The Hops Project: Developing Transformational Government Services

Marta GATIUS[1], Alasdair MANGHAM[2], Leonardo LESMO[3]
[1]*Talp Research Center, Jordi Girona 1-3, Bacelona, 08034, Spain*
Tel: +34 93 413 7797, Fax: +34934137786, Email:gatius@lsi.upc.edu
[2]*London Borough of Camden, Argyle Street,London, WC1H 8NG,England*
Tel: +0044 79744704, Email: alasdair.mangham@camden.gov.uk
[3]*University of Torino, Torino,Italy*
Tel: +390116706708, Email: lesmo@di.unito.it

Abstract: In this paper we propose the creation of a blended channel of both voice and text to leverage existing E-Government investments and break down the digital divide. For this purpose, in the context of the European project HOPS we have developed a multilingual dialogue system for guiding the user when accessing the local administration public web contents. The current implementation of the dialogue system supports speech (through the telephone) and text mode (through the web) in several languages (English, Spanish, Catalan and Italian). An evaluation methodology has also being developed as part of the project.

Keywords: E-government, web usability, dialogue systems, multilinguality, evaluation methodologies

1. Introduction

The term E-government is one that has existed since the late nineties and has come to be associated with a European wide effort to put government services online by a generally agreed date of 2005.

The agreed evolution of E-government over the past five years suggests that it follows three main phases:

1) **Promotion of access and connectivity**: This phase was largely concerned with the promotion of the technology and the development of infrastructure, both legislative and physical, to support the adoption of the new technologies. The success of this phase was measured by the number of people online whether or not they were accessing government services. The political justification for such an approach was often that in order for the nation to compete in the global knowledge economy it required the necessary infrastructure.
2) **Provision of services online:** E-government is created as an addition to existing services, whereby information that was previously obtained only from an expert or from written literature is made available online. In this phase some transactional services were e-enabled, as were some paper forms. However in the case of the forms these were only e-enabled in the front office and required manual intervention for processing within the back office. The success of this phase was judged by the number of services available online, in many European countries the number of enabled services and the dates for achieving enablement were set as performance

targets. This provision of services was completely supply driven as opposed to demand led; consequently the take up of many online services was low.

3) **Efficient Electronic Government:** In this phase governments will seek to take advantage of the developments in infrastructure and the supply of online information to develop demand driven automation of existing services. This will enable applications for services to be processed electronically from the front office through to the back office. As back office integration, often with legacy systems is expensive these new customer channels will only be sustainable if they can achieve efficiencies either through reducing demand on an existing channel, such as a call centre, or through an overall reduction in back office staff.

In the research paper 'Beyond E-government' commissioned by the UK presidency in 2005 Booz, Allen and Hamilton propose what they describe as a fourth wave called next generation government:

Traditional channels are ICT – enabled in the back office and offered alongside electronic channels in a seamlessly integrated fashion. Merging of channels, for example, internet telephony, allows government to offer blended channel solutions as part of their channel management strategies such that electronic channels become mainstream.

One of the challenges that has existed throughout all three phases of E-government, each of the municipalities involved can justifiably claim to have entered the third phase, has been to include those people who do not have access to the internet. Over the past five years a variety of solutions has been proposed to break down what has been commonly referred to as the digital divide. These have included WAP enabled telephones, digital televisions and inexpensive Internet devices; for the most part these have had a limited success. The most successful means of breaking down the digital divide has been an economic one with both the price of connectivity and hardware falling by more than sixty percent over the five-year period since 2000. However, in a social capital survey the London Borough of Camden established that thirty five percent of its population still have no access to online services and its section of London is recognised as having the second largest knowledge economy, per capita, in the UK. However, the penetration of telephones, either mobile or fixed, in all of the municipalities is near ubiquitous. Therefore there are clear advantages in taking some of the developments in voice portals and applying them to the public sector.

In doing this the municipalities did not want to create another channel that would have to be populated with its own specific information and have its own interactions with the back office but one that would integrate seamlessly with the existing information systems and back offices. Furthermore, it would have to achieve the creation of the new service without any perceived degradation of quality over the existing voice and text channels, as failure to do this would mean a loss in any potential efficiency gains.

It was to address exactly these types of challenges that three of the leading E-government municipalities in Europe, the City of Barcelona (Spain), London Borough of Camden (UK) and the City of Turin (Italy) created the ***Enabling an Intelligent Natural Language Based Hub for the Deployment of Advanced Semantically Enriched Multi-channel Mass-scale Online Public Services***, (HOPS) project (http://www.bcn.es/hops/).

In this paper we focus on how a blended channel of both voice and text can be created to leverage existing E-Government investments and break down the digital divide. Next sections describe the multilingual dialogue system developed for the European project HOPS. This system was designed for guiding the user when accessing the local administration public web content. Section 3 describes the methodology followed when developing the multilingual multimodal dialogue system. The technologies used are described in Section 4. Section 5 describes the evaluation methodology being developed as part of the project. The final section draws some conclusions.

2. Methodology

The HOPS platform has been designed to support an open, interoperable and re-configurable e-government platform. Although the current system was developed to support speech (through the telephone) and text mode (through the web) the system design will support extending the number and complexity of available channels. An overview of the first prototype is shown in Figure 1. This figure shows the components of the current system as well as the data used by each of these components (in the left column). When incorporating a new service only the data (not the components) need to be modified. The process of adapting the system to a new service consists basically of two steps: representing the service tasks in an ontology and generating the domain-restricted linguistic resources. The representation of the service tasks in ontologies can improve the communication as well as the knowledge management. In the system we have developed the ontologies describing the service are used both for dialogue management and for generating the linguistic resources.

The HOPS platform is composed of three layers: Presentation, Dialogue Management and Data. Components on the presentation layer interact with the user through several channels. The current implementation these channels are web and telephone. Components in the dialogue management layer control dialogue flow. These components can interact with the presentation and data layers but they do not directly interact with the user. Components in the data layer manage data for the service.

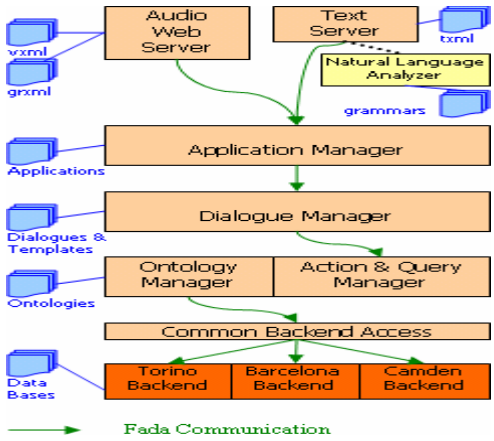

Figure 1. The architecture of the first prototype

3. Technology Description

This section describes the technologies used in the multilingual multimodal dialogue system we have developed: voice, text processing, dialogue management, knowledge management and integration technologies.

3.1 The Presentation Layer

In current implementation those components on the presentation layer control are: the Audio Web Server, supporting voice interaction through the telephone and the Text Server and the Natural Language Analyzer, supporting text interaction.

The voice components of the system are those of the multi-lingual Loquendo platform ("http://www.loquendocafe.com"), based on VoiceXML. We have selected VoiceXML because it is a standard widely adopted and it allows a rapid deployment of spoken dialogue systems to Internet-data. VoiceXML allows creating a voice application separating the content and business logic from the presentation layer. Thus, developing a voice application

becomes easy; the programmer can write it without having to know anything about the voice hardware on which the application runs.

The VoiceXML interpreter is in charge of controlling the interactions with the user defined in a VoiceXML file. The interpreter is in charge of providing help to the user and handling possible errors (there is no answer, the answer has not been recognized,…). The automatic speech recognition system uses grammars (following the Speech Recognition Grammar Specification standard) to recognize user interventions. The main reason for using grammars instead of statistical language modelling is that a significant corpus is needed for statistical language modelling. Besides, grammars can also be easily adapted when the service contents change. They can also be adapted to the language used for different types of citizens (different ages, countries, education). Grammars can also model sentences mixing two languages, that is, sentences in a specific language that include the pronunciation of specific proper names in another language (i.e., asking in Spanish information about an English movie). Furthermore, grammars can incorporate semantic information to recognize and process user interventions.

The textual input is processed by a logic-based language analyzer performing syntactic-driven semantic analysis ([1]). In order to facilitate the system management, the grammars used by this parser are obtained from those used by the voice recognizer. In future implementation of the system we will compare the results obtained when using this parser to those obtained when using a statistical parser.

3.2 The Dialogue Management Layer

The components controlling the dialogue flow are the Application Manager and the Dialogue Manager. The Application Manager is responsible for the session management and it controls the beginning of the conversation, when the user is asked to choose the language and the service. Each session is described by the user identifier, the channel type, the language and the service. The Dialogue Manager is responsible for sending the files describing the next interaction to the front-end components. Then, these components process the next user intervention and return the resulting semantic interpretation to the Dialogue Manager.

The platform we have developed is focused on task-oriented dialogues. In task-oriented dialogues the service-dependent knowledge appearing in communication is related basically to the tasks the service can perform, finding a particular piece of information or executing a particular transaction. Usually, this task can be decomposed into subtasks. In our system the task-specific knowledge appearing in communication is represented in ontologies in the data layer. Mainly, this information consists of the input/output task parameters.

During communication, the Dialogue Manager uses the tasks description in the ontology as well as the dialogue context to plan next action. If the next action is a new interaction with the user then the files describing this interaction are dynamically created.

This file contains the text representing the system messages and, if necessary, the information to be obtained from the user, that is, the attribute(s) needed by the application. If the next action is a query to the service back-end the Dialogue Manager accesses the specialized query module in the data layer, the Action Query Manager.

3.3 The Data Layer

As mentioned before the service-specific knowledge involved in the communication is represented in the ontologies. These ontologies are represented in the standard language OWL. We have selected the OWL language to represent ontologies for several reasons: It is a standard language, it is more expressive than other standard languages, its reasoning capabilities are enough for the project requirements, and it has tool support.

The service tasks are represented by a set of ontology concepts. These concepts are described by attributes (slots). These attributes represent the information required/resulting when performing the service tasks (including the input/output parameters). Dialogues consist mainly of asking/giving the values of these conceptual attributes.

In order to obtain the order of the sequence of interactions (in which context a particular slot value must be asked/given to the user) preconditions are associated with the ontology concepts. Preconditions on service attributes are a useful mechanism to support differet types of communication. In a system-driven dialogue the system asks the user the value of each of the attributes representing the information required to perform a specific task. In a mixed initiative dialogue, the user can take the initiative and the system will only ask for additional information. That is, depending on the value given to a particular attribute, information about new attributes will be required

Domain ontologies providing domain-specific semantics can also be used. Ontologies describing domain concepts and relation can improve communication for specific services, such as those related to complex domains.

We are also planning to integrate in the data/knowledge level the Electronic Service Delivery (ESD, http://www.esd.org.uk/standards), a structured category lists related to web services that are used as standards by local governments in the United Kingdom. This family of lists may be seen as lightweight ontologies providing governments with a common reference. The terms appearing in these government lists se lists are represented the Integrated Public Sector Vocabulary (IPSV), with 2732 preferred terms and 4230 synonyms. This vocabulary will be used to facilitate the integration of new services into the system.

3.4 Integration

In the system we have developed the different components are interrelated between themselves composing a service oriented architecture (SOA). We followed SOA because it favours modularization, composability, interoperability between different platforms and languages, loose coupling between modules and error recovery.

We have selected FADA technology (Federated Advanced Directory Architecture, http://fada.techideas.info) to support the SOA principles. Fada technology and its related libraries are part of an Open Source project (http://sourceforge.net/projects/fada). FADA supports interoperability between components distributed over LAN or WAN connections, mass-deployment of HOPS services for the citizens, multiple languages and operating systems, synchronous communication, interoperability behind corporate firewalls and the incorporation of new services in run-time. Although using a new technology implies an extra implementation effort, the use of FADA presents relevant advantages. In FADA, the service registry is a network where several nodes keeping information about the services are interconnected. This characteristic provides to the system a more redundant and fault tolerant nature.

4. Defining the success of voice enabled online public services

The opportunities offered by the new delivery channels present local government organisations with issues that relate to both the efficiency of service delivery and the effectiveness of the new delivery channels that methodologies for evaluating the first two waves of E-government have not been designed to address. The methodologies for evaluation of the third wave of E-government are not yet firmly established and suggestions for evaluating services that fall into the fourth wave, such as HOPS, are in the research phase.

The evaluation model proposed in HOPS goes beyond the traditional method of evaluating the delivery of E-government, and considers specific metrics that relate to Service Adequacy, Effectiveness and Efficiency. The purpose of the HOPS method is to create an evaluation methodology that can be used to evaluate projects that span the gap between what can be considered third and fourth wave E-government services.

The UK government's Office of the E-Envoy channels framework (HMSO 2002) identified the telephone as being the most popular channel for interaction with government services. User surveys in all of the local governments involved in the HOPS project also indicate that the Telephone is the most popular channel for citizens to access services. Service adequacy which is best defined as being does the service meet the individual's needs, will therefore form a key indicator in any evaluation of a system that diverts or supplements the existing call centre as the primary destination for the telephone channel. Any measurement of service adequacy is dealing with an aggregation of subjective viewpoints, consequently it is important to understand the key drivers behind citizens' perception of service adequacy.

Research carried out by the Canadian Institute for Citizen Centred Services (Citizens First 3, 2003 iccs) identified five key drivers for which are primarily concerned with citizen satisfaction in relation to service adequacy:

- Timeliness - I was satisfied with the amount of time that it took to get the service
- Knowledge Competence - Staff were knowledgeable and competent
- Staff were Courteous - Staff were polite and made an effort to help me
- Fairness – I was treated fairly
- Outcome – I got what I needed

It is important to note that out of the five indicators two are based upon the quality of customer service or human-to-human interaction. If the call centre is considered as another channel that repurposes the presentation of electronic information then it is the human-to-human interaction that distinguishes it from the other delivery channels. In moving the voice channel from a human to human interaction to a human to machine interaction it is necessary to consider which weight to put on this part of the evaluation.

4.1 Effectiveness

Whilst service adequacy is the key indicator of the Citizen's perspective the main indicator for the management of the call centre is the effectiveness of the new technology channel.

In call centre management terms effectiveness can be considered in the following two categories.

1. The total number of calls that can be diverted through to the HOPS agent
2. The number of calls that are completed by the HOPS agent

The HOPS system will also only be considered effective if the citizen is able to complete their transaction in that environment. If a significant number of citizen's find that they have to exit the system and speak to a call centre representative then this will have a negative impact on the overall effectiveness of the call centre. This is because those citizen's who have exited the HOPS system are likely to be already dissatisfied with the adequacy of the service that they have received and consequently theirenquires will take longer to deal with.

The evaluation process will record the number of calls that are existed to the call centre. The framework will also seek to capture the number of calls that are 'hung up' within a specified time frame. This will be done to capture the number of calls where the user is

dissatisfied to the point where they exit the call completely and seek to carry out their enquiry through an alternative channel.

4.2 Efficiency

Understanding the impact of efficiency gains in the public sector is more complex than within the private sector that has the sole business driver of profit. In the public sector any efficiency gain must be weighted against the potential impact on the service delivery. The other key factor in the public sector is how the costs can be reduced, as the public sector is primarily a service industry the only opportunity for realising cashable benefits from efficiency savings is to reduce the costs that are associated with people.

In order to present a holistic view of what efficiency gains an organisation might hope to obtain from the HOPS or other advanced E-government platforms it is important to consider the data from a number of perspectives. Each of these perspectives is equally valid in terms of making an investment decision but equally all of them offer a different kind of return.

In Figure 2 the matrix places the technology in one of the four corners of an input versus output model. The ideal situation would be to place the HOPS platform in the bottom right hand corner of the diagram but political considerations within any local government organisation could well lead to any of the quarters being considered to be a valid place.

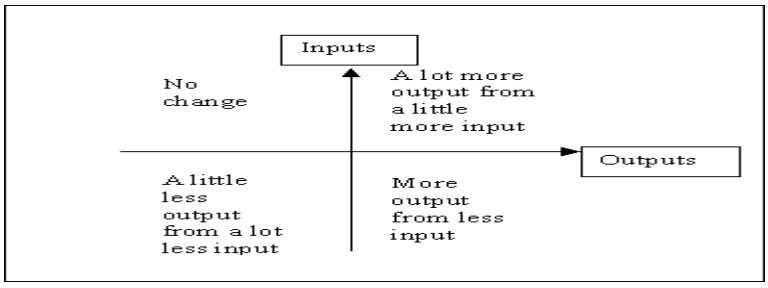

Figure 2: Input vs. output model

One of the key factors in understanding the return on investment for E-government projects is the relationship between cashable and non-cashable efficiency gains.
Whilst non-cashable benefits will improve the service delivery they will not pay for the investment required to install the system. As part of the project deliverable 6.4 a return on investment model has been produced that will baseline the cashable efficiency gains created through the installation of the HOPS platform as part of a call centre.

In the cases where either no call centre exists or a service that is included as part of the HOPS platform is one that is not currently delivered through the call centre. The ROI model will have to be adapted to contain both cashable and non-cashable efficiency gains. Essentially the difference between the cashable and non cashable gains can be explained in the diagram below:

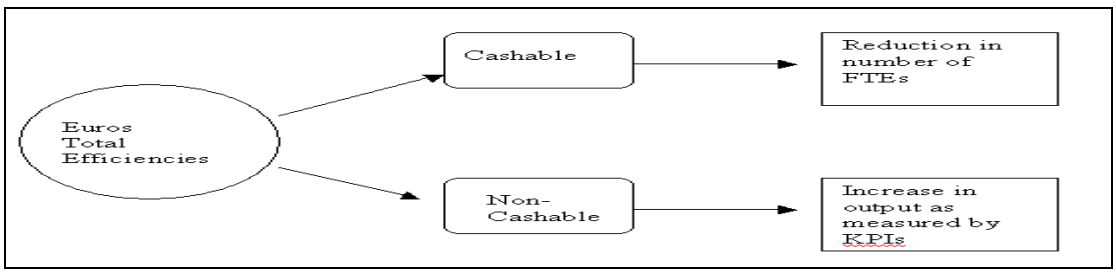

Figure 3: Cashable vs. Non Cashable gains

In terms of cashable benefits a further analysis of these is proposed as part of the evaluation framework. It is a common mistake in business process engineering to overestimate the efficiencies generated through time saved by the introduction of new processes or technologies. In order to understand the capacity for efficiencies created by the HOPS platform it is proposed to evaluate time saved against the following framework.

From the diagram above it is clear that the middle portion of the triangle can generate either cashable or non-cashable gains.

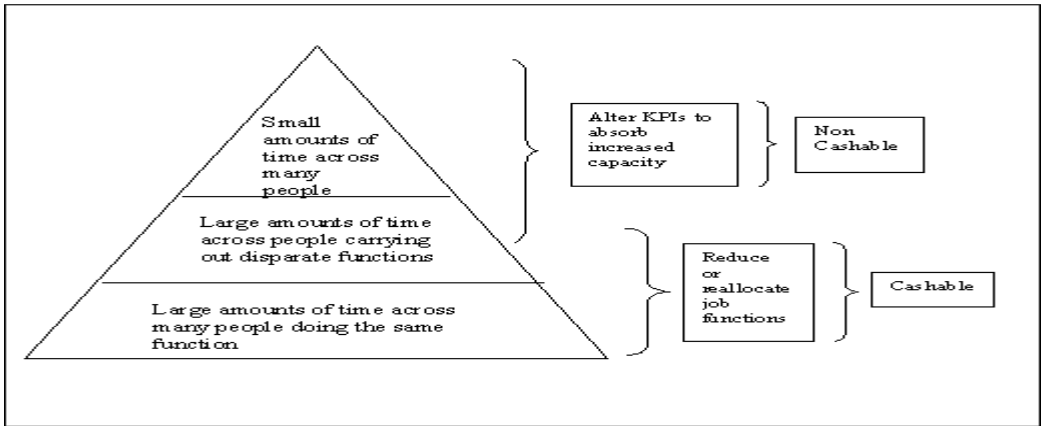

Figure 4: Distribution of gains

Conclusions

In this paper we propose the creation of a blended channel of both voice and text to leverage existing E-Government investments and break down the digital divide. For this purpose, we developed a multilingual dialogue system for enhancing the usability and accessibility of the public administration web contents. Future work will include the evaluation of the dialogue system following the evaluation methodology being developed as part of the project.

An important difference between the system we have developed and other work focused on the use of NL to improve the usability of administration web contents is in the type of voice interaction supported. While the DS we have developed supports dialogues through telephone (and through the web), the voice interaction in other systems for e-government, such as that developed for the European SAFIR project (http://www.safir-fp6.org) is limited to simple sentences thru easy devices as TV.

The HOPS project is focused in the development of a dialogue system supporting telephone access because, although voice technologies present technical limitations, the telephone channel can facilitate the access to public administration web services to all segments of population, including people not computer skilled and people living in less favoured areas. For this reason, telephone seems specially appropriate for e-government in developing countries. Besides, as the dialogue system was designed to be easily adaptable to new technologies, services and languages, the enhancement of the system capabilities incorporating other semantic web techniques (i.e., service discovering) and other languages could be a good opportunity for research cooperation between Europe and Africa.

References

[1] M. Gatius. Using an ontology for guiding natural language interaction with knowledge based systems. Ph.D. thesis, Technical University of Catalonia. 2001. Available at http://www.lsi.upc.edu/~gatius/tesis.html

[2] J.T. Richards & V.L. Hanson. Web accessibility: A broader view. In the Proceedings of the World Wide Web Conference 2004. New York, USA.

[3] D21 First User requirements. Internal technical document of the HOPS project, 2005.