Computer Engineering and Intelligent Systems
ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online)
Vol.9, No.7, 2018

Semantic Retrieval System for Nigerian Onion Domain Ontology

Aliyu Rufai Yauri Mustapha Abdullahi
Department of Computer Science, Kebbi State University of Science and Technology, Nigeria

Abstract
With the recent Nigerian Government's effort to diversify its economy, several efforts have been put in place by the Government to revitalize Nigerian Agricultural sector as part of the diversification effort. Part of effort put in place to revitalize Agriculture it to encourage local and international investment into Agriculture in the country. However, due to lack of enough information on the agricultural activities in the country remains a challenge in the revitalization of the country's agriculture. Although Google and other search engines provide access to data on the Web, these search engines are base on traditional keyword search and therefore lack semantics which affect the effectiveness of the retrieved results. This paper presents semantic document retrieval of Nigerian Onion domain. The work involves developing Ontology for Nigerian Onion domain. Protégées Ontology editor is then used to store the Ontology and OWL-DL query is used to retrieve the documents semantically. To measure the effectiveness of the proposed work, precision and recall evaluation metrics were used.
Keywords: Ontology, Onion, Information Retrieval, Semantic Retrieval, Domain

1. Introduction

On daily basis huge amount of heterogeneous data are deposited on the Web at exponential rate. This has led to the increase in the quest for effective information retrieval systems. Information retrieval systems such as Google, Yahoo has played a major role on the access of data on the Web. However these retrieval systems are based on traditional keyword search which lacks semantics in the query processing and retrieval process and as a result, allot of irrelevant results are returned. To overcome the shortcomings of the traditional keyword search systems, the concept of the Semantic Web was introduced by the W3C consortium. In semantic web approach, data is given a well-defined format that models the meaning of information on the Web, as well as applications and services, so as to discover, annotate, process and publish data that is encoded in them (Zou, Finin, & Chen, 2004.). Data in semantic web is represented in ontology triple representation (RDF) format.

In simple terms ontology can be seen as objects that may exist in a particular domain and the relationships that may exist between those objects. Data represented ontology representation (RDF triple) format is stored in the knowledgebase in order to facilitate manipulation and querying semantically (Ciccarese, Ocana, Garcia, Sudeshna & Clerk , 2011). Therefore in order to build a knowledgebase for a particular domain, ontology for that such domain needs to be built.

Although several works has been done on the retrieval of documents semantically, semantic document retrieval systems in Nigeria are still at its infant stage. In fact to the best of our review, no work exists currently on semantic document retrieval in Nigeria Agriculture domain. In this work, semantic document retrieval for Nigerian Agricultural domain is presented. The work involves developing Ontology for Nigerian Onion domain, in particular Onion. The developed ontology is then used to develop semantic document retrieval systems. The research makes of the developed Ontology to develop a semantic document retrieval system that enables retrieval of Onion data semantically. Nigeria in one of the largest producers of Onion, therefore provide access to Onion information semantically will not only provide information about onion but, also attract foreign investors.

2.0 Literature Review

The main building block of the semantic web is ontology, which transforms web content into a machine-readable format that can be manipulated (Ahmed & Gerhard, 2007). Ontology is the main building block of the semantic web which transforms web content into a machine-readable and format that can be manipulated (Ahmed & Gerhard, 2007). Ontology, in other words Web Ontology

Language (OWL), is commonly defined as formal, and explicit specifications of shared conceptualization. Formal signifies ontology as a machine-readable format. Whereas, the concepts or entities used are explicitly described, shared, and displayed, ontology is concept that captures knowledge in a widely acceptable standard, and its conceptualization reflects ontology as a notion that identifies entities in the real world (Hu, 2004). In other words ontology can simply be seen as the study of entities that exist in the real world, and the things they have in common (Lawson, 2004). Ontology facilitates standards for integrating and sharing data in a conceptual schema. Objects, entities or concepts are identified and annotated with the relationships that exist between them.
In the concept of ontology, an entity or object is referred to as the same thing. This research will be using 'concept' to denote an entity or object, while 'relationship' is seen as the things concepts have in common, known as properties. Properties can be classified into object properties and data properties. Object properties represent the semantic relationship between concepts, while data properties define the relationship between a

concept and its literals. Annotation of concepts enables better descriptions of the concepts in the form of metadata, facilitating greater meaning for human and machines to easily process and share.

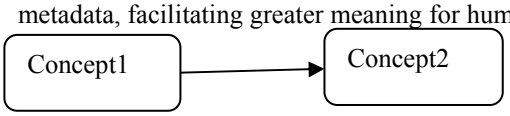

Figure 1: Example of object property

Figure 1 is a graphical representation of an object property, where semantic mapping between two concepts (Concept1 and Concept2) is provided. The semantic mapping gives a better description of the concepts.

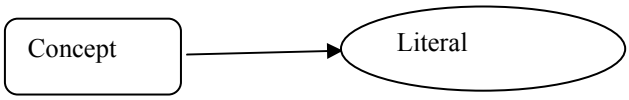

Figure 2: Example of data property

Figure 2 is a graphical representation of data property, where a concept is mapped with its literal. A literal is a mechanism for describing a concept itself. It gives an additional description of a concept. Figure 3 denotes an example of ontology representation for students.

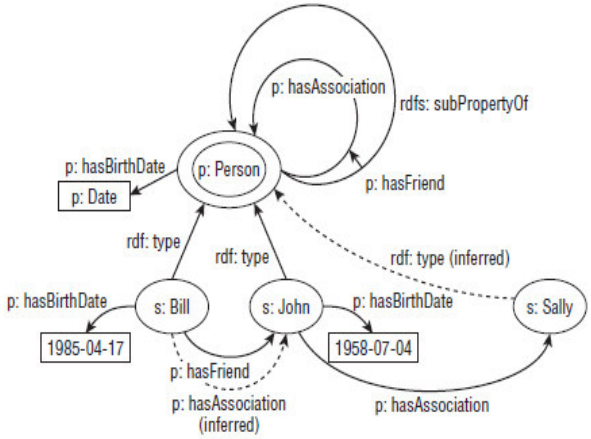

Figure 3: Example of student ontology representation

Ontology can be created automatically, semi-automatically, or manually (Erdmann, Maedche, Schnurr, & Staab, 2000). Automatic creation of an ontology involves using an automated tool to automatically generate the ontology from a domain (Balakrishna & Srikanth, 2008). Semi-automatic ontology creation involves a combination of human effort and automated tools (Balakrishna, Moldovan, Tatu, & Olteanu, 2013). Manual ontology is usually complex and time-consuming especially when dealing with a great deal of data (Ahmed & Gerhard, 2010). Manual ontology creation involves the design and creation of an ontology completely by a human expert (Tao, Embley, & Liddle, 2009).

Semantically annotated ontology are stored in the knowledgebase where a semantically formulated query is used to retrieve the relevant information. Computers can retrieve more relevant documents by focusing on the underlying meanings in queries.

There are several works done on semantic technology which comprises of ontology development and semantic information retrieval in recent Years. However, in Nigeria Onion domain, to the best of our knowledge, we identify one research that focuses on ontology in Nigeria agriculture which a work in (Syed Malek F. D Syed Mustapha and Emmanuel Ukpe, 2013). In their work they reported ontology development methodology in Nigeria Onion domain but, their work fails to clearly describe the methodology and shows prove that the ontology has been developed and evaluated. Therefore the effectiveness of their methodology cannot be guaranteed.

The primary objective of this study is to develop ontology Nigeria Onion domain and use the ontology to create semantic information retrieval systems in Agricultural domain. The propose methodology will be evaluated using precision and recall metrics.

3.0 Proposed Semantic Document Retrieval for Nigerian Agricultural Domain Ontology

This section provides a step by step procedure for the propose Semantic Document retrieval of Nigeria Agricultural domain. Figure 4 presents graphical representation of the proposed Approach.

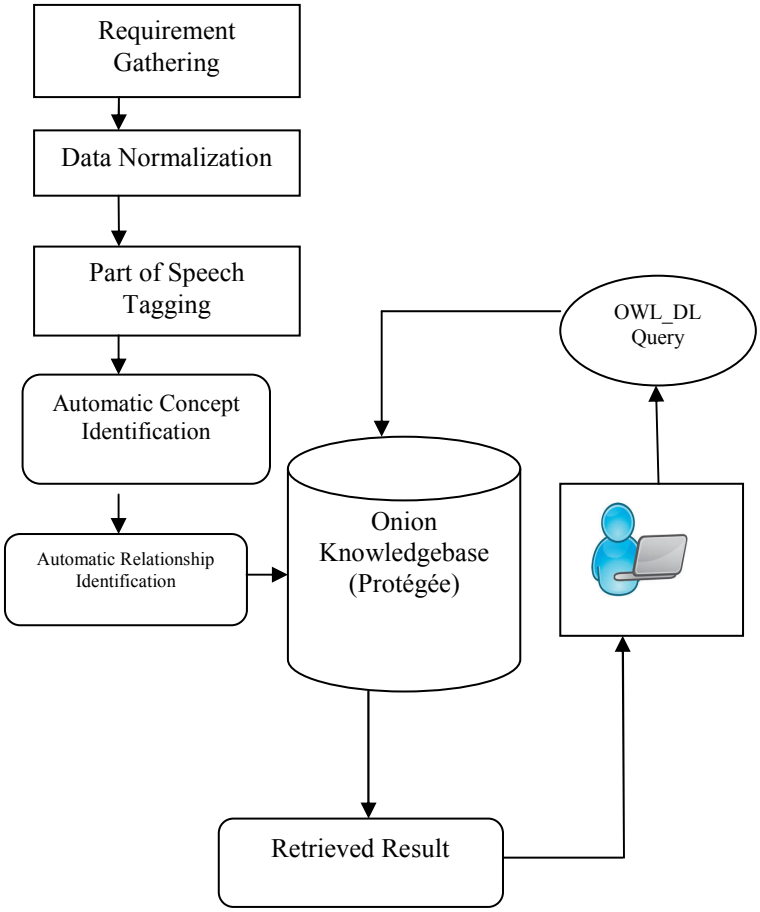

Figure 4. Frame Work of the Semantic Document Retrieval of Nigerian Onion Domain Ontology

Figure 4 shows the step by step procedures that will be involved in the implementation of this research. The research is divided into two module comprising of:

- Onion Ontology Development Module
- Semantic Retrieval of Onion Ontology Module

3.1 Onion Ontology Development Module

This section is concern with the development of Onion ontology which will serve a backend for the semantic retrieval part. This module involves engaging expert of agriculture domain to get the necessary data. Then normalization of the data, then tagged the document with part of speech using Stanford part of speech tagger. Then automatic concepts identification is carried out. Again agriculture expert are engage to identify possible relationships between the earlier identified concepts to form triple (subject, predicate, object).

3.1 Semantic Document Retrieval of Onion Ontology

In this section, the developed onion ontology is stored in a knowledgebase. Knowledgebase is a storage location for semantically structured data such as ontology. The knowledgebase to be used in this research is protégée. Prototogee is an ontology storage and retrieval tool that enables the storage and retrieval of ontology concepts. The research used protégée built in reasoner which is used to access an external DIG compliant reasoner, thereby enabling inferences to be made about the classes and individuals in the ontology. The DL Query tab in protégée provides a good feature for searching a classified ontology. It is a standard Protégé plug-in. The query language that is supported by the plug-in is based on the Manchester OWL syntax. The Manchester OWL Syntax is a new syntax that has been designed for writing OWL class expressions (Ontology of Quran Concept). It is influenced by both the OWL Abstract Syntax and the DL style syntax, which uses description logic symbols, such as the universal quantifier (∀) or the existential quantifier (∃). A quantifier is “an operator that limits the variables of a proposition (Aliyu Rufai y, Rabiah A. Azreen A, Masrsah A.). DL query is used to retrieve relevant document semantically.

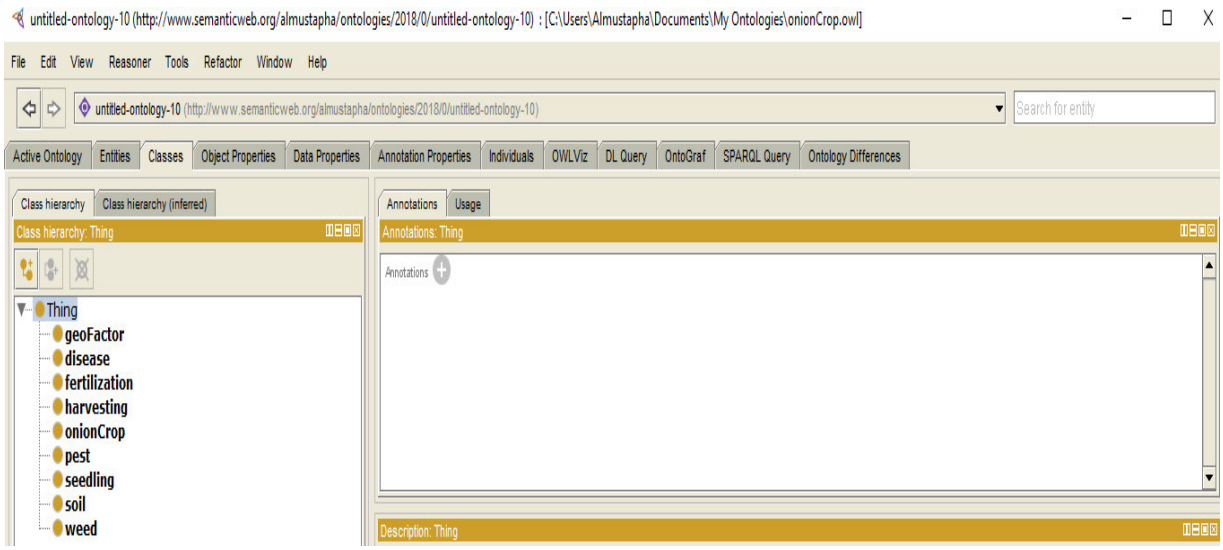

Figure 5. Onion Ontology in Protégée

With store Onion ontology, user can query the knowledgebase semantically and have a more precise answer. For example user may ask, what are the diseases of Onion crop?. All the diseases of Onion crops stored in the knowledgebase are then returned to the user.

4.0 Evaluation

The effectiveness of the proposed semantic document retrieval of Nigerian Onion Domain was measured using precision and recall. Precision and recall are popular methods used for evaluating document retrieval approaches in terms of the relevancy of the retrieved information. Precision is used to measure retrieved documents that were relevant, as seen in Equation 1, while recall measures the percentage of relevant documents retrieved, as seen in Equation 2.

$$\text{Pr}\,ecision = \left| \frac{\{relevant\ \ documents\} \cap \{retrieved\ \ documents\}}{|\{retrieved\ \ documents\}|} \right| \quad 1$$

$$\text{Re}\,call = \left| \frac{\{relevant\ \ documents\} \cap \{retrieved\ \ documents\}}{|\{relevant\ \ documents\}|} \right| \quad 2$$

Table 1. Results of Precision and Recall

Precision	95
Recall	100

Table 1 show that 95% of the retrieved documents were relevant and 100% of the entire relevant documents to the query were returned. The shows that proposed method has increase the chances of getting more efficient and effective result using semantic retrieval instead of the current traditional keyword search systems such as Google and Yahoo.

5.0 Conclusion

The research presents semantic document retrieval for Nigerian Onion domain. The research involves ontology development that combines the use of computer and human expert to identify concepts and relationships in Nigeria Onion domain. The developed ontology is then used to develop a knowledgebase for semantic retrieval. The research was evaluated using precision and recall metrics to show the effectiveness of the proposed methodology.

6. 0 References

1. Zou, Y., Finin, T., & Chen H. (2004). F-OWL: An inference engine for the semantic web, Proceedings of the Third International Workshop (FAABS), April 16–18, 2004, USA.
2. Ciccarese, P., Ocana, M., Garcia Castro, L. J., Das, S., & Clark, T. (2011). An open annotation ontology for science on web 3.0, Journal of Biomedical Semantics, 2041-1480-2-S2-S4.
3. Ahmed, Z., & Gerhard, D. (2007). Web to Semantic Web & Role of Ontology, In the proceedings of

National Conference on Information and Communication Technologies, (NCICT-2007), Pakistan, 9th May 2007.

4. Lawson, T. (2004) A Conception of Ontology, University of Cambridge, Faculty of Economics, Cambridge.
5. M. Erdmann, A. Maedche, H. Schnurr, and S. Staab. (2000). From manual to semi-automatic semantic annotation: About ontology-based text annotation tools. In, Proceedings of the COLING 2000 Workshop on Semantic Annotation and Intelligent Content, August 2000.
6. Balakrishna, M., & Srikanth, M. (2008). Automatic Ontology Creation from Text for National Intelligence Priorities Framework (NIPF).In Proceeding of 3rd International conference for Ontology for the Intelligent Community (OIC), Fairfax, VA, USA, Dec 3-4.
7. Tao, C., Embley, D. W., & Liddle, S. W. (2009). FOCIH : Form-based Ontology Creation and Information Harvesting. In Proceedings of the 28th International Conference on Conceptual Modeling , Gramado, Brazil.
8. Syed Malek F. D Syed Mustapha and Emmanuel Ukpe, (2013). Agriculture Ontology for Sustainable Development in Nigeria. Journal of Advances in Computing 2013, 3(3): 57-59 DOI: 10.5923/j.ac.20130303.04
9. Aliyu Rufai Yauri, Rabiah Abdul Kadir, Azreen Azman and Masrah Azrifah Azmi Murad(2013) Quranic Verse Extraction base on Concepts using OWL-DL Ontology. Research Journal of Applied Sciences, Engineering and Technology 6(23): 4492-4498, 2013
10. Jean Vincent, Fonou-Dombeu, & Magda Huisman, (2011). Combining Ontology Development Methodologies and Semantic Web Platforms for E-government Domain Ontology Development. International Journal of Web & Semantic Technology (IJWesT) Vol.2, No.2, April 2011