SMARTCITY: CUSTOMIZED AND DYNAMIC MULTIMEDIA CONTENT PRODUCTION FOR TOURISM APPLICATIONS

Spadoni F.   **Tatarelli, A.**   **Loparco L.**   **Rossi R.**

Rigel Engineering S.r.l.,
Livorno, Italy

(spadoni@rigel.li.it)   *(tatarelli@rigel.li.it)*   *(loparco@rigel.li.it)*   *(rossi@rigel.li.it)*

Tariffi F.   **Sassolini E.**   **Ongaro P.**
Space S.p.a.,   CNR ILC,   Meta Srl,
Prato, Italy   Pisa, Italy   Lucca, Italy
(flavio.tariffi@spacespa.it),   *(eva.sassolini@ilc.cnr.it),*   *(paolo.ongaro@gruppometa.it)*

Abstract - This paper presents the SMARTCITY project experience: customized and dynamic multimedia content production for professional tourism applications.

INTRODUCTION

This paper presents the final results of the SMARTCITY project, co-funded by the Tuscany Region under the POR CREO 1.d program. The project proposes an innovative methodology as well as advanced technologies enabling professional services for cultural tourism applications in urban areas as well larger archaeological sites.

As of today, many scientific and technological efforts in the specific sector focused on the introduction of technology and all the advantages it brings in cultural and tourism scenarios, rather than content. In addition, the lack of thorough analysis of the needs and behavioural pattern of the new "experience tourists" when performing real or virtual tourist routes, strongly limits the impact that new technologies have on the development of innovative products and services capable of completely fulfil user expectations.

The SMARTCITY project defines new ways of technological and content oriented "cognitive mediation" of the cultural and tourist experience, by implementing in a simplified but pragmatic and effective way, the paradigm of the so-called ambient intelligence. The project developed methodologies and solutions to meet the emerging demand of fruition of the cultural space and cognitive mediation directly or implicitly expressed by tourists. To this end, the project tries to replace old manual techniques and "handmade" manufacturing solutions, characterized by high intellectual labour intensity and very low flexibility in reuse and adaptation of the content, with industrial oriented methodologies suitable for large scale production, aimed at different fruition target, flexible in reuse and content cross-referencing. Such innovative methodologies allow to realistically enforce the business assumption of systematic, efficient and flexible market coverage of the national landscape of cities and places of culture, starting from the Tuscany regional scenario.

The first part of the research was dedicated to accurately and concretely modeling of user needs and behavior during fruition of cultural tourism events, and to the analysis and development of an innovative methodology for information retrieval and indexing of a large unstructured knowledge base.

The second part of the project was focused to finalize a loosely-coupled system architecture integrating the different system components (knowledge extraction, indexing, authoring editor for guides) and to design and implement the authoring tool for preparing multimedia tourist guides. Finally, the authoring tool has been evaluated with real professional tourist guide authors.

To validate the proposed methodology, the project developed a collaborative platform where non technical users can easily design and develop dynamic multimedia content for the proactive and personalized fruition of tourist routes and cultural paths both in physical (e.g. in the context of art cities) and in virtual locations.

In the next Section, we will give an overview of the general methodology developed in SMARTCITY for the dynamic production of tourist guides based on a selected corpus of texts. Then, we will provide some details about the descriptive, topological and semantic thorough indexing methods used to extract meaning from for unstructured content resources. In the main Section, we will present in detail how the general methodology developed in SMARTCITY has been implemented in an authoring system dedicated to non-expert tourist guide editors, for the creation by an iterative process of search - refinement - composition - enrichment. Finally we will propose some conclusions and further work indications.

GENERAL APPROACH

The project started by analysing the behavioural patterns as well as the use of information "experience tourism". Then, the project focused on information retrieval and aimed to develop the methodological and technical foundations for a new approach to content management systems for supplying tour guides and focused on two aspects: (i) descriptive, topological and semantic thorough indexing methods for content resources, and (ii) interactions between audio and physical space, and between audio and Virtual Reality Spaces.

The main idea of the SMARTCITY approach involves the development of new technologies to improve both creation and access to tourism content. Today's tourism content engineering is characterized by a *handcraft* approach: every content is built from scratch for each specific field of application. That means that if the same content has to be published both to a paper book and a Web site, the two publications will require formatting, styling and revisioning twice. It also means that a publisher will always choose only the traditional publication formats, which can be reached by the most tourists and do not require extra work, time and money.

An *industrial* approach instead would optimize the content production process, through the management of digital content repositories. This kind of approach provides structured tourism information to each publisher. The user will not create new guides from scratch but he will be driven through a step-to-step publication, selecting the required information potentially among thousands of different sources. The available data will be automatically parsed, categorized, indexed and easily reached through categorization and intelligent suggestion of contents. Each piece of work by the publisher will not generate a single content, as in the *handcraft* approach. It instead produces a digital repository including all the required tourism information, which can now be published in several formats, without any effort from the publisher itself. The same guide can be published both to paper book and Web content, it can be available for Internet navigation, as a mobile devices application, virtual 3D tour or combination of some of them. The dynamic content generated can also be used by the next-generation of audio guides as well as for off-line and on-line interactive visit supports.

All these functionalities will be provided from the SMARTCITY environment. Moreover, an old guide can be used as a starting point to create new contents, both saving time and providing better quality to the end user.

On the other side, the industrial approach meets the requirements of the undergoing radical transformation of cultural consumption. Due to recent changes to the touristic behaviour, the tourism experience has become personal, unstructured and focused to individual interests both in terms of location and cognitive orientation. The *industrial* approach allows the publisher to provide custom content, dynamic guides, interactive experiences that can be enforced just reorganizing the original guide and choosing different publication formats. For example visiting a museum, a tourist that is interested in some specific pictures will be pleased to choose his own itinerary through the museum, and listen the description of each picture while walking through the rooms. The same guide will also be available to conventional tour guides, not requiring custom itinerary and audio description.

Another aspect involves the variety of the available information. As a very big part of the touristic targets include minor destinations, it becomes difficult to arrange together information about many less known touristic targets. Consequently it is also difficult to find similar attractions or interesting places near to a single location. The SMARTCITY approach allows finding that kind of information with automatic suggestion of similar contents and restriction of the geographic search area. As the entire database is available to the publisher, he will require no direct knowledge of the place to write about neither hours of research for documentation and cross-referencing. The arranged material can always be published in many formats and reused for future work, spending even less time.

Within the framework of the project, we created a specific domain reference (text) corpus that contains a network of knowledge, automatically extracted, concerning the city of *Empoli* and its neighbourhoods, which offers authors semantic syntheses useful for our specific search needs.

SEMANTINC ANNOTATION AND INDEXING TECNIQUES

The project devised a methodology for the creation of semantic metadata specific to the text materials, i.e. the creation and production of analytical, descriptive and semantic metadata, not only for the entire text unit (text, the volume digitalized, etc.), but also for the *atomic* textual units (at level of single page, paper, or section). The approach developed is based on a *generalized* text enrichment, which is able to identify all forms of knowledge in the text, through the integration of available technologies for statistical and linguistic analysis, without the use of hypotheses, or predefined structures and ontologies. All information extracted, are associated with the text in a *paratextual* formalism made with XML tags. The resulting well-annotated corpus is then appropriately indexed to improve search performance.

To characterize the various parts of our Corpus of text, which do not necessarily correspond to the original units, we provide the system developer the ability to identify specific parts of the Corpus, at user's request. For example, in the characterization of Points of Interest (POI), it is important to define the particular subdomain. Generally, the descriptions of POIs can be traced to known categories such as: works of art (painting, sculpture), and museum objects or archaeological excavations. By analyzing a text, there is no automatic way to distinguish public from religious monuments on the basis of the original document or by type of content. There is not enough information for the categorization of documents, particularly in text materials downloaded from the web. For this reason we are able to identify macro-areas only, on which we are able to ensure a correct attribution. The concept of *artwork* is too broad and contains within it, objects of different nature. The terminology typical of each subdomain should be sufficient to identify the relevant parts of the text, but everything is still conditioned by the size of the subdomain.

Based on past experience we have developed procedures for the generation of semantic filters (or fields of interest) to assign a score to each textual unit, representing its significance with respect to each domain of interest for each unit of text and for each volume, and a statistical summary of the semantic information extracted, such as terminology, multi-words as named entities, verbs.

The development and integration of results led to categorization of some paragraphs and / or volumes. The subdomains tested are *"Archaeology (Archeologia)"* and *"Production of glass art (Produzione del vetro artistico)"*.

For areas where automatic semantic filters have not produced accurate results, because they were not sufficiently characterized by terminology, such as: *"Biographies of famous people (Biografie di personaggi celebri)"* and *"History (Storia)"*, we have experimented a different strategy that takes into account the extracted terminology and the type of the texts. For example, there are texts relating to biographies of famous people in the Empoli area, such as Alessandro Marchetti, Ferruccio Busoni, Jacopo Chimenti, etc. that have allowed certain attribution directly from the structure into chapters.

AUTHORING SYSTEM

The previous chapter explained how the information is parsed and how the indexing process works. The usage of the indexed material is allowed by the graphical user interface and by the functionalities provided by the authoring system, that presents a step to step procedure that helps the publisher to choose the information, arrange it and choose the preferred publication features.
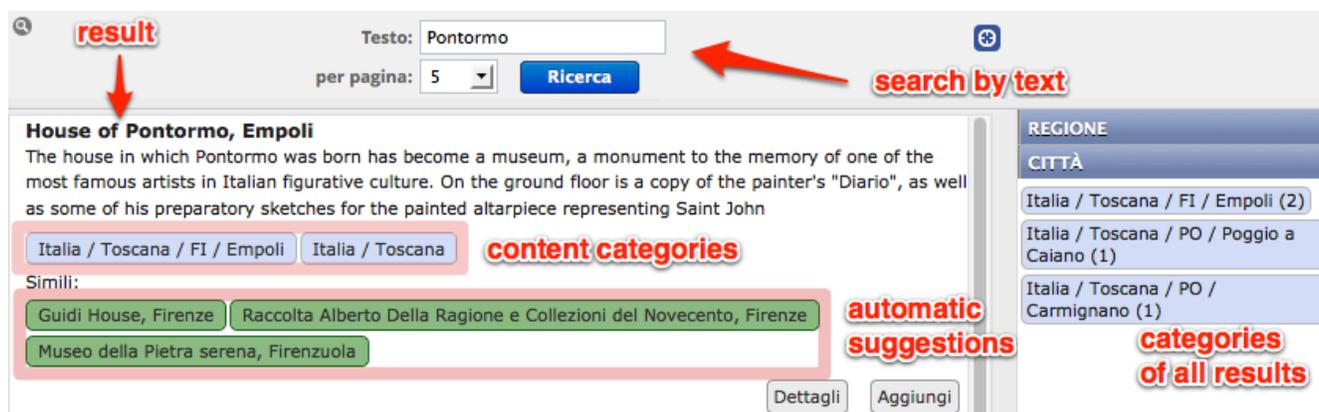

Figure 1 - Simple text search

The authoring system has a very important role, as it directly offers functionalities to the end user. If it is well structured it increases productivity. Moreover the publication process all happens here: the raw information selected using the graphical interface are automatically formatted from the back-end, and published using the user preferences.

Figure 2 - Category and map area selection

As we can see in Figure 1 - Simple text search, new material can be easily found through a text query. For each result the system will automatically give an introduction, the matched categories and a short list of similar documents.

The initial search can be further restricted to more specific results by selecting one or more categories or by map area selection, as we can see in Figure 2. In a few steps the publisher can collect a great amount of information.

After the collection stage the user can organize the content in a tree structure, adding personal modifications and searching for multimedia to associate to each content as we see in Figure 3 – (multimedia may not be published depending on the chosen format: obviously video and audio would

not be available to paper guides).

Figure 3 – Edit text and add multimedia

While organizing the guide structure, the publisher can use the map. Each point of interest is automatically located and his position is shown, in order to allow an easy tour planning. If needed, the user can also show path and directions of the itinerary (Figure 4 – Itinerary and directions).
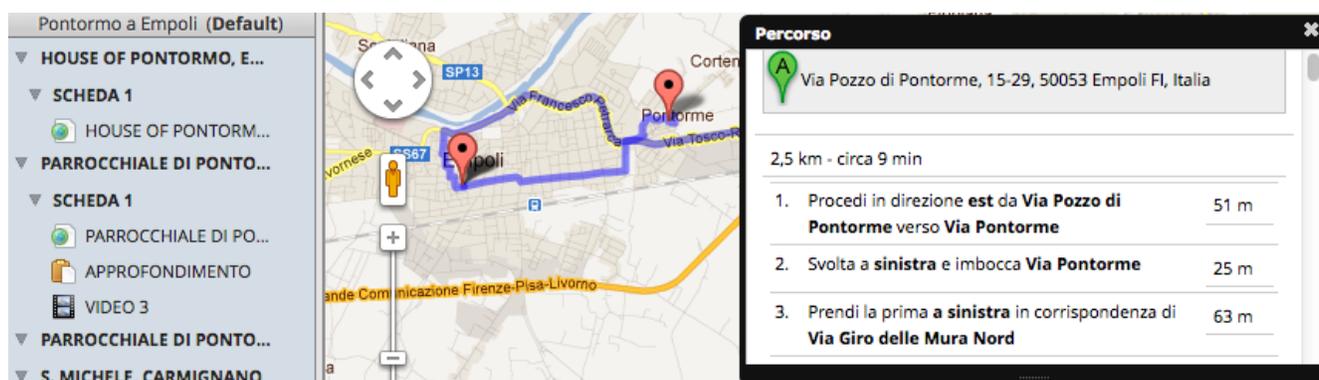

Figure 4 – Itinerary and directions

When finished to edit guide contents, the publisher will choose generator and options to publish current guide. The chosen generator will automatically format contents, insert media and create the user interface of the current publication. An example of generation can be found in Figure 5.
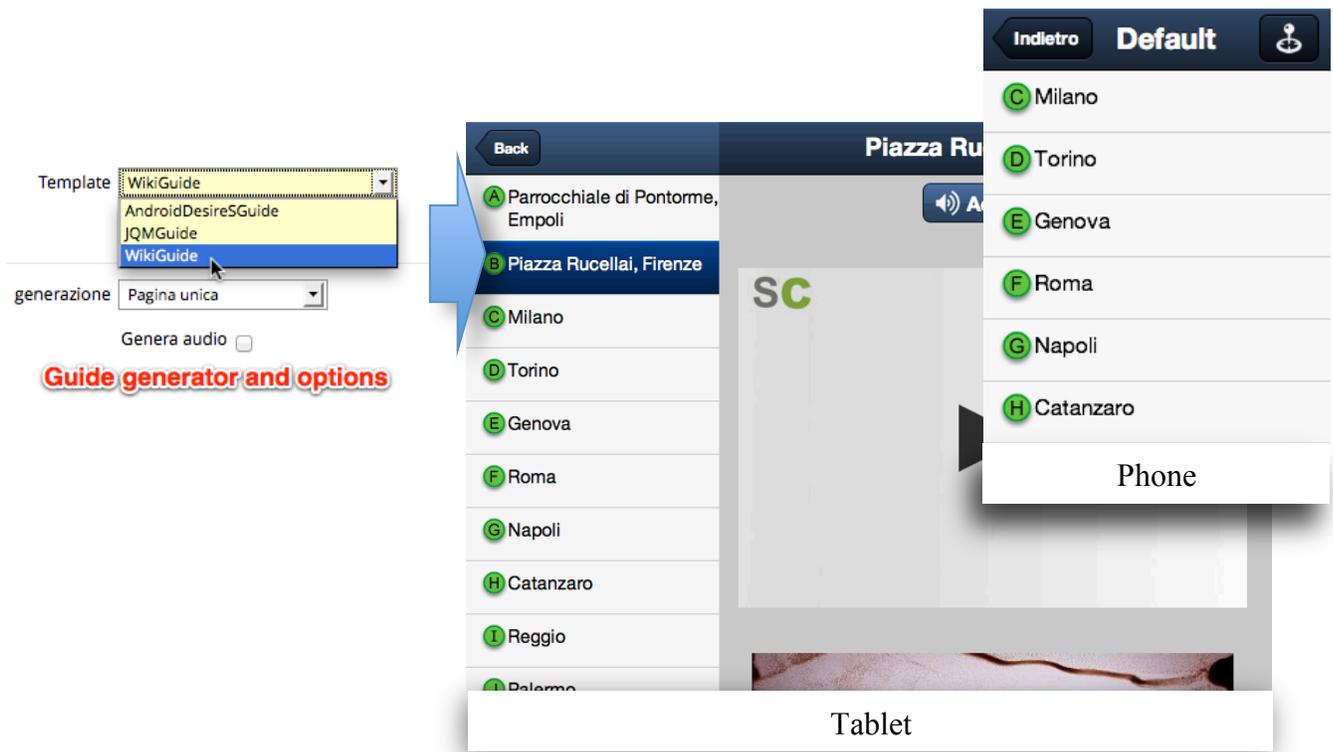
Figure 5 – Publication options and examples

ACKNOWLEDGMENTS, CONCLUSIONS AND FURTHER WORK

The project SMARTCITY developed advanced tourist guide editing processes taking advantage of industrial practice, research progress and emerging web 2.0-style collaborative technologies. The involvement of real users, i.e. professional tourist guide editors, in the development of the methodology and tools, and in the subsequent validation effort is essential for its acceptance in the target markets.

The first part of the research was dedicated to accurately analyze and model needs and behaviors of users during the fruition of cultural tourism events, and to the development of an innovative methodology for annotating and indexing a large unstructured knowledge base. Then, the project implemented the authoring tool for preparing multimedia tourist guides, integrating the different components described in this papers (knowledge extraction, annotation and indexing, authoring editor for guides). The authoring tool interface was designed to accommodate the large amount of heterogeneous data (text, multimedia, maps) and to allow operators to efficiently manage the knowledge intensive process of search – selection – use – refinement and publication activities defined by the SMARTCITY methodology.

Further work is envisaged in the industrialization effort to bring the current prototypes to the market, including internationalization (the current semantic annotation is only for the Italian language) as well as for extending the methodology to different applications such as theatrical productions, theme parks or videogames.