Textile Engineering Domain Specific Optimized Web Search System

Varsha S. Kshirsagar
Dept. of CSE, NK Orchid College of Engineering & Technology, Solapur, Solapur University, Solapur, Maharashtra State, INDIA

Prof. S. A. Khuba
Dept. of CSE, NK Orchid College of Engineering & Technology, Solapur, Solapur University, Solapur, Maharashtra State, INDIA

Abstract - The Semantic Web is next evolutionary paradigm of present web. This evolved paradigm is tagged as Web 3.0. Semantic Web is new form of web where content of the web is meaningful to computers. This will unleash a revolution of new possibilities. Ontology is at the core of the Semantic Web. Ontology is used to explicitly describe and represent the domain specific conceptualization. Useful information is extracted from large unstructured data source and the general-purpose as well as domain specific ontology have been developed. This paper describes system architecture for developing ontology in Textile engineering domain. This domain specific ontology is used for query optimization. Textile engineering domain ontology defines taxonomy. Taxonomy is publishable on internet. Definitions of basic concepts are included in the taxonomy which is machine-interpretable. The advantage of using Textile Engineering Domain ontology is demonstrated with the strategy of query optimization. Three kinds of query optimization techniques are used. These are respectively based on is-a relation, part-of relation, and equivalent-class relation in the domain ontology. By using this strategy, the user query is optimized by using a domain-specific ontology. The optimized query is processed using Google API for web information search.

Keywords - Semantic Web, Ontology, Query Optimization, Information Retrieval and web search.

1. INTRODUCTION

Paper focuses on the two phases that are involved in development of project:
1) Developing OWL Textile Ontology.
2) Building domain specific information searching system based on GoogleAPI and domain ontology, where optimized user query is processed.

Developing OWL Textile Ontology

The World Wide Web is a resource of huge information with virtually unlimited potential. However, this potential is relatively untapped as it is difficult for machines to process and integrate this information meaningfully. Therefore, researchers have begun to explore the potential of associating web content with explicit meaning, in order to create a Semantic Web [7].

So as to integrate information from different sources, there is need to have shared understanding of the domain and organizing the knowledge. Knowledge representation formalisms provide structures for organizing knowledge.

In Artificial Intelligence, the most cited definition of ontology is given by Tom Gruber: “Ontology is an explicit specification of a conceptualization” [4]. Conceptualization refers to an abstract model of some phenomenon in the world by having identified the relevant concepts of that phenomenon. Explicit means that the type of concepts used, and the constraints on their use are explicitly defined [3]. Formal refers to the fact that the ontology should be machine-readable. Shared reflects that notion that an ontology captures consensual knowledge that is, it is not private of some individual, but accepted by a group”.

Domain Ontology is an ontology Model which provides definitions and relationships of the concepts, major theories, principles and activities in the domain. Domain ontology provide shared and common understanding of a specific domain [4].

Query Optimization

Whenever user submits a query to any search engine, problem occurs with retrieval system such as low precision and low recall. The reason behind the low precision of retrieval system is that users can hardly express their requirements in a formal way which is essential for the computer to understand the user’s intention. On the other hand, the low recall is because of that the search engine retrieves information mechanically with the technique of keyword matching and returns only those results fully or partially matching the keywords. So, in traditional information searching system, there will be no result returned for two reasons. First reason is if none of the keywords is matched during the retrieval, and second is that the truly desired information is listed at the far end of the result for its poor matching with the
Therefore we can improve the quality of information retrieval by optimizing the user query. And for query optimization, domain ontology is used as a middleware between people and computer. The strategy includes three kinds of optimization based on is-a relation, part-of relation and equivalent-class relation [2].

2. RELATED WORK

Some of the existing knowledge base in textile engineering specific domain area is as follows:
1) RIKIPEDIA - The rieder textile knowledge base offers specialists, students and other interested parties a reference work with a practical focus on the technology of short staple spinning.
2) COMAPRE-Cleaner production textiles knowledge base, which is primarily aimed at the textile industry to assist in benchmarking performance through national, international and corporate norms and standards for cotton textile production.
3) FDAS-Fabric Detect Analysis System is intended to be an identification and diagnosis system for defects encountered in woven fabrics.
4) TEXMEDIN - Textile and apparel EuroMEDiterranean heritage for Innovation, is a transnational project, whose objective is to create a transnational cluster to increase the competitiveness of the partners territories in Textile and clothing sector focusing on quality, design and innovation.

Knowledge base which exists in the textile engineering domain area are all based on specific process in textiles and are not internet accessible. Therefore aim of this project is to develop generalized knowledge base using OWL ontology language and this knowledge base will be accessible through internet. Due to inference capability provided by OWL, a new knowledge can be derived from knowledge described in OWL ontology and use of created Textile domain ontology is demonstrated by building domain specific information searching the system based upon Google API and domain ontology, where optimized user query is processed[2].

3. BLOCK DIAGRAM

As shown in Figure 1, there are three important blocks namely information in the form of text, textile engineering OWL Ontology and query optimizer.

1) Information in the form of Text: This block represents the knowledge about the Textile Engineering Domain formally. Knowledge can be obtained from different resources like Textile Engineering Domain Dictionary or you can take help of any Textile Expert to get knowledge or from internet.

2) Textile engineering OWL Ontology: An ontology “consists of concepts, hierarchical (is-a) organizations of them, relations among them, axioms to formalize the definitions and relations. The aim of this project is therefore to develop Textile knowledge base that can be shared and processed by machines as well as people.

3) Query Optimizer: By adopting Textile Domain Ontology as a semantic base, we implement an information searching system, which firstly optimizes the users query according to the relations including is-a, part-of, and equivalent-class in the domain ontology and then process the optimized query and finally merges the searching results to the user.

4. RESULTS AND DISCUSSION

Here, Textile Engineering domain ontology is developed by using Protégé Ontology Editor. Developed Domain Ontology contains different classes, attributes, instances, relations and axiom and can be deployed on ontology repository so as to make it internet accessible.
The query portal provides a unified interface for the users to enter their query and to show the searching results passed from query optimizer. Figure 3 shows the query page.

5. CONCLUSION

Contemporary information searching systems have little semantic inferring ability to search the hidden information in the web. Google uses some techniques of natural language processing, for example, extension to synonyms, but does not utilize the ontological relations between concepts. This results into low precise and recall of truly required information. That’s why user’s search requirements are seldom satisfied. On the other hand, we have noticed that information required by users is often strongly dependent on a specific domain. Therefore, in this paper we have devised system architecture where a Textile Engineering ontology is developed and three kinds of query optimization respectively based on is-a relation, part-of relation, and equivalent-class relation in the domain ontology are investigated. By using this strategy, a domain-specific information searching system building upon GoogleAPI and domain ontology is built, where optimized user query is processed.

6. REFERENCES