Sharing concepts and experience through a knowledge base : the benefits that actors of the mud building European community are getting from developing the domain ontology

Christian Bois

Université de Toulon

Laboratoire L.E.P.O.N.T.

B.P. 132 83957 La Garde Cedex France

christian-l.bois@chello.fr

Author : Bois Christian

ABSTRACT :

As most of the research and practice communities, the mud building actors are developing the quality and the quantity of their exchanges of information and knowledge.
The development of the Semantic Web has attracted the attention as a mean for making knowledge widely available after a phase of knowledge base creation that can be run by actors spread over a continent or more.
Challenges for the domain include its inscription in time back to the origins of man –large amount of developed practices- and the new interest for a technique fitting well within sustainability, comfort and health concern –wide extension of the related concepts.
The text is aimed at having mud builders understand a little more ontologies and ontology specialists to understand a little more the specificities of mud building. Other stake holders may be interested by the dimensions of complexity coming from the fact that mud (cob, rammed earth, adobe) is a continuously "living" material.
Challenges are both technical –new tools, articulation with existing systems- and social –ability of very different actors (engineers, architects, masons, etc.) to collaborate for building a common knowledge tool.

Introduction

The Semantic Web will soon become a practical reality when a consensus will be reached about the final content of the OWL – Ontology Web Language necessary for developing knowledge bases.
Meanwhile, researchers at the Stanford University School of Medicine (1) have developed a practical ontology development tool called "Protégé 2000" (2).
We (3) have been interested by the informational and human dimensions of this development of the Semantic Web. For historical reasons, our relationship with actors of the mud building community gave raise to the idea of exploring how this community could take advantage of building the ontology of the discipline.
Actors of this community -architects, engineers, masons, etc.- and Semantic Web specialists are working together to identify :

- what is the available information on mud building on the Internet.

- what is the most needed information, the core set of reference texts.
- how these texts can be analysed and annotated to transform information into sets of knowledge organized in an ontology with shared vocabulary and links between facts and concepts.
- how this working around an ontology tool could add value to individuals and group knowledge.
- how this improvement may lead to e education processes.

The explored issues are :

- the technical necessary conditions for realising the domain ontology and make the knowledge available in a smart way.
- the dynamics of a working group faced to this challenge.

1._Finding the useful information in a reasonable time

The need for improvement of how information is accessible through the Internet, on how this information needs to be validated to become knowledge is experienced by most actors.
For example, if one uses Google search engine with the key words *"mud building" Europe science* 22 links are found to documents. But these documents are about Islam, Africa, Indo European language, kibbutzim and science fiction. No document relates to the researched topic.
If *"Cob",* which is a sub category of *"mud",* is used, a list of documents is found but no pertinent documents in the first pages of the list (2500 items).
Since the early development of the World Wide Web, this problem of the "ocean of data" where useful information can't be found has been raised by Berners-Lee (1998) and others.
The developers of the Semantic Web are tackling this problem, developing generations of tools adapting the technologies of Description Logic to the Internet context (4).
Mud building European stake holders may take advantage of these tools for :

- identifying the core/common elements of knowledge and practice.
- keeping track of the regional/cultural specificities articulated with the common core.
- articulating the ontology of mud building with the ontologies of other traditional or modern ways of building.
- exploring the multilingual possibilities.
- documenting the characteristics of sustainability, relation to health concern, etc.

There is a challenge at having the millenary art and knowledge enter the Semantic Web era.

2._Mud building : art, practice and science together

"Building is building" may say a beginner. During centuries the dynamic qualities of mud building have been known. Then came a time where the motto was to build "light, quick and immediately cheap". Mud building is heavy, slow ; and cheap only on the long run. The qualities as well as part of the know how was lost and then rediscovered with concern to comfort, health and sustainability.
When one thinks of sustainable architecture and building, the adobe, cob and rammed earth building approaches are the very examples of low cost (real social and environmental cost), long lasting material – Marcom Alain (2002).
Having both multiple and dynamic characteristics, mud building should be viewed as a complex system.
Knowledge bases are more adapted than traditional databases to represent this complexity.
Some characteristics of a mud made building can be represented on matrix 1.
The mud building we are talking about has its dynamic qualities fully functional, i.e. outside nor inside walls are covered by any concrete, plaster or insulating material. As few as possible lime coating may be used.

Matrix 1 Heat transfer, moistness regulation and natural cooling, true building cost ratio.

The cost issue is to be considered in terms of ratio. The cost evaluation for a mud building is supposed to include factors like :

- user's quality of life -comfort and health.
- energy savings : energy needed for producing other types of material is not used.
- use of available energy from work force, worldwide.
- planet quality of life -no pollution for producing the material.

On the second axis, mud building is advocated for it's moistness comfort in dry days added with an evaporation/cooling effect (this affects cost through savings for cooling).
The third factor is the heat transfer quality of clay walls. In spring and autumn, thick clay walls take calories in during the day and give them back to the inside of the house during the night. This is added with a condensation/warming effect in night time (this affects costs through savings for heating). (5)

3._Some characteristics of knowledge : the sound of quality

In highly industrialized processes, it has been discovered since the sixties that it was very difficult to have the know how described and written down. Actors speak of the "sound" of the correctly adjusted machine (6). Are adobe, cob and rammed earth to become industrialized practices? If the light industrialization is considered, the process has already begun.
Whether mud buildings will be individually crafted or somehow industrialized, quality will have to be :

- explained to customers and regulation bodies.
- realized when building.
- controlled.
- taught.

Fine descriptions of mud building materials and techniques is needed for this purpose.
The following matrix exemplifies some parameters of mud building.

Much work has been done by researchers to link the different qualities of earth with the different preparation techniques. (7) The matrix 2 gives a simplified view of what can be measured for the different mixtures and fabrics.
Every region of Europe where mud building is the traditional way - sometimes every practitioner- has a part of the whole set of information.

Matrix 2 Clay, water and straw, molding or tamping

There are many exchanges between actors in the mud building area.
But shall we imagine three typical situations.
Situation 1_Helena is a young student in architecture or urbanism or so in a school with no tradition in mud building practices. She wants to have a rather complete view of the existing resources in Europe : today (2002), it will take her ages for getting the result.
Situation 2_His name is Just and he is a professional in masonry and he has recently discovered, as well as some of his clients, the qualities of mud building. Just wants to find the best training and documentation on the different techniques. He faces some difficulties to find the needed tools, particularly if Just is Spanish and his customers are Dutch and they want to share on the matter.
Situation 3_Gil is a doctor and he is interested by the qualities of mud building for health. He may find articles through the Internet but lost between articles about missionaries, paper mud and vampires (8).
They all need easy to reach information that can be transformed into knowledge by it's quality of matching with the learner's needs.

4._Making knowledge accessible : metadata and ontologies

From a collection of documents, there are different ways for having some emergence of knowledge.

Among them :

- analysing the document and annotating it. It has long been done by teachers and researchers but the idea is to have these annotations made available easily to documents users.
- during the production of the document to use a “controlled” vocabulary, i.e. to refer to a set of words which are both shared and defined precisely enough. This is the classical approach that can be found in scientific databases and their thesaurus organisation. (9)

Classical scientific databases are powerful tools but can be enhanced by the World Wide Web capabilities in different ways (10) :

- a high variety of documents i.e. multiple formats like HTML, Word point doc, PDF text files, pictures and sounds.
- instant link between documents i.e. the HTML pages facilities to include Internet addresses pointing to texts, pictures and sounds available within the same document or on the Web.
- in terms of content, the immediate expression of any author.

One key aspect of the Internet is its openedness, the ability for authors to publish instantly documents that can be indexed by search engines.
“Data” is the content of the document i.e. text, matrixes, pictures, sounds, etc..
From the data of one document, it is possible to write metadata i.e. data about the data. This can take the forms of tagging or of annotation on a separate document linked to the first one.
Metadata existed at early stage of the Internet in a simple way. For example, this document has a tagging header automatically generated by the Word software. This tag is readable through the “File” “Properties” menus of Word where it is possible to see information about the document.
Within HTML, typical tags are “title”, “author”, “template”, “last author”, “revision”, “version”, etc..
This shows that inventors of the Web already thought about searching problems.
But we are at the very beginning of the Semantic Web paradigm. Right now, few authors fill in the tags purposely. The consequence is that the Google search engine manages the “author” and “title” tags only.
So, it is yet possible to find an article authored by Sir *Clay Cob* out of the articles talking about *clay* and *cob*.
The Semantic Web, which is a “super” way of tagging will need people to tag their documents. This change in editing habits will be possible because the authors either are themselves the users of the kind of document they annotate or must really respond to the needs of their stakeholders.
The World Wide Web Consortium is actually working on extensions of the tagging system.
The idea is to take advantage both of :

- the flexibility, opened ness, etc. of the Web.
- the power of organized thesaurus like systems.
- the skills acquired by Artificial Intelligence researchers in Description Logic.

HTML is now to be replaced by XML which is completed by RDF which is extended by DAML-OIL (12).
What is represented in matrix 1 is somehow an ontology i.e. a classification system.
In the preceding matrix, the ontology had a diagram form.
In the DAML-OIL language –then OWL-, it should be represented through textual statements. (13)
This is rather transparent to domain experts. Tools like Protégé (2) make it “easy” to create and maintain ontologies. Ontologies can be articulated one to an other. A system including one or several ontologies can be made accessible through the Web.
The development of a universal format, the Dublin Core (14) allows to work on common, shared bases, with other ontology developers.

5._Making it happen : the mud building European knowledge network

Mud building actors are both :

- working on the organisation of the knowledge gained both in field realisations, ethnologic observations and laboratory research.

- conscious of the emergence of the Web new possibilities –i.e. that the Semantic Web tools will soon be available and that they need to take this opportunity.

Our idea is to merge these two concerns, with a very practical approach : build a first version of an ontology of the mud building domain with the Protégé tool.
This is the work in progress with a mutual discovery :

- domain experts discover the ontologies principles and tools.
- ontology experts discover mud building.
- information and communication expert establishes the links.

There are global technical and human issues in developing ontologies.
Key issues include the "border/extension", "multi lingual", "enoughness", "outside links" dimensions of the ontologies.
The experience gained through this work could be transferred to other areas of the construction field.

Conclusion

At the same moment -this beginning of a century- two apparently distant communities are creative.

Mud building actors receive growing interest from individuals as well as organisations because of the qualities of the material and the technique –comfort, health and sustainability.

Semantic Web actors are urged to develop standards for ontologies and surrounding systems for easing the access to information, for the assessment of knowledge and the emergence of the Web of Trust.

These two dynamics are merged in an experiment for building the ontologies for the mud building domain.

The way is not easy. Engineers, architects, ethnologists, masons have their concerns, their languages.

Tools have their limits, the standards are not yet fully stabilized. There is both game and play. Game in the familiarisation with rules. Play in the necessary creativity. Our hypothesis is that most actors will increase their knowledge and a palette of related abilities.

Notes

(1) http://www.smi.stanford.edu/

(2) The ontology creation and management tool is available through free easy download at :

http://protege.stanford.edu/index.html

A first draft version of the ontology is available at :
http://bat710.univ-lyon1.fr/~cbois/ontomud/index.html

(3) The author is a senior researcher in Information and Communication Sciences as defined in French academic organisation. Keywords of the sub discipline are :

- interface between Information Technology props and the individual or group.
- migration of information from classical scientific databases (e.g. Medline) to information spread within different bodies, in different forms within the Internet realm.
- educational and political issues arising from the technological evolutions.
- transformation of information into knowledge, building the "Web of trust".
- evolution of public and private services within this new information/knowledge paradigm.

Among issues related to sustainable development, the author has been interested by mud building which is the traditional practice in the region of Lyon/Grenoble where he has been mostly living in the past decades.
On the other hand, he has a long term working fellowship with researchers in Artificial Intelligence in this same geographical area.
He has made links with mud building and AI people at a European and worldwide level.

(4) The DAML language is being merged with the OIL language to give raise to the OWL Ontology Web Language.
http://www.w3.org/2001/sw/WebOnt/

(5) There are different models for energy saving. One is known as the "thermostatic bottle model" where :
- wall insulation comes first.
- as few exchanges of air with the outside is managed. Many inhabitants simply can't stand it.

The mud building model relies on wall thermal inertia.
A simple description explains the difference between the two models. If one opens for a while the windows of a mud house in a very cold day, only the heat in the air is lost. After closing the windows, the house is very quickly warm as the heat is in the walls. With an insulated house, the heat is mostly in the air. If the inhabitant lets fresh air in, it takes a lot of time and energy to be back to temperature comfort.
Most "simple" person understand that. Many young architects who have been raised in the "insulation, mechanically controlled heat and moisture paradigm" have great difficulties to jump to the "inertia, heat transfer and evaporation/cooling paradigm".

(6) Examples of "subtle" knowledge
"The sound and vibrations of the machine tell me immediately whether the process is correct or whether I have to intervene. With time, you can just feel it." (interview with a metal worker, CeA-research project 1991)
"Our research indicates that, in manufacturing firms where production is tailored to meet specific customer requests, craft workers are often unable to describe how they go about their work. The remark of a worker at a paper manufacturing plant - "We know the paper is right when it smells right"- is not unusual." (B. Victor, A.C. Boynton 1998)
http://www.um.ipk.fhg.de/ccwm/impl%20knowledge/implicites%20knowledge.htm

(7) For example : http://www.craterre.archi.fr/craterre/Francais/Recherche/theme2.html

(8) The search was done with *"mud building" health*.

(9) Since 1972, Dialog was created as the on line data system for digitalized databases. http://www.dialog.com/about/

(10) We make it simple and don't consider exceptions to the main stream.

(11) Metadata
Searching with Google "meta data" gives 149000 answers.
"meta data" + *definition* gives 30500 answers.

(12) http://www.w3.org/TR/daml+oil-walkthru/
(13) Here is a simplified example of how what is said in natural language is translated in DAML-OIL language.
This is just a matter of syntax (what is the noun, what is the verb, what is the punctuation), very simple.
Authors of articles and/or annotations and/or ontologies will not have to know this syntax. They will work with user friendly tools. This example is to show how it will be behind and the logic of it.
"Mud building is a class" is expressed by :

```
<daml:Class rdf:ID="Mud building">
<rdfs:label>Mud building</rdfs:label>
"Adobe is a subclass of mud buiding is expressed by :
<daml:Class rdf:ID="Adobe">
<rdfs:subClassOf rdf:resource="#Mud building"/>
</daml:Class>
"Ramed earth is a subclass of mud building" is expressed by :
```

```
<daml:Class rdf:ID="Rammed earth">
<rdfs:subClassOf rdf:resource="#Mud building"/>
<daml:disjointWith rdf:resource="#Ramed earth"/>
</daml:Class>
```

"Nothing can be both Adobe and Ramed earth" is expressed by the "disjointWith" statement.
This makes the information in the matrix that was only readable by a human actor readable by a computer.
The DAML-OIL allows to define properties of a matter and the value for these properties.
"A property of mud building is to include components" is expressed by :

```
<daml:ObjectProperty rdf:ID="hasComponents">
```

"One of the properties may be to be wet or moist" is expressed by :

```
<daml:ObjectProperty rdf:ID="hasWater">
  <rdfs:subPropertyOf rdf:resource="#hasComponent"/>
  <rdfs:range rdf:resource="#Moist"/>
</daml:ObjectProperty>
```

(14) http://dublincore.org/documents/dces/

References

Berners-Lee Tim *Semantic Web Road map* September 1998.
http://www.w3.org/DesignIssues/Semantic.html
Marcom Alain *Enquète sur la productivité dans la construction en terre* Mémoire de DPEA Ecole d'Architechure de Grenoble Juin 2002 To be published soon Under the direction of Vincent Rigassi.
Salomon Thierry Bedel Stephane La maison des négawatts Terre vivante Ed. 1999
Wason Thomas D. **Dr. Tom's Meta-Data Guide**
http://www.imsproject.org/drtommeta.html

Many other documents are available within the websites cited in the Notes.