

Representing an application of ontology for the indexation of semi-structured data on Protégé, based on the ontological model of the Porter's Five Competitive Forces

Organizations generally have a bulk of information in their computers out of which they make little use, because they do not possess an efficient way to access those information. The major difficulty in having access to those data is caused by their structuring nature. They are non-structured data. Considering this fact, this article aims to find a solution to permit the access to non-structured data. A strategy of competitiveness is adopted as a factor for the indexation of those data. By means of the application of ontologies, a model was built to allow the structuring of those data, making them accessible for computational processing. The results suggest that it is possible to transform non-structured data into semi-structured data, and then apply existing computational facilities to make those data available for the organization.

An ontology was developed in this work, using Protégé, by domain modeling, about Porter's Five Competitive Forces.

The figure 1 illustrates the result of method applying the method developed by NOY and MCGUINNESS, 2001).

To comply what indicates the methodology in step one, three questions ought to be prescribed. a) Which domain the ontology will take part in? b) For which matters the ontology will be used? c) What kind of questions or information the ontology will have to answer for?

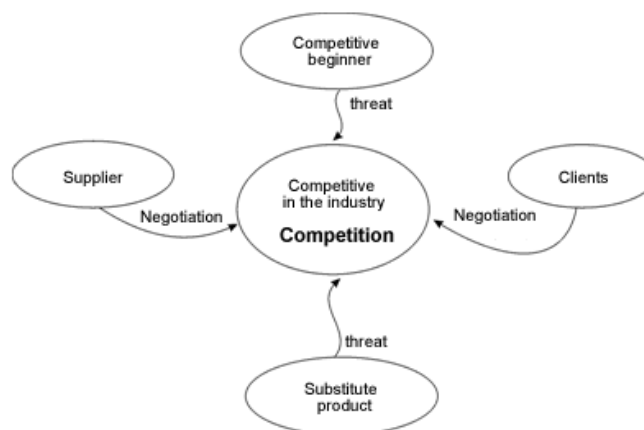


Figure 1 – Domain of the knowledge that will be modeled by the ontology

The competitive potential beginner proceedings have some interesting aspects related to the entrance of new companies in the market. Some of them are: a) Scale economy; b) The product's contrast; c) The capital's need; d) The altering cost; e) The access to the distribution ways; f) The disadvantage of scale independent costs.

But the competitive within the market proceedings broach the following rules: a) plentiful competitive or well balanced; b) The industry's low enlargement; c) The storage high cost; d) The lack of differentiating or the altering cost; e) The capacity grown with many increases; f) The competitive diverging; g) Many strategic interests; h) The use of strategies to establish high egress limits.

The substitute products aspects, in other words, those that can take possession of the existing products, it is important to detach the following aspects: a) The functions to be fulfilled by the services or products; b) The products or services that fulfill equivalent functions.

For the matters prominent of the supplier, it is important to detach: a) The market is controlled by few supplier companies, being more tight than the industries that makes it's own sales; b) The

Author: Antonio Pereira Candido (apec@cefetsc.edu.br); Jorge Hermenegildo (jorge@cefetsc.edu.br); Isabella Friedmann (isafg@csmtelecom.com); Diego Alberto Moretto (diegomoretto@brturbo.com.br)

supplier does not compete with substitute products; c) The supplier's product is very important to the buyer business; d) The products are distinguished or the group developed altering cost; e) The group of supplier are a deep threat to the integration in the future.

In the case related to the clients, it is important to detach: a) The convergence or acquisition of a big roll from one seller. This aspect can reduce the negotiation's capacity and establish a fragility aspect in the client – supplier terms; b) The products purchased by the client represent a significant fraction of its own costs or acquirements. c) The products acquired by the clients are standardized and non-differentiated; d) The client does not have many costs related to altering of any kind; e) The client get low profits; f) The possibility of beyond integration; g) The industry's products are not important in the matter of quality of the products or buyer's service; h) The access to the information.

In the step two a research begins to check the possibility of remaking use of the ontologies already existents in the company, or that are located in ontology's web servers. Therefore, the company ought to have documents of ontologies, to allow the access to prior projects, or references to developed projects that can afford data to begin at a point straight ahead, related to the point where the company is.

In the step three the tackle is to enumerate important conditions of the ontology. In this sense a list is presented, related to the important conditions considered by the ontology to be projected. a) To tackle in information non-structured; b) Identify related aspects with the theoretic reference, tackled in step one.

Regarding to the size of the ontology and the objective of the article, is given attention to the aspect related to the magnitude of the client. This particular aspect of the ontology, was consider in a more detailed way. Therefore, bellow are reported the relevant conditions for this competitive dimension, focusing this stage, proposed by the methodology: a) In which conditions would be to say something? b) Any document quoted to a client is not supposed to escape one's notice; c) Every document (is to be identified, containing threats to the strategy of competitiveness under the focus of clients); d) The directors of the company ought to know in the case of documents containing the name of clients of determined levels, contain subjects that can compromise the company's competitive strategy; e) It is ought to have conditions to identify documents by clients, documents by subjects, types of documents by clients and types of documents by subjects; f) The competitiveness strategy is supposed to be concerned to establish identification and selection of documents; g) The marking process is supposed to happen by non-specialized staff, who would not know the competitiveness strategy and procedure of selecting documents.

In the step four classes and hierarchy must be defined. Therein, it is being mentioned to the communication between company and client, specifically to the documents issued by the clients, it is notable that every document has the following concepts: a) Every document has one specific type; b) Every document has a sender; c) Every document has a receiver; d) Every document has one kind of subject; e) Every receiver has one type of classification; f) Every sender has um level of classification that fits.

In the step five, to establish an association between the classes defined above, the following relation is presented: *hasName x belongs to*. Considering the properties and the classes, it is possible to form, for example, the following sentences: Document has a subject; Subject belong to a document; Level has a client; Client belongs to a level; etc.

In the step six is performed the slots facets definition. The definition of the slots or facets ponders a classification that settle a relation between domain and image of this taxonomy. The types of classification are: Object property: One object property connects individuals. b) Datatype: A Datatype property connects an individual to a literal; Annotation property: An Annotation property connects classes to a literal or string.

In the step seven, is related to the instance creation. In this process is involved the fulfill of the properties described above. In many situations, the operation is performed in the data level. Therefore, this step is developed in the software making use to implement the ontology. To perform an example, one element of the ontology is presented, named Occurrence: One document with a **type of letter**, with a **complaint**, sent by **Alfa**, who is a client **Classe A**, for the department of **Sells**.

After the development of the computational model, its application allows the organization to keep tracking the elements that were identified as being members of the context of the domain

Author: Antonio Pereira Candido (apec@cefetsc.edu.br); Jorge Hermenegildo (jorge@cefetsc.edu.br); Isabella Friedmann (isafg@csmtelecom.com); Diego Alberto Moretto (diegomoretto@brturbo.com.br)

considered by the directors. To exemplify, consider that one of the domain's context is: "It is necessary to maintain a better relation with the clients". This element of the context is supported in the ontology through the document modeling resort, which is a class that links to the clients. Thus, it is considered that a contact made with the company, that has as a client a sender and the subject of the contact is a complaint, care for the assumption proposed by the context "It is needed to maintain the most good possible relation with the clients". Therefore, it must be identified and classified.

This property, identified in the contact with the client, is a loss or a symptom that the domain's context "It is needed to maintain the better possible relation with the clients" is being prejudiced. In this case, an action is established, a behavior to be followed by this company. This behavior can be the act of sending the client a mail, or sending to the channels that attends the client.

As a propose to practice and authenticate, a rule was developed indicating that, a client of the class "A", make contact with the company through a way of non-structured communication, as being a complain, the document will be identified and a message "There is a complain of the client *Name of the Client*" is shown.

As a propose of modeling, the creation of this phrase means that the event (receiving the mail for the client that is complaining) can be managed in a way that can be considered suitable when developing the tool using the solution created.

The non-structured documents usage process is made through the marking process by the employee who works in the bureaucratic sector of the company. The employee does not need to know about the process's result of its marking, but the indicated action accomplishment in the ontology's model will perform so that one application, that is not part of this article, starts a procedure, which can be a report to the director of the sales department, or the immediate sent of the mail indicating receipt and elucidation for the aspect at issue.

Therefore, the attention to the client who is complaining, does not depend on the sending of the mail for the director, or on the employee remembering of this routine implanted by the competitiveness strategy.

It is important to emphasize some aspects about the administration's support advantages. The responsible for the marking process does not need to know the policy that leads the ambience's conduct. In other words, the advertisement that one complaining client has arrived will be sent to the director without the interference of the employee who is performing the marking process.

The document containing non-structured data, through the marking process and the ontology support, can be converted in a semi-structured document, allowing the computational usage in several levels and for many proposes. One of them is to grant succession to the lines of direction as of the competitiveness strategy established by the company. The effectiveness on the supervision of the lines of direction of the competitiveness of the strategy is a fact, because it counts on the computational support that, besides it is quick, it is also exact and constant. The document's location in this case, by the directors, is immediate. This aspect helps the directors in the matter of supporting the competitiveness strategy, because the time to begin the corrective activities is diminished.

The models validation occurred by the application of the resources based in formal logic principals and the conjuncts theory of the groups, executed in a computational level, using resources of the tool that allows to develop the ontology (MUSEN et al., 1995).

As of the ontology's functioning, it is possible to establish validation rules that will grant supported rules by the domain's context. These rules are the lines of direction of the ontology's conduct.

Established the proposed model it was possible to conclude that the application allowed to index files from a theoretic reference, in this case, Porter's Five Competitive Forces.