulary of words to recognize: this operation is identified with the definition of the so called "grammar". A grammar is a group of (thousands of) words or sentences which define each language, together with the recognizing modalities of such words used to train the recognition engine.
The text returned by the recognizer, besides the recognized text, may contain some labels, called "semantic lables" that allow to associate a meaning to the various parts of a sentence.
In the prototype case, the resource module uses the grammar studied for every QALL-ME domain, which provides for natural and free speech for the request of an information.

The architecture is modular and is based on a manager component and a number <n> of resource modules which perform the recognition itself. The number of resource modules depends on how many recognition sessions one has to satisfy at the same time.
The architecture has been appropriately studied to permit a suitable scalability, Fault-Tolerance and Load-Balancing functionalities.

The manager module coordinates all the activities between the applicative modules which ask for system services and resource modules that perform such activities.
To work properly, the manager module must always know which modules are active, what they are doing and which are their features, such as the language used for the recognition.
In this way the manager can ask the proper resource an English or Italian recognition session.

The resource modules perform the recognition asked by the applicative modules, calling, in the special case of the prototype, the functions of the ASR ITC-IRST Spinet engine.
In any way, independently from the service supplied and the associated engine, a resource module supplies a certain number of channels for the service. A channel can be interpreted as an instance of the engine wich can be asked for a recognition, using properly codified instructions.

TTS Translation module

The speech synthesis system supplies its applicative modules with the capability of obtaining from a character string a sound file containing the synthetised speech.

From an architectural point of view, the synthesis system structure reflects exactly the general structure described in the previous paragraph. In this case too, the resource module performs its own activities depending on the specific engine and for the synthesis may use engines of different producers, according to the features and quality guarantees they can supply. At present Nuance and Loquendo TTS have been integrated.

Unlike the ASR service, which, considering the complexity of the recognition phase and the size of grammars, must run on a server platform, the speech synthesis program may also be installed on a client device.

This allows to improve the quality of the service, avoiding the voice sample traffic that would be generated from the server in a "server based situation".

In this case, when *QALL-ME Mobile Client* asks for a text synthesis, the library will use software installed on the device.

For the prototype we decided to use Loquendo TTS for symbian platform.

A selection of the required contents

The *QALL-ME Mobile Client* offers the GPS position as a contextual data, i.e. the user's coordinates. It allows to add to the enunciated sentence some information about the place in which the question is formulated. Even if the sentence doesn't refer to the place where the user is enunciating it, the coordinates could given by hand if in the semantic analysis the place complement lacks a proper definition.

In this phase of the project, the domain we are working on regards cinemas' showings and pharmacies' shifts in the province of Trento. The user could make questions about a film or a cinema belonging to this domain.

Cinema's data are organized in cinema → auditorium → file → showing and the user can ask for:

- Title
- Original title
- Duration
- Year
- Genre
- Distribution
- Cast (actors)
- Direction
- Nationality
- Plot

There are two categories of questions:

- Informative questions, when the user asks for an information using the natural language and an answer in natural language extracted from the DB is expected;
- Proximity questions, when the user is looking for a place near him or in proximity of a given address.

Proximity questions

Cinemas are geo-reported by the addresses contained in the Tnet Consulting feed, then the geographical coordinates are assigned (WGS-84) and saved in the QALL-ME server DB. In this kind of questions, where the query starts from a given geographical point or address, a filter based on the air distance, can be added to the data extraction, to calculate the Euclidean distance between the point given by the user (GPS or address) and the point of interest (cinema or pharmacy) in the DB. A distance ordered list will be the result, which will be on hand inside the client. Each object in the list will have a couple of coordinates, which could be used to ask for a circumscribed map or the way to get to a destination.

4. Desktop Demo

The desktop version of the QALL-ME demo is represented by a web application that can be found at the web address http://qallme.itc.it/server/demo. The application takes as input a question in natural language, and outputs the answer formatted as an HTML page. The user also selects a predefined location, in the four cities of the academic partners (Alicante, Saarbrücken, Trento and Wolverhampton) and a predefined date relevant to the data available to each partner. Figure 7 represents the architecture of the desktop application.

HTML form – The HTML form presents the user with a field in which he/she types his/hers question and a predefined list of location and a list of predefined dates, compatible with the available data.

Web Application – The web applications handles the communication with the main web service of the QALL-ME Prototype. It sends a request as a question in natural language and temporal and special constraints expressed by the predefined location and date, and receives as output the answer to the question.

Question Representation Module – The question representation module delivers the answer to the user as an HTML page. A screenshot of the demo can be seen in Figure 1.

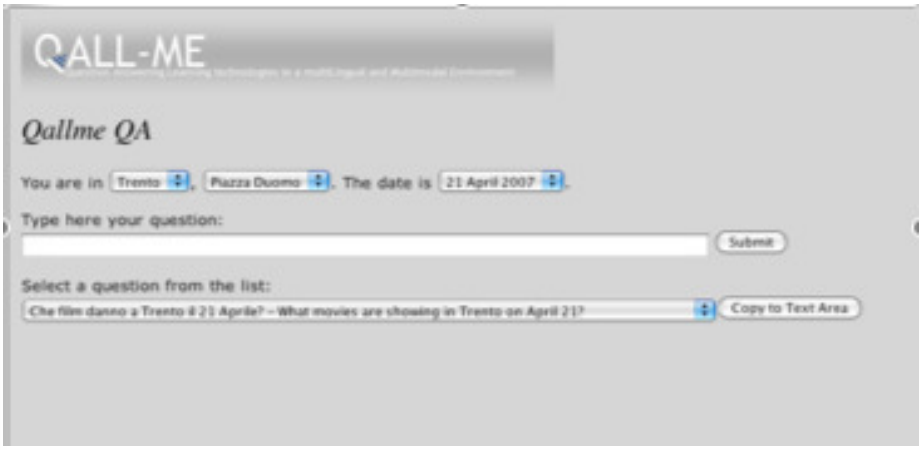


Figura 1 Screenshot of the Desktop Version of the Qallme Demo

During the Desktop Demo, the system is able to answers the following kind of questions (divided by languages)

Spanish Questions:

- ¿Cuándo dan la película [MOVIE] en [DESTINATION]?
- ¿Cuándo puedo ver la película [MOVIE]?
- ¿Hasta cuándo emiten la película [MOVIE] en el cine [CINEMA]?
- ¿Dónde dan la película [MOVIE] en [DESTINATION]?
- ¿Dónde puedo ver la película [MOVIE]?
- ¿Cuál es la dirección del cine [CINEMA]?
- ¿Qué películas puedo ver en el cine [CINEMA]?
- ¿De qué trata la película [MOVIE]?
- ¿Qué cines hay en [DESTINATION]?
- ¿Qué cines de [DESTINATION] dan la película [MOVIE]?
- Quisiera saber el teléfono del cine [CINEMA].

German Questions:

- Wo läuft in [CITY] der Film [MOVIE]?
- In welchem Kino läuft der Film [MOVIE]?
- Wo ist das nächste Kino, in dem der Film [MOVIE] läuft?
- Wo ist das nächste Kino in [CITY], in dem der Film [MOVIE] läuft?
- Wo ist das Kino [CINEMA]?
- Welche Postleitzahl hat das Kino [CINEMA]?
- In welcher Straße ist das Kino [CINEMA]?
- Wo läuft der Film [MOVIE]?
- Wann beginnt der Film [MOVIE] im [CINEMA]?
- Was für ein Film ist [MOVIE]?
- Wer führte Regie bei dem Film [MOVIE]?
- Wer spielt in dem Film [MOVIE] mit?
- Wie lautet der Originaltitel von [MOVIE]?
- Wie lange dauert der Film [MOVIE]?
- Worum geht es in dem Film [MOVIE]?
- Wer schrieb das Drehbuch zu dem Film [MOVIE]?

Italian Questions:

a. rel(Movie,Cinema) – A relation between cinema and a movie
   i. Pattern 1: dove danno [MOVIE]
   ii. Pattern 2: [MOVIE] a [CINEMA]

b. rel(Movie,Title) – A relation between a movie and its title
   i. Pattern 1: qual é il titolo del film
   ii. Pattern 2: come si chiama il film

c. rel(Movie,StartingHour) – A relation between movie and its starting hour
   i. Pattern 1: quando comincia [MOVIE]
   ii. Pattern 2: [MOVIE] a che ora inizia

d. rel(Movie,Date) – A relation between movie and a date of a show
   i. Pattern 1: quando daranno [MOVIE]
   ii. Pattern 2: [MOVIE] [DATE]

e. rel(Movie, Director) – A relation between a movie and its director.
   i. Pattern 1: film diretto da [director]
   ii. Pattern 2: regista di [MOVIE]

f. rel(Movie, Genre) – A relation between a movie and its genre.
   i. Pattern 1: film di [GENRE]
   ii. Pattern 2: [GENRE] diretto da [DIRECTOR]

g. rel(Movie, Description) – A relation between a movie and its description.
   i. Pattern 1: avvenimenti principali di [MOVIE]
   ii. Pattern 2: trama di [MOVIE]

h. rel(Cinema,Name) - A relation between the cinema and its name
   i. Pattern 1: dove danno il film
   ii. Pattern 2: in quali sale proiettano [MOVIE]

i. rel(Cinema,Phonenumber) – A relation between the cinema and its telephone number
   i. Pattern 1: numero di telefono di [CINEMA]
   ii. Pattern 2: telefono di [CINEMA]

j. rel(Cinema,Address) – A relation between cinema and its address
   i. Pattern 1: indirizzo di [CINEMA]
   ii. Pattern 2: [CINEMA] dove si trova

k. rel(Cinema,City) - A relation between cinema and its city
   i. Pattern 1: cinema di [CITY]

ii. Pattern 2: dove e` uscito a [CITY] [MOVIE]

English Questions:

- In which cinema from [CITY] the movie [MOVIE] is shown?
- In which cinema can I see the movie [MOVIE]?
- Where is [CINEMA] cinema?
- What is the post code for [CINEMA] cinema?
- What is the address of [CINEMA] cinema?
- What time does the movie [MOVIE] start at [CINEMA] cinema?
- Who directed the movie [MOVIE]?
- Who acts in the movie [MOVIE]?
- Who wrote the script of the movie [MOVIE]?
- How long is the movie [MOVIE]?
- Are there any films on [DATE] starring [ACTOR]?
- Can you tell me a name of a film that's on in [CITY] written by [WRITER]?
- Are there any films made by [STUDIO NAME] preferably in [CITY]?
- Are there any [GENRE] films out at the moment?
- Can you tell me which films are certified [RATING]?
- Can you tell me name of cinemas which have [FACILITY]?