RAPPORT DE RECHERCHE

L R I

EXPERIMENTAL ASSESSMENT OF THE TARGET ADAPTIVE ONTOLOGY-BASED WEB SEARCH FRAMEWORK

GUELFI N / PRUSKI C / REYNAUD C

Unité Mixte de Recherche 8623
CNRS-Université Paris Sud – LRI

10/2008

Rapport de Recherche N° 1504

CNRS – Université de Paris Sud
Centre d'Orsay
LABORATOIRE DE RECHERCHE EN INFORMATIQUE
Bâtiment 490
91405 ORSAY Cedex (France)

Experimental Assessment of the TARGET Adaptive Ontology-based Web Search Framework

Nicolas Guelfi
LASSY – University of Luxembourg
Nicolas.guelfi@uni.lu

Cédric Pruski
LASSY – University of Luxembourg
and
LRI-Univ. of Paris-Sud-CNRS & INRIA Saclay – Île-de-France
Cedric.pruski@uni.lu

Chantal Reynaud
LRI-Univ. of Paris-Sud-CNRS & INRIA Saclay – Île-de-France
Chantal.reynaud@lri.fr

Abstract

Finding relevant information on the Web is often difficult for most of the users. Although Web search applications are improving, they still need to be more intelligent to adapt to the search domain targeted by users, the evolution of this domain and users' characteristics. In this paper, we present an experimental assessment of the TARGET framework for improving the relevance of the documents when users are searching the Web by using adaptive ontologies. This is done first by introducing the TARGET approach. We will briefly present the used ontologies and their ability to adapt to domain evolution. We will then detail the TARGET tool used in our experimentations. This includes its architecture, its ability to carry out the ontology adaptation process as well as the way it searches the Web and ranks the returned results. Finally, we discuss the results obtained using the tool through the presentation of our case study devoted to the retrieval of scientific articles.

1. Introduction

Since the advent of the WWW in the early nineties, the ever-increasing number of documents constituting the Web requires the development of a new generation of intelligent tools in order to assist users to find relevant Web information. Actually, one of the main difficulties for common Web users lies in the construction of good queries. The choice of adequate keywords for targeting the right search domain is often hard as they are often ambiguous. Consider the query "publications on trees". It is difficult to decide whether the term *tree* refers to graph theory or to botany. In addition, by virtue of knowledge evolution, the knowledge of the search domain is constantly changing over time what makes the selection of the right keywords even more complex for users as they must be aware of these evolutions.

Besides, the domain targeted by the query is often huge. In fact, a given domain contains information for many different kinds of users and it is hard for usual search engines to decide what information is relevant for a particular user only by considering his initial query. In consequence, the person who entered the query must skim the results returned by common Web search engines to keep only the information relevant to his profile. Consider for instance the scientific research domain. An academic researcher may not be interested in the same kind of information as an industrial researcher may be even if the same initial query, which aims at targeting the same search domain, is entered.

As a result, we believe that the new generation of Web search tools should, among other things, assist users to target the right search domain, integrate domain evolutions and offer users the possibility to select the category of people they belong to. These elements (search domain, evolution and user profile) contain much information for filtering unwanted Web documents. However, domain evolution is a complex problem, which has rarely been integrated in Web search techniques. Hence, we propose the TARGET framework. It implements an original approach based on OWL [9] ontologies for improving Web search by adapting to both search domain evolution and user's profile. This is done through the coherent integration of adaptive ontologies for modelling search domain and users' characteristics, domain adaptation technique, query enrichment rules and advanced data structures for Web data extraction and ranking purpose. Furthermore, we propose an experimental assessment of this framework through a case study.

The remainder of this paper is structured as follows. Section 2 briefly introduces the TARGET approach. Section 3 deals with the tool implementing the proposed approach. We present in Section 4 the experimental assessment of the framework. Section 5 proposes some related work in the field of adaptive Web search. Finally, last section wraps up with our concluding remarks and outlines future work.

2. The TARGET approach

The TARGET approach, illustrated in figure 1, aims at improving (in terms of relevance) the results of a Web search.

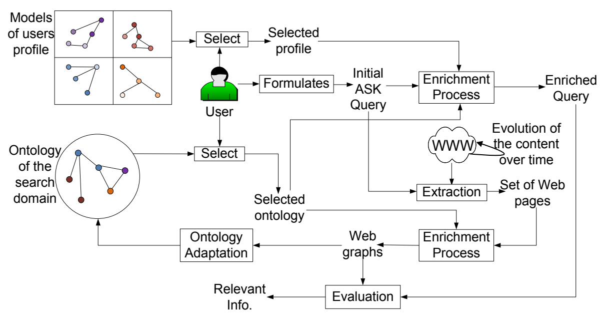

Figure 1. The TARGET approach

The development of the overall process can be divided into two phases. During the first one, the user plays the main part. Actually, as it is the case in most of existing Web search applications, the user enters his query (here using the ASK query language [4]). Then, he has to select both the targeted search domain and the profile it characterizes him best (models of user profile in figure 1). In fact, these two elements correspond to OWL adaptive ontologies loaded in the system. The remaining, which consists in the second part, is completely automatic and transparent for the user. The system will extract Web pages by submitting the query to a common Web search engine. In parallel, the system enriches the emitted query with the selected ontologies according to rigorously defined query enrichment rules [5]. The set of pages returned by the engine is then transformed into WPGraphs and W^3Graph [4] (Web graphs on figure 1) using the domain ontology for ranking purpose. Finally, the enriched query is evaluated and the most relevant Web pages are extracted from the Web graphs and ranked.

Although the TARGET approach seems to be quite similar to existing ones, it derives its originality in the use of adaptive ontologies. These are not only able to represent the knowledge of different domains but they also have the capability to adapt to the evolutions of their associated search domain. Therefore, we have to zoom in the characteristics of such ontologies to understand their properties as well as their contribution in the enhancement of the quality of Web search.

2.1. Adaptive ontologies

TARGET is an adaptive approach because the implemented ontologies have the ability to reflect the evolution of the knowledge of their respective domains in order to represent them as faithfully as possible. To make our ontologies adaptive, we had to define a set of features that allows the characterization of knowledge evolution. The proposed features are inspired from general knowledge representation and management. They have been experimentally assessed using a study of ontology evolution for a specific domain [6]. We have highlighted the following evolution features. First, we observe the **emergence** and/or **deletion** of knowledge that occurs when new concepts emerge in the domain or in the contrary when obsolete ones are removed from it. In fact, these aspects can result from a specialization or generalization of knowledge when the evolving domain requires more or less precision. We also observed that some concepts **persist** in the domain over time, different concepts do not have the same **semantic weight** (i.e. importance) in the domain and the **semantic distance** that separates concepts (i.e. concepts directly linked by a given ontological relation such as equivalence or subsumption) can vary according to their usage in the domain. Lastly, we identified the ability of knowledge to resist to ongoing changes. This **resistance** prevents, in some words, the deletion of concepts of the domain or the modification of their semantic weight or semantic distance.

Concerning the representation of these features, we advocate the use of OWL *annotation property*. The emergence of concept is represented using a date that corresponds to the moment when the concept integrates the domain. The persistence duration as well as the semantic weight and the semantic distance are represented using integers whereas the resistance to change is represented using a real number. Observe that the use of annotation properties requires OWL DL because we annotate only classes and object properties, which is important for reasoning purposes, since OWL DL is decidable.

We define ontology evolution as semi-automatic incremental process. An evolution step aims at fixing the new ontology definition by applying adaptation rules on its elements. At each evolution step, we must:

1. Integrate every new concept of the domain resulting from the automatic analyze of the domain (see section 3.2). This implies the definition of new concepts (i.e. OWL class) in the ontology and the definition of relations (i.e. OWL properties, equivalence or subclass) to link the new concepts to the existing ones. Moreover, the newly added concepts must be annotated with the evolution features.
2. Revalue the persistence duration for each existing concept and remove concepts whose persistence duration reaches zero.
3. Revalue the semantic distance for each existing relation of the ontology (i.e. Object properties, equivalenceClass and subclassOf).
4. Revalue the semantic weight for each existing concept of the ontology.
5. Possibly reassign the resistance for existing ontological relations.

This adaptation process only requires user intervention for linking the newly detected concepts to those already existing in the ontology. The tool implementing the TARGET approach presented in section 3 has been built in order to assist users in this task. The remaining of the process is automatic and is based on the use of a corpus of documents that is built as one goes along with the Web documents returned to the user after a search. We will detail this aspect more precisely section 3.2 as well as the various metrics that intervene in the revaluation of the various evolution coefficients. Furthermore, we propose an experimental assessment of the rules. The results we obtained are presented in section 4.

2.2. Adaptive Web search

The adaptive aspect of Web search in the TARGET approach is done using the adaptive ontologies presented earlier. In fact, this model of adaptive ontologies combined with the adaptation rules serve in our approach to represent both the search domain targeted by users and the different views they could have on the domain (see figure 1) and to make them adapt to domain evolution over time. Then, these ontologies are used for two different purposes.

The first one deals with the query enrichment process. Actually, the adaptive ontologies used to represent the search domain and the user's profile serve as basis for query enrichment. In our previous work [5], we proposed a set of ontology-based Web query enrichment rules. These rules rely on various existing ontological relations and aim at extracting the concepts of the ontology that best characterize the domain with respect to the terms of the initial query. The so-extracted concepts are then added in a rigorously defined manner to the initial query. Nevertheless, depending on the size of the ontologies, the number of extracted concepts can become problematic which in turn can affect the quality of the enriched query. This is why we refine the list of concepts taken from the ontology with respect to the evolution features extracted concepts are annotated with. To this end, we favour first the semantic distance, then the semantic weight and last, emergence date. It means that we keep the concepts that are close (from the semantic distance point of view) from those of the initial query, then if too many concepts remain, we keep those whose semantic weight is the highest and eventually filter the oldest (regarding the emergence date) ones. Then we add the most important concepts (in the sense of the semantic weight) of the selected user's profile. The so-enriched query contains information related to search domain and user's characteristics and will allow to target the right documents at query interpretation time.

The second facet concerns the ranking of the relevant pages. This is done based on the structural properties of the Web graphs that are constructed using adaptive ontologies. This aspect will be detailed much more precisely later in section 3.3.

3. Implementing the approach: the TARGET tool

This section presents the specificities of the tool that implements the TARGET approach. Since the approach integrates ontology adaptation and Web search, we will base the presentation of the tool on these two perspectives.

3.1. Tool architecture

The architecture of the tool (http://se2c.uni.lu/tiki/tiki-index.php?page=TargetTool) that implements the TARGET approach is depicted in figure 2 hereafter.

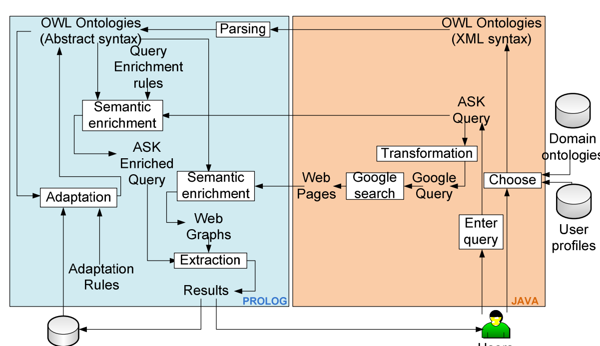

Figure 2. Tool architecture

The tool is made up of two parts. A front-end developed in Java that acts as an interface with the user and the second part that contains the PROLOG inference engine that manages both the ontology adaptation and the adaptive Web search tasks. Actually, the OWL ontologies selected by the users are first parsed using the THEA PROLOG (http://semanticweb.gr/TheaOWLLib/) parser and loaded as PROLOG facts in the inference engine.

3.2. Ontology adaptation

The adaptation process is based on a corpus (see figure 2) incrementally built of pages returned to users at search time. We believe that these pages are significant for the search domain so they can be exploited for ontology adaptation purpose.

To explain the adaptation process performed by the tool, we will follow the steps given section 2.1. We start with the new concepts of the domain. In order to detect them, the system analyzes the corpus. In fact, the system counts the words contained in the documents of the corpus. We assume that the terms that are not part of the ontology yet and that appear the more frequently on these pages must be considered as relevant for the domain. In consequence, they must be proposed to the user as labels of new concepts of the domain at evolution time. The system allows the selection of the concept users want to add to the ontology. It also permits the definition of the type of relation (i.e. equivalence, subsumption …) that links the new concept to other existing concepts. Moreover, the evolution features' values of the newly added concept can be set up (if the user does not agree the proposed default values) and those of the existing concepts can be modified if needed.

In the other steps of the process, we need several metrics for adjusting the coefficient according to domain evolution. These metrics are defined using statistics applied to the documents of the corpus and concerns the semantic distance, semantic weight and persistence duration.

The updated value of the semantic distance between two concepts c and d at time τ, $SD_{(c,d)\tau}$, is defined as a function of its former value, $SD_{(c,d)\tau-1}$, the value of the resistance to change, $R_{(c,d)}$, and a value computed on the documents of the corpus that have been added between τ-1 and τ namely $\Sigma_{(c,d)\tau}$. This latter value represents the average of the number of words separating consecutive occurrences of c and d in the newly added documents. So for instance if in a document the first occurrence of c and the first occurrence of d are separated by 10 words and if there are 20 words between the second occurrences of both terms the corresponding value of $\Sigma_{(c,d)\tau}$ will be (10+20)/2 = 15. As a result,

$$SD_{(c,d)\tau}=SD_{(c,d)\tau-1}+\frac{\Delta SD_{(c,d)}}{R_{(c,d)}} \text{ with } \Delta SD_{(c,d)}=\Sigma_{(c,d)\tau}-SD_{(c,d)\tau-1}$$

We assume that the more frequently two concepts are jointly cited and the less words separate these concepts in the relevant documents of the domain, the closer they are. Therefore, we take into account the variation of the semantic distance engendered by the evolution weighted by the resistance to change. The semantic weight of a concept c ($SW_{(c)\tau}$) is updated in the same manner but here the value computed on the corpus ($I_{(c)\tau}$) is the frequency of concept c in the newly added documents. Hence:

$$SW_{(c)\tau}=SW_{(c)\tau-1}+\frac{\Delta SW_{(c)}}{R_{(c)}} \text{ with } \Delta SW_{(c)}=I_{(c)\tau}-SW_{(c)\tau-1}$$

Here we assume that the importance of a concept is linked to the frequency this concept is cited in the relevant documents of the domain. The resistance to change plays an important part in the evolution process since it has an influence on the impact of the evolution of the domain. In consequence, a high resistance will slow down the modifications of the semantic distance and semantic weight. In the contrary, a low resistance will speed up the evolution process.

Concerning the persistence duration, the resistance is applied in a different manner. In fact, the value associated to the resistance tells how many steps it will take to reduce the persistence duration. For instance, a resistance of 2 means that the persistence will decrease every 2 evolution steps while a resistance of 0.5 means a decrease of 2 units in 1 evolution step. A value set to 0 for the resistance associated to a concept implies a direct deletion of the concept from the ontology since its persistence duration will be set to 0.

The ontology adaptation process we proposed is consistent in the sense that the resulting ontology remains coherent (i.e. reasoning can still be performed) with respect to the modifications that have been done mainly the addition and/or deletion of new concepts and relations that engender structural modifications on the ontology. This property can be checked easily using the predicates offered by the THEA library.

3.3. Web search

The tool uses Google as an entry point to the Web. Actually, the query is sent to Google, then in parallel the query is enriched using the selected domain ontology and user profile according to the rules mentioned before. The pages returned by Google are enriched using the domain ontology into Web graphs and loaded into the PROLOG inference engine as PROLOG facts. Finally, the enriched query is processed as a PROLOG formula over the PROLOG

Web graphs. The system re-ranks pages, which fully match the enriched query, and displays it, in a friendly manner, on user's screen. The ranking we propose relies on the structure of the Web graphs. In fact, to each Web page corresponds a WPGraph [4]. The vertices of the graph represent the label of concepts (or terms) of the page and edges represent the semantic link between concepts. This set of edges is built according to the domain ontology. Consider two terms of a page, if these terms are part of the domain ontology and if there is a path between these concepts in the ontology, then there is an edge between these two concepts in the WPGraph. As a result, the more edges the WPGraph will have, the most relevant (regarding the domain ontology) the associated Web page will be. Hence, we rank the results with respect to the number of edges of the WPGraphs that fully match the enriched query.

4. Experimental assessment

In previous sections, we have presented the different concepts of the TARGET approach. As explained, the adaptation is made with respect to domain evolution and user's profile and using ontology-based query enrichment rules. In order to strengthen these aspects, we made several experimentations using the tool detailed section 3.

4.1. The case study

The case study on which the assessment of the approach is based, deals with the retrieval of scientific articles published at the World Wide Web series of conference. To this end, we have gathered all the research papers accepted over a decade (i.e. from 1996 to 2006). This set of documents forms our "micro" Web (observe that these documents can be found as HTML pages on the Web). Moreover, from the call for papers of the conference we have built the (initial) ontology representing the domain of the conference [6]. From that point, we made this ontology evolved and we used it to build various Web graphs. Then, we reuse both the obtained graphs and ontologies for query enrichment purpose.

Our micro Web has the advantage to be finite and of suitable size (i.e. about 600 documents). Thus we will be able to measure the recall and the precision of the returned results, which is of utmost importance for information retrieval approaches. Moreover, because of this micro Web, we will shortcut the use of Google as the entry point to the Web for our experimentations.

Concerning the user profiles, we defined two various profiles. The first one reflects the basic knowledge of a fundamental researcher (i.e. someone who is interested in fundamental research). In consequence, the ontology representing this domain is made up of concepts like *formal*, *mathematics*, *algebra*, *proof*, *theorem* and so on. The second profile describes the knowledge of an applied researcher, which corresponds, in some words, to an industrial researcher. Therefore, concepts of this ontology are rather: *tool*, *prototype, application, industry*, etc.

4.2. Scenarios

Our main objective is to show that both our adaptation process and our query enrichment rules really improve the relevance of the results of a Web search (i.e. improve the precision and the recall). In consequence, we propose four various scenarios based on our case study.

The first one (scen1 on figures 3 and 4) consists in the basic case. In fact, this case corresponds to the Web search process offered by usual Web search engines. Therefore, we neither perform any query enrichment nor use any user profiles or domain ontologies. Observe that we restrain the search space to our micro web. This consists in a first filter, which is not the case when we search the Web using a Web search engine. In consequence, the precision of the returned results will already be enhanced.

The second set of experimentations (scen2) highlights the benefit of using ontologies for improving Web search. It consists in implementing basic OWL ontology (without evolution feature) for representing the search domain and uses it as input for query enrichment rules in order to enrich the initially submitted query.

The third scenario (scen3) put the stress on the benefit of adaptive ontologies. We use adaptive ontologies (with evolution features) for representing the search domain but we do not use any user profile in the query enrichment process.

The last one (scen4) implements the full TARGET approach, which includes the use of adaptive ontologies obtained using the adaptation process presented section 2.1, the use of user's profile and advanced query enrichment rules.

For all scenarios, we have tested a hundred of queries implementing all the constructors of the ASK query language and we measure manually the precision and the recall of the documents returned by our tool.

4.3. Experimental results

In order to build the set of initial queries, we consider ourselves as basic Web users. For this reason, we had to decide two things. The first one deals with our focus (i.e. what will be the subject of the documents we are looking for). The second point concerns our corresponding profile (i.e. if we are rather fundamental or industrial researchers). The results of our experiences are contained in the two figures hereafter. Figure 3 represents the measurement of the precision of the results whereas figure 4 illustrates the recall of the same obtained results.

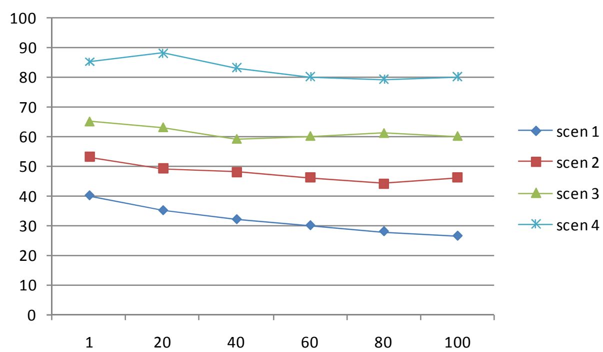

Figure 3. Precision of the search results

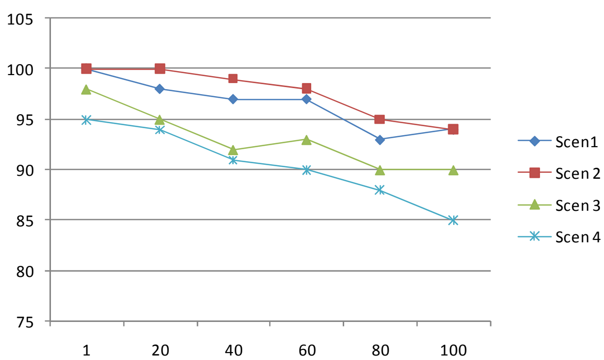

Figure 4. Recall of the search results

As illustrated in the two graphics the TARGET approach as described in this paper gives better results as usual search applications or basic ontology-based ones. This is all the more so true concerning the precision of the returned documents. Actually, the TARGET approach (Scen. 4 on the graphics) reaches an average of 80% in precision. The high precision underlines the quality of the adaptation process defined in section 2.1 and the comparison of the precision of scen2 and scen3 shows the interest to consider adaptive ontologies. Furthermore, the difference of the precision between scen3 and scen4 highlights the benefit of using user's profile for searching the Web. In fact, the adaptive ontology representing the search domain allows to target the right search space while the profile of the user permits to filter among the relevant pages of the domain the remaining pages that will not interest the user. This explains the high precision.

Recall is less affected. Actually, as aforementioned in the paper, our main objective is to increase the relevance of the returned results and our measures of the precision reinforce our position but the precision increases to the detriment of the recall. Because we filter many unwanted pages (by virtue of the query enrichment rules where we favour the use of the conjunction operator of the ASK query language), it happens that some pages that could be considered as relevant, slip among hence affecting the recall. The other scenarios have a better recall because they return a huge number of documents and among them of course there are the most relevant ones. This, in turn, reduces drastically the precision of the returned results.

Our approach is particularly efficient when the keywords of the initial query are imprecise, ambiguous or few, which is the case for most of Web queries submitted to usual Web search engines [11]. This is the result of combining adaptive ontologies representing the search domain that allows the disambiguation of the query and user's profile to reinforce the filtering effect of the query.

From a complexity in time point of view, our approach is not that efficient because of the construction of the Web graphs, which is time consuming. However, as a guide, during our experimentations the building of Web graphs, representing about six hundred of huge (i.e. 8000 words per page) Web pages, took less than five minutes, which means less than one minute for hundred pages (this is important as we apply this approach on the hundred first pages returned by Google). This construction cost is largely compensated by the quality of the search results compared to the very low precision of usual search application (scen1 on figure 3). We believe that it is acceptable for a user to wait few seconds to have the relevant information he is really interested in.

5. Related work

In the field adaptive Web search, various interesting approaches were proposed. They differentiate mainly in the way the adaptation is performed.

First of all, many approaches claim that Web search systems should directly adapt to the characteristics of the user who initiates the search. Such an approach has been proposed in [12] where the authors build a profile of the user based on information related to his Web navigation history. The approach proposed by Liu et al. [8] advocates the study of user's previous queries to build his associated profile. Although having shown great results, this kind of approach raises several problems. The first one concerns the quality of the so build profile. Actually, the behaviour of Web users can influence the quality of the profile in particular when users are not focusing on a given domain but have numerous and completely different focuses. The profile will be vague and can hardly be reuse for Web search. The second major problem deals with data privacy issues because personal information are stored and treated without users' agreement.

Second, existing approaches put the stress on the context of the search [7]. Most of them apply query expansion techniques using various artefacts like linguistic tools such as ontologies [1], statistics heuristics or machine learning techniques based on text collections [10], or feedback [3]. This consists roughly in extracting domain specific keywords and appending them to the initial query. This leads to the disambiguation of the query. Moreover, Chirita et al. [2] combine personalization and query reformulation by analyzing the so-called Personal Information Repository (i.e. the personal collection of text documents, emails, cached web pages, etc). However, not any of the evoked approaches take into account the domain evolution problem. This can be problematic mainly for Web search approaches that rely on the use of linguistic tools.

To wrap up, we try to make the most from both families of approaches. Hence, our work is original from existing one in the sense that we use ontologies for representing both the search domain and user's profile that have the ability to adapt to the evolution of their respective domains for enhancing Web search. Moreover, we use these adaptive ontologies as input of query enrichment rules and ranking technique that have shown great performances in Web search.

6. Conclusion

In this paper we have presented the TARGET framework for optimizing the relevance of the results when searching the Web. Our approach, which integrates domain evolution and user profile in order to make the search adaptive, has shown interesting results as highlighted by the serious experimental validation we have carried out. Nevertheless, some aspects of the approach need improvements. Actually, the construction of the Web graphs is time consuming which could be considered as a drawback. This is why our future work will be devoted to enhance the construction of these data structures. We also plan to improve the adaptation process by making it as automatic as possible. The aspect of the user interface dealing with the query capture and the display of the results is also under improvement.

7. References

[1] J. Bhogal, A. Macfarlane, and P. Smith, "A review of ontology based query expansion," *Information Processing and Management*, vol. 43, pp. 866-886, 2007.

[2] P. A. Chirita, C. S. Firan, and W. Nejdl, "Personalized query expansion for the web," in *SIGIR '07: Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, (Amsterdam, The Netherlands), pp. 7-14, ACM, 2007.

[3] Z. Chen, X. Meng, B. Zhu, and R. H. Fowler, "Websail: From on-line learning to web search," *Knowledge and Information Systems*, vol. 4, no. 2, pp. 219-227, 2002.

[4] N. Guelfi and C. Pruski, "On the use of ontologies for an optimal representation and exploration of the web," *Journal of Digital Information Management (JDIM)*, vol. 4, September 2006.

[5] N. Guelfi, C. Pruski, and C. Reynaud, "Les ontologies pour la recherche ciblée d'information sur le Web: une utilisation et extension d'OWL pour l'expansion de requêtes," in *Proceedings of the Ingenierie des Connaissances 2007 (IC07) french conference*, (Grenoble, France), July 2007.

[6] N. Guelfi, C. Pruski, and C. Reynaud, "Understanding and Supporting Ontology Evolution by Observing the WWW Conference," in *Proceedings of the International Workshop on Emergent Semantic and Ontology Evolution (ESOE 07)*, (Busan, South Korea), November 2007.

[7] S. Lawrence, "Context in web search," *IEEE Data Engineering Bulletin*, vol. 23, no. 3, pp. 25-32, 2000.

[8] F. Liu, C. Yu, and W. Meng, "Personalized web search by mapping user queries to categories," in *Proceedings of the eleventh international conference on Information and knowledge management*, pp. 558-565, ACM, 2002.

[9] D. McGuinness and F. van Harmelen, "OWL web ontology language overview." W3C Recommendation, February 2004.

[10] S. Oyama, T. Kokubo, and T. Ishida, "Domain-specific web search with keyword spices," *IEEE Trans. on Knowl. and Data Eng.*, vol. 16, pp. 17-27, 2004.

[11] C. Silverstein, H. Marais, M. Henzinger, and M. Moricz, "Analysis of a very large web search engine query log," *SIGIR Forum*, vol. 33, no. 1, pp. 6-12, 1999.

[12] K. Sugiyama, K. Hatano, and M. Yoshikawa, "Adaptive web search based on user profile constructed without any effort from users," in *Proceedings of the 13th international conference on World Wide Web*, pp. 675-684, ACM, 2004.