Volume 2, Issue 6 June 2016 ISSN NO: 2454 - 7514

North Asian International Research Journal Consortium

North Asian International Research Journal

Of

Science, Engineering and Information Technology

IRJIF IMPACT FACTOR: 3.01

Welcome to NAIRJC

ISSN NO: 2454 -7514

North Asian International Research Journal of Science, Engineering & Information Technology is a research journal, published monthly in English, Hindi. All research papers submitted to the journal will be double-blind peer reviewed referred by members of the editorial board. Readers will include investigator in Universities, Research Institutes Government and Industry with research interest in the general subjects

Editorial Board

Address: - Dr. Ashak Hussain Malik House No. 221 Gangoo, Pulwama, Jammu and Kashmir, India - 192301, Cell: 09086405302, 09906662570, Ph. No: 01933-212815,
Email: nairjc5@gmail.com, nairjc@nairjc.com, info@nairjc.com Website: www.nairjc.com

IRJIF IMPACT FACTOR: 3.01

SPECIFIC RANKING USING SEMANTIC WEB FOR INCREASING THE EFFICIENCY OF THE WEB CRAWLER

MOHIUDDIN ANSARI & ASST. PROF. MR. YATIN CHOPRA

CBS Group of Institutions, Maharshi Dayanand University, Haryana

ABSTRACT:

A Web Crawler is an internet bot that downloads data from World Wide Web for search engine and Indexing. Web information is constantly changing and is updated without prior notice. The Web Crawler checks the World Wide Web for the updated information. People visiting the website frequently are actually the familiar websites and this makes it high ranked website. In this paper the network traffic solution is used to get desired information. This paper will be actualize Ontology Based Topic Specific Search Using Semantic Web. The strategy for web crawling with filters is utilized. It is a query based approach with Jena API. The proposed approach takes care of the issue of re-visiting web pages by crawler. The Semantic Web is an extended version of the present Web that permits the meaning of data to be decisively described regarding all around characterized vocabularies that are comprehended by individuals and computers. As Topic based pursuit is an inquiry interface paradigm taking into account a long running library tradition of faceted classification. Furthermore, effective search systems frameworks have proven that they are both capable and instinctive for end – users, especially in drafting complex queries. In this way, topic – based search shows a promising path for semantic searching interface design only if it gets successfully combined with Semantic Web Technologies. Topic based web search-engines is different from other search engines like Google, MSN/BING, Yahoo! as it only integrates information, indexes it and answers the queries of user.

Keywords: ***Semantic Web, Web Crawler, World Wide Web, Information.***

1. INTRODUCTION

Keyword or Topic-Based-Search is helpful particularly to a client who realizes what keywords are utilized to list the image and in this way can do the formulation easily and quickly. This methodology is problematic, in any case, when the client does not have a clear objective at the top of the priority list, does not comprehend what there is in the database, and what sort of semantic concepts are included in the area. The target is to make a Topic-

Based-Semantic web search tool which is highly user-friendly that helps in providing advanced search options with topics. A client shouldn't know about the concepts supporting the semantic web to utilize it. The user must feel that whatever they are searching on this semantic web right now is similar to what they do normally with search engines daily.

Topic based search is an interface based on a running library based on traditional faceted classification. An efficient search system has proved that it powerful and intuitive for end users in drafting any queries. Thus the topic based search engine will use semantic web documents like RDF, XML, and OWL instead of human readable documents like PDF, DOC or HTML. The main part of this topic based engine is to describe query in the form of topic description. Let us say that a user gives a query "Give me the names of student studying in MDU University and are below 25". The queries are not made using natural language but with a very easy to use interface that help users to build the target queries.

OBJECTIVES

The objective of the work is about providing information about the hotel domain. The objectives are as follows:

- Using SPARQL providing hotel relevant information.
- Hotel Ranking
- Representation of hotels using Knowledge-Base

SPARQL provides better performance than SQL. Hence SPARQL is used. Here Hotel Ontology is developed using RDF since RDF is a scheme language as it has pointer at the top of the document in which RDF scheme is being used. Anyone can easily make a new scheme document. General representation is for Knowledge Representation as in RDF, a document assumes that specific things have properties that has values.

2. LITERATURE REVIEW

Gateway of Semantic Web is to provide access to ontologies and Sematic Data. Hence it has 3 main roles:

- Collection of Semantic Content from web that are available.
- Extract metadata and index which can be useful
- Implementing the query access to the data.

Raman Kumar, Goyal Vikas Gupta, Vipul Sharma, Pradeep Mittal in June 2015 described **"Tourism Ontology and Semantic Management System: State of the Arts Analysis"** stating the global importance of steadily rising tourism creating a new opportunity in many countries. The solution for information management for complex tasks of tourism are still at an early stage from a semantic point of view. This paper aims applying, evaluating and concretizing semantic web technologies such as ontologies, semantic search to information rich tourism domain and semantic annotation of contents. Identification of seven tourism ontologies are suitable as a base for creating problem specific ontologies.

P. Jourlin, R. Deveaud, E. Sanjuan-Ibekwe, in the year 2012 has described a web crawler based on GNU/Linux and Postgre SQL which is a novel, focusable, scalable and distributed web crawler. They have released a GNU public license. The report shows the use case related to the analysis of Twitter's stream about the presidential election of French in 2012 in its URL.

SA Patel and JM Patel in the year 2012 has described web crawler as an intelligent agent in which pages available in internet are growing tremendously day by day and in this case searching the required information in the internet becomes very hard work. Generally, the searching of information in WWW can be done by searching the list of the link.

SS Vishwakarma, A Jain, in the year 2012, described a **web crawler algorithm with query based approach with increasing efficiency**. In his paper the network traffic solution is proposed in which approximately 40% of the web traffic is by web crawler and the web crawling filter is used which is query based approach. The approach actually solves the problem of revisiting web pages by a web crawler.

Marc Najork has described Web Crawler Architecture in which a web crawler, being an important component of search engines, is given one or more seed URLs, download the web pages associated with that URLs. It also extracts the hyperlinks and recursively continues to download the pages by that identified hyperlinks. They are also used in many applications like data mining, shopping engines, indexing of web pages within search engines and so on. Their two main data structures, one which is set to be crawled URLs and second the content providers without overloading any particular web server.

A. Agarwal, D. Singh, A. Kedia, A. Pandey, V Goelin the year 2012 described **a design of a parallels migrating web crawler** that poses certain drawbacks such as generations of large amount of redundant data and

wastage of bandwidth of the network due to transmission of such unwanted data. In order to overcome these drawbacks with traditional crawler techniques, they have proposed a parallel migrating web crawler and presented with detailed requirements along with the Crawler Architecture.

N. Singhal, A. Dixit, RP Aggarwal in the year 2012 described **"Regulating Frequency of Migrating Web Crawler based on Users Internet"** shows that due to the lack of efficient techniques, the crawlers add up unwanted traffic to already overloaded Internet. Optimization of frequency of visiting sites can be done by calculating the refresh time. This helps in optimizing the effectiveness of the crawling system by managing the revisiting frequency. An alternate approach for the optimization of frequency of visiting websites can be done based on user's interest.

RK Rana and N. Tyagi in the year 2012 described **"A Novel Architecture of Ontology based Semantic Web Crawler"**. They presented that searching meaningful information from billions of resources of data on web is difficult task by looking at the growing popularity of internet. The future of WWW is semantic web where ontologism are used to give a valid meaning to the web content. On Web Semantic the data will be linked to different ontologies and processing of information becomes very hard without proper knowledge of sematic mappings between different ontologies. The Architecture can exploit the semantic Meta data to discover and take out information from semantic web.

D. Khurana and S. Kumar described **"An Improved Approach for Captain Based Image Web Crawler"** in which the WWW is a global, read write information space. All the information in the form of text documents, images, multimedia and other sort of information are targeted to be resources by short and unique identifiers knows as Uniform Resource Identifiers so that it can be easily found and accessed in the simplest way. It becomes a very big reservoir of information providing unrestricted access to a Hugh unending reservoir of information present in the form of hypertext markup language. These documents contains hyperlinks to each other documents.

Web Crawler Design Issues: A Review present that the big and dynamic nature of web increases the requirementforupdating for web based information for retrieving desired information from the system. Crawlers encourage the procedure by taking after the hyperlinks in Web pages to automatically download a partial preview of the Web. While a few systems depend on crawlers that comprehensively crawl the Web, others concentrate on subject particular accumulations. In present paper the different sorts of crawlers are talked about. The paper likewise talks about a few web crawler design issues alongside their answers.

3. ONTOLOGY

An ontology is a model of the global world, displayed as a tangled tree of associated concepts. Concepts are language-independent abstract entities, not just words. These are portrayed in this ontology using English content only as a simplifying convention. Whereas machines wouldn't care if concepts were referenced by, say, numbers - to make sure they are language independent - such a naming convention would make the ontology completely opaque to the individuals who have to create and use it. So we use English names for ideas and, both in the ontology and in every writings of the ontology, use capital letters/ words to distinguish ideas, like DOG, from words in confirmed language, like English "dog" or French "chien". The main purpose of Semantic Ontology is to make use of automated text-processing by using knowledge based representation of abstracts of the world. Ontology displays how concepts are related like DOG and CAT are somewhat closely related as they are MAMALs and properties of each has TAIL, FUR, but CAT could be the AGENT of HISSING where DOG could be AGENT of BARKING. We can take an example of TABLE referring a horizontal flat surface with 4 legs and being made of WOOD/METAL and is located at BUILDING/ROOM and so on.One sense of a word "table" is being linked to this concept. However in day-to-day meaning of "table" has different meanings as well. One is arrangement of data in a particular order where as other meaning of table is physical rectangular form of word. Thus giving different unambiguous concept of the same word "table". We understand the meaning of the word "table" on context form: say if you put a cup of tea on table, the chances are the "table" becomes the instance of TABLE where as if you give related information in table, the "table" becomes the instance of "CONCEPT". While making ontology, we tend to record information about things and typical concepts that helps o disambiguate the meaning of the words like "table".

Running Example: Hotel Ontology: Using Hotel Ontology examples using ontology API, the principles of using ontology AP is illustrated. This ontology presents a very simple model for showing the concepts & activities related to academic conference. This ontology is serialized in RDF/XML included with Jena. There are many was for writing this ontology and a verity of opinions as to what kinds of definition should be shown. In Jena, a minimal component of ontology is taken for a particular view. Below is the

Table: Jena API Packages

com.hp.hpl.jena.query.*;
com.hp.hpl.jena.query.larq.IndexBuilderString;
com.hp.hpl.jena.query.larq.IndexLARQ;
com.hp.hpl.jena.query.larq.LARQ;
com.hp.hpl.jena.rdf.model.Model;
com.hp.hpl.jena.rdf.model.ModelFactory;
com.hp.hpl.jena.sparql.util.StringUtils;
com.hp.hpl.jena.sparql.util.Utils;
com.hp.hpl.jena.util.FileManager;

4. WEB CRAWLER

A Web crawler is an Internet bot that deliberately searches the World Wide Web, ordinarily with the end goal of Web indexing. A Web crawler may likewise be known as a Web spider, a subterranean insect, a programmed indexer, or a Web scutter. Web indexes and some different destinations use Web creeping or spidering programming to redesign their web substance or lists of others locales' web content. Web crawlers can duplicate all the pages they visit for later preparing by a web search tool that indexes the downloaded pages with the goal that clients can look them a great deal all the more rapidly. Crawlers can approve hyperlinks and HTML code. They can likewise be utilized for web scratching.

Crawling arrangement: The conduct of a Web crawler is the result of a mix of policies:

- a choice approach that states which pages to download,
- a return to arrangement that states when to check for changes to the pages,
- a amiability approach that states how to abstain from over-burdening Web destinations, and
- a parallelization approach that states how to organize circulated web crawlers.

5. WEB SEARCH ENGINES

A web index is a product framework that is intended to look for data on the World Wide Web. The indexed lists are for the most part introduced in a line of results frequently alluded to as internet searcher results pages (SERPs). The data might be a blend of website pages, pictures, and different sorts of records. Some internet searchers additionally mine information accessible in databases or open catalogs. Not at all like web indexes, which are kept up just by human editors, internet searchers additionally keep up constant data by running a calculation on a web crawler.

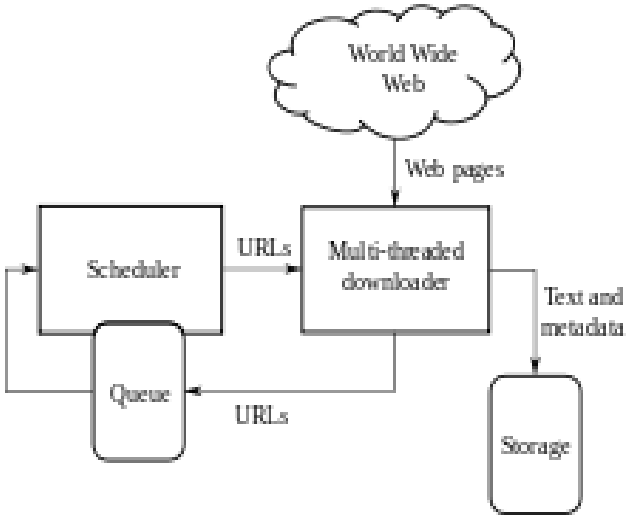
Working of Search Engine: A web crawler works in the accompanying request:

1. Web slithering
2. Indexing
3. Searching

Web crawlers work by putting away data about numerous site pages, which they recover from the HTML markup of the pages. These pages are recovered by a Web crawler (in some cases otherwise called an arachnid) — a computerized Web crawler which takes after each connection on the website. The site proprietor can prohibit particular pages by utilizing robots.txt.

The web crawler then examines the substance of every page to decide how it ought to be filed (for instance, words can be removed from the titles, page substance, headings, or unique fields called meta labels). Information about website pages are put away in a record database for use in later questions. An inquiry from a client can be a solitary word. The file discovers data identifying with the inquiry as fast as possible. Some web crawlers, for example, Google, store all or part of the source page (alluded to as a reserve) and additionally data about the site pages, though others, for example, AltaVista, store each expression of each page they find. This reserved page dependably holds the real pursuit content since the one was really ordered, so it can be extremely helpful when the substance of the ebb and flow page has been redesigned and the hunt terms are no more in it. This issue may be viewed as a gentle type of linkrot, and Google's treatment of it builds ease of use by fulfilling client desires that the inquiry terms will be on the returned website page. This fulfills the rule of minimum amazement, since the client typically expects that the inquiry terms will be on the returned pages. Expanded hunt importance makes

these reserved pages exceptionally valuable as they may contain information that may never again be accessible somewhere else.



High-level architecture of a standard Web crawler

At the point when a client enters a question into an internet searcher (regularly by utilizing catchphrases), the motor looks at its list and gives a posting of best-coordinating website pages as indicated by its criteria, as a rule with a short rundown containing the report's title and in some cases parts of the content. The file is worked from the data put away with the information and the strategy by which the data is indexed. From 2007 the Google.com web crawler has permitted one to look by date by clicking "Show seek devices" in the furthest left section of the underlying list items page, and afterward selecting the sought date range. Most web search tools bolster the utilization of the Boolean administrators AND, OR and NOT to assist determine the hunt inquiry. Boolean administrators are for exacting inquiries that permit the client to refine and amplify the terms of the hunt. The motor searches for the words or expressions precisely as entered. Some web indexes give a propelled highlight called vicinity look, which permits clients to characterize the separation between keywords. There is likewise idea based seeking where the exploration includes utilizing measurable examination on pages containing the words or expressions you hunt down. Too, common dialect inquiries permit the client to sort an inquiry in the same structure one would ask it to a human. A site like this would be ask.com.

6. PROPOSED WORK

Here in this proposed framework we plan and build up a semantic web engineering that can ease the clients from the overburden of doing a ton of watchword based inquiry before getting the relevant result. This framework

takes the normal language client question as subjects in an easy to use environment. In this proposed framework the subject based web interface must be produced for finishing the objective of semantic scanning in a profoundly heterogeneous semantic web environment. The proposed framework shows the refinement with higher exactness and in programmed way. Dissimilar to the current frameworks in this examination work to manage ontology a typical interface must be planned.

The proposed framework engineering must be divided into three distinct modules: First module takes the client inquiry as subject depiction and the second module gives a method to the client to upgrade list items by giving different image filtering mechanism. The third module is to display the results according to ranks. Be that as it may, for the framework to comprehend client inquiry and give best result, it ought to be prepared first. RDF document structure utilized as a part of the work.

SPARQL Query used

```
@prefix  :<http://example/> .
@prefix  xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix  dc:        <http://purl.org/dc/elements/1.1/> .
@prefix  rdfs:   <http://www.w3.org/2000/01/rdf-schema#> .
:1dc:title "Hyatt Regency Mumbai in Mumbai (Bombay)" ;
rdfs:urls "lonelyplanet.com/india/mumbai-bombay/hotels/hyatt-regency-mumbai";
rdfs:imgurls "media.expedia.com/hotels/3000000/2590000/2589300/2589285/2589285_28_b.jpg";
rdfs:continent "Asia";
rdfs:country "India";
rdfs:typeofroom "Two Twin";
rdfs:typeofhotel "Five";
rdfs:rate "Cheap";
rdfs:facility "Fitness Bowling Archery Casino Fitness Center Golf Swimming Pool";
rdfs:facility2 "Swimming Pool";
```

Table: RDF used in Hotel Ontology

7. MATERIALS AND METHODS

Jena is a programming tool which is used in this proposed system with the help of Java programming language. With Ontology API, Jena aims to give a consistent interface for programming for ontology application programming that is independent of ontology language used in our programs. Jena API for Ontology is language neutral. The java class names used are actually not specific to the language used. For example, the java class name 'OntClass' represents 'OWL" class or RDFS class. We need to have a required language to design ontologies. In this proposed system of RDFS, the Jena language is used which is supported by ontology language. RDFS helps the oncologists to make simper hierarchy concepts and similar properties.

SPARQL is pronounced as "sparkle" is an acronym for SPARQL protocol and RDF query language is also used in this proposed system. RDF is a query language that is used for databases. SPARQL query consists of triple pattern - conjunction, disjunction and optional pattern. It helps in retrieving and manipulating data stored in Resource Description Framework (RDF) format. This format was made standard format by Data Access Working Group (DAWG) of WWW Consortium and is also known as one of the important key technology of semantic web. SPARQL 1.0 became official W3C Recommendation in January 2008 and SPARQL 1.1 became in March 2013.

8. IMPLEMENTATION

Here in this proposed framework we plan and build up a semantic web design that can diminish the clients from the overburden of doing a considerable measure of watchword based hunt before getting the relevant result. This framework takes the common language client inquiry as themes in an easy to understand environment. In this proposed framework the point based web interface must be produced for fulfilling the objective of semantic skimming in an exceedingly heterogeneous semantic web environment. The proposed framework displays the refinement with higher precision and in programmed way. Not at all like the current frameworks in this examination work keeping in mind the end goal to manage metaphysics a typical interface must be composed.

The proposed framework design must be partitioned into 3 unique modules: First module takes the client question as theme depiction and the second module gives a component to the client to improve indexed lists by giving different picture filtering apparatuses. The third module is to display the images according to their rank. Be that as

it may, for the framework to comprehend client question and give best result, it ought to be prepared first. The SPARQL inquiry utilized as a part of the work:

```
@prefix :       <http://example/>.
@prefix xsd: <http://www.w3.org/2001/XMLSchema#>.
@prefix dc:       <http://purl.org/dc/elements/1.1/>.
@prefix rdfs:   <http://www.w3.org/2000/01/rdf-schema#>.
:1 dc:title "Hyatt Regency Mumbai in Mumbai (Bombay)" ;
   rdfs:urls "lonelyplanet.com/india/mumbai-bombay/hotels/hyatt-regency-mumbai";
   rdfs:imgurls
"media.expedia.com/hotels/3000000/2590000/2589300/2589285/2589285_28_b.jpg";
   rdfs:continent "Asia";
   rdfs:country "India";
   rdfs:typeofroom "Two Twin";
   rdfs:typeofhotel "Five";
   rdfs:rate "Cheap";
   rdfs:facility "Fitness Bowling Archery Casino Fitness Center Golf Swimming
Pool";
   rdfs:facility2 "Swimming Pool";
  .
```

Table RDF used in Hotel Ontology

Complete RDF File

```
# Data for LARQ example 1
@prefix  :<http://example/> .
@prefix  xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix  dc:       <http://purl.org/dc/elements/1.1/> .
@prefix  rdfs:   <http://www.w3.org/2000/01/rdf-schema#> .

1. dc:title "Hyatt Regency Mumbai in Mumbai (Bombay)" ;
rdfs:urls "lonelyplanet.com/india/mumbai-bombay/hotels/hyatt-regency-mumbai";
```

rdfs:imgurls "media.expedia.com/hotels/3000000/2590000/2589300/2589285/2589285_28_b.jpg";
rdfs:continent "Asia";
rdfs:country "India";
rdfs:typeofroom "Two Twin";
rdfs:typeofhotel "Five";
rdfs:rate "Cheap";
rdfs:facility "Fitness Bowling Archery Casino Fitness Center Golf Swimming Pool";
rdfs:facility2 "Swimming Pool";

.

2. dc:title "Shangri-La Hotel" ;
rdfs:urls "lonelyplanet.com/hotels/";
rdfs:imgurls "media.expedia.com/hotels/6000000/5310000/5301700/5301691/5301691_11_b.jpg";
rdfs:continent "Asia";
rdfs:country "India";
rdfs:typeofroom "Two Twin";
rdfs:typeofhotel "Five";
rdfs:rate "Cheap";
rdfs:facility "Fitness Bowling Archery Casino Fitness Center Golf Swimming Pool";
rdfs:facility2 "Swimming Pool";

3. dc:title "Phoenix Park Inn Resort in Candolim, India - Lonely Planet" ;
rdfs:urls "lonelyplanet.com/india/goa/hotels/";
rdfs:imgurls "media.expedia.com/hotels/2000000/1040000/1039900/1039834/1039834_32_b.jpg";
rdfs:continent "Asia";
rdfs:country "India";
rdfs:typeofroom "Two Twin";
rdfs:typeofhotel "Five";
rdfs:rate "Cheap";
rdfs:facility "Internet Fitness Bowling Archery Casino Fitness Center Golf Swimming Pool";
rdfs:facility2 "Balcony Side Heater Beach Side King Bed Kitchenate";

.

4. dc:title "Radisson BluHotel Greater Noid" ;

rdfs:urls“expedia.com/Ghaziabad-Hotels-Radisson-Blu-Hotel-Greater-Noida.h4480236.Hotel-Information";

rdfs:imgurls "media.expedia.com/hotels/5000000/4490000/4480300/4480236/4480236_9_b.jpg";

rdfs:continent "Asia";

rdfs:country "India";

rdfs:typeofroom "Two Twin";

rdfs:typeofhotel "Five";

rdfs:rate "Cheap";

rdfs:facility "Fitness Bowling Archery Casino Fitness Center Golf Swimming Pool";

rdfs:facility2 "Internet Balcony Side Heater Beach Side King Bed Kitchenate Refrigerator Big Tub";

.

5 dc:title "Regal International East Asia Hotel, Shanghai - Hotel Reviews ..." ;

rdfs:urls "ebookers.com/hotels/China/Shanghai/Regal_International_East_Asia_Hotel.h330724/";

rdfs:imgurls "www.tnetnoc.com/hotelphotos/724/330724/2631759-Regal

International-East-Asia-Hotel-Hotel-

Exterior-1-DEF.jpg";

rdfs:continent "Asia";

rdfs:country "China";

rdfs:typeofroom "Two Twin";

rdfs:typeofhotel "Three";

rdfs:rate "Expensive";

rdfs:facility "Fitness Bowling Archery Casino Fitness Center Golf Swimming Pool";

rdfs:facility2 "Internet Balcony Side Heater Beach Side King Bed Kitchenate Refrigerator Big Tub";

INPUT OUTPUT ARCHITECTURE FOR PROPOSED SYSTEM

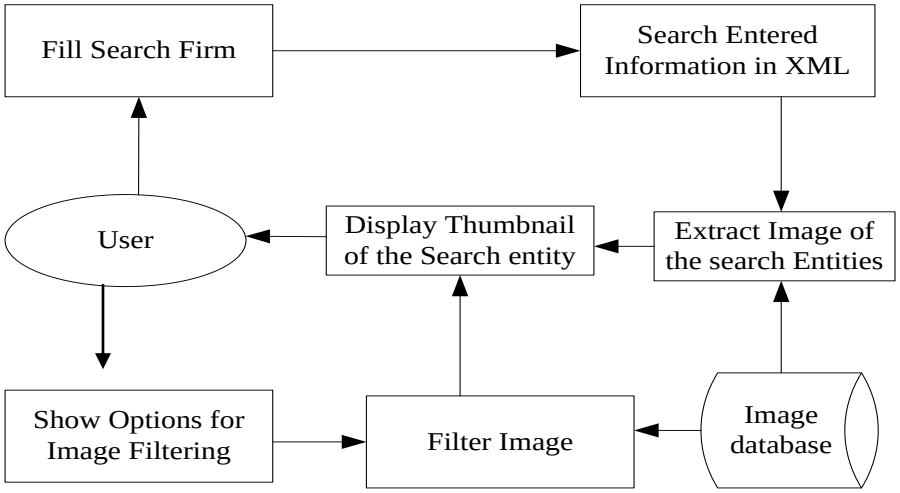


Figure Input/ Output Architecture

The aforementioned conditional figure speaks to the info yield architecture for proposed web crawler.

Preparing the System: The information of different clients is gathered and composed around ontology of clients. The framework has an expansive database of pictures having a place with different classes. These pictures are passed into a calculation which separates different metadata of picture, for example, document sort, record size, document measurement, date made on and so on. A calculation "Closest neighbor addition" technique was utilized to compute the normal shade of the picture by resampling the picture to a 1 x 1 measurement. All the points of interest alongside the URL of picture record and its classification is put away in a database. The class of a picture is recognized physically and it can be anything like age, spot where he works, area and so on.

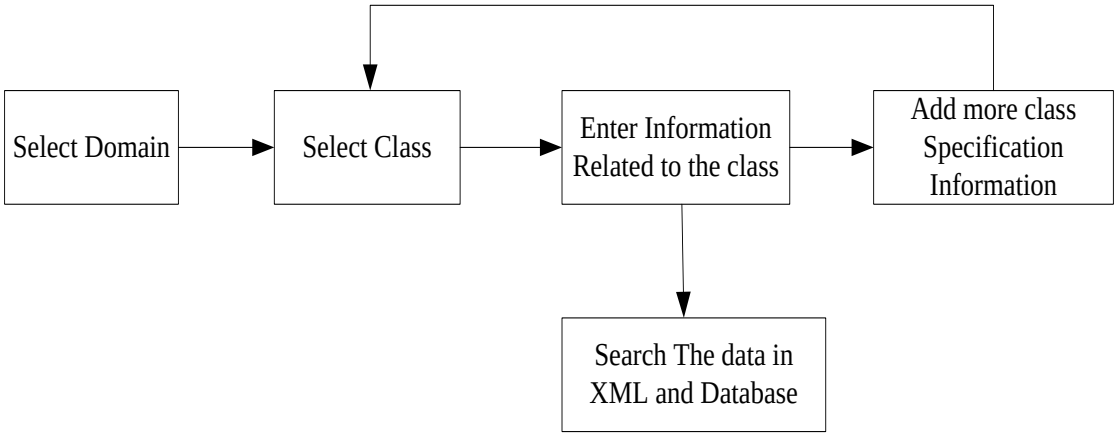


Figure Structure of Topic based search form

As client enters his normal dialect inquiry by selecting class and subjects and afterward entering the points of interest he has. He can enter various questions correspondingly. Once the client has manufactured his inquiry, it has gone through a pursuit instrument where the information is initially checked in the xml document then pictures are recovered from the database

Client Interface for picture show and sifting

Client interface is the system that client can see and utilize. For a specific space, client enters pertinent hunt catchphrases. These watchwords are then sought in the database utilizing SQL question. Figure 5 shows the metadata extraction from the picture.

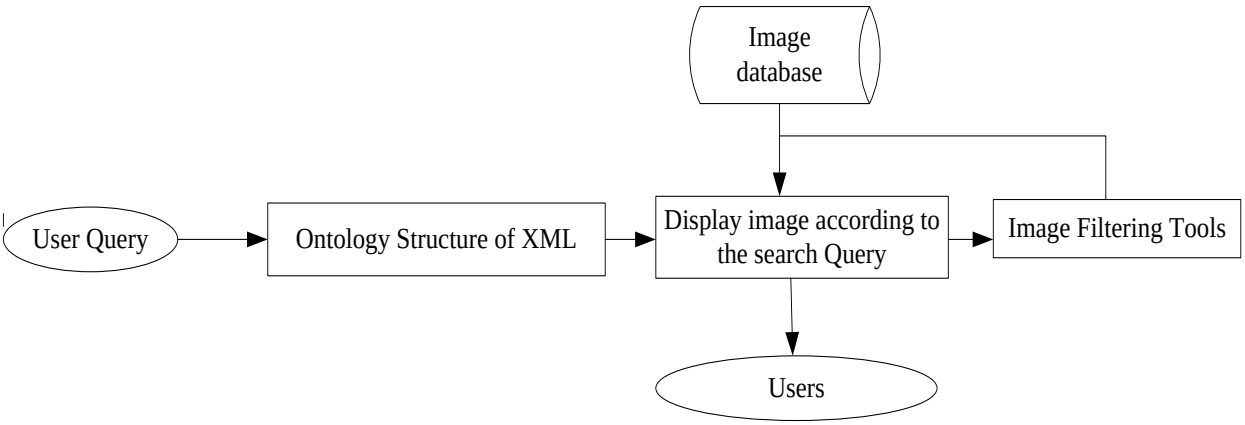


Figure Extracting metadata from image

Select the requirements:

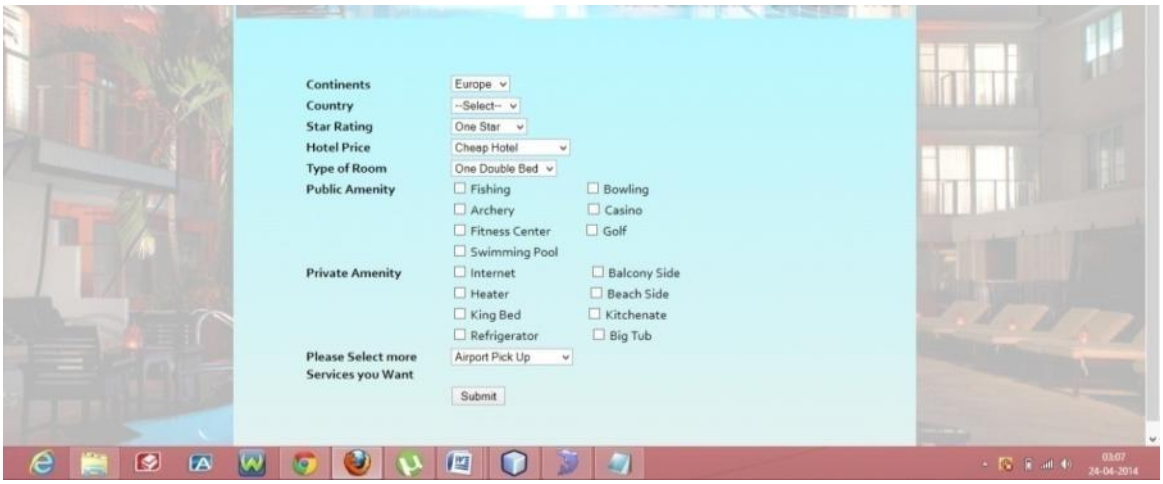


Figure 4.3.1: Interface for Semantic Search

According to the requirements the related information of hotels is collected from the RDF File:

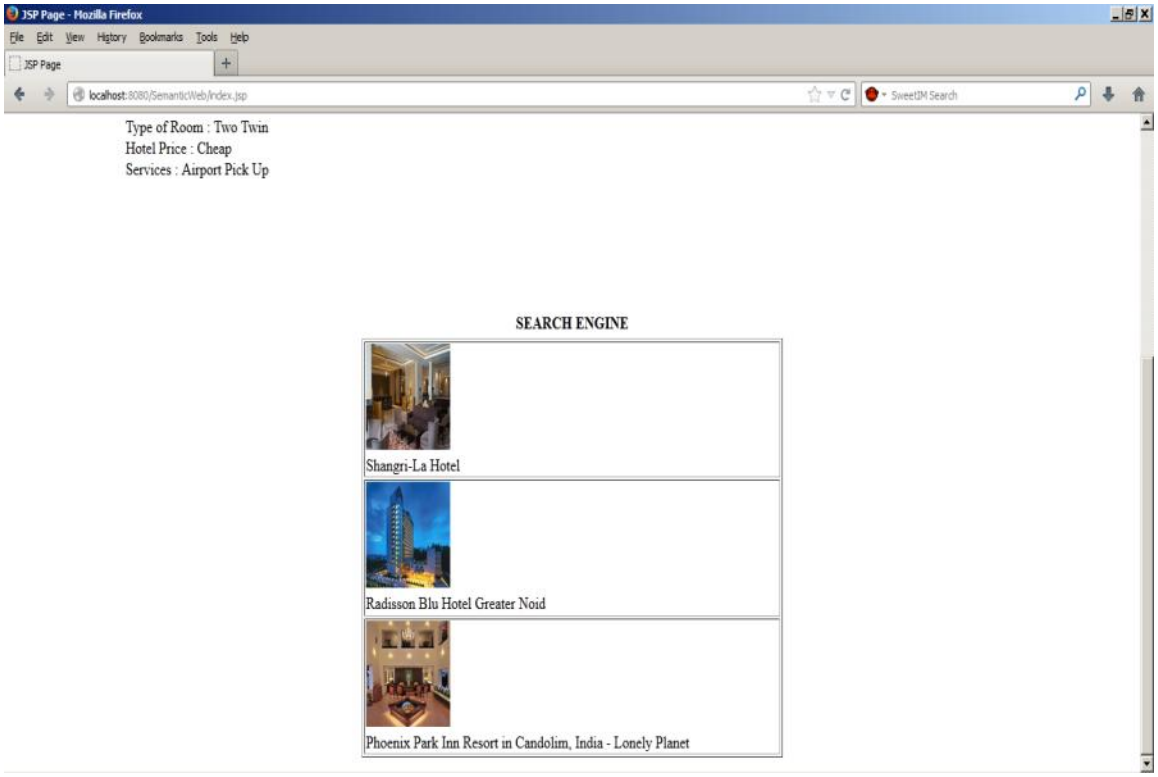


Figure 4.3.2: Interface for Search Results

If we want more details for the hotels, on clicking the images it opens the website of the hotel, maintaining the ranking of the urls. The predicates**rdfs:urls**and**rdfs:imageurl**are used.

On clicking Radisson Blue Hotel Greater Noida again and again it will increase the ranking of this hotel. If it has ranking greater then Shinghai Hotel, it will show Radisson Blue Hotel Greater Noida before Shinghai Hotel.

This work is using some Data Models

Fields	**Type**
Continents	Char(30)
Countries	Char(30)

Table: Continents and Countries included

Fields	Type
Hotel_id	Number
URL	Char(100)
ImageUrl	Char(150)
Title	Char(150)

Table: Urls are stored for further linkings

Fields	Type
Hotel_id	Char(30)
Rank	Number

Table: Ranks are stored for further rankings.

9. CONCLUSION

This work executes the Semantic Web. It is an extended version of the present web in which data is given well defined significance, making computers and individuals to work in collaboration. The Semantic Web gave a typical system that permits information to be shared and reused crosswise over application, enterprise, and community limits. Semantic Web methods will help structuring more and more information and will integrate it at Webpage level. This will help avoiding reinvention of data description. We need to stop worrying about how to address 25 percentage of problem when 75 percentage is already addressed with present Semantic Web Standards. Committing more new standards, 90 percentage of it will be addressed soon. Semantic Web Methodology helps in describing different types of information. When the pieces of information are defined, system will describe even more. Initially it will help us describing ourselves which is critical to personalize. Further it will help in determining places and things that are linked in different. Semantic Web techniques will simply help us structure more information and incorporate it at Web scale. It will abstain from reexamining the information depiction wheel. We have to quit wrangling about how to address 20 percent of the issue, when 80 percent of it is now addressable with current Semantic Web gauges. With more duty to existing guidelines, 90 percent of it could soon be addressable. Semantic Web techniques help us portray different piece sorts of the data. Once those pieces are depicted, machines can help us more. In the first place it will help us portray ourselves on

the web. That is basic to personalization and significance. At that point it will help us depict spots and things in connected ways we haven't possessed the capacity to sometime recently.

10. FUTURE SCOPE

HOTEL ONTOLOGY structure is produced utilizing RDF.RDF offers a system for the development of sensible language for coordinated effort in the semantic web. It is a XML based language speaking to trade of information. It is giving data on the significance of data. In RDF an archives makes presumptions that particular things have properties which have values. Thus, positively it is a general component for information representation. The meaning of the system is area nonpartisan: the semantics of particular areas is not altered, but rather the component is usable for the depiction of data from any domain.

RDF is a plan language. A RDF record has a pointer to its RDF plan at the top. This is a rundown of the terms of information that are utilized as a part of the archive. Anyone can make another plan record.

Semantic Web methodology can be utilized as a part of an assortment of use spaces. It can be utilized:

- In asset revelation to give better internet searcher capacities.
- In listing for depicting the substance and substance accessible at a specific site, page or computerized library.
- By wise programming operators to encourage learning sharing and trade.
- In content rating as utilized as a part of the work ror rating the lodgings on snaps.
- In portraying accumulations of pages that speak to a solitary intelligent 'record'.
- For portraying protected innovation privileges of pages.
- For communicating the protection inclinations of a client and also the security approaches of a site.

Ontologies are the building obstructs for the semantic web. Ontologies can have critical impact on the web as they permit the preparing, sharing and re-utilization of information between projects. An ontology is a characterization framework for ideas and their basic associations inside a particular area of information as we have utilized Hotel space as a part of this work. It is a sort of proto-hypothesis, demonstrating which components exist inside a particular space and how these components can be identified with each other. They bolster the combination of heterogeneous and conveyed data assets.

11. REFERENCES

[1] "Ontology Based Web Retrieval", Raman Kumar Goyal, Vikas Gupta, Vipul Sharma, Pardeep Mittal, UIET, Panjab University, Chandigarh, 4AP (CSE), BFCET, Bathinda, June 2015.

[2] "Ontology-Based Crawler for the Semantic Web"-Faculty of Science, Department of Applied Computer Science, by Felix Van de Maele, May 2006.

[3] "Ontology Focused Crawling of Web Documents"--Marc Ehrig, Alexander Maedch.

[4] "Ontology Based Information Retrieval"--Department of Cybernetics and AI, Technical University of Kosice, --by Jan Paralic, Ivan Kostial, Slovakia.

[5] "Crawling the Hidden Web"-by Sriram Raghavan, Hector Garcia Molina, --Computer Science Department, Stanford University, USA.

[6] "An Efficient Adaptive Focussed Crawler Based on Ontology Learning" --By Chang Su, Yang Gao, Jianmei Yang, Bin Luo, Proceedings of the Fifth International Conference on Hybrid Intelligence Systems- 2005 IEEE.

[7] "A New Approach to Design Domain Specific Ontology Based Web Crawler", --By Debajyoti, Arup Biswas, Sukanta, 10th International Conference on Information Technology, -- 2007 IEEE. Ringeet. al.

[8] "Ontology Based Web Crawler"--By Ganesh S, Jayaraj M., Aghila G., --Information Technology-Coding & Computing, 2004 volume 2, 337-341-IEEE Sept 2012.

[9] "An efficient scheme to remove crawler traffic from the internet.", --By Yuan X., H. Macgregor and J. Harms-"Proceedings of the 11th International Conference on Computer Communications and Networks" Oct-2002. 14-16,-IEEE-CS Press(pp: 90-95).

[10] "The Ethicality of Web Crawlers"-By Sun. Y, Council G. Isaac and Giles C. Lee, in the proceedings of 2010 "IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology"-Toronto Canada Aug-2010.(pp: 668-675) .

[11] 'Web Crawler' From "Wikipedia"- Link: http://en.wikipedia.org/wiki/Web_crawler

[12] 'World Wide Web' From "Wikipedia"- Link: http://en.wikipedia.org/wiki/World_Wide_Web

Publish Research Article

Dear Sir/Mam,

We invite unpublished Research Paper,Summary of Research Project,Theses,Books and Book Review for publication.

Address:- Dr. Ashak Hussain Malik House No-221, Gangoo Pulwama - 192301 Jammu & Kashmir, India
Cell: 09086405302, 09906662570,
Ph No: 01933212815
Email:- nairjc5@gmail.com, nairjc@nairjc.com , info@nairjc.com
Website: www.nairjc.com