THE PROJECT OF THE ITALIAN CULTURE PORTAL. A STANDARD BASED MODEL FOR INTEROPERABILITY AMONGST CULTURAL HERITAGE DATA SOURCES.

M. Emilia Masci [a *], I. Buonazia [a], D.Merlitti [a]

[a] SNS, Scuola Normale Superiore di Pisa, Piazza dei Cavalieri 7, 56124, Pisa, Italy
(e.masci, i.buonazia, d.merlitti) @sns.it

KEY WORDS: Classification - Cultural Heritage – Data structures – Interface – Interoperability – Metadata – Retrieval – Standards

ABSTRACT:

The paper describes the scientific and technical project carried out by SNS for the Italian Culture Portal, promoted by Italian Ministry of Cultural Heritage and Activities, as a media for knowledge and enhancement of Cultural Heritage. Its main mission is to communicate the ensemble of Italian culture, making available existing data and giving wider visibility to cultural websites. The Portal won't duplicate data in a central repository, on the opposite it will harvest metadata from external data-sources. The adoption of OAI-PMH will allow metadata migration from content providers to one or more harvesters, adding services as indexing and automatic classification. As the Portal must join very diverse contents assuring system scalability, it seemed unsuitable to use a data-model with predefined Entities. Thus, accomplishing with world-wide used Standards, a Metadata Schema based on Dublin Core has been designed. As DC Element Set is very restricted and groups many information into one Element, Qualified DC, which refines DCES using Refinements and Encoding Schemes, has been preferred, in order to assure interoperability together with a minimum loss of information and more significant retrieval.

Following DCMI recommendations, a DC Application Profile based on the domain of Italian Culture has been designed, adding new types of resources ('Corporate Body', 'Physical Person', 'Project'), Refinements to define specifications and relations amongst Element and Encoding Schemes as controlled vocabularies and thesauri adopted by Italian cataloguing Standards. Finally the paper describes the design of the user interface, specially criteria for search and browse functionalities.

1. INTRODUCTION

In September 2004 the Italian Ministry of Cultural Heritage and Activities (MiBAC) committed the scientific and technical project for the Italian Culture Portal to Scuola Normale Superiore di Pisa (SNS), which delivered the project during 2005. At the moment SNS is working as a consultant for MiBAC to flank the company which is carrying out the Portal, which will be named "CulturaItalia".

The main mission of the Italian Culture Portal is to communicate to different kinds of users the whole ensemble of Italian culture, as a media conceived for the diffusion of knowledge, promotion and enhancement of cultural heritage. Thus, CulturaItalia will offer access to the existing resources on cultural contents and will give more exposure to the vast amount of websites pertaining to museums, libraries, archives, universities and other research institutions.

Users will access resources stored in various repositories browsing by subjects, places, people and time. It will be possible to visualise man information about the single item and to further deepen the knowledge directly reaching the websites of each institution.

The Portal will harvest metadata from different repositories and will export metadata to other national and international Portals. It will also provide contents created and managed by an editorial office, to offer updated news on the main cultural events and to provide thematic itineraries for a guided navigation through the harvested contents.

Resources originating from various data-sources will remain under the control of institutions responsible for their creation, approval, management and maintenance: data will not be duplicated into the Portal's repository and will be retrievable through a unified and interoperable system.

In order to guarantee the interoperability of various kinds of cultural resources and to allow retrieval and indexing functions on their contents, a specific Dublin Core Application Profile has been designed on the basis of the complex domain of "Italian Culture". The PICO AP (so called from the Project's acronym), which will be exposed in this paper, has been currently reviewed and improved according to the first mapping experiences made by SNS on some repositories, whose contents have been chosen to be harvested by CulturaItalia. The PICO AP will be soon published on a PURL (Persistent Uniform Resource Locator).

2. ANALYSIS OF USERS, MISSION AND DOMAIN

The project for CulturaItalia has been developed through the following steps:

- users and domain analysis;
- definition of user scenarios and use cases;
- overall architecture design;
- content analysis;
- analysis of the state of the art on descriptive metadata standards;
- design of the metadata schema (PICO AP);
- design of the user interface;
- project prototype.

The project moved from the analysis and definition of the expected users' target, consequently on the identification of users' needs and requirements.

The identification of potential users of the portal moved from the requirements issued by MiBAC, which pointed out that the Culture Portal should be distinguishable in its domain and functionalities both from the official web site of MiBAC (http://www.beniculturali.it/), oriented to people in charge of management and preservation of Cultural Heritage, and from the Portal for Italian Tourism. Moreover, potential addressees of a cultural portal have been identified with the analysis of some of the most important European and international portals (as the

French Culture Portal http://www.culture.fr/, or the British http://www.24hourmuseum.org.uk/), and of websites of cultural institutions such as museums, theatres, universities, etc.

The target of the Portal has been identified in Italian and foreign users, such as tourists and people interested in, and passionate of, culture; business users (publishers, merchandising agents, etc.); young people, from primary to high school; culture professionals such as scholars, museums curators, researchers, etc. Special contents and services (thematic paths, newsletters) will be created for each kind of user. It is important to notice that users can be people with physical or cognitive disabilities, and full accessibility must be assured, according to Italian law (n. 4 issued 2004-01-09) and W3C recommendations.

The mission of CulturaItalia is to promote Italian culture and heritage in Italy and abroad, attracting web users toward cultural themes and giving visibility to Italian cultural public and private institutions and to activities and projects focused on culture. Moreover, CulturaItalia aims at identifying, promoting and integrating existing resources, such as websites, databases, digital libraries, offering an index of Italian cultural resources and heritage, in order to create flexible and scalable relations between resources and to allow interoperable queries on indexed subjects, places, events, and people.

The Domain of Italian Culture is a wide concept, conceived in different ways. MiBAC is responsible for preservation, management, research and exploitation of the Italian cultural patrimony, which is composed by tangible and un-tangible heritage. Tangible heritage consists in architectural and environmental objects; artworks and collections; manuscripts, edited books as well as the current literature; archaeological and demo-ethno-anthropological objects; contemporary art and architecture. Un-tangible heritage deals with events pertaining to music; dance and theatre, circuses and street performances; cinema, humanities and scientific culture.

3. PROJECT'S METHODOLOGY

On the basis of domain and users analysis, and on possible sustainability strategies, some user scenarios have been written, describing different approaches to the Portal, by different kind of users. Scenarios described the following users, functionalities and inter-actions:

- Foreign tourist: language selection, access from the map, browsing and e-booking;
- General user: disambiguation of query results, use of contents suggested by the editorial staff and linked to results of user's query;
- Italian teacher with partial visual deficit: accessible set up, simple and advanced search, registration to the Portal, submission of a comment to the editorial staff;
- Foreign researcher: free and advanced search, access to the web site identified through the Portal;
- Journalist: search amongst cultural events, purchase of printable pictures, registration, download;
- Publishing house: search amongst images, contact for banner exchange;
- Tourists with motion deficit: browsing from place and events, visualization on the map, participation to forum;
- Italian high school student: simple search, print function and e-commerce tools.

Adopting UML (Unified Modelling Language), such descriptive scenarios have been transformed in use cases diagrams, identifying actors, both human (different final users of the front end and of the back office) and IT components, and inter-

actions between actors and the system from the first query to the final result.

Adopting UML methodology has been useful also to improve cooperation in a staff composed by IT developers as well as cultural domain experts, overcoming the gap of different languages.

On the basis of main functionalities identified by the user requirements, the core components of the System Architecture have been designed; as the project should be used as a - non mandatory - feasibility study for the final development, costs and benefits of some existing systems and components have been considered.

Moving from the analysis of the contents foreseen for the Portal, the best solution has been identified in a harvesting procedure. A study of the state of the art on different metadata standards, categories and taxonomies, controlled vocabularies and ontologies for describing cultural resources, suggested to define a specific DC application profile, which will be in depth described later.

Finally the user interface has been designed, specially focussing on the functionalities of searching and browsing.

4. METADATA HARVESTING

CulturaItalia will give integrated access for information pertaining to the domain of "Italian Culture", as it has been defined in a previous chapter. Resources coming from various data-sources will not be duplicated into the Portal's repository. On the contrary, it will offer an index of those contents by harvesting metadata pertaining to their data.

Before harvesting, metadata will be mapped into one metadata schema, which permits the indexing, browse and query functions on the whole ensemble of harvested contents. Metadata will be harvested using the Protocol for metadata Harvesting by Open Archive Initiative. This protocol allows the metadata migration from content providers to one or more harvesters, adding services as indexing system or automatic classification. OAI-PMH uses HTTP protocol for data transfer and XML for data coding.

Each institution responsible for contents to be harvested will establish, together with MiBAC, which data will be accessible from the Portal, as some resources or part of them could contain confidential information that shouldn't be published. Contents coming from external data-sources will be imported in the Portal trough the harvesting of metadata and the mapping in one metadata schema.

Contents to be harvested by CulturaItalia, as pertaining to a domain very wide as the one above described, are extremely heterogeneous. Moreover, new content providers will join the project, offering data which presently could be not predictable. Therefore, it seemed unsuitable to use a data model with predefined entity types: on the opposite, in order to guarantee system's scalability, the designing of one metadata schema has been preferred as a more flexible solution.

In the first phase of the project, an analysis of the state of the art of the most used models for data description has been carried on. Amongst the most used standard must be mentioned: VRA - Visual Resources Association (VRA), CDWA - Categories for the Description of Works of Art (Getty, 2006), the CIDOC-CRM, which focuses mainly on art objects (ICOM, 2000); EAD - Encoded Archival Description (LC, 2002) and MARC - MAchine-Readable Cataloguing format (LC, 2007), used in the Archives and Libraries domain. Also, most relevant thesauri concerning cultural domain have been taken into consideration, from the cultural section of UNESCO thesaurus (UNESCO, 2003) to ULAN – Union List of Artists Names (Getty, 2000b) and AAT – Art and Architecture Thesaurus (Getty, 2000a),

issued by Getty Research Institutes. These models are very useful for a specific domain, but they are not completely reusable for such a wide range of contents. Dublin Core standard seemed more suitable, as a more general metadata model, useful to assure interoperability into a system which has to join contents codified in very different data models. Moreover, adopting DC would allow the possibility of introducing more specific refinements. In order to be compliant with more used international standards, Dublin Core metadata scheme has been adopted.

5. THE PICO APPLICATION PROFILE

As the Portal will join different kinds of contents, it seamed unsuitable to use a data model with predefined entity types. For guaranteeing system's scalability, a flexible solution has been preferred, which consists in the designing of a unique metadata schema: to respect world wide used standards, the Italian Culture Portal will adopt a metadata set based on Dublin Core standard.

This standard is very used because it consists in one scheme that can be applied to every kind of resource, distinguished by the element 'Type'. Anyway, it is not really efficient for cultural resources because, as the DC Element Set (the so called 'Simple DC') is very restricted, many different information must be grouped into one element (DCMI, 2006a). For this reason, in the last years Dublin Core Metadata Initiative (DCMI) divulged the Qualified DC schema, which refines DC Element Set using Element Refinements and supporting Encoding Schemes, to attribute to a given property the value selected from a controlled vocabulary, a thesaurus, or an ontology (DCMI, 2006b).

Thanks to Dumbing Down algorithms, now developed in XML, in the data sharing with a system that supports Simple DC, it is possible to reduce Qualified DC values into Simple DC values. With this process, there is a minimum loss of information and more possibility to obtain a significant retrieval. At the same time, interoperability between repositories based on Dublin Core is assured.

DCMI suggests to institutions and research groups to develop DC Application Profiles for specific applications and domains, designing schemas which can join:
- All, or a selection of, DC Elements and Refinements;
- Elements from one or more element sets;
- Elements from locally defined sets (Heery – Patel, 2000).

A DC Application Profile has been designed for the Portal of Italian Culture on the basis of recommendations, documents and samples published by DCMI, in order to define further extensions specially conceived to retrieve information pertaining to Italian culture. This application profile could be further expanded for harvesting eventually unexpected contents in the future, by adding Refinements and Encoding Schemes that could be necessary for data retrieval.

The PICO AP has been designed by I. Buonazia, M. E. Masci and D. Merlitti (SNS working group on metadata, supervised by U. Parrini). It has been recently improved on the basis of the first mappings performed on some data-models or metadata schemas related to contents to be harvested by CulturaItalia. An official publication is currently under development. It will be edited on a PURL, following the DC AP Guidelines.

This DC Application Profile joins in one metadata schema all DC Elements, Element Refinements and Encoding Schemes from the Qualified DC and other refinements and encoding schemes specifically conceived for the CulturaItalia domain.

Therefore, the following namespaces are included into this metadata schema: 'dc:', 'dcterms:', 'pico:'. This application profile could be further expanded for harvesting eventually unexpected contents in the future, by adding Refinements and Encoding Schemes that could be necessary for data retrieval. It has been recently improved on the basis of the first mappings performed on some data-models or metadata schemas related to contents to be harvested by CulturaItalia.

The PICO AP will be soon published according to the Dublin Core Application Profile Guidelines, issued as the CEN Workshop Agreement CWA 14855 (CEN, 2003). Schemas used for the PICO AP is going to be published and related namespaces will be made available under the PURL top level domain 'pico'.

In the current activity of generating metadata for items coming from data-sources which were originally not provided with metadata, the use of Dublin Core Structured Values syntax has been adopted (Cox – Iannella, 2000), in order to express, within the single DC Element or Element Refinement, simple or structured values, also providing a label which can be useful for a better understanding by the user.

6. PICO TYPE VOCABULARY

One of the most relevant extensions of PICO AP is the introduction of three more types. The DCMI Type Vocabulary, conceived for the description of digital resources, seemed not completely fitting for the above mentioned domain foreseen for CulturaItalia. For this reason the PICO AP extended the DCMI Type Vocabulary (DCMI, 2006c), introducing the PICO Type Vocabulary, a controlled vocabulary which adds the types 'Corporate Body', 'Physical Person' and 'Project'. These three new types have been introduced also for an efficient mapping into the PICO AP of the entities described in the data model of the European Project MICHAEL (Multilingual Inventory of Cultural Heritage in Europe) (MICHAEL, 2006); MICHAEL aims at describing digital cultural resources at a collection level, and will have a mutual relationship of content providing with CulturaItalia. The PICO Type vocabulary defines types as follows:

'Corporate Body': any public and private, Italian and foreign Institution, in charge of preservation, management, research, enhancement of cultural heritage and activities. Corporate Bodies can be mutually related; a Corporate Body can be responsible and/or owner of physical objects, events, collections, products and services;

'Physical Person': any dead or alive physical person, related anyhow to other resources (artists, novelists, scholars, performers, publishers, etc.). Physical Persons can be in various ways responsible for the creation, publishing, management of a resource, a project, an event, etc.;

'Project': Italian and foreign initiatives for research, preservation, management, enhancement of cultural heritage and activities, in which Corporate Bodies and Physical Persons can be in different ways implied. Projects could include, or be related to, other projects, be the context in which resources are digitized, services or events are created, etc.

7. PICO AP QUALIFIERS AND ENCODING SCHEMES

A list of Element Refinements and Encoding Schemes introduced into the PICO follows, with a brief definition of each.

7.1 PICO AP qualifiers for the element Creator

Author (Element-Refinement)= Any dead or alive physical person, any corporate body and institution, responsible for the creation of a resource. It can be a writer, a painter, an architect, a musician, a photographer, a collector (as the author of the collection).

Commissioner (Element-Refinement)= Any dead or alive physical person, corporate body and institution, responsible for the commission, the order and/or the funding of the design of a resource.
ULAN - Union List of Artist Names (Encoding Scheme)= Controlled vocabulary by The Getty Research Institute.

6.2. PICO AP qualifiers for the element Subject

Thesaurus PICO (Encoding Scheme)= Thesaurus composed by hierarchically structured keywords for indicating the topic of all the resources included into CulturaItalia. This ontology includes terms for assigning the resources to the index and to the themes menu of the Portal.
UNESCO Thesaurus (Encoding Scheme)= Thesaurus for indicating the topic of resources on education, culture, natural, human and social sciences, communication and information. Multilingual: English, French, Spanish.
AAT - Art and Architecture Thesaurus (Encoding Scheme)= Thesaurus defined by Getty Research Institute for indicating the topic of resources pertaining to art and architecture objects.
ICONCLASS (Encoding Scheme)= Taxonomy of the iconographic subjects for the Western Art, from Medieval to the Contemporary Art. Multilingual: English, German, Italian, French, Finnish (KNAW, 2006).

7.2 PICO AP qualifiers for the element Description

Information (Element-Refinement)= Information about the resource, as opening and closing hours.
Contact (Element-Refinement)= Information about contacts related to the resource.
Services (Element-Refinement)= Services offered by the resource. E.g. cafeteria or restaurant services, services for unpaired people, laboratories and activities, extra.

7.3 PICO AP qualifiers for the element Publisher

Distributor (Element-refinement)= Any dead or alive physical person, any corporate body and institution, responsible for the distribution of an edited or published resource.
Printer (Element-refinement)= Any dead or alive physical person, any corporate body and institution, responsible for the print of an edited or published resource. This term comprehends both printers of physical (books, journals, images, etc.) and digital (CD, DVD, etc.) resources.

7.4 PICO AP qualifiers for the element Contributor

Editor (Element-refinement)= Any dead or alive physical person, any corporate body and institution, responsible for the making, editing or organisation of the resource. E.g. the editor of a volume of proceedings or of an exhibition.
Performer (Element-refinement)= Any dead or alive physical person, which contributes to the execution of the resource by acting a performance, with reference to some entertaining events in particular. E.g. an actor, dancer, singer, musician, etc.
Producer (Element-refinement)= Any dead or alive physical person, any corporate body and institution, responsible for the artistic and/or economic production of the resource. This term is used for producers of cinema, music, theatre, etc.
Responsible (Element-refinement)= Any dead or alive physical person, any corporate body and institution, responsible for the management, organisation, administration, etc. of the resource or of a part of it. In some cases it coincides with the contact person, whose contacts are indicated for people who are looking for information about the resource. E.g. the responsible of a

project or of one of its work packages, a museum director, the director of a university or of a department, etc.
Translator (Element-refinement)= Any dead or alive physical person who made the translation of the resource.

7.5 PICO AP qualifiers for the element Type

PICO Type Vocabulary (Encoding Scheme)= Controlled vocabulary which includes some resource types specifically conceived for the Italian Culture Portal domain: Corporate Body, Physical Person, Project. Those types are not foreseen by the DCMI Type Vocabulary.
CD Type - Collection Description Type Vocabulary (Encoding Scheme)= A list of types that categorize a collection (DCCD, 2006).

7.6 PICO AP qualifiers for the element Format

Material And Technique (Element-refinement)= The material of the object and of its support and the technique of execution of a resource of type ' Physical Object'.

7.7 PICO AP qualifiers for the element Identifier

ISBN - International Standard Book Number (Encoding Scheme)= The International Standard Book Number is an uniform and persistent identifier for a given title or for the edition of a title pertaining to a given publisher. See: http://www.isbn.org/ or http://www.isbn.it/.
ISSN - International Standard Serial Number (Encoding Scheme)= The International Standard Serial Number is the international identifier for serial publications such as printed or digital newspapers and periodicals. See: http://www.issn.org/.Identifier.

7.8 PICO AP qualifiers for the element Relation

Preview (Element-refinement)= Any form of abstract, reduction, image, video streaming used as anticipation of the resource.
Promotes (Element-refinement)= The described resource promotes and/or organizes the referenced resource.
is Promoted By (Element-refinement)= The described resource is promoted and/or organized by the referenced resource.
Manages (Element-refinement)= The described resource manages with different responsibilities (scientific, administrative, technical, etc.) the referenced resource.
Is Managed By (Element-refinement)= The described resource is managed with different responsibilities (scientific, administrative, technical, etc.) by the referenced resource.
Is Owner Of (Element-refinement)= The described resource owns the referenced resource.
Is Owned By (Element-refinement)= The described resource is owned by the referenced resource.
Produces (Element-refinement)= The described resource produces in its physical, or administrative, or any other issue, the referenced resource.
Is Produced By (Element-refinement)= The described resource is produced in its physical, or administrative, or any other issue, by the referenced resource.
Performs (Element-refinement)= The described resource performs, directly participating (e.g. as actor or musician) to, the referenced resource.
Is Performed By (Element-refinement)= The described resource is performed by the referenced resource.

Is Responsible For (Element-refinement)= The described resource is anyhow responsible for, or is the contact person of, the referenced resource.

Has As Responsible (Element-refinement)= The described resource has as responsible and/or contact person the referenced resource.

Contributes To (Element-refinement)= The described resource contributes anyhow to the referenced resource.

Has As Contributor (Element-refinement)= The described resource is produced, managed, organized with the contribution of the referenced resource.

Digitises (Element-refinement)= The described resource is responsible of the digitisation of the referenced resource.

Is Digitised By (Element-refinement)= The described resource is digitised by the referenced resource.

Anchor (Element-refinement)= Reference to the URL of the web-page publishing the resource described by the metadata record.

7.9 PICO AP qualifiers for the element Coverage

Date of Birth (Element-refinement)= Date of Birth pertaining to resources with type: Physical Person.

Date of Death (Element-refinement)= Date of Death pertaining to resources with type: Physical Person.

Place of Birth (Element-refinement)= Place of Birth pertaining to resources with type: Physical Person.

Place of Death (Element-refinement)= Place of Death pertaining to resources with type: Physical Person.

ISTAT Code (Encoding Scheme)= Code assigned by Istituto Nazionale di Statistica italiano (Italian National Institute for Statistics), which identifies inhabited places in the Italian territory. See: http://www.istat.it/strumenti/definizioni/comuni/

Postal Address (Encoding Scheme) = Postal address of a resource with type: Physical Object or Corporate Body.

8. USER INTERFACE

The final phase of the project design, and preliminary basis for the development of a project prototype, has dealt with user interface and content management. CulturaItalia will publish different kinds of contents:

static contents: Head and logo, access to multilingual versions, credits, contact information, mission, site map, copyright;

dynamic contents, from CMS: news, itineraries, focus, press release, forum, FAQ, newsletter, specific areas (e.g.: young users);

dynamic contents, from harvesting: metadata harvested from external repositories;

business logic contents, depending on the user session: search results, bookmarks, etc.;

user inputs: layout personalization controls (font, contrast, colour), registration area to access in a private area to save bookmarks, annotate events in agenda, etc.

The interface will allow data retrieval on those contents trough different possibilities for searching and browsing. User will access contents through three kinds of searches:

free search: user composes one or more words, using boolean syntax;

advanced search: user refines the query in the catalogue, selecting if the item to be retrieved is "place", "person", "event", or "object", refining through the selection of the resource type or of the data provider (if known);

geographic search, selecting a place on a list or on a map related to a GIS system.

It will be possible to browse the catalogue through the Main Menu or the Theme Menu. According to the 4 High Level

Elements of DC Culture, defined by Aquarel project and approved by MINERVA project (MINERVA, 2003), the Main Menu of the catalogue is structured in:

- Who: people, institutions, administration offices, museums, archives, libraries, universities, etc.;
- What: art objects, monuments, documents, books, photos, movies, records, theatre and music productions, etc.;
- When: contents retrievable through temporal periods;
- Where: browse by region, province, town, on a controlled list or directly on a GIS.

User will browse the catalogue using a 'faceted' system: he can start the query from one of the four elements and further refine the results range, browsing according to another criterion.

A simplified alternative for browsing is the Themes menu. It groups the resources according to the following arguments: Archaeology, Architecture, Visual Arts, Environment and Landscape, Cinema and Media, Music, Entertainment, Traditions, Humanities, Scientific Culture, Education and Research, Libraries, Literature, Archives, Museums, Exhibitions.

The Portal will not publish only resources harvested from external repositories, but will produce also new contents: an editorial office will prepare and manage contents to provide interesting relations between resources and make the user discover them through links among different kinds of information. Those new contents will be tailored on different users' targets and will be distributed into the following sections of the Portal:

Itineraries: articles focused on a theme, aimed at suggesting a virtual tour through some resources selected from the catalogue;

Focuses: short monographs on a single argument;

Events: information on cultural events (exhibitions, concerts, theatre, conferences);

News: selected news on Italian culture.

Finally the Portal should provide services, to be eventually implemented in a later phase after the first realization, such as multilingual versions, newsletter and forum to attract a recurrent access by users and an area specially devoted and conceived for young users.

9. CURRENT WORK AND CONCLUSION

The Portal is currently under development. Reply S.p.A., an Italian IT enterprise selected in an European context, is developing the technical system, while the editorial staff, under MiBAC supervision, is preparing contents and identifying new providers. SNS is flanking MiBAC in testing functionalities and interfaces of the system, and works as consultant for identifying new content providers and data sources, analysing the data models adopted by each provider, defining mappings to the PICO AP, monitoring and improving results of harvesting procedures.

As a matter of fact, while interoperability through different data systems is everywhere an increasingly growing need, in Italy the practice of providing data with standardized metadata is still at its very beginning and data sources provided with metadata (even 'Simple DC') that can be directly harvested are currently very rare. Therefore one of the main activities carried out by SNS for flanking the Portal development is to generate automatically metadata, on the basis of a mapping between the data model adopted by each data provider and the PICO AP.

This activity also aims at making cultural institutions, such as museums, libraries and archives, more conscious that the use of metadata is a basic issue in the creation process of cultural digital resources, in order to assure an efficient retrieval (thus,

knowledge and use of relevant, authoritative and scientific information) through national or international networks.

References from websites:

All the following websites have been accessed on 2007-05-21.

Cox, S. – Iannella, R., 2000. *DCMI DCSV: A syntax for writing a list of labelled values in a text string*, http://dublincore.org/documents/2000/07/11/dcmi-dcsv/.

CEN, 2003. European Committee for Standardization, *CWA14855 - Dublin Core Application Profile guidelines*, http://www.cen.eu/cenorm/businessdomains/businessdomains/iss/cwa/cwa14855.asp/.

DCCD, 2006. Dublin Core Collection Description Working Group, *Dublin Core Collection Description Application Profile*, http://www.ukoln.ac.uk/metadata/dcmi/collection-application-profile/#cldCLDT.

DCMI, 2006a. *DCMI Metadata Element Set, Version 1.1*, http://www.dublincore.org/documents/dces/.

DCMI, 2006b. *DCMI Metadata Terms*, http://www.dublincore.org/documents/dcmi-terms/.

DCMI, 2006c. *DCMI Type Vocabulary*, http://www.dublincore.org/documents/dcmi-type-vocabulary/index.shtml/.

Getty, 2000a. Getty Research Institutes, *Art & Architecture Thesaurus*, 2000, http://www.getty.edu/research/conducting_research/vocabularies/aat/.

Getty, 2000b. Getty Research Institutes, *Union List of Artist Names*, http://www.getty.edu/research/conducting_research/vocabularies/ulan/

Getty, 2006. Getty Trust - College Art Association, M. Baca - P. Harpring (edited by), *Categories for the Description of Works of Art*, http://www.getty.edu/research/conducting_research/standards/cdwa/index.html/.

Heery, R. – Patel, M., 2000. *Application profiles: mixing and matching metadata schemas*, Ariadne Issue 25, http://www.ariadne.ac.uk/issue25/app-profiles/.

ICOM, 2000. ICOM / CIDOC Documentation Standard Working Group, *CIDOC object-oriented Conceptual Reference Model (CRM)*, http://cidoc.ics.forth.gr/index.html/.

KNAW, 2006. Royal Academy of Arts and Sciences Amsterdam, ICONCLASS, on line edition, http://www.iconclass.nl/.

Kokkelink, S – Schwantzl, R., 2002. *Expressing Qualified DC in RDF/XML*, http://dublincore.org/documents/dcq-rdf-xml/.

LC, 2002. Library of Congress, *EAD 2002 Schema*, http://www.loc.gov/ead/eadschema.html.

LC, 2007. Library of Congress, *MARC Standards*, http://www.loc.gov/marc/.

MICHAEL, 2006. *MICHAEL Project Data model*, http://www.michael-culture.eu/doc.html.

MINERVA, 2003. MINERVA PROJECT, *Handbook for quality in cultural Web sites: improving quality for citizens*, http://www.minervaeurope.org/publications/qualitycriteria.html.

UNESCO, 2003. UNESCO Thesaurus, on line edition http://www2.ulcc.ac.uk/unesco/.

VRA, 2007. Visual Resource Association, *VRA Core 4.0*, 2007, http://www.vraweb.org/projects/vracore4/index.html/.