



The 7th International Conference on Ambient Systems, Networks and Technologies  
(ANT 2016)

## Cross-Domain Semantic Web Model for Understanding Multilingual Natural Language Queries: English/Arabic Health/Food Domain Use Case

Ahmed Al-Nazer, Saeed Albukhitan, Tarek Helmy\*

*Information and Computer Science Department, College of Computer Science & Engineering,  
King Fahd University of Petroleum & Minerals, Dhahran 31216, Mail Box 413, Saudi Arabia,  
[alnazera, albukhitan, helmy]@kfupm.edu.sa*

---

### Abstract

With the growth of the Semantic Web and its applications, the need to use it in different languages, such as Arabic, is becoming more important. Two of the challenges with the Semantic Web technologies are the lack of multilingual support and the complexity of integrating multiple ontologies used by this technology. The objective of this paper is to present efforts that will help users who use the Arabic language to ask natural language questions and then get their semantic representation in SPARQL that allows them to be executed and get the relevant semantic results. This natural language interface makes more use of the cross-domain ontologies and hence improves the understanding of their inquiries, which is needed in some critical domains such as health and food where precise advice is essential. The approach we followed is multilingual and overcomes the limitations in the published relevant systems. With the proposed approach, users who speak Arabic can use the widely published ontologies in English without concern for the translation of their questions. The proposed approach will take care of matching the entered questions with the relevant ontologies to produce their semantic Web queries. The proposed approach has been implemented and empirically evaluated. The experimental results are promising, which will help in improving the awareness and usage of the Semantic Web by different lingual and cultural users.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of the Conference Program Chairs

*Keywords:* Semantic Web; SPARQL; Arabic Natural Language; Question Answering; Semantic Web Service.

---

---

\* Corresponding author. Tel.: +966-3-860-1967; fax: +966-3-860-2174.

*E-mail address:* [helmy@kfupm.edu.sa](mailto:helmy@kfupm.edu.sa)

On leave from College of Engineering, Tanta University, Egypt

## 1. Introduction

The Semantic Web is an extension of the normal Web where data are linked and reasoned. Semantic Web technologies help in building knowledge bases that contain linked vocabularies along with rules that organize the relationship between these concepts. While there are many applications and domains that have used Semantic Web technologies, there are growing challenges when it comes to using these technologies, especially when it requires average users to link different languages and domains. A lot of Semantic Web resources and data sets are available in common languages, such as English, while users from different languages are challenged to utilize this wealth of information on the Web. Although there are some efforts to translate the queries on the common search engines, translation wouldn't work on the Semantic Web because of the nature of the technology that relies on the meaning of the question. Thus, we address this challenge in this paper by proposing a model that deals with this barrier in a better way to allow users from different languages, such as Arabic, to use the available resources in English. Correlating different domains is one of the challenges once we start looking into the Semantic Web repositories, as each domain has its own experts who are focusing on the domain vertically. Because there is a need to deal with the domains horizontally and manage cross-domain questions, we address this challenge through this paper, which deals with critical domains. In this paper, health and nutrition domains were chosen as a case study because they are critically important. Also, there is a growing demand on the Web for answers to health-related questions, while the common search engines still depend on the keywords and hence mislead the end-users. In the health and nutrition domains, we focus on the relations between “*diseases*,” “*body parts*” and “*body functions*” with the food and nutrients.

The Natural Language Interface (NLI) deals with the Natural Language (NL) queries and questions. It takes care of the complexities behind the used technology and eases the input required by the common users. We implement an NLI that can help in reaching out to the Semantic Web repositories and ontologies to produce the matching Semantic Web queries to the user's questions. This is not limited to the specific language the ontology is in; rather, it extends to be used in cross-language use cases. In this paper, we address these challenges with the objective to have an easy-to-use tool that can convert the natural language queries into Semantic Web queries using the relevant technology, namely, Simple Protocol and RDF Query Language (SPARQL). We aim to translate and match the NL input into the relevant formal query written in SPARQL. This query can then be executed in the data store to get the matching Semantic Web results. Any SPARQL query consists of a number of triples, in the form of subject-predicate-object, where the query reasoning engine matches the triples of the SPARQL query with the stored Resource Description Framework (RDF) triples in the knowledge base. These knowledge bases are created in an earlier phase during the annotation process.

The rest of this paper is organized as follows. Section 2 reviews the existing related work on the NLI for the Semantic Web and their Arabic language support. Section 3 presents the ontologies we have used in our use case. Section 4 presents the architecture of the proposed approach. Section 5 discusses the process of semantic query manipulation. Section 6 presents the experimental results and discussion. Finally, we summarize the paper and highlight the future work directions in Section 7.

## 2. Related Work

In this section, we present a literature review on the related work to the NLI for the Semantic Web and their Arabic language support. Let's discuss first the support for Arabic language in the Semantic Web technologies. Based on our findings, there is limited work on the use of the Semantic Web technologies to build applications in the Arabic language. Yet, there is some work on the development of Arabic ontologies, which can be found in the literature<sup>1,2,3</sup>. There are also some works related to one stage in the Semantic Web, namely the entity extraction, in Arabic using ontologies<sup>4,5</sup>. There are also a number of papers that employ the ontology representation for Islamic knowledge<sup>6,7,8</sup>. In addition, there are some publications that use ontology on cross-language information retrieval<sup>9,10</sup>. Most of the use for ontologies in the information retrieval is focusing on the query expansion where the results were pulled from the Web<sup>11,12</sup>. Our work takes a different route, where we address the query interface based on the existing ontologies for RDF stores that are used in the Semantic Web to hold the knowledge bases. There is a noticeable growth in the efforts to build Arabic ontologies. A good example is the Arabic Ontology project<sup>13</sup>, which

Download English Version:

<https://daneshyari.com/en/article/485408>

Download Persian Version:

<https://daneshyari.com/article/485408>

[Daneshyari.com](https://daneshyari.com)