

DATA RETRIEVAL FROM A CULTURAL KNOWLEDGE DATABASE

G. Marcos, H. Eskudero, C. Lamsfus, M.T. Linaza

Asociación Centro de Tecnologías de Interacción Visual y Comunicaciones, VICOMTech, Spain

ABSTRACT

This paper describes a domain ontology which has been designed within the European art-E-fact project (IST-2001-37924) to represent the associated knowledge and to enable the description, exchange and sharing of multimedia added-value content for the creation of artistic expressions both within and between cultural institutions. Within the art-E-fact project, authors access and retrieve all the content stored. Since authors do not mainly know how to deal with a database management, a mapping method has been developed. This paper presents the details of the developed mapping and content retrieval methodology through the domain ontology implemented for the project.

1. INTRODUCTION

Semantic Web technologies can enable the Cultural Heritage sector to make its information available in meaningful ways to researchers, its own curators, the didactics departments and even the general public. Content generators will benefit from the ability to define an exhibition and have the entire process from the identification of the pieces to be shown in the exhibition to the final production of the exhibition.

Delivering these technologies to the heritage sector depends on the syntactical and semantic mark-up of content, the development of better knowledge analysis and modelling tools, widespread adoption of interoperable knowledge representation languages and the implementation of suitable ontologies. In most of this, the heritage sector is lagging behind.

For the semantic technologies to succeed in the field of information exchange and interoperability between cultural institutions, there is a great need not only to gain interoperability using standard ontologies, but to provide the cultural heritage community with user-friendly tools for inferencing and semantic information retrieving of the content stored in cultural databases. One should always keep in mind that the personnel from cultural and artistic institutions do not have usually a technical background and most of them are not very familiar with dealing with databases.

This paper presents a contribution concerning a simple methodology for non-technical users in order to map or relate the cultural content with the storing databases. Section 2 describes in more detail different aspects of the art-E-fact ontology, including a brief description of the project, the scope of the ontology, the description of the metadata and the applied form. Section 3 presents the methodology applied to the mapping and finally, Section 4 provides some conclusions and further works.

2. THE ART-E-FACT ONTOLOGY

2.1. Description of the project

The aim of the art-E-fact project (IST-2001-37924) is to create a generic platform for Interactive Storytelling in Mixed Reality that allows artists to create artistic expressions in an original way within a cultural context between the virtual and the physical reality.

With the means of Interactive Storytelling and Mixed Reality of virtual autonomous characters, multimedia, physical props and devices, and multimodal human-oriented interactions for artistic expression are enabled [2]. The target platform is both a new medium for the communication of informational content, and a new form of act. The main objectives of the art-E-fact project are:

- to develop a generic platform for Interactive Storytelling in Mixed Reality that allows artists to create artistic expressions in an original way, within a cultural context between the virtual ("new") and the physical ("traditional") reality,
- to use the platform to actually build a compelling Mixed Reality installation that facilitates the access to a knowledge base of inspirational material of art history-reflecting the way humans created art since at least 4000 years,
- to involve artists and the analysis of artistic methods, on from the beginning of the project through all project phases, as well as
- to create a showcase within an interdisciplinary team that can be used for the evaluation of artistic methods, as well as for the diffusion and exploitation of the

results, leading to more accessible tools for artistic expression in the future.

Artists can create a Mixed Reality exhibit by using the generic system to shape a specific instance of expression. They make choices of specific interaction devices and physical props to be used for anthropomorphic interactions, as well as corresponding interaction metaphors; they define dialogues with a degree of autonomy and behaviour of virtual characters, and they create multimedia elements to be accessed during run-time.

art-E-fact aims at the provision of Mixed Reality technology addressing two directions: firstly, to provide a generic platform for artistic expression to enable interactive exploration of artworks, and secondly, to ease the task of artwork creation by providing a standard VRML compatible VR framework enhanced with interaction and sensor features.

The Mixed Reality generic platform for Interactive Storytelling serves as an experimental platform allowing authors with artistic or humanistic backgrounds to make design decisions that go beyond the state-of-the-art creation systems of digital media. In summary, it is possible for artists to include anthropomorphic interactions such as gestures, body poses into their design of Mixed Realities, and to direct lifelike avatars in order to act.

2.2. The art-E-fact ontology

The art-E-fact ontology is part of the outcome of the art-E-fact project. Since the target of the art-E-fact project is to create stories about artworks and thus create art, we have to be taught by the experience that was gained the last 4000 years of civilization.

The domain metalevel ontology conception leads the author to assimilate the internal world of the creator of an artwork, create and tell stories.

As noted above, this is not just a conception of the experts performing the scientific diagnosis, but it is also a tool for artists, authors and content generators. Artists using the ontology have to create stories or experiences concerning one or more selected artworks, including its main features, technical data, historical context, etc. All this information will be included within the Cultural Content concept. The domain metalevel ontology conception will lead them to assimilate the internal world of the creator of an artwork, and create and tell stories.

In order to build the art-E-fact ontology, the main focus has been put on the target group that will use the Generic platform, that is, the artists. Among the three possible alternatives to define the classes (top down, bottom up and a combination of both development process) a combined development process has been used.

This means that we have first defined the most representative concepts, and then we have generalised and specified them appropriately in order to get an accurate representation of the knowledge stored in the database.

Scoping the ontology has been mainly based on two brainstorming sessions with the artists and the content providers. Having these brainstorming sessions allowed us to produce most of the potentially relevant terms and phrases. At this stage, the terms alone represented the concept, thus concealing significant ambiguities and differences of opinion.

A clear issue that arose during these sessions was the terminology differences among different art styles, between the Greek traditional iconography and the traditional European painting schools. The concepts listed during the brainstorming sessions were grouped in areas of work corresponding to naturally arising sub-groups. Most of the important concepts and many terms were identified. The main work of building the ontology was then to produce definitions. During scoping, most of the important concepts and many terms have been identified.

This methodology of the ontology building process also includes information about mapping the structure of the content database and the representation of the conceptual schema represented in the database.

3. MAPPING THE ART-E-FACT DATABASE USING THE SINGLE CONSTRAINT MAPPING METHODOLOGY

The task of integrating heterogeneous information sources put ontologies in context. They cannot be perceived as standalone models of the world, but should rather be seen as the glue that puts together information of various kinds. Consequently, the relation of an ontology to its environment plays an essential role in information integration. In this paper, we use the term mapping to refer to the connection of an ontology and the information it describes [1].

3.1. The single constraint condition

One necessary and sufficient condition for the database to be mapped through the ontology is that there must be a field in one of the tables within the structure of the database (called hereby "DB entry point") that, once it has been specified, it allows accessing all the fields on a table due to the structure of the database.

In this way, it is possible to guarantee that once this field is correctly identified, queries to retrieve information from each of the fields in every table of the database are allowed.

3.2. The database architecture

The main idea concerning the design of the architecture of the ArtWorks Database (AWDB) system corresponds to the conception of the ontology. In this architecture, the main levels of knowledge are treated as main "thematic entities" for the database. The entity of the identification-work is the table "ITEM" (DB entry point) and all the tables are related to this table. There can never be a table with no connection to at least one more table, since that would automatically mean that there is some data that is not reachable by the querying engine.

3.3. The mapping

The most obvious application of mapping is to relate the ontologies to the actual contents of an information source. The art-E-fact ontology is related to the cultural content provided by the cultural participants within the art-E-fact project. Among the different approaches used to establish the connection between ontologies and information databases, we have chosen a definition of terms. In order to make the semantics of terms in a database schema clear, it is not enough to produce a copy of the schema. The definitions of the ontology do not correspond to the structure of the database. These are only linked to the information by the term that is defined.

The ontology is not a "mirror" of the structure of the database. Moreover, it should act as a "semantic index" representing the data and information stored in the database. Therefore, a class of the ontology can not be identified with a table in the database. The mapping has been carried out at the property level, so that a property in the ontology represents a field on a table of the database.

To map the database through the ontology, properties can be divided into two different groups:

- Properties made up by "Class" and "Instance". As this group is used to define relations between concepts, the associated properties cannot be used for the mapping, since they only show information about the structure of the ontology; and
- Properties made up by "Literals" (e.g. string, boolean, float, etc.). This second group can store information about the information stored in the database (e.g. the name of a person), so it is used to take part on the mapping.

The properties of each Class in the ontology have to be matched with a field on the database. As the aim of this mapping is to facilitate the access to the content of the database, we can have two properties of a Class pointing at the same field, but not vice versa, since it

could cause some conflicts when trying to retrieve data from the database.

In order to be able to access and retrieve the content, we have designed a "middle" ontology that records the information of the structure of the ontology, the database and the relations between them. This ontology allows artists to retrieve data about a concept of the original ontology without using the syntactic of a database language.

The mapping tool has been implemented as a part of the art-E-fact project. The problem is that most of the cases the artist is not an expert him/herself on the topic, since the profile of the user is more of a story-teller or an artist rather than that of an expert in History of Art.

Artists could just explore and get the information of the database with an interface, or even directly. Since this is a very complex and tough task, the mapping tool provides an easier way of accessing to the mysterious internal world of artists.

4. CONCLUSIONS AND FURTHER WORK

The art-E-fact project aims at developing a generic platform for interactive storytelling in Mixed Reality that allows artists to create artistic expressions in an original way. This platform should facilitate the access to knowledge bases which content will be used as an inspirational material for them.

The domain ontology that has been implemented within the art-E-fact project gives artists a general overview of the content stored in the database provided by the other members of the consortium without having to deal with a rough structure of a database.

The methodology for the mapping engine has been developed in a generic way, as well as both the database and the ontology. The mapping engine constitutes an easy way of accessing the interchanged content without the need of any other interfaces nor software, just through the knowledge representation (i.e. ontology).

As further improvement of the system, we are developing a multiple constraint mapping methodology. This means that we are having more than only one only DB Entry Point. This is rather more complex than the single constraint method, but it will fasten the speed at which the access to the information on the database is achieved.

Both the implementation and the ontology conception within the art-E-fact project are generic. The use of the middle ontology for mapping the database and retrieving information allows a generic method for accessing the database. Thus, if the content of the database is changed and a new ontology for the new domain is defined, the mapping method could be applied to infer the database.

5. REFERENCES

- [1] H. Wache, T. Vögele, U. Visser, H. Stuckenschmidt, G. Schuster, H. Neumann and S. Hübner, "Ontology-based Integration of Information- A Survey of Existing Approaches", *Proceedings of IJCAI 2001 Workshop 'Ontologies and Information Sharing'*, 2001.
- [2] I. Iurgel, "From another point of view: art-E-fact", *Proc. TIDSE '04*, Vol. 1, pp. 26-35, 2004.