MACE: Connecting Architectural Content Repositories to Enable New Educational Experiences Inside a Collective External Memory

Vittorio Spigai[1], Massimiliano Condotta[1], Elisa Dalla Vecchia[1]
Stefan Apelt[2] and Christian R. Prause[2]

[1] University IUAV of Venice, Faculty of Architecture, Department of Construction Building, Venice, Italy
[2] Fraunhofer FIT, Mobile Knowledge Group, Sankt Augustin, Germany

Abstract: In the practice and learning of Architecture and Civil Engineering, it is fundamental to access a big amount of learning material. A considerable part of the knowledge which once was written in books is now being moved to digital media. Today, most of the contents are produced and presented in digital format only. Spread around the world, digital content repositories containing a big amount of notions exist, but are oftentimes unknown and disjointed. As a consequence, they are not very efficient resources for learning at the moment.

The European research project MACE (Metadata for Architectural Contents in Europe) aims at connecting digital architectural repositories by harvesting their metadata and enriching it through the integration of content and domain, context, competence and process, and usage and social metadata. The network created will allow for federated access and search over all connected repositories, allowing a new way of exploring notions and knowledge in the architectural domain, using the web as a "*collective external memory*" [4].

Keywords: MACE; architectural repositories; browsing; facetted glossary; architectural taxonomy

Introduction

Since the start of solving design problems, when visiting for the first time the physical environment hosting their "spatial problem solving", architects conduct information seeking activities based on "a process of scanning and glancing" [1]; "eye scan the visual field" and by a "pre-attentive perception identify objects or message of interest": the elements that most of all appear important and influence the design approach.

The necessity to rely on a variety of notions and information therefore starts from the preliminary phases of the design process and continues in all following steps. In some cases it is characterized by a *purposive information-seeking behavior*, when this is a conscious and focused activity; in other cases by *seemingly non-purposive information behaviour* [2] when it is an instinctive, almost involuntary action of drawing information from the most disparate fronts, an action sometimes disguised as creation or invention, but actually a work of referencing to them.

In any case, whether the goal is to support a designer, may he be a professional or a student, active in sectors as architecture, city planning and civil engineering design (*design problem solving activity*) or to support a researcher interested in architecture's world, documentation and technical knowledge, but without any design-applied goal (*documentation activity*) [3], the possibility to access a big number of information, examples, notions, etc. is fundamental. It is true that this is an important factor in all the disciplines, not only in architecture and civil engineering; nevertheless, it has to be considered that in architecture, a large amount of information is held in visual media (images, photos, sketches…), which are generally hard to index and to find by means and tools currently available.

A large amount of this information is stored online in specific existing architectural repositories; however, these repositories are often closed to non-registered users and insufficiently known, not connected among each other and not always easy to navigate.

One aim of MACE is to create a common indexing strategy to structure these “rambling pieces” of architectural information. The final goal is to allow learners to have, through strengthening and optimising the on–going knowledge digitalisation process, new ways of exploring notions and knowledge: a multiplication of the learning opportunity held using the web as a “collective external memory” [4].

To fulfil this scenario, the starting point is to classify and index a large amount of learning objects (LOs) from connected repositories with metadata developed for this purpose. These metadata are subsequently used to trace LOs during the user seeking activity.

1 Introduction to MACE metadata

There are different “levels” of data. On the lowest level (L0), there are real objects (buildings or famous architects that one can touch or talk to) or non-real architectural contents (theories, concepts). One level above (L1), there are (digital) photos or texts that describe or refer to buildings, architects, concepts and projects. On the level above (L2), metadata containing information of digital objects (creation date or author of photos and texts) or content-related information (keywords about the subject) are managed.

L0	L1	L2
Real world	Digital object	Metadata

Table 1 - Levels of data

In digital knowledge organisation, metadata are very important for describing digital contents in machine-understandable ways so that contents can be categorised, filtered, searched and retrieved via their associated metadata.

A lot of existing metadata can be read from content repositories and stored in a central metadata store. This procedure is called “metadata harvesting” and is in MACE implemented through OAI-PMH[1] [5]. Only metadata describing learning objects are transferred; the learning objects themselves are not copied and remain in control of their owner without changing access conditions for contents.

Bridging the gaps between information from several heterogeneous repositories created independently is another problem. Most often these repositories do not use the same metadata standard for describing their contents. In order to close these gaps, MACE uses one application profile consisting of different kinds of metadata.

2 MACE metadata

2.1 Content and domain metadata

Content and domain metadata contain information about the digital objects’ contents as well as the real-world objects they refer to. Some of them are generated automatically, like creation date or author’s name, while other information needs to be entered manually, for example the classification of concepts and real objects.

Each learning object in one of the MACE root repositories has a set of metadata attached to it. This metadata, however, often follows a proprietary information schema. To enable merging of content and domain metadata, a common schema is needed. This schema is the MACE Application Profile (MACE-AP). For this profile, we identified those kinds of content and domain metadata which are either common to all schemas or for which enrichment is feasible and worthwhile with regard to time and cost.

The content and domain metadata are encoded and transferred in a Learning Object Metadata (LOM) [6] based schema with extensions specific to the requirements of architecture and MACE.

2.2 Context Metadata

Context Metadata describe the context of an object (or subject) depicted in a digital content, e.g. GPS position, relation to other objects, surrounding terrain etc. They are generated by correlating different data sources available online.

Context information is relevant to the interaction

[1] http://www.openarchives.org/pmh/

between a user and a computer [7]. It characterises the situation of a person, place or object. In MACE, we connect objects and contents that, at first glance, have little in common [8]. We do this by creating relations between unlinked digital contents on the basis of a similar contextual background. For that, using Boolean information retrieval or temporal context analysis seems to be a promising approach [9].

The relations between digital objects have attributes [10]. By keeping virtual tokens for things outside the address range of computer representation (like real world objects/subjects and users), we are able to save context relations between any two digital objects. This idea of storing attributes at relations allows for a very flexible approach in connecting digital contents like learning objects with geo-information systems, building materials databases and seemingly unrelated data.

2.3 Competence and process metadata

Competence metadata are needed to show which digital contents can be used to obtain a competence. They describe competencies required to interpret a learning resource or give qualities a learner can obtain. Competencies can be described in various ways [11]. In a next step, MACE users can be searched for by their competencies; this helps finding experts for topics. Competence metadata are generated manually by experts or the users themselves.

Process metadata help shaping architectural learning processes. The following kinds of learning design methods are most prominently used in architecture [8]:

- Problem-based learning, which focuses on the problem solving aspect;
- Case-based instruction, which teaches the learner by giving a corpus of specific instances of architectural precedents and desirable outcomes;
- Discourse-based learning, which aims at creating knowledge by interaction, discourse and discussion.

Enriching content with competence and process metadata makes reuse of teaching constructs easier, as teachers use existing instructional design and fit it to their classes.

Learning processes are modelled in reusable designs using the IMS Learning Design (IMS-LD) specification. For competencies we use a competence card metaphor based on the competence standards IMS RCDEO [12] and HR XML [13].

2.4 Usage related and social metadata

Usage metadata describe events and user activities. They are generated automatically by using event- and user-tracking mechanisms. We then calculate metrics to highlight most frequently requested contents and most active users. *Social metadata* include free-form user annotations (commonly referred to as “tags”) and user feedback, e.g. ratings of contents.

Usage related and social metadata can be obtained from MACE content providers as well as from MACE tools, based on logs provided by different applications.

By using the Contextualised Attention Metadata (CAM[2]) schema, we normalise the usage information to enable derivation of new knowledge about the use of learning objects by correlating usage data from different sources [14]. Once captured, CAM does not change. Instead, it represents a continuous stream of new instances. Therefore we use the lightweight RSS protocol [15] to transport the metadata from the providing repositories to the central metadata store.

In case of data captured by content providers, we use RSS over HTTP[3] connections to exchange usage data. Once retrieved, every instance is deleted at provider side and stored in the MACE CAM store without further processing to ensure safe storage and availability of the original instance [8].

3 Content and domain metadata glossary

As previously mentioned, Content and Domain metadata concerns the content of digital objects and are therefore related to the description of architectural facts (e.g. a building or an urban area).

In MACE, the activity of enriching LOs with content and domain metadata has been started by a group of experts and, on the long run, will be operated by user communities. To avoid loss of information during these phases, we decided to use a prearranged glossary for the indexing activity.

3.2 The MACE taxonomy

[2] http://tinyurl.com/2kbaxk
[3] http://tools.ietf.org/html/rfc2616

The glossary used for content and domain metadata, has to envisage various aspects of the architectural domain in which different knowledge fields – may they be connected to the poetic-artistic side (ideas, cultural and social message of a project) or to the technical one (functionality, living wellness, building ease) – are called to simultaneously gather a project.

Moreover, it needs to be based both on *objective data* that should cover all the aspects of the domain and on *personal and intuitive data*, which will intercept interpretation peculiarities of the *Design Problem Solving Activity*.

With *objective data* we mean all those aspects of architecture that refers to objective (non-interpretational) aspects and for this reason aren't influenced by architectural trends or by theoretical and personal concerns. The main challenge in this case is to develop a standardized and shared taxonomy, able to cover all the aspects of the discipline featured in the architectural and engineering domain.

It is, on the other side, more complicated to develop a taxonomy able to classify personal and intuitive data. At first glance, trying to classify non-objective data may seem to be an oxymoronic task, but in this field, semiotics theory states the basis to help us perceiving and understanding messages in art, and thanks to those studies and methodologies we can try to develop strategies to classify non-objective data. Among the complex and variegated interpretative models offered by the current state of this discipline, we decided to rely on the Hjelmslev's interpretative model and on its following interpretation and adaptation held by A. J. Greimas[4]. This model, based on the double opposition of *contents/expression* and *substance/form*, has then been first extrapolated and enlarged to a system of categories and levels devoted to the reading of any visual work, and then reduced and focused on architectural works[5].

Following a conceptual factorization of the architectural domain based on these studies, the MACE taxonomy has been organised in 27 "facets", grouped in 6 "categories" (see table 2). The taxonomy features categories related to both *objective data* used in the design and documentation activity (e.g. Materials, Structural Profile) and *personal and intuitive data* used in the design activity (e.g. Perceptive Qualities, Project Cue) [3].

Identification	Intervention type, Project type, Functional typology, Form typology
Context	Location, Geographic and Urban context
Technical Design	Materials, Construction form, Building element, Technological profile, Structure profile, Systems and equipments, Technical performance, Maintenance and conservation
Constructing	Construction management, Construction phase, Construction activity, Machinery and equipments
Theories and Concepts	Styles, periods and trends, Theoretical concepts
Conceptual Design	Project cues, Project actions, Form characteristics, Perceptive qualities, Relation with the context

Table 2 - The MACE taxonomy faceted classification

The taxonomy terms have been gathered and re-grouped, starting from existing architectural thesauri (UniClass[6], ISO12006[7], the AAT Getty Vocabulary[8], and the Ci/SfB[9]) which have been reviewed on the basis of several other sources[10].

Relying on pre-existent and known thesauri was two-fold: (1) it was important to collect as many concepts, keywords, definitions and names as possible to keep a balanced and neutral position and to allow the expression of any point of view; and (2) it was important to collect a glossary of already defined and approved terms for efficient re-use.

The developed taxonomy consists of facets; each of them seen as an independent axis along which documents can be classified, and "addresses a conceptual dimension or feature type relevant to the collection" [18]. A facet contains a number of

[4] It's not our intention to summarize here a balance of the semiotics studies' results. To deepen see [16] [17] [3].

[5] The model reworking has been operated by IUAV of Venice (Prof. V. Spigai), in collaboration with UNIVPM in Ancona (Prof. M. De Grassi), during the period 1994 – 2004; see [3] [17].

[6] http://www.connet.org/uk/esc/classification.jsp

[7] http://www.iso.org

[8] http://www.getty.edu/research/

[9] http://www.ascinfo.co.uk

[10] *Literature:* Hadid, Z.: "Figures" – Eisenman, P.: "Conceptual and Formal tools" – Neufert, E. Bauentwurfslehre 32°ed. - Martí Arís, C.: "Le variazioni dell'identità. Il tipo in architettura" - Crema L.: "Manuale di storia dell'architettura antica" - Istituto Centrale per il Restauro (ICR) "The Risk Map of Cultural Heritage".

hierarchical categories, which contain a number of terms (see figure 1). Through this structure and the resulting associative relationships between concepts (e.g. by looking at the parent of a term), a semantic map becomes visible for searchers as well as indexers, helping to create a mental image of the overall topic and to select the most appropriate term.

A big part of the values is composed by a number of synonym terms. These are automatically assigned to a LO once a user has chosen one term during manual tagging. This way, finding content is simpler for users, since they can rely on an architectural language similar to the accustomed one.

Furthermore, users are able to add keywords to a specific LOM record, even if it does not exist in the application profile yet. Such keywords are stored in a freeform text field to be included for search and other usages, and also for later reviewing by experts. If the keyword is commonly used and approved by experts, it is added to the classification vocabulary. This hybrid of a pre-defined top-down hierarchy and a bottom-up folksonomy allows utilising “wisdom of the crowds” in a controlled manner to profit from existing personal knowledge. In this quality assured way, our taxonomy is extended and improved over time, thus having the flexibility to adapt to gradually emerging changes.

4 Example of MACE usage

MACE offers the possibility to navigate contents with different approaches. For example, an architect following the suggestion of transparency during a *design activity*, can search for design cases featuring a “transparent” perceptive quality (see figure 1). Furthermore, the system also offers many possibilities to gather information related to objective data using categories such as: geographic location (“Milan”, “Rome”), artistic trend (“Futurism”, “Rationalism”), author (“Wright”), and so on.

By triggering the search with simple keywords, the system offers the possibility to experience interesting navigational paths. For instance, a student searches for cases to solve a design problem related to covering a big public space. Following the metaphor of the sky vault, he queries the system with the keywords “roof” and “glass”.

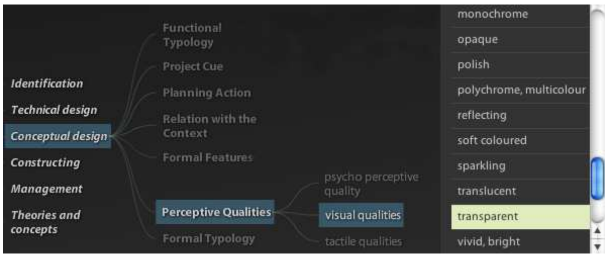

Figure 1 - The user navigates through the taxonomy to the final term. The results are continually updated, enabling search refinement.

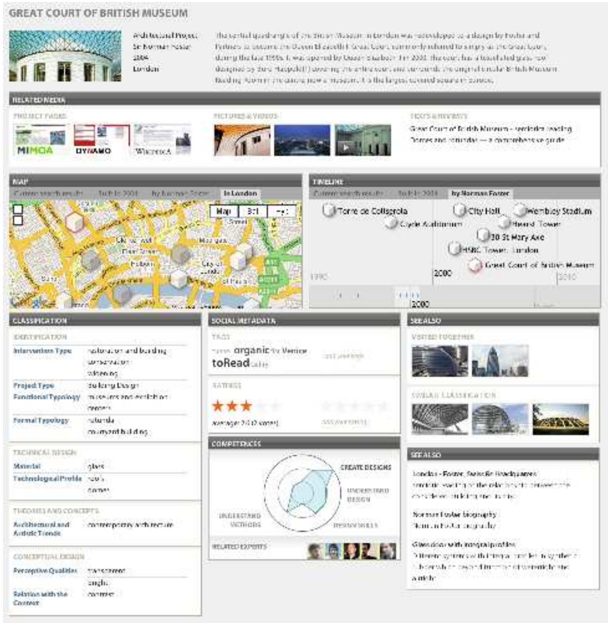

Figure 2 - A draft MACE sample page, displaying various widgets run with described metadata kinds.

From the results list the student realises that his search matches many contents, and that other users have been interested in the same topic, too. He decides to browse hot topics, expecting these to have a higher importance for his work. He can do so because MACE provides the required usage metadata. The system displays the detail page (see figure 2) where, through a map widget, he realises the location and contextual situation of projects, and also finds other important projects located in the same area.

Using a timeline widget, he browses other projects of the same architects according to their construction date. Usage and social metadata enables him to explore other related contents, and one of this (a picture of a dome designed by Nervi) captures the attention of the student who recognises it as relevant information. Thus, he is able to advance his research based on an unexpected but fitting example found by serendipitous browsing inside the MACE collective external memory.

5 Conclusions

By enriching and connecting existing portals and their contents through different types of metadata described, we provide a unique single access point to all the architectural domain repositories federated by MACE.

This way, heterogeneous and previously isolated repositories are accessible, connected and visible to a large community of users who can browse them in an intuitive way, which reflects the most common logical associations and mental reasoning paths followed by architects and engineers.

Moreover, the connected repositories are continuously enriched. The enrichment activity will be carried on during the use of the system. Users can themselves index the LO repositories in a “voluntary” and “manual” way by adding keywords chosen from the Content and Domain metadata taxonomy, rating the various contents, or adding comments, etc.; but also in an “automatic” way by enriching the LOs with usage related and social metadata, so the system is able to analyze the navigation paths of different users.

References

[1] Arthur, P., and Passini, R.: Wayfinding: People, signs and architecture. New York: McGraw Hill (1992)

[2] Rice, Ronald E., McCreadie, Maureen M., & Chang, Shan-Ju L. 2001. Accessing and browsing information and communication. Cambridge, MA: MIT Press.

[3] Spigai, V., Condotta, M., Dalla Vecchia, E., Nagel, T. (2008). Semiotic based facetted classification to support browsing architectural contents in MACE. In C.P. *Performance and Knowledge Management*. Joint CIB Conference. Naaranoja M., Otter A., Prins, M., Karvonen A., Raasakka V. (Eds). Helsinki, Finland

[4] Stefaner, M., DallaVecchia, E., Condotta, M., Wolpers, M., Specht, M., Apelt, S., Duval, E. (2007). MACE – enriching architectural learning objects for experience multiplication, in: C.P. *Creating new learning experiences on a global scale*, Springer LNCS; Germany.

[5] Open Archives Initiative: Protocol for metadata harvesting 2.0. online (June 2002). http://www.openarchives.org/OAI/openarchivesprotocol.html

[6] Institute of Electrical and Electronics Engineers Learning Technology Standards Committee: IEEE standard for learning object metadata (draft). IEEE standard 1484.12.1 (2002)

[7] Dey, A.K.: Understanding and using context. Personal Ubiquitous Computing 5(1) (2001) 4–7

[8] Prause, C.R., Ternier S., de Jong, T., Apelt, S., Scholten, M., Wolpers, M., Eisenhauer, M., Vandeputte, B., Specht, M., Duval, E.: Unifying Learning Object Repositories in MACE. In: Proceedings of the First International Workshop on Learning Object Discovery & Exchange. (2007)

[9] Prause, C., Kuck, J., Apelt, S., Oppermann, R., Cremers, A. B.: Interconnecting documentation - harnessing the different powers of current documentation tools in software development. In Cardoso, J., Cordeiro, J., Filipe, J., eds.: Proceedings of the Ninth ICEIS. Volume ISAS., INSTICC (2007) 63–68.

[10] Zimmermann, A., Lorenz, A.: An operational definition of context. In: CONTEXT 07. (2007)

[11] Cheetham, G., Chivers, G.E.: Professions, Competence and Informal Learning. Edward Elgar (2005)

[12] IMS Global Learning Consortium: Ims reusable definition of competency or educational objective specification (rcdeo). online (October 2002) http://www.imsglobal.org/competencies/

[13] HR XML Consortium: Competencies (meas-urable characteristics). online (August 2004) URL: http://ns.hr-xml.org/2 3/HR-XML-2 3/CPO/Competencies.html

[14] Wolpers, M., Najjar, J., Verbert, K., Duval, E.: Tracking Actual Usage: the Attention Metadata Ap-proach. International journal Educational Technology and Society 7(7) (2007) 1176–3647

[15] UserLand Software: RSS 2.0 specification. online (August 2002) http://www.rssboard.org/rss-2-0

[16] Greimas, A. J., Courtes J. (1979). Sémiotique. Dictionnaire raisonné de la théorie du langage, Paris, Hachette.

[17] Spigai, V., Condotta, M., Stefanelli, C. (2006). Collaborative e-learning in engineering and architecture: intelligent systems for knowledge sharing in on-line design laboratories. In C. P. H. Rivard, E. Miresco, H. Melhem (Eds.) “*Joint International Conference on Computing and Decision Making in Civil and Building Engineering*”. Montreal.

[18] Quintarelli, E., L.Rosati, and Resmini, A. 2006. Facetag: Integrating Bottom-up and Top-down Classification in a Social Tagging System. EuroIA 2006, Berlin.