

Querying the Dataset from the Developed Ontology for Swineflu Disease

Radhika Pathi¹, Suresh Verma Penumatsha², Lakshmi Kalyani Neerukonda³, P.Rama Krishna⁴

¹Research Scholar, Rayalaseema University; India; ²Professor, Adikavi Nannaya University; India; ³Assistant Professor, VNR VJJET; India; ⁴Assistant Professor, VNR VJJET; India

Abstract

Ontology design and querying the data set to obtain accurate results has been a topic of research. The aim of the paper is to represent ontological knowledge in the field of Medical Information Systems to store the knowledge about Swine flu disease and thereby querying the data and formalize the knowledge base development. The developed knowledge from RDF is converted to rules and then querying is done through different methods and performance analysis has been done on the retrieved query. The results predicted determine the suitability of the method where querying gives the accurate result for swine flu disease ontology.

Keywords: Querying; Ontology; Swine-flu; Protégé; SPARQL;

Introduction

Semantic Web a layer of existing web enables machines to operate better and share meaningful knowledge. Semantic Web provides technology stack where key part is standardized information exchange. As part of Semantic Web, Computers should have access to structured collection of documents and sets of inference rules so that they conduct automated reasoning. Semantic web can otherwise be referred as Linked Data. Aristotle defined the term Ontology^[1,8,9] in his Metaphysics. In 1998, Studer et al.^[2] defined the term Ontology as : “An ontology is a formal, explicit specification of a shared conceptualization.” Conceptualization denotes abstract model being representing some phenomenon identifying Knowledge in detail with individual relationships. In our paper , We present our application for searching on appropriate semantic information on Swine flu in respiratory disease ontology which is a domain ontology .This we have done using Classes and Objects and the inter individual relationships which gives the information to patient or any person searching the Web such as what drug is suitable for a patient given his symptoms. Explicit states that type, constraint and their use is explicitly stated. Shared echoes the perception that Ontology represents consensual knowledge refers to some group and not individual. Features are described by Classes and attributes are represented by data properties. Instances are represented by individual

classes and querying to retrieve semantic information is done with the help of SPARQL. The main purpose of the paper is to communicate information retrieval ,querying through SPARQL and make an analysis of performance of web based applications to predict the query evaluation that gives accurate results when queried. The remaining part of the paper is organized as follows .Section 2 describes about the different tools used for Semantic Web and Section3 describes about the framework for the development of swine flu ontology and section4 presents analysis of NASA TLX INDEX^[3,4] on applying SPARQL^[5], Ontograf^[6], SWRL^[7] Queries on the ontology that is retrieved and the evaluation is presented in the form of graph and Section 5 concludes our work.

Tools used for Semantic Web in our Case Study

Respiratory diseases are the diseases caused due to the disorder of the air and lungs that affect the human respiratory system. These respiratory diseases may cause damage to the organs and the other internal structures that deal with the breathing, and may include nasal cavities, the throat, the bronchitis and many other organs present. Identification of these respiratory diseases and taking the correct predicate is a difficult task for a normal individual affected. Respiratory diseases cause the malfunction of the internal organs that block and cause damage to the lungs which may cause severe health issue. There are many diseases which are caused on the respiratory

system. Some of them namely, Asthma, Pneumonia, Upper respiratory tract infection, Infant respiratory distress syndrome, Cystic fibrosis, Shortness of breath, Obstructive lung disease, Bronchitis, Obstructive sleep apnea, Influenza-like illness, Acute severe asthma. These diseases may affect the human structures and organs directly. Knowledge of these diseases to the individuals is very much less rather than the experts in that domain. To get the knowledge about the diseases is very much difficult and that may happen only when the expertise help is provided to the user. To get the preventive measurements for a user without expertise help may end up in the wrong cause. To understand the content and to provide the better knowledge about the diseases ontologies provide an accurate representation to the user without the help of the expertise in that domain.

There may be various characteristics and different issues related to the diseases. Each of them are classified into the specific categories in the domain those are affecting different organs or the structures. Sufficient knowledge about the diseases and the corresponding precautions or about the diseases through the ontologies can help in removal of unnecessary damage and less spread of the wrong information through the web. Technologies that help to build a content oriented system is possible using semantic web^[10,11]. By using semantic web which provide information using the ontology help to show the related issues in the respiratory diseases, how it is caused and how it can be prevented. Much of the related information can be seen on the web rather than having the misguided information.

Methodology

Tools for Ontology Development

Protégé^[12,13] - Protégé is developed by Stanford University as a tool for Semantic Web and it is a framework for building Intelligent Web Systems. It can be adapted to build both simple and complex ontological applications where the concepts used are classes, objects and the inter individual relationships established among them. Protégé provides us meta modeling an the best part

is that there is no syntax involved and domain ontologies can be developed easily with this.

Jena – Apache Jena is open source framework based on Java. It provides framework to extract data from and write to RDF and graph which is represented as abstract model. Data can also be populated from databases, Files, URLs and also it has the API for creating and accessing RDF data.

RDF – A Resource Description Framework, is used as model for interchange of data on the Web. URIS are assigned to data objects and then RDF is used to create statements about them. RDF is used for serializing triples and RDF consists of Subject-Predicate-Object triple. This makes data easier to store and interchange.

SPARQL – An Acronym for SPARQL Protocol and RDF Query Language is used for querying, retrieving and manipulating data stored in RDF format. It allows us to search the Web and discover relationships among the classes and individual objects. SPAQRL takes text in the form of queries and return the results. SPARQL consists of SELECT queries which are similar to SQL thereby it is easy for the user to write if he has knowledge on the SQL queries. SPARQL produces results in three formats XML,JSON and CSV formats.

Ontograf – Ontograf provides support for navigating relationships of OWL ontologies. The various relationships that can be supported are subclass, individual, domain/range object properties, and equivalence. It provides zooming, spring layout, searchable relationship filters, configurable tooltips, pinning tooltips, OWL Imports View for better enhancement of views.

SWRL – SWRL rules can be executed through a tab in protégé SWRL Tab. It supports interoperating with other rule engines and there are other user-defined libraries and various built-in libraries are also provided mathematical, temporal and string operators. This language provide a powerful means extracting OWL Ontologies.

Framework For Swine Flu Disease Ontology

A framework for Swine Flu disease ontology is developed where ontology is created with the help of Protégé. Protégé helped in modelling of ontology classes. Classes used for development of Swine Flu disease are demonstrated in Fig1 .

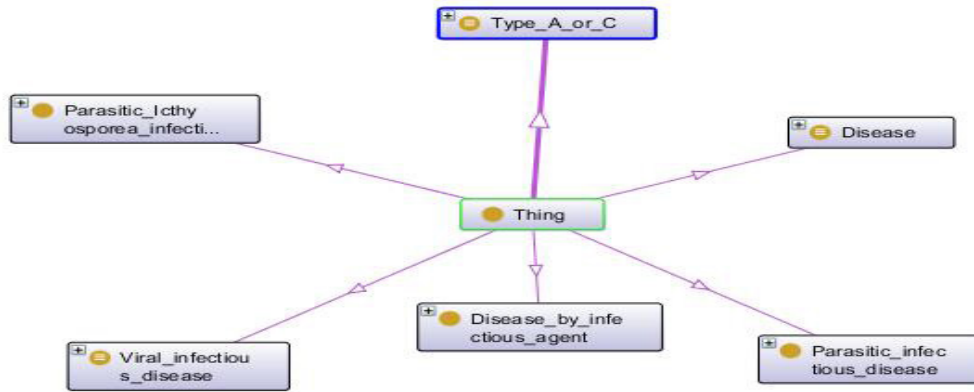


Fig1 : Display of Term Candidates in an Unstructured List

Classes used are highest occurrence of Disease and its diagnosis, Symptoms of individuals, Causes of disease, Its prevention and Vaccines, Etiology. Relationships used are Named_Disease, Caused_By, Used_to_diagnose, Types, Has_severity, Managed_To, Disease_Name, Managed_by, Used_to_cure, Inter individual relationships are represented in the following Fig2.

diseases.

After SwineFlu disease ontology is developed, the next step is we evaluated our ontology using a Fact++ Reasoner which verified the ontology by executing each class and their properties.

Results & Discussion

Performance Evaluation

Our paper deals with SwineFlu ontology to disseminate knowledge to the user regarding SwineFlu the symptoms risk factors vaccines drugs Type_A/ Type_c. we conducted one experiment to gather information regarding SwineFlu on two sets of users. One set of users has awareness on SwineFlu and the other does not have any awareness on SwineFlu. Taxonomy on SwineFlu ontology was shown in fig 1 where it comprised of classes, objects and instances. Various queries are posed on ontology and performance was evaluated as such which query gives the better performance. The type of queries used here are SPARQL Ontograp and SWRL.

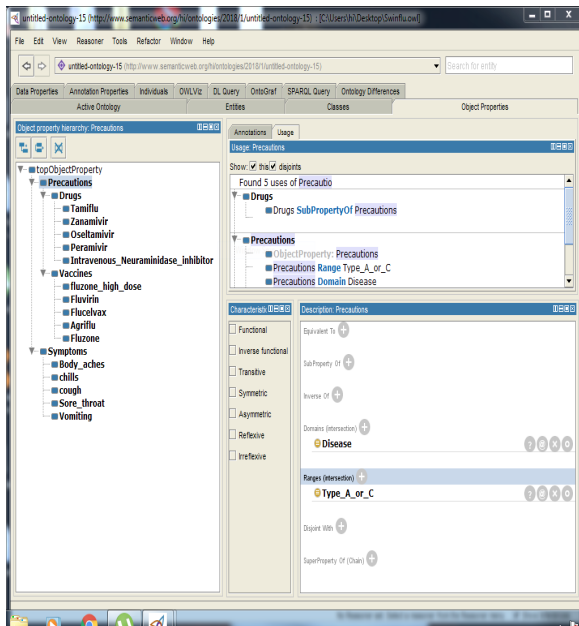


Fig 2: Snapshot of inter individual relations in swine flu

1. User is asked to find the term “Fever” from Disease ontology

```
SPARQL query:
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX ab: <http://semanticweb.org/chandana/ontologies/2018/2/untitled-ontology-54#>
PREFIX ns: <http://semanticweb.org/chandana/ontologies/2018/2/untitled-ontology-54/swineflu#>
SELECT ?subject
WHERE {
?subject ?predicate ?object FILTER
regex(str(?subject), "Fever", "i")
} order by ?subject
```

Fig 3: Query 1

Fig 3: Query 1

2. User is given the SwineFlu details and is asked to find the drug which is used to cure “Cough”.

```
SPARQL query:
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX ab: <http://semanticweb.org/chandana/ontologies/2018/2/untitled-ontology-54#>
PREFIX ns: <http://semanticweb.org/chandana/ontologies/2018/2/untitled-ontology-54/swineflu#>
SELECT DISTINCT ?object
WHERE {
?subject ?predicate ?object
FILTER regex(str(?predicate), "curedBy", "i")
FILTER regex(str(?subject), "Cough", "i")
} order by ?subject
```

Fig 4: Query 2

3. Query for the user about the year of attack of disease like “In which year particular patient affected with swineflu disease”.

```
SPARQL query:
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX ab: <http://semanticweb.org/chandana/ontologies/2018/2/untitled-ontology-54#>
PREFIX ns: <http://semanticweb.org/chandana/ontologies/2018/2/untitled-ontology-54/swineflu#>
SELECT DISTINCT*
WHERE {
?subject ab:PatientID=50.
?subject ?predicate ns:hasSymptom swineflu
}
```

Fig 5: Query 4

Different users tried to find out the information regarding swineflu disease on the basis of results retrieved by SPARQL, Ontograf, SWRL. There after NASA TLX(Task Load Index) As been used for query analysis and time need to perform execution.

A comparative study on the results obtained by using different queries is done from this study. The SPARQL query has given higher efficiency and performance in retrieving the results when compared to other tools like ontograf, SWAL Query. This comparison states that SPARQL is far better suited than other tools in giving optimal results.

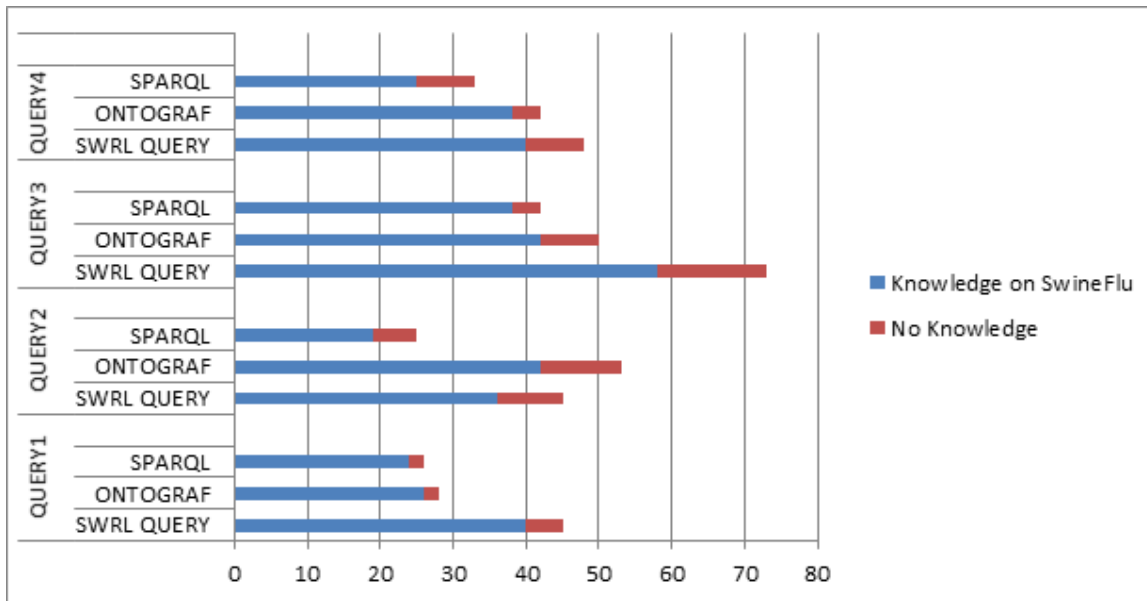


Fig 6. SPARQL, Ontograp, SWRL Query TLX and Time Comparison.

Conclusion

This paper gives knowledge on the creation of domain ontology for SwineFlu Disease and the queries are done using SPARQL. This gives performance analysis on ontological evaluation based on ontological tools. The analysis results state that SPARQL method is efficient method and it outperforms the other querying methods. Semantic web being intelligent web is used for developing ontology for swineflu that is useful for medical information systems that provides a linked data for the benefit of health care systems. This can be extended with machine learning to train the patterns so that ontology for other respiratory diseases can be developed in future.

Ethical Clearance- This research was approved by the Research and Ethics Committee of the ESI Hospital, Pulmonology center, Chest Diseases, ESI, Hyderabad,India.

Source of Funding- This study was not funded by any Funding Agency.

Conflict of Interest - None declared

References

1. Gruber TR. A Translation Approach to Portable Ontology Specifications. Knowledge Acquisition, 1993;6(2):199–221.
2. Studer R, Benjamins R, and Fensel D. Knowledge engineering: Principles and methods. Data &

- Knowledge Engineering, 1998;25(1–2):161–198.
3. Hart, Sandra G. NASA Task Load Index (NASA-TLX). v. 1.0. Paper and pencil package (Instruction manual):NASA Ames Research Center; 1986 Jan. 26p. Report No. 2000-002-1488.
4. Sandra GH, Lowell ES, Development of NASA-TLX(Task Load Index): Results of empirical and theoretical Research. Hum-ment Work load 1, Amsterdam: North Holland Press. 1988;239-250.
5. Perez J, Arenas M, and Gutierrez C. Semantics and Complexity of SPARQL. ACM Transactions on Database Systems, 2009; 34(3)
6. Swaminathan V, Sivakumar R. A comparative study of recent ontology visualization tools with a case of diabetes data. International Journal of Research in Computer Science, 2012; 2 (3):31-36.
7. Connor OM, Tu S, Nyulas C, Das A, Musen M. Querying the Semantic Web with SWRL. Advances in Rule Interchange and Applications,2007; 4824:155–159.
8. Keßler C, Raubal M, and Wosniok C. Semantic Rules for Context-Aware Geographical Information Retrieval. Springer-Verlag Berlin, 2009; 5741: 77–92.
9. Katifori A, Halatsis C, Lepouras G, Vassilakis C, and Giannopoulou G. Ontology visualization methods – a survey. ACM Computer Surveys,2007;39(4).
10. Berners-Lee T, Hendler JA, and Lassila O. The semantic web. Scientific American,

2001;284(5):34-43.

11. Wang TD and Parsia B. CropCircles: Topology sensitive visualization of OWL class hierarchies. *The Semantic Web – ISWC,2006*; 4273:695-708.
12. Gennra JH , Musen MA. The evolution of Protégé: an environment for knowledge-based systems development. *International Journal of Human-Computer Studies*, 2003;58(1):89-123.
13. Knublauch H, Fergerson RW, Noy NF, and Musen MA. The Protégé OWL Plugin: An Open Development Environment for Semantic Web Applications. *Third International Semantic Web Conference*, 2004; 7-11.