Towards an Ontology for Annotating Degradation Phenomena

T. Messaoudi, L. De Luca

UMR 3495 MAP CNRS/MCC

Marseille, France

tommy.messaoudi,livio.deluca@map.cnrs.fr

P. Véron

UMR CNRS 7296 LSIS, Arts et métiers ParisTech

Aix-en-Provence, France

Philippe.veron@ensam.eu

Abstract—**In the field of built heritage, a huge amount of data describes the state of monuments: documentary data (texts, sounds, images) as well as analytic data from sensors, provides historical, archeological and constructive information. These data, produced by experts coming from several fields, are the foundation for the creation of new information through scientific observations. The problem is that these data are not related nor spatialized. Nowadays we are able to generate extreme accurate 3D model with throughout images-based or laser scanner acquisition. However these 3D models do not carry information regarding their morphological complexities. The design of an ontology for the conservation domain seems to be the best solution in order to obtain understandable entities thanks to their own data related between them. This article present the first attempt of a development of a web information system based on spatialized images semantic annotations tool, related to a domain ontology describing knowledge regarding stone degradation phenomena. Our approach is to produce a domain ontology able to document and therefore provide a framework able to help the decision-making process of experts in the cultural heritage conservation domain. The information annotated with the use of this ontology can enrich not only the scientific observations, but also to help to create new knowledge. In this way, it is possible to link and gather quantitative and qualitative aspect into only one information system. We will propose several example and queries able to exploit the reasoning power of the above information system.**

Index Terms—**Conservation, Ontology, Data management, digital modeling, semantic annotation, knowledge, cultural heritage**

I. INTRODUCTION

The analysis of cultural data is one of the biggest challenges in the cultural heritage field. Masonry structures, which is our object of analysis, are documented by a lots of heterogeneous data, like documentary (text, sound, historical images, etc.) or analytic data from sensors. These data provides constructive, historical and archeological information about the studied building. These are also produced by a lot of various experts coming from different domain (architecture, mechanics, computer science, etc.). According to this context, we can conclude there are various digital information acquisition. Nowadays it is possible to produce accurate 3D model but it is not sufficient to completely analyze its morphological complexity. This is the reason why we want to create and define a novel information processing chain from data collection to tool validation. On the one hand conservation experts use conventional technic to describe building morphology by using simple images or even translucent paper to create cartography documentation. On the other hand, different experts realize a lot of acquisition like point cloud from laser scanning, scientific imagery (IR thermography, UV imagery), a structured modeling to make mechanical behavior analysis. These two things respectively represented by qualitative and quantitative aspects can be related around a 3D multilayered analysis support with different level of complexity. This 3D support based on image-based modeling, allow us to realize annotations giving heterogeneous information about masonry building conservation from integrated and interpreted data by scientific observations. The goal is to integrate and structure different heterogeneous information into an understandable representation that take in consideration the multiple factors describing building conservation state. Then we introduce "informative continuum" notion thanks to the link between the whole of spatialized data to create information. Scientific observation is possible thank to domain ontology, formalized and integrated into an information system. This paper is divided into 4 part. A state of the art has been realized and I've selected two relevant example of them that will be presented in part 2. And then, our approach concerning the way to design our domain ontology dedicated to degradation phenomena annotation will be presented in part III. Finally, section IV will list on-going-works and bring some perspective for further development.

II. STATE OF THE ART

An ontology is used to describe, share and reuse knowledge and data between software and humans. Several ontologies emerged in several knowledge fields (medicine, mathematic, law etc.) in order to analyze and formalize terms. Our research is oriented toward heritage building monitoring, this is the reason why we have selected two relevant example to underlining our research orientation.

A. Domain ontology definition

"Ontology is defined as an explicit and formal specification of a shared conceptualization" [4]. The term is borrowed from Philosophy domain, where an Ontology is a systematic account of existence. This definition is probably difficult to understand

but if we explain it term by term, we can easily understand its meaning. Conceptualization mean modeling an abstract phenomena in the world. "Explicit" refer to the concepts and constrains that need to be clearly defined. An ontology is "formal" because it is understandable and readable by a computer. And finally the most important for us, it can be "shared" that is to say it gather consensual knowledge agreed by a specific expert group. In other word, domain ontology is a shared knowledge formalization dedicated to a specific domain. Our domain is masonry building conservation domain.

B. In Museal domain

The first example is an information system dedicated to museum domain realized by the research group SeCo which stand for Semantic Computing [6]. Their ontology named Museoalan Ontologia structures a huge amount of heterogeneous data from several domain: like painting, building, cultural sites, archives, library, etc. Kulturisampo application, created by the same research group using this ontology, propose an interface for organizing and filtering a significant amount of data, giving to the users the possibility to overlap the several information stored in their knowledge base. For instance we can overlap a specific photograph showing an old map to a current map, scalable at the same time and in the same way. Each viewable element from database is linked with other various elements. The methodology is very interesting but this ontology don't manipulate 3D data type, but only basically 2D data.

C. MONDIS

MONDIS which stand for MONument Damage Information System dedicated to heritage building degradation also manages alteration phenomena. MONDIS is based on an automatic process, crossing cultural heritage alterations documentation, its diagnosis and interventions for potential restoration operations [2]. Several information types are connected like: "architectural component description" giving information about the cultural heritage object construction (physical and functional property), "event cultural heritage object" corresponds to changes of condition to the object studied. "Damage diagnosis and intervention" informs us about the relationships between alteration, active agent, and mechanisms and provides a way to intervene on one of them to stop the alteration. And then "risk and measure assessment" represents the prevention. This system is very interesting for us, however MONDIS doesn't use spatialized annotations, therefore it is mostly focused on the qualitative aspect. This is the reason why we would like to enhance it, using it as a starting point of our work.

III. MAIN APPROACH

The main aspects to analyze the building conservation state is in one hand acquisition that is to say a quantitative aspect. On the other hand the Human-based identification for the interpretation of these data to create information and knowledge. The main issue is the lack of relationship between these two aspects. Hence our hypothesis is: our Information system structured by an ontology allow us to relate quantitative and qualitative aspects to analyze cultural heritage conservation through Scientific Observation. The question is: how to design ontology?

And how do ontology can manage that? Several method have been found to design ontology, and we have selected one of them. In addition a solution has been found in order to manage these different aspects.

A. Semantic annotation

Our work aims to define a domain ontology, allowing us to relate acquired data, with formalized shared knowledge in relation to our domain which is the heritage building conservation. Hence to combine qualitative and quantitative aspects. Our future web tool will mainly be based on referenced spatially photographs annotation [8] produced by an automated image-based modeling. Thus architectural and state conservation analysis can be performed directly on images. The annotation process generates 3D information inherent to images. This means that a 2D/3D bijective link is introduced between pixels and their 3D coordinates. This relation allows us to transfer annotations between images. An image of the set is annotated and the 2D-to-3D relation extracts the corresponding 3D point cloud of the drawn area. The corresponding areas on other images of the set is then retrieved by using the 3D-to-2D relation *"Figure 1"*. The latter relation conversely permits to retrieve positions in other images where 3D coordinates of the annotation are detected and highlight the corresponding areas [11].
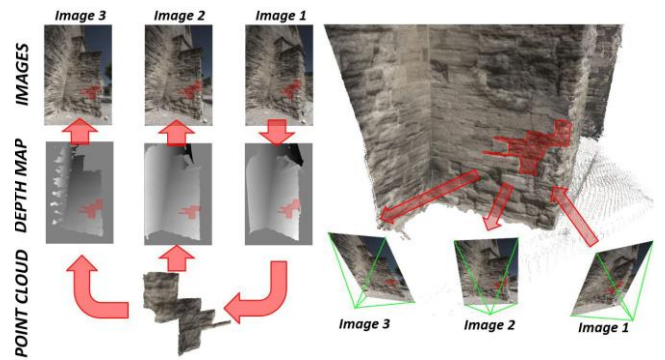

Fig. 1. Transfer of annotation of a set using 2D/3D relation

This method provides a simple and fast way to annotate images all at once instead of one by one. Annotations will be related with the data stored in database giving them a quantitative aspect. Indeed these regions will be used to extract dimensional values such as extension, position in 3D space, orientation, magnitude giving information about material loss, etc. And then to realize spatial distribution of same region type in order to extract similarities.

B. A domain ontology for annotating degradation phenomena on oriented images

Annotations related to data, will be linked with scientific observation and knowledge from a domain ontology. Thanks to the relationship between shared knowledge and scientific observation, annotation will be carrier of a qualitative aspect. Shared knowledge include the whole of concepts concerning stone degradation description and their links, retrieved from the

ICOMOS Glossary [1]. Knowledge from this glossary and concepts from CIDOC-CRM scientific observation [3] has been integrated. The latter is an ontological model able to document scientific practices in our domain and to manage both quantitative and qualitative aspects. In this case shared knowledge can be used and enriched by other case studies. The conservation experts 'goal is to provide a complete description of studied masonry building. Many observable information can be retrieved from the building. By crossing these information, it is possible to produce a line of reasoning induced to these links. Mainly, the expert need to obtain credible information and knowledge to undertake possible restoration operation. Annotations perform on spatialized images represents not only a shape but also the way to spatialize scientific observation and attach data. The user (expert) will have the possibility to annotate three thematic description layers: architectural components, materials and then alterations. Related to this different layers, we have consensually defined formal descriptor table allowing us to accurately characterize each them with their own specificities. Scientific observations will be done by using these data attached to layers with embedded information like environment (weather, building context for instance), acquisition (direct or indirect), temporal states, documentary resources, directly integrate into a dedicated data base. Thanks to these three linked layers, it will be possible to generate a lot of information about the studied building conservation state. The building's annotation set is representative of the whole structured knowledge concerning building conservation. To realize that it is important to exploit a relevant methodology in ontology design. The Lassila's method [5] [7] appears to be the best to fit our needs concerning its design method and therefore it was used to create our ontology. The above methodology required a domain expert to validate a controlled vocabulary which terms, as well as their meaning, will remain unchangeable during the whole process: by using an existing glossary. Following this step a thesaurus have to be built, in order to precisely define the complex and structured relationships between the terms and then obtain a taxonomy. At this step it's possible to create a lightweight ontology up to obtain a heavyweight ontology by adding more rules, constrains, and axioms. Two glossary types are exploited in our project: Illustrated glossary on stone deterioration patterns [1] and architecture vocabulary [10].

Indeed expert can perform different annotations on layers *"Fig. 2"*. We can count two annotations (red and white) giving us a quantitative aspect through extension, position, orientation (data) automatically extract from database or informed by the expert. In addition we have a qualitative aspect through a tag related to concept from ontology: **buttress** and **alveolization**. Thanks to the ontology reasoning, our system can inform us how the "alveolization", has been created. Our ontology is based on this follow structuration: "Masonry Building is composed of architectural component build with Material. And active agent acts on material and generates degradation mechanism inducing alteration".

According to domain ontology, to create this alteration, we need some components. This is a typical encountered alteration
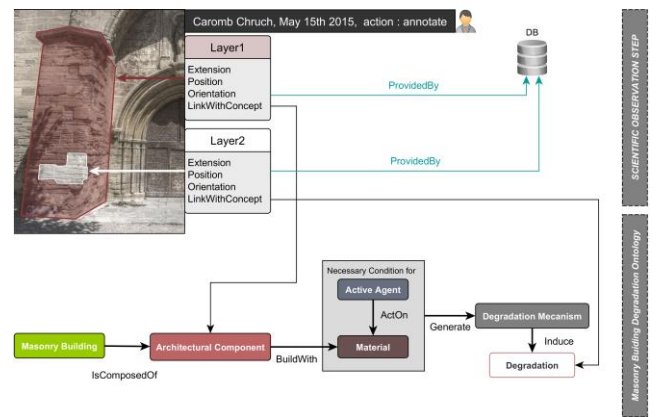

Fig. 2. Example of use linking scientific observation and domain

and it allows us to explain the domain ontology usefulness *"Figure 3"*. A masonry like church is composed of several architectural entities such as buttress, built with porous stone containing soluble salts (chemical and physical properties). Due to the effect of the rain or underground circulation, water (active agent) seeps in stone by an alteration mechanism named rising damp. As a consequence a chemical reaction between water and soluble salts creates tensions because of salt crystallization. This internal pressure is cause of a typical feature by inducing material loss: an alteration named alveolization. We have used the Lassila's method in which the use of glossary is very important [5][7][9]: therefore two glossary type are exploit: ICOMOS illustrated glossary on stone deterioration pattern [1] to describe degradation and alteration phenomena and architecture vocabulary [10] to identify masonry structure architectural components. Protégé is the ontology editing tool used to integrate these two taxonomies and relationships between concepts to design our domain ontology.

C. Typical queries

This future web application will be able to answers to the following queries (query examples used come from our case study, the Caromb church (south of France)):

Query 1:
- Input: selected architectural element class: left buttress
- Query: composing material and referenced alterations
- Query results: left buttress; material => yellow stone, grey stone; alterations => alveolization, spalling

Query 2:
- Input: selected alteration type AND selected architectural element class => alveolization and left buttress
- Query: composing material the most altered by selected alteration type
- Query results: material => yellow stone *"Figure 4"*

Query 3:
- Input: selected mineralogical components list
- Query: Some materials matching to this composition
- Query results: yellow stone of Caromb, stone 2, stone 3

Fig.4. viewable result of query 2

IV. CONCLUSION

To conclude, thanks to annotation tool, it will be possible to cross several 3D regions and to generate information. Indeed, a scientific observation can identify a spatial region related to a concept like "alveolization". This concept will be linked with other regions that will allow us to realized similarities, to extract common orientations, extensions and other parameters that we are integrating into our future system. Our work is to make a link between scientific observations and shared knowledge, and use such information to create new knowledge. Moreover we can use these acquired knowledge in order to guide new observations. Then through scientific observation, we can enrich the data and provide information about studied building conservation state. The notion of temporal monitoring is still to be integrated. The main goal is to monitor masonry building degradation therefore other future acquisitions will be done to compare different conservation state. The difficulty is to readjust these acquisitions with the previous temporal state and compare two annotations analysis and generate other information. This is an on-going work and the ontology is still to be completed, because we are searching the most relevant methodology to combine our domain ontology with the CIDOC-CRM extension dedicated to scientific observation (CRMsci).
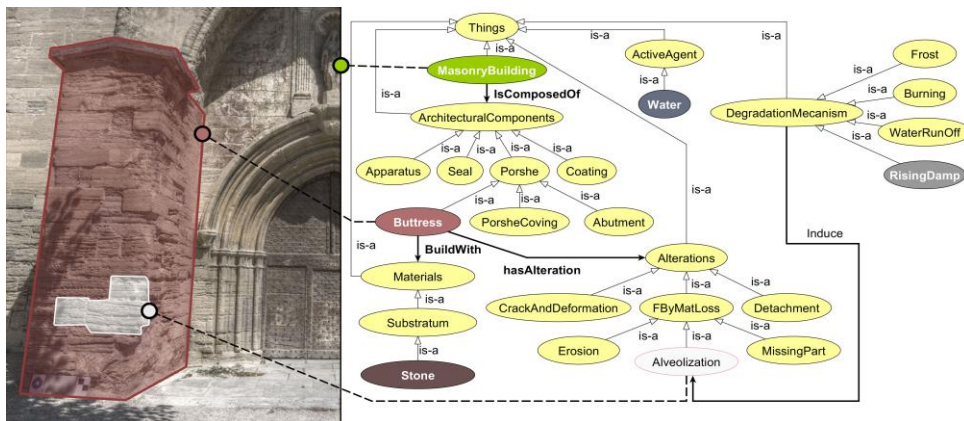
REFERENCES

[1] P. Bromblet, J. M. Vallet, and V. Verges-Belmin: Illustrated glossary on stone deterioration patterns. Monuments and sites, 2008.

[2] R. Cacciotti, J. Valach, P. Kunes, M. Cernansky, M. Blasko, and P. Kremen: "*Monument damage information system (MONDIS): An ontological approach to cultural heritage documentation*". ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume II-5/W1 (2013).

[3] M. Doerr: "*The CIDOC CRM, an Ontological Approach to Schema Heterogeneity*". In: Proceedings of Semantic Interoperability and Integration, 2005

[4] T. R Gruber: "*Toward Principles for the Design of Ontologies Used for Knowledge Sharing*". International Journal Human-Computer Studies 43, p.907-928, 23 August (1993).

[5] N. Hernandez: "Ontologie de domaine pour la modélisation du contexte en recherche d'information".

[6] E. Hyvönen: "*Publishing and Using Cultural Heritage Linked Data on the Semantic Web*", Synthesis Lectures on the Semantic Web: Theory and Technology, Morgan & Claypool, pp.1-119, (2012).

[7] O. Lassila, R. R. Swick: "*Resource description framework (rdf) model and syntax specication w3c recommendation*", (1999).

[8] A. Manuel, C. Stefani, L. De Luca, P. Véron: 2D/3D "Semantic Annotation Towards a Set of Spatially-Oriented Photographs". Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (2013).

[9] T. Messaoudi, A. Manuel, E. Gattet, L. De Luca, P. Véron: "Laying the foundations for an information system dedicated to heritage building degradation monitoring based on the 2D/3D semantic annotation of photographs" EUROGRAPHICS Workshops on Graphics and Cultural Heritage (2014)

[10] R. Navigli, P. Velardi: "*From Glossaries to Ontologies: Extracting Semantic Structure from Textual Definitions*", Series information for Frontiers in Artificial Intelligence and Applications, IOS Press, pp. 71-87, (2008).

[11] J. M. Pérouse de Montclos: "*Architecture : méthode et vocabulaire*", Monum, éditions du patrimoine, 2004.

Fig. 3. Domain Ontology sample dedicated to masonry building alteration monitoring: alveolization example