Framework for UNESCO Intangible Cultural Heritage

Maria Teresa Artese; ITC - CNR (National Research Council); Milan; Italy
Isabella Gagliardi; ITC - CNR (National Research Council); Milan; Italy

Abstract

UNESCO, in its 2003 Convention for the Safeguarding of the Intangible Cultural Heritage, obliges State Parties to inventory cultural heritage and making those inventories accessible. It allows a great flexibility to determine how the inventories should be prepared. In this paper we will present the framework for the inventorying, cataloging, searching and browsing of multimedia digital objects related to ICH (Intangible Cultural Heritage), for preservation and safeguard, that we used to design and implement the AESS (Archivio di Etnografia e Storia Sociale Lombardy Region) archive: the data concern mainly popular traditions handed down generation by generation, such as traditional fairs, popular songs, and customs. We will present for the first time the features that characterize our approach, describing the framework we developed, within a methodological approach for designing applications, pointing out the characteristic features of the data and of the online system.

Introduction

According to the 2003 Convention for the Safeguarding of the Intangible Cultural Heritage of UNESCO, the intangible cultural heritage (ICH) or living heritage is defined as the practices, representations, expressions, as well as the knowledge and skills, that communities, groups and, in some cases, individuals recognize as part of their cultural heritage[1]. The Convention states that the ICH is manifested in the domains of oral traditions (including languages), performing arts, social practices, knowledge, and traditional craftsmanship. The process of inventorying intangible cultural heritage and making those inventories accessible is one of the specific obligations outlined in the UNESCO 2003 Convention for the Safeguarding of the Intangible Cultural Heritage. UNESCO allows enough flexibility for a State Party to determine how it will prepare its inventories. Moreover, intangible heritage elements should be well defined in the inventories to help put safeguarding measures into practice. Online various inventories and archives for ICH objects are available, at different degrees of deepness of information and interaction [3].

In this paper we will present the framework for the inventorying, cataloguing, searching and browsing of multimedia digital objects related to ICH (Intangible Cultural Heritage), for preservation and safeguard. We used the framework to design and implement the AESS (Archivio di Etnografia e Storia Sociale Lombardy Region) archive[2]. AESS has been founded to preserve, study, and enhance the value of documents and images of life, social transformations, literature, oral history, material culture, and anthropic landscapes of the Lombardy territory: the data concern mainly popular songs and other audio/video records describing the popular traditions handed down generation by generation, such as traditional fairs, and customs.

The database, together with software tools for cataloging, searching and browsing constitutes a de facto standard for the cataloging of cultural heritage legacy in Italy, and is used not only by the Lombardy Region, but adopted also by the Puglia and Emilia Romagna Italian Regions, thus satisfying the principle of reusability and interoperability required by the guidelines for the construction of a good digital collection [14].

In the paper we will present for the first time the features that characterize our approach: what are the data, users, aims, place and time for the system use and how the data is shown in a granularity suitable for the device, user and aim, what are the tools (standard or innovative). We will describe the framework and the different environments we implemented, within a methodological approach for designing applications. We will focus on technological aspects related to framework. We will omit aspects related to semantics of the cataloging, which are the responsibility of experienced ethnographers.

The paper is structured as follows: after a description of the framework, with a detail view of the features we investigated in our approach, the environments of the framework are described. Then, we conclude with a discussion of the system presented and future developments.

The Framework

In designing and implementing the intangible heritage framework, we adopted a methodology to eliminate or almost minimize the design-errors and to ease the maintenance phase.

Over the years a number of methods have been proposed for the development of hypermedia (the hypermedia paradigm is based on the idea of organizing information as chunks freely browsable by users selecting links and making use of other advanced navigation tools, such as indexes or maps) and web applications, from RMM [10] (and its extension) [11] to ADM [7], including OOHDM [15], HDM [9], and WebML [6] and later on for WIS (Web Information Systems) applications [4] [17].

The first phase has led to the identification of requirements, data structure, different environments of the framework, with intended target users and aims, place and time of its use. Among the basic requirements that we intend to comply, are indicated:

- Usability: the experts must be able to enter or edit data without any technical knowledge or prior training phase. The search engine must be interrogated in a simple and intuitive way [13];
- Accessibility: the information should be accessible to everyone, even and especially, to people with disabilities. In particular, the website and the services provided should not de-

[1] http://portal.unesco.org/en/ev.php-URL_ID=17716&URL_DO=DO_TOPIC&URL_SECTION=201.html/

[2] http://aess.itc.cnr.it/

pend on the browser used (independence from the browser). The Web Accessibility Initiative (WAI) of the World Wide Web Consortium (W3C)[3] has defined the rules and made available some tools for assessing the accessibility of Web sites;

- Collaborativeness: the system should allow specialists to enter, expand, modify and delete any entries, even on the same information at the same time. The system must be designed to allow concurrent access to data, keeping information safe and secure;
- Scalability: the data structure must be designed according to the continuous updating of the web and technology, to adapt to new needs and the availability of new technological solutions;
- Interchange and compatibility: the structure must allow the export of data, choosing the appropriate level of semantic description. It should also take into account the existing standard, such as metadata standards [5], SKOS [12] etc...;
- Multilingual: The system must implement an information architecture that is independent of language. The list of languages of interest can be updated at any time, depending on the needs of the project, for which the same entry can be inserted and made available in multiple languages, through links from one language to another, as well as the data entry and search interface.

AESS (Archivio di Etnografia e Storia Sociale Lombardy Region) database has been structured having in mind the principal aims of the system, that is to store and maintain all the information the ethnographic catalogers considered important not only for the web browsing to disseminate cultural heritage but, above all, to preserve and safeguard the traditions through the production and storage of digital documents, by ethnographers and specialists. In addition to protecting and preserving the intangible cultural heritage, it serves to keep traditions, customs, and languages alive which are used in niche fields or in small groups that might otherwise be lost with the loss of the people who handed them down from generation to generation. It is used to circulate practices, knowledge, and dances to enhance the area in order to increase tourism and to diffuse good traditional practices.

AESS data structure (shown in figure 1 and described in more details in [1]) is rather complex, having to cope with raw data constituted by interviews, transcription of lyrics, transcription of music, scores, manuscripts, typographical fonts, photographic documents, cinematographic documents in form of sound files, images, videos and text representing events, fair, tradition etc. These are the result of focused ethnographical and anthropological field works.

The ITEM is the core entity: it is the minimal information unit that can be cataloged (for example, a song, an interview, a drawing), and is described by title, incipit, metric, keywords. In most cases it is accompanied by audio/video/texts etc. This information is stored on physical media (its SUPPORT, for example the book or the digital file), connected to physical objects in the library through the signature. The entity TOKEN is the physical representation of the item, specifying one of the 5 categories provided: audio, video, image, text, pentagram, and stores the relationship between the ITEMs and their SUPPORTs: the number of the track on the CD, the page number of the book. Its digital file, if available, is uploaded and stored in the MULTIMEDIA CONTENT entity together with ancillary data (author, date, place,). An ITEM can be related to another ITEM through a self-relation. A FUND is a comprehensive set of SUPPORTs and therefore of documents. The recordings of the songs/dances/interviews (each cataloged as an ITEM) are conducted during FIELDWORKS, at festivals, fairs, events or special recording registration campaigns, part of more general ethnographical RESEARCHes. A RESEARCH, on its turn, can be part of another RESEARCH. The PERSON/GROUP and BIBLIOGRAPHY are connected to all entities, to records all the persons/groups/organizations who are involved, in any role, in cataloging intangible heritage, and bibliographic references. The entity LEXICON allows to translate specific dialect terms into Italian and to report variants of the same term in different villages.

The ICH entity stores information on cultural heritage as a whole. If a cataloger wants to enter detailed information about songs, interviews and videos recorded during a Carnival, for example, he has to compile the related ITEM/TOKEN/SUPPORT/MULTIMEDIA CONTENT cards, otherwise, if he is interested only on the Carnival as an event, MULTIMEDIA CONTENT can be connected directly, supplying information such as title, author, data and locality for each content. ICH card is maintained alive adding year by year new information concerning its temporal repetition, for instance new images or videos of the last event. All information linked to ICH card allows to reconstruct and display in a web page the evolution of the event, showing both its textual descriptions and its multimedia timeline representing the event history.

Data stored in the database belong mainly to the Lombardy Region, but Puglia and Emilia Romagna Region data, together with data related the cross-border project AESS-E.CH.I. (Ethnographies Italy Switzerland) with partners as Piedmont Region, Valle d'Aosta and Trentino Alto Adige, including neighboring countries such as Switzerland and France, flow in the AESS Central Database. It is necessary to maintain the virtual breakdown of such corpus data, to be able to allow the identification and extraction of "virtual databases" at any time.

The framework is composed by the following three environments (see fig. 2):

1. *Data entry*, which deals with the insertion and management of data, according to AESS data structure. In order to keep them consistent and up-to-date and to avoid useless duplication, all data (belonging to all organizations) are stored in a single database (AESS Central database);
2. *AESS Search*, which displays the data related to AESS Lombardy Region database or other virtual databases to the target users through paths, prepackaged query or free text search;
3. *IntangibleSearch*, which enables users to navigate in ICH heritage, through indexes and free text search, displayed on a map.

The following users have been identified, according to their specific aims:

Data entry

- supervisors users: people who, with special permits, control

[3] http://www.w3.org/

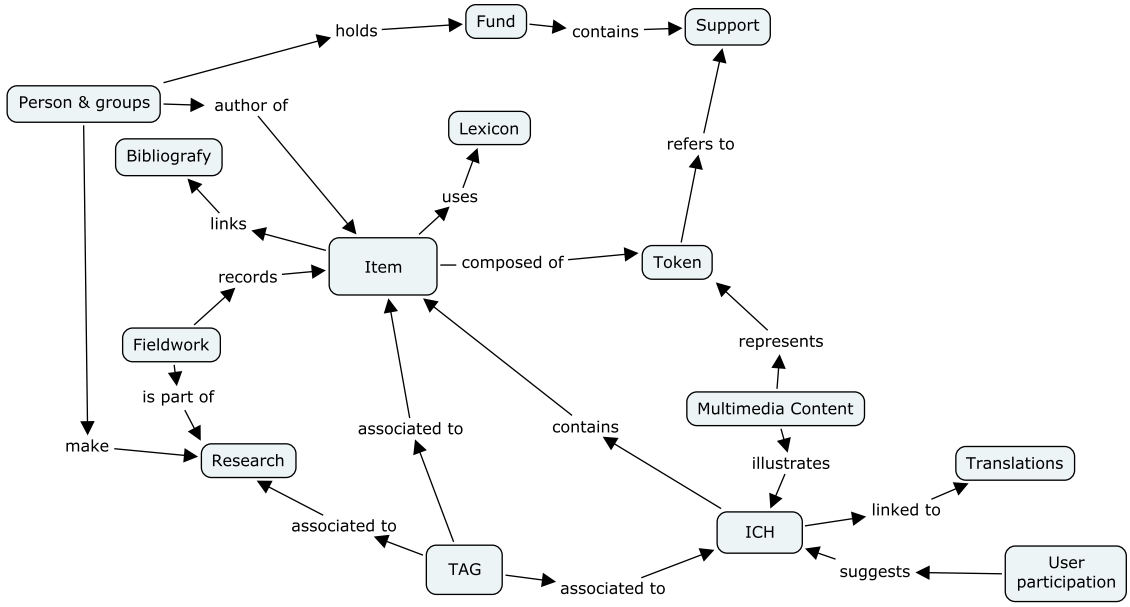

Figure 1. *AESS data structure.*

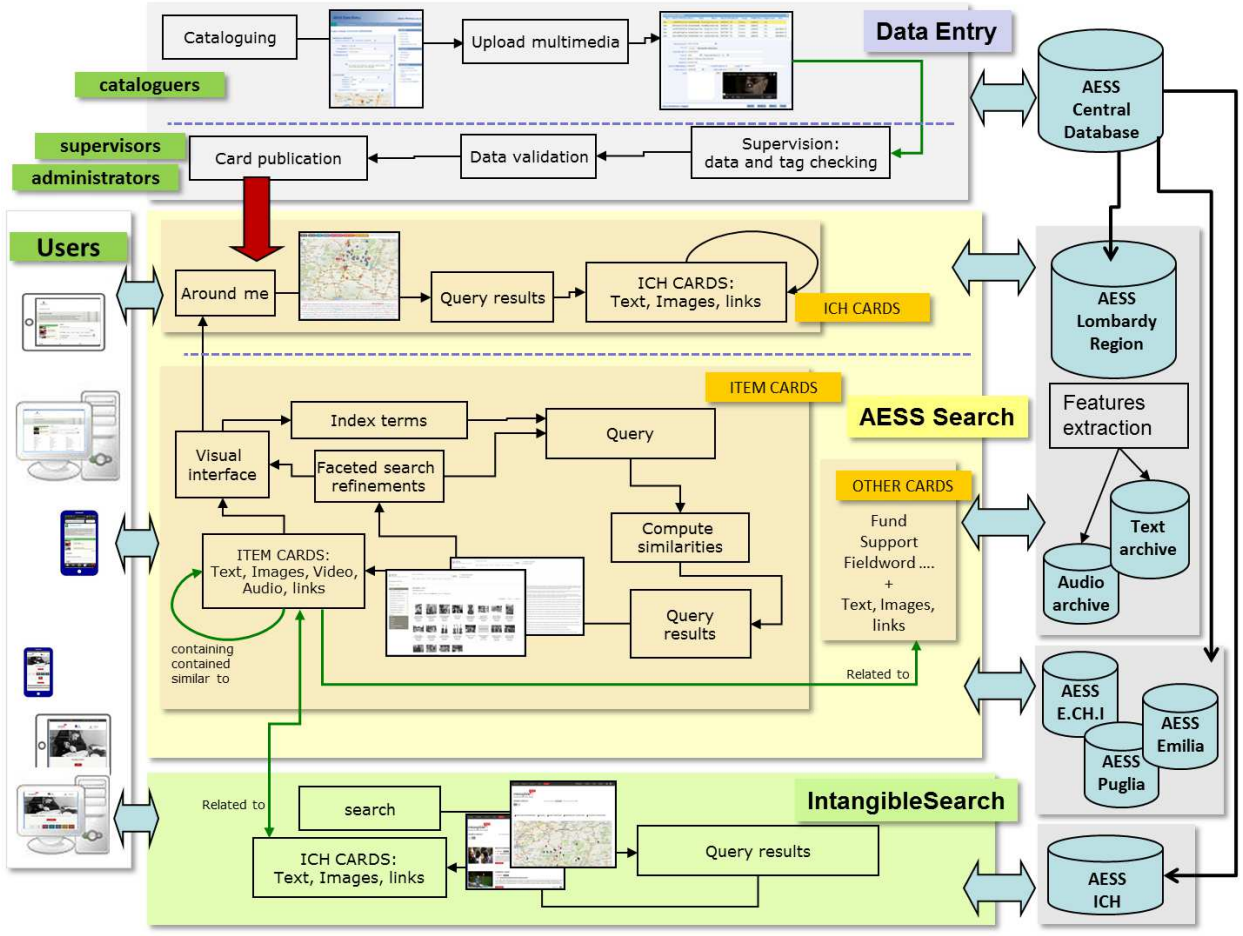

Figure 2. *AESS framework environments.*

and correct data already entered, validate the cards, manage vocabularies and lists of tags, and create new users;
- catalogers users: people who access the database to insert new data, documents and digital media;
- Region staff: experts or specialists, working at Lombardy/Puglia/... Region, who are in charge of the cataloging process;
- experts: people who access the database only for specialist consultation, without any intervention on the content of the cards;

AESS Search

- web users: people who consult the on-line database, looking for the material of interest, browse through the information available jumping from one topic to another;
- Region staff and experts as described above;

IntangibleSearch

- community: people who know the cultural intangible heritage and help to keep them alive and to suggest new ones;
- tourists: people who want to be informed about events, fairs, carnival around them, to enrich their knowledge and live new experiences.

Each user category has a preferred device/place/time:

- Catalogers/supervisors use the system preferably in office, with a wide screen computer to compare and organize information;
- experts, Region staff use the system either during work time to analyze and study or during fieldwork to identify information;
- web users use the system to enrich their knowledge, either during work time or spare time;
- tourists use the system on holiday with smartphone or tablet to enjoy or have information of events around;
- community use the system to protect ICH they are aware of, by suggesting them to the experts.

The system adjusts the granularity of the content shown, according to the different devices (such as tablet, smartphone, pc), automatically detected, and to the action taken in the three environments. For example results on a small device will be focused on multimedia and interactive elements, otherwise complete data will be displayed.

Data Entry

Data Entry[4] environment is thought for catalogers, supervisors, administrators, Region staff and experts, and it is protected by two levels of password access. Specific tools have been developed for catalogers, to facilitate the integration of data and multimedia content, the connections between data, the detailed view of inserted cards, as well as for supervisors to control, clean, validate and publish data in a simplified manner. For experts, who need only consult the card, there are a number of restrictions to allow full view without being able to alter the data in any way. The Data Entry handles all types of cards provided by the AESS Data structure and, regarding to ICH cards, is possible to insert their translations in the 5 languages (Italian English, French, German and Ladin). Database data currently inserted count 15.612 audio, 2.226 videos, 28.160 images, 16.785 text transcripts and 4.753 PDF files with staves. Regarding ICH cards, currently there are 352 Italian cards, with translations into other languages, each accompanied by its own multimedia materials and possibly related to ITEM cards already in the database.

The system has been designed to handle the fact that various organizations enter their data to form a single database, to avoid unnecessary duplication, but need to query and view only their own data. To make this separation possible, automated procedures have been developed that perform a data processing to create virtual databases. The data of Emilia Romagna Region, for example, are located physically in *AESS Central Database*, but need to be published, accessed and search separately: thus the procedure creates overnight *AESS Emilia* database. In this way AESS Central Database is never exposed to public access from outside, while remaining protected and available only for supervisors/catalogers.

AESS Search

AESS Search[5] offers tools and searching/browsing modes to cope with the complexity of AESS data structure, and to allows experts and Region staff to perform queries to retrieve the objects of interest.

The ITEM (in fig. 1), the leading actor of this module, is shown here through its physical representation: audio, video, image, text, or pentagram. For example, the document describing the popular song ”La Cecilia” can live simultaneously in multiple formats such as the audio file in MP3 format, the text transcription of the book page containing song words, the PDF file containing the musical score. When the user searches for ”Cecilia” he will find the same document in the audio section with the MP3 file of the song, in the text section with the transcription of song words, in the stave section with the musical score.

The system offers a two steps search mode, with a faceted query refinements. Faceted search [18] aims to combine navigational and direct search to leverage the best of both approaches. In a typical faceted search interface, users start by entering a query into a search box. The system uses this query to perform a full-text search, and then permits navigational refinement on the results of that search.

The home page of AESS Search contains different search mode:

- Guided tours, designed by supervisors to guide users to discover particular subset of data within the database, automatically extract the documents and show them together with an abstract describing their particular features. Examples of these pre-packaged searches are Puppets and puppet, Carnival of Schignano, The Lombardy of peasants, Sounds of tradition; supervisors can modify them or add new ones;
- Search indexes are shown directly to users, divided into 6 categories, defined by experts. Inside each category terms are sorted in descending order with respect to the number of linked documents.

[4] http://aess.itc.cnr.it/data_entry/index.htm
http://aess.regione.lombardia.it/data_entry/index.htm

[5] http://aess.itc.cnr.it/ricerca
http://aess.regione.lombardia.it/ricerca

- Free text search: users can search the database for any free text. It is possible to restrict search within one of the 5 sections: audio, video, image, text and pentagram.
- Search on map, through "Map of the Intangible Heritage" specific for ICH cards, opens a map showing all ICH cards available, with their tags and UNESCO categories. The list of tags are displayed with different sizes depending on the number of cards referred, computed in real time. By selecting the tags and categories, users can filter ICH cards and open it directly through the available marker on the map. If the search is performed through a smartphone or a tablet, with a GPS, heritage around me are shown.

In the second step, it is possible, by selecting one or more options, to refine the current research, and to specialize it further by narrowing the set of documents retrieved, using textual indexes associate. Moreover, methods and tools for the content-based retrieval have been developed and integrated into the system: similarity-based textual description [16]; similar audio retrieval and cluster of audio files [8].

Custom instances of AESS Search can be realized for the virtual archives, in a seamless and semi-automatic way, through a procedure that allows to specify their own logo, header, title, data, and even graphic style of web site. Figure 3 shows the homepage of Lombardy Region and Puglia Search engine: it can be noticed that, while maintaining the same structure, color, text, images have been changed.

In the designing phase, particular attention has been paid to make content and graphics independent each other so to easily adapt contents to new needs and new devices, such as smartphones and tablets, automatically recognized.

IntangibleSearch

IntangibleSearch[6] [2] is the search engine related to ICH database, the inventory of the intangible cultural heritage, and make it simple, usable, easily accessible the navigation for the categories of users identified in the analysis phase. ICH cards have associated multimedia and/or any specific document (ITEM) in the database for a detailed description. The navigation flow of the web site is:

- Search: Predefined queries related to the UNESCO categories, protagonists (people who play a role in the intangible heritage), tags (with a description) and text search are available to users.
- Results: Users are shown the results in a map, in alternative to the standard display in list form.
- Details: A detailed description of the ICH card is provided: a great importance is given to the multimedia, shown as a timeline on the top of the page. Timeline provides the user with the history of the living heritage, its evolution over time: it tells the content directly through the voice of people living a ritual, practicing a knowledge, constructing an object, singing a story. For a in-depth navigation, a link to the corresponding card in aess.regione.lombardia.it is provided.

IntangibleSearch offers its interface in 5 languages: the documents are shown in the same language of the interface, if exists, otherwise in Italian. Web 2.0 tools are integrated, with links for RSS, Facebook, Twitter and suggestion of Cultural Heritage.

[6] `http://www.intangiblesearch.it`

Conclusion

In this paper we have presented a framework for the inventorying, cataloguing, searching and browsing of multimedia digital objects related to ICH (Intangible Cultural Heritage), according to UNESCO guidelines. The framework is composed of three environments, whose main features are the following.
Data Entry allows:

- to tailor the interface for the different organizations;
- to enter information and multimedia related to any entity in the AESS data structure;
- to create virtual databases to be shown and searched separately from AESS search.

AESS search principal strengths are related to data complexity management; in particular it:

- works on the whole complex AESS structure;
- offers different search modes;
- can be tailored for virtual databases;
- integrates similarity search tools.

IntangibleSearch offers a standard three-steps search, with particular attention to

- Multilingual user interface and cards (if available);
- Multimedia fruition;
- Web 2.0 characteristics: user suggestion, comments, Twitter and Facebook link.

Future developments will include the integration of tools for annotation and retrieval of similar images, as well as for more interactive display modes. Tools for export data in standard format (rdf/xml/rss) will also be developed to make the information available to others for re-use, e.g. for open linked data.

Acknowledgments

This study was performed as part of a Lombardy Region research contract, the project: ECHI Etnografie italo-svizzere per la valorizzazione del patrimonio immateriale- INTERREG. Thanks to the Region, and in particular to the Archivio di Etnografia e Storia Sociale for permission to publish this paper. Thanks also to Assetweb, for their precious work in the implementation of intangiblesearch.it

References

[1] M. T. Artese, I. Gagliardi, Cataloging intangible cultural heritage on the web. Proc. Progress in Cultural Heritage Preservation, pg. 676. (2012).

[2] M. T. Artese, I. Gagliardi, Browsing and searching UNESCO Intangible heritage on the web: two ways, Proc. Digital Heritage International Congress (DigitalHeritage), pg. 417. (2013).

[3] ASPACI, Identificazione partecipativa del patrimonio culturale immateriale, Progetto, E., & di Cooperazione Transfrontaliera, P. O. a cura di ASPACI Associazione per la Salvaguardia del Patrimonio Culturale Immateriale. (2011).

[4] P. Barna, F. Frasincar, G.J. Houben, and R. Vdovjak, Methodologies for web information system design, Proc. Information Technology: Coding and Computing 2003, pg. 420. (2003).

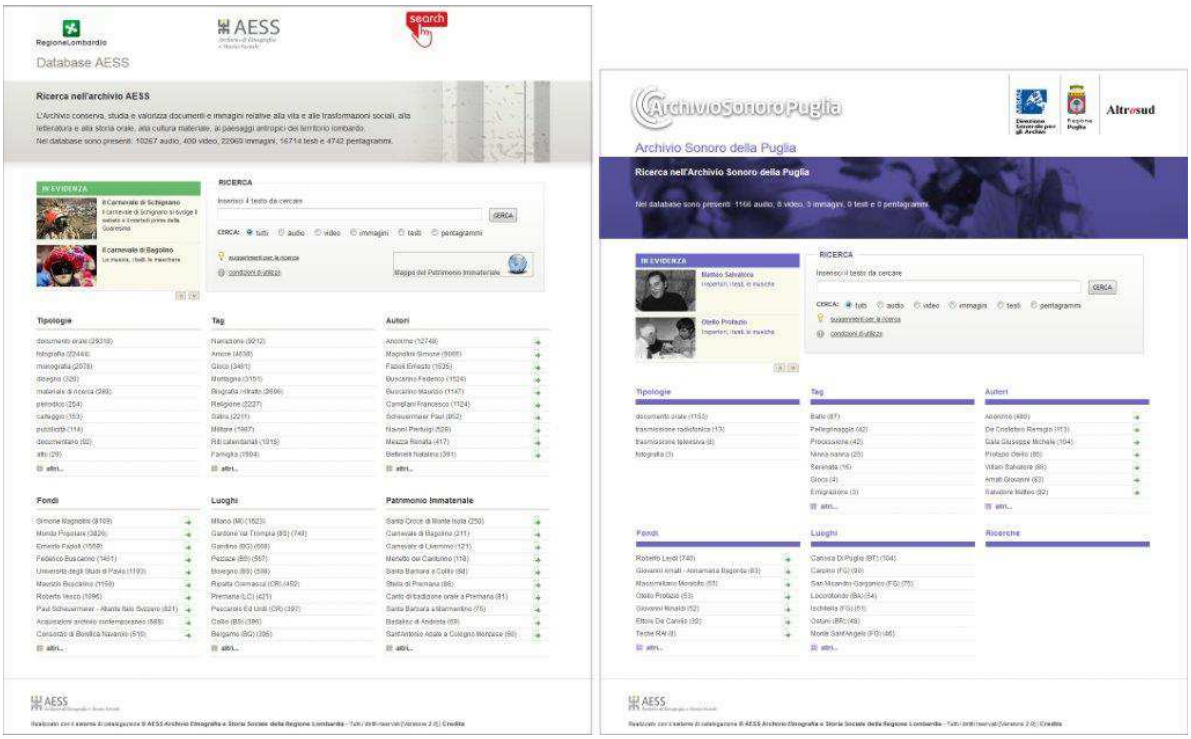

Figure 3. *AESS search homepage for Lombardy Region and Puglia Region.*

[5] W. Cathro, Metadata: an overview. National Library of Australia Staff Papers.(2009).

[6] S. Ceri, P. Fraternali, and A. Bongio, "Web Modeling Language (WebML): a modeling language for designing Web sites", Computer Networks, 33, 137 (2000).

[7] P. Diaz, S. Montero, and I. Aedo, "Modelling hypermedia and web applications: the Ariadne Development Method", Information Systems, 30, 649 (2005).

[8] I. Gagliardi, and P. Pagliarulo, Audio information retrieval in HyperMedia environment. Proc. the sixteenth ACM conference on Hypertext and hypermedia, pg. 248. (2005).

[9] F. Garzotto, P. Paolini, and D. Schwabe, "Hdm: a model-based approach to hypertext application design", ACM Transactions on Information Systems (TOIS), 11, 1 (1993).

[10] T. Isakowitz, E.A. Stohr, and P.Balasubramanian, "Rmm: a methodology for structured hypermedia design", Communications of the ACM, 38, 34 (1995).

[11] T. Isakowitz, A. Kamis, and M. Koufaris, The extended rmm methodology for web publishing, New York University (1998).

[12] A. Miles, S. Bechhofer, SKOS simple knowledge organization system reference. Technical report, W3C (2009).

[13] J. Nielsen and H. Loranger, Prioritizing web usability (Pearson Education, 2006).

[14] NISO Framework Working Group A Framework of Guidance for Building Good Digital Collections, NISO Baltimore, Maryland (2007), ISBN (13): 978-1-880124-74-1 (2007).

[15] G. Rossi, D. Schwabe, F. Lyardet, Web application models are more than conceptual models. Proc. Advances in Conceptual Modeling, pg. 239. (1999).

[16] R. Schettini, G. Ciocca, A. Colombo, I. Gagliardi, M.T. Artese, Intangible heritage management and multimodal navigation. Proc. Axmedis, pg. 45. (2008).

[17] W. Schwinger, W. Retschitzegger, A. Schauerhuber, G. Kappel, M. Wimmer, B. Pröll, C.C. Castro, S. Casteleyn, O. De Troyer, P. Fraternali, "A survey on web modeling approaches for ubiquitous web applications", International Journal of Web Information Systems, 4, 234 (2008).

[18] R. Senjuti Basu, W. Haidong, D. Gautam, N. Ullas, M. Mukesh, Minimum-effort driven dynamic faceted search in structured databases,, Proc. of the 17th ACM conference on Information and knowledge management, pg. 13. (2008).

Author Biography

Isabella Gagliardi *works at Construction Technologies Institute of the National Research Council (ITC-CNR) Multimedia Information Technologies section. Her major areas of re-search include Hypermedia Information Retrieval models and methodologies, automatic generation of hypertextual links between text/audio/image, dynamic web based database design and implementation, and clustering algorithms. Recently he has focused his research on the development of multimedia information systems available on the web. On these topics she has several national and international publications, and she has been working in national and international research projects.*

Maria Teresa Artese *is a researcher at Construction Technologies Institute of the National Research Council (ITC-CNR), took her degree (Laurea) in Computer Science at the University of Milan in 1990, and since then till 2000 she has been working in a private as programmer. Since 2000 she has been working at the Institute of Multimedia Information Technologies of the Italian National Research Council, where her research has focused on information systems and the development of new methodologies and algorithms for multimedia data fruition on the web.*