



RDF Based Framework for Multi Image Retrieval using Semantic Web

A. Gauthami Latha*

Dept of CSE VIIT,
Duvvada, India
gauthamilatha@gmail.com

Y. Srinivas

Dept of IT,
GITAM University, India
ysrinivasit@rediffmail.com

E. Sirisha

Dept of CSE,
GVP College of Engineering (Autonomous)
India
eedupuganti13@gmail.com

Ch. Satyanarayan

Dept of CSE, JNTU K,
Kakinada, India
chsatanarana@yahoo.com

Abstract: Retrieving number of images from the Web, based on image content is a challenging field which is currently receiving lot of attention. The images of interest are retrieved from the database, based on the context query. The proposed model shows the architecture to integrate knowledge from dataset and semantic techniques. To extract the features like texture and color of images, we use semantic tools like Jena and Sparql Protocol and RDF Query Language (SPARQL). The model is developed using Resource Description Framework (RDF), where the initial search is based on keyword, then extended to knowledge based and SPARQL is used to retrieve relevant images from the corresponding RDF. The aim of the model is to explain the advantages of using Semantic Web languages and technologies to create, store, manipulate and process image metadata.

Key Words: Jena, RDF, Semantic Techniques, Semantic Web, SPARQL

I. INTRODUCTION

Most of the present Web content is suitable for human utilization. Web content which was generated automatically from databases is presented without the original structural information. Today, search engines like Google, Yahoo are exploited in searching the relevant information based on a keyword. These search engines drive the web towards success. The present web has certain limitations such as

- High recall, low precision.
- Retrieval of irrelevant documents
- Results are highly sensitive to vocabulary

For Example, if we want to search the “musical instruments of size 16 GB”, it retrieves all the type of data where it finds the word 16 GB i.e. it doesn't consider 16 GB as the capacity of the system but rather it considers it as a simple word, here the search is being done based on word. To overcome these limitations, we propose a system that searches the queries based on the semantics. In section-2 the proposed system is discussed, architecture of the system is explained in section-3 and the implementation of the proposed model is presented in section-4 and section-5 shows the results and comparisons between the current web and semantic web are discussed in section-6.

II. PROPOSED SYSTEM

A. Semantic Web:

The term "Semantic Web" is often used more specifically to refer to the formats and technologies that enable it [1]. These technologies include the Resource Description

Framework (RDF), RDF Schema (RDFS) and the Web Ontology Language (OWL), all of which are intended to provide a formal description of concepts, terms, and relationships within a given knowledge domain [1]. The Semantic Web is generally built on syntaxes, which use URIs (Uniform Resource Identifiers) to represent data, usually in triples based structures. These syntaxes are called "Resource Description Framework" syntaxes.

Semantic web can provide the ability to tag all content on the Web and describe each piece of information and give semantic meaning to the content. Using semantic web, searching becomes more effective and users can find the precise information [2], [3], [4], [15].

B. Objectives:

The main objectives of the proposed system are

- To provide efficient search mechanism.
- To formulate the exact query dynamically.
- To express queries across diverse data Sources and get the appropriate results.
- To provide only relevant data.
- To maintain interoperability and re-usability of the code.
- To support Query answering over several documents.

C. Scope:

The system is developed using RDF Jena and Net beans frame work. The front-end is GUI based and is simple and attractive. User has to enter the query along with the category so that the relevant data is displayed. The system maintains the dataset about Babies, Animals, Birds, Bikes and Roses with

certain predicates .In our paper, the relevant results are shown to prove the efficiency of the search mechanism.

III. SOFTWARES

A. Jena:

Jena is a Java framework for building Semantic Web Applications. It provides a programmatic environment for RDF, RDFS and OWL, SPARQL. The Jena Framework includes an RDF API, reading and writing RDF in RDF/XML, N3 and N-Triples (Subject, Predicate, and Value), parsers and writers for the RDF formats, a rule-based inference engine and an OWL API [5]. The key RDF package for the application developer is *com.hp.hpl.jena.rdf.model*. This package contains interfaces for representing models, resources, properties, literals, statements and all the other key concepts of RDF for creating models. So that application code remains independent of the implementation. Jena provides an API to extract data from and write to RDF graphs. The graphs are represented as an abstract "model". A model can be sourced with data from files, databases, URLs or a combination of these. A Model can also be queried through SPARQL and updated through SPARQL [6].

a. Rdf:

RDF is extensively recognized as an excellent choice for representing and processing semi-structured data. The Resource Description Framework, or RDF, enables data to be decentralized and distributed. The RDF data model is similar to classic conceptual modeling approaches such as Entity-Relationship or Class Diagrams, as it is based upon the idea of making statements about resources (in particular Web resources) in the form of subject-predicate-object expressions. . The subject denotes the resource, and the predicate denotes traits or aspects of the resource and expresses a relationship between the subject and the object.

RDF is a simple data model and ability to model disparate, abstract concepts has also led to its increasing use in knowledge management applications unrelated to Semantic Web activity. The RDF is a language for describing information and resources on the web [9]. Putting information into RDF files, makes it possible for computer programs to search, discover, pick up, collect, analyze and process information from the web [7, 8]. RDF help us to integrate data from different sources without custom programming, offer your data for re-use by other parties and allow to decentralize data in a way that no single party "owns" all the data [5].

Example to represent the Subject, Predicate and object:

- i. A Subject is anything that can have a URL, such as "http://doc".
- ii. A Predicate is a Subject that has a name, such as "author".

An Object is the value of a Predicate, such as "Siri".

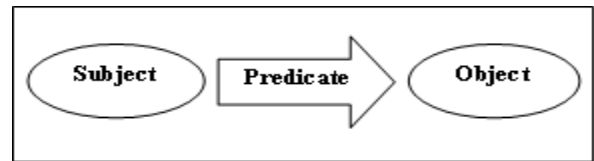


Figure 1: Representation of Subject, Predicate and Object

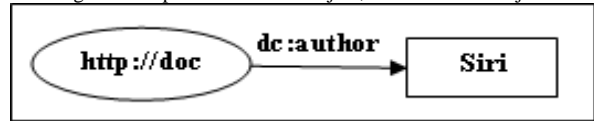


Figure 2: Represents an example that describes Subject, Predicate and Object
Table1 shows the RDF generated for the example shown in Figure2 where http://doc is the subject, dc: author is the predicate and Siri is an object.

Table1: Generated RDF

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf=
http://www.w3.org/1999/02/22-rdf-syntax-
ns#
xmlns:dc="http://doc#"
<rdf:Description about="http://doc/">
<dc:author> Siri </dc:author>
</rdf:Description>
</rdf:RDF>
```

b. Sparql:

SPARQL is a query language for pattern matching against RDF's. SPARQL query consist of triple patterns, conjunctions, disjunctions and optional patterns. SPARQL is a RDF query language. The syntax resembles SQL, making for a friendly learning curve, but SPARQL is far more powerful, enabling queries spanning multiple disparate data sources containing hetero - geneous semi-structured data. SPARQL takes the description of what the application wants, in the form of a query, and returns that information, in the form of a set of bindings or an RDF graph. The results of SPARQL queries can be result sets or RDF graphs. SPARQL support in Jena is currently available via a module called ARQ. The SPARQL language specifies different query variations for different purposes.

The various commands associated with SPARQL query are

a) **Select Query:** Used to extract raw values from a SPARQL endpoint and the results are returned in a table format.

b) **Describe Query:** Used to extract an RDF graph from the SPARQL endpoint, the contents of which are left to the endpoint to decide based on what the end user requires. Each

of these query forms takes a WHERE block to restrict the query although in the case of the DESCRIBE query the WHERE is optional. The SPARQL query processor will search for sets of triples that match the triple patterns, binding the variables in the query to the corresponding parts of each triple [10, 11].

USING SELECT-FROM-WHERE

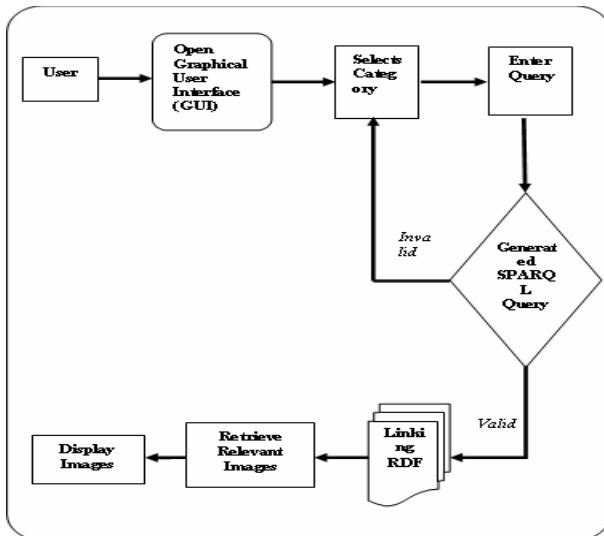


Figure 3: Architecture for Image Retrieval using Knowledge Based System

As in SQL, the queries in SPARQL also have a SELECT - FROM - WHERE structure. SELECT specifies the *projection*: the number and order of retrieved data. FROM is used to specify the source being queried. This clause is optional; when it is not specified, we can simply assume we are querying the knowledge base of a particular system and WHERE imposes constraints on possible solutions in the form of graph pattern templates and Boolean constraints [12].

B. Netbeans:

NetBeans refers to both a platform framework for Java desktop applications and an integrated development environment (IDE) for developing with Java, JavaScript, PHP, Python, Ruby, Groovy, C, C++, Scala, Clojure, and others. The NetBeans Platform allows applications to be developed from a set of modular software components called modules. Applications based on the NetBeans platform can be extended by third party developers. The platform offers reusable services common to desktop applications, allowing developers to focus on the logic specific to their application.

IV. ARCHITECTURE

Whenever the user opens the GUI, it displays the message to select the appropriate category like Babies, Animals, Birds, Bikes and Roses.

Then the user can select the required category and pass the query to retrieve relevant images from the corresponding RDF [13], [14], [16]. When the query is passed by the user, the corresponding SPARQL query is generated as in case1 and case2 below.

Case1: The generated SPARQL query is invalid, i.e. if the user passes the query is not appropriate or else the desired

RDF does not present in the dataset.

Case2: The generated SPARQL query is valid, then it links to the corresponding RDF and from the RDF relevant images are retrieved and finally displayed to the user.

V. IMPLEMENTATION

In this article we have considered various images under the categories like Babies, Animals, Birds, Bikes and Roses. In Babies category the dataset is maintained with baby female faces and baby male faces and babies with hair and without hair. In Animals category, we have considered wild animals, reptiles, animals with stripes and with specific color. The Birds database category consists of all those birds that – can fly, cannot fly and by birds name are maintained. In Bikes category the dataset is maintained basing on model and color of the bike. In flowers category, Rose flowers are considered, it is possible to retrieve a single rose, double roses, bunch of roses, and we can even retrieve a rose flower based on color. In this paper we concentrate only on the categories Babies and Roses.

Category 1: Babies

When the query is passed to display the babies images the RDF is represented as in Table2.

Table2: Babies RDF

```

<?xml version="1.0" encoding="windows-1252"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:Babies="http://Babies#" >
  <rdf:Description rdf:about="http://Babies/1">
    <Babies:rdf> s </Babies:rdf>
    <Babies:face> female </Babies:face>
    <Babies:path>E:\Semantic Web1\Final\SemanticWeb1\imagessiri\babys\withhair\6.jpg </Babies:path>
    <Babies:hair> with </Babies:hair>
  </rdf:Description>
  <rdf:Description rdf:about="http://Babies/2">
    <Babies:rdf> s </Babies:rdf>
    <Babies:face> male </Babies:face>
    <Babies:path>E:\Semantic Web1\Final\SemanticWeb1\imagessiri\babys\withouthair\12.jpg </Babies:path>
    <Babies:hair> without </Babies:hair>
  </rdf:Description>
  <rdf:Description rdf:about="http://Babies/3">
    <Babies:rdf> s </Babies:rdf>
    <Babies:face> female </Babies:face>
    <Babies:path>E:\Semantic Web1\Final\SemanticWeb1\imagessiri\babys\withouthair\8.jpg </Babies:path>
    <Babies:hair> without </Babies:hair>
  </rdf:Description>

```

Table2 contains information of babies. For each baby it lists the face, hair, and path.

The query for babies with female faces and with hair and their path is shown as follows.

Table 3: SPARQL query for Babies RDF

```

PREFIX Babies: <http://Babies#>
SELECT ?face ?path
WHERE { ?x Babies:face "female"; Babies:hair "with"; Babies:path ?path. }

```

Case 1: The result for the babies female with hair query is as in figure 4 & 5

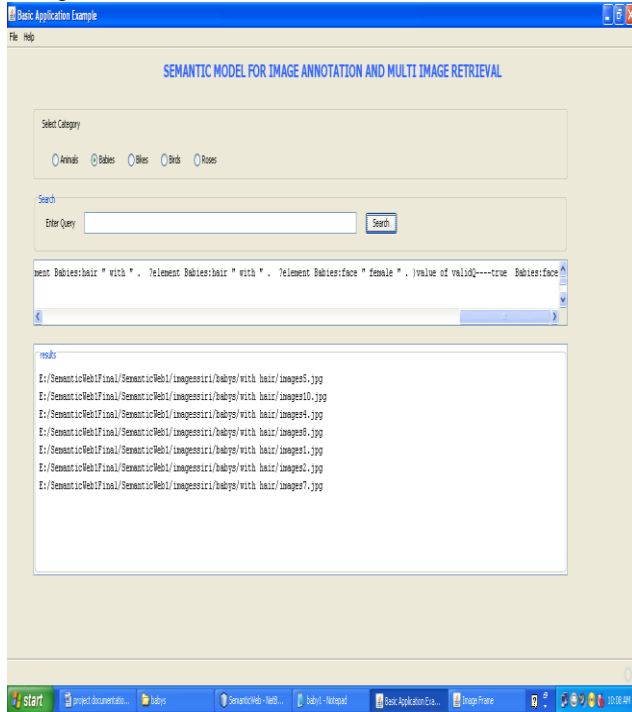


Figure 4: Screen shot to enter the Query for babies female with hair

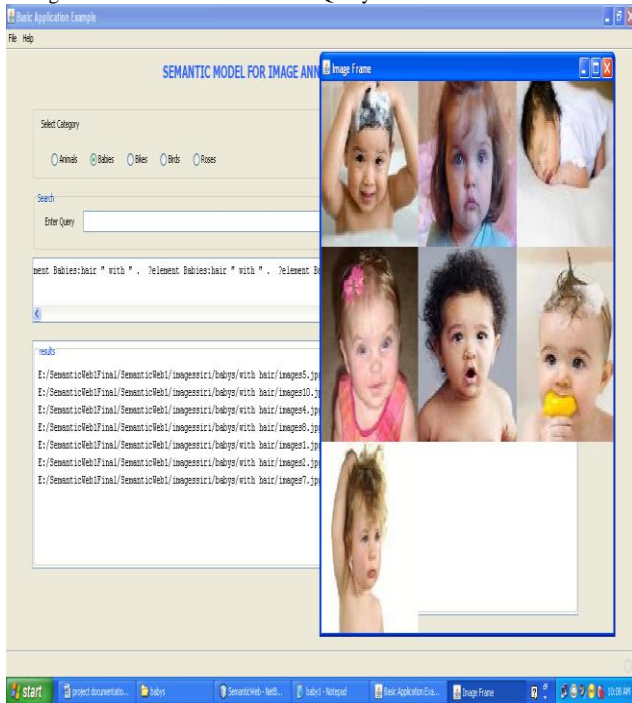


Figure 5: Screen shot of the images for babies female with hair

Case2: The result for the babies without hair query is as in figure 6&7

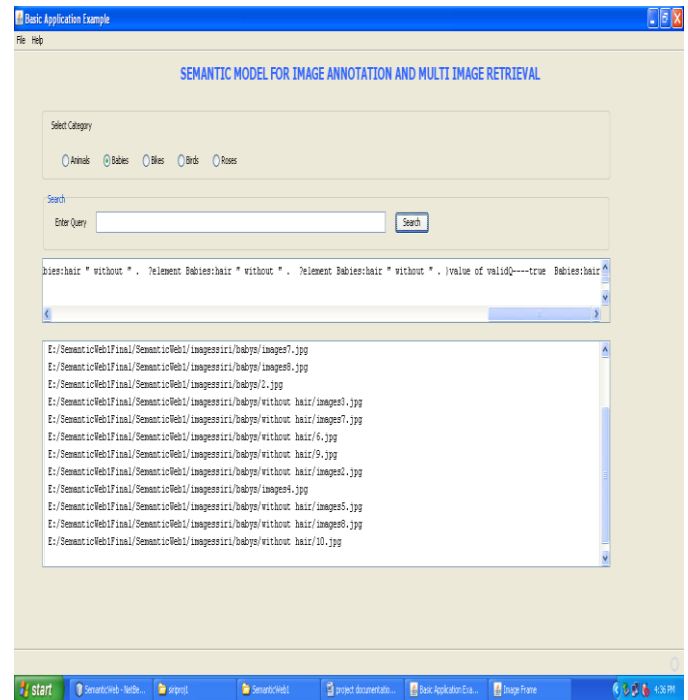


Figure 6: Screen shot to enter the Query for babies without hair



Figure 7: Screen shot of the images for babies without hair

Category 2: Roses

When the query is passed to display the roses images the RDF is represented as in Table4.

Table 4: Roses RDF


```

<?xml version="1.0" encoding="windows-1252"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:Roses="http://Roses#" >
  <rdf:Description rdf:about="http://Roses/Roses1">
    <Roses:rdf> s </Roses:rdf>
    <Roses:type> single </Roses:type>
    <Roses:path>
      E:\Semantic Web1\Final\Semantic Web1\image\roses\red rose1.jpg </Roses:path>
    <Roses:color> red </Roses:color>
  </rdf:Description>
  <rdf:Description
    rdf:about="http://Roses/Roses114">
    <Roses:rdf> s </Roses:rdf>
    <Roses:type> bunch </Roses:type>
    <Roses:path>
      E:\Semantic Web1\Final\Semantic Web1\image\roses\rosesmix18.jpg </Roses:path>
    <Roses:color> mixed </Roses:color>
  </rdf:Description>
  <rdf:Description
    rdf:about="http://Roses/Roses115">
    <Roses:rdf> s </Roses:rdf>
    <Roses:type> bunch </Roses:type>
    <Roses:path>
      E:\Semantic Web1\Final\Semantic Web1\image\roses\rosesmix20.jpg </Roses:path>
    <Roses:color> mixed </Roses:color>
  </rdf:Description>

```

Table4 contains information of roses. For each rose it lists the type, path and color.

Table 5: SPARQL query for Roses RDF

```

PREFIX Roses: <http://Roses#>
SELECT ?path
WHERE { ?x Roses:path . }

```

Case 1: The result for the roses query is as in figure 8&9

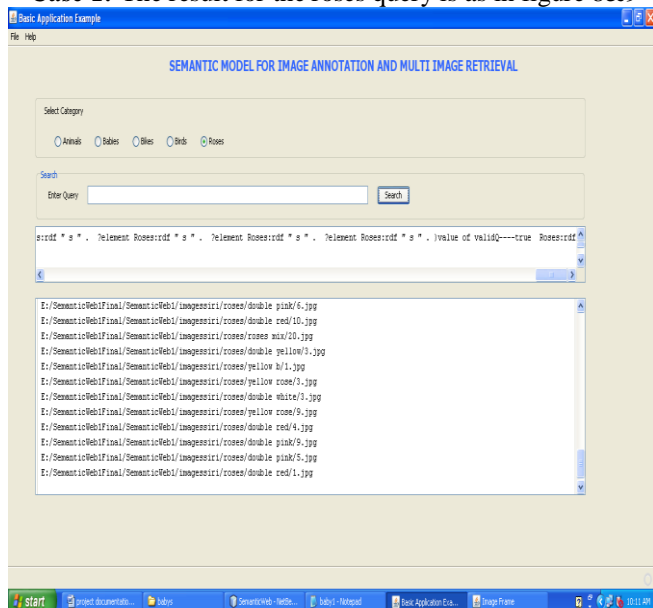


Figure 8: Screen shot to enter the Query for Roses

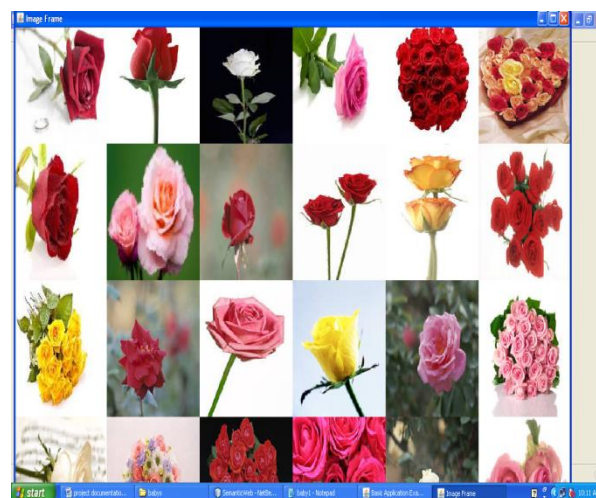


Figure 9: Screen shot of the images for Roses

Case 2: The result for the rose single with white color query is as in figure 10&11

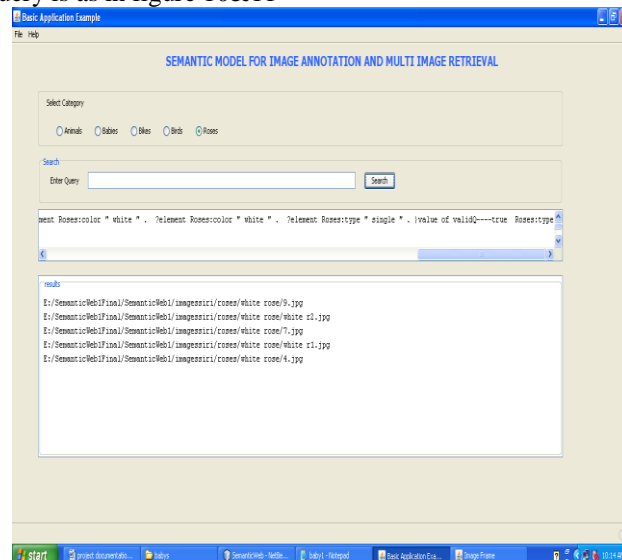


Figure 10: Screen shot to enter the Query for rose single with white color

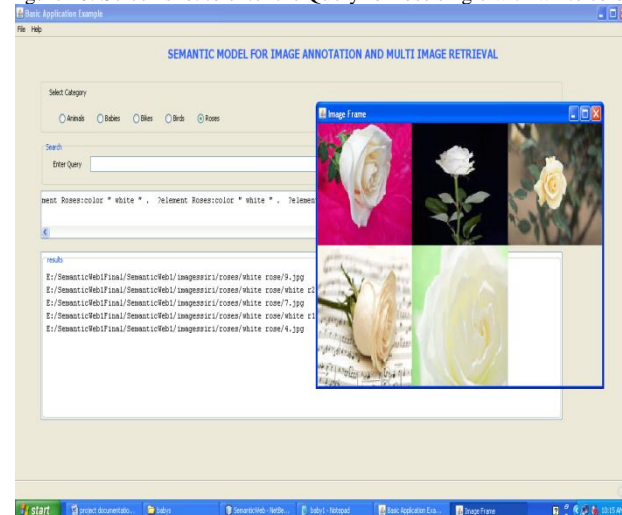


Figure 11: Screen shot of the images for rose single with white color

VI. CURRENT WEB Vs SEMANTIC WEB

As the results shown in section5, the semantic web retrieves the relevant images when compared to the current web. Consider the two categories like Babies and Roses, in current web also.

Discussion-1: The query that we have shown in category1 and case1 in section5 – “babies female with hair” is shown as in figure 12&13 for current web.

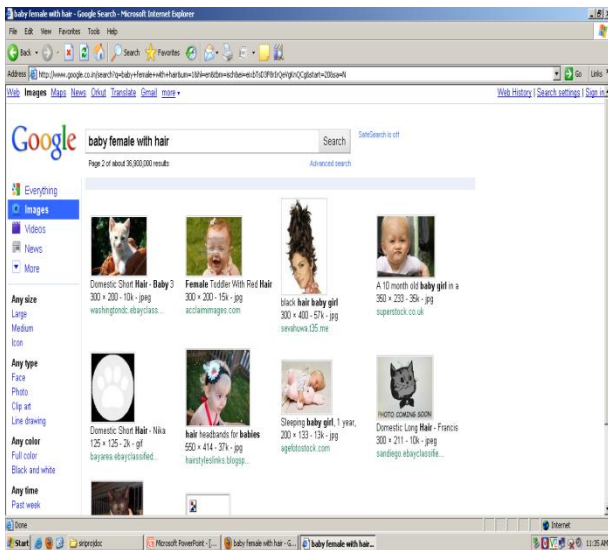


Figure 12: Screen shot of the images for “babies female with hair” from the current web

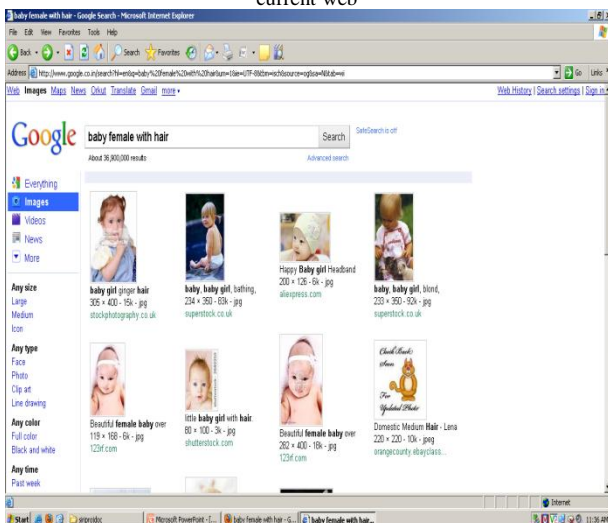


Figure 13: Screen shot of the images for “babies female with hair” from the current web

From these figures we can show that the irrelevant images are also displayed, when we make use of the current web. But by making use of RDF we can retrieve relevant images, their by we can say that semantic web is more helpful to make use of, than the current web.

Discussion-2: The query that we have shown in category2 and case2 in section5 – “rose single with white color” is shown as in figure 14&15 for current web.

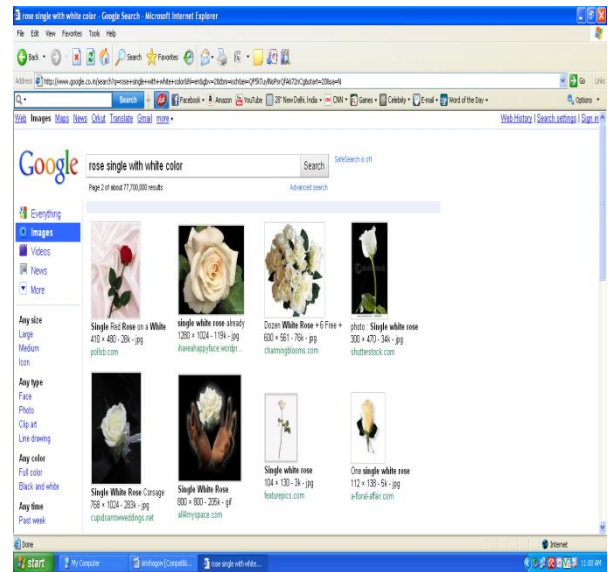


Figure 14: Screen shot of the images for “rose single with white color” from the current web

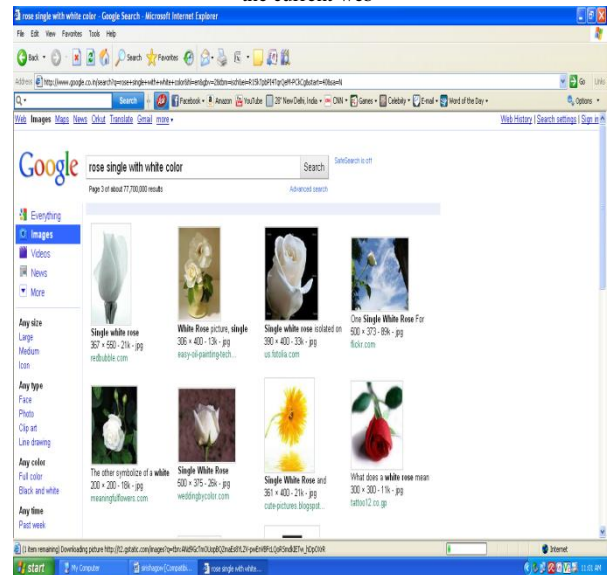


Figure 15: Screen shot of the images for “rose single with white color” from the current web

VII. CONCLUSION

Semantic web has brought a drastic change in the current searching process. It helps in creating new knowledge and new applications. Using semantic web, searching process is fast and efficient. Semantic web improves the accuracy of web searches, decrease the irrelevant data display, success in profound understanding of the web content and display the result according to the query passed. In this paper we described a framework which represents the semantic model of images. The main aim of the paper is to shown the advantages of semantic web over current web for image retrieval. The experimental results show that the knowledge-based technique significantly improves the retrieval performance compared to a keyword-based search technique.

VIII. FUTURE SCOPE AND CONCLUSION

In the future, this can be extended by increasing the dataset by adding any number of categories and any number of images in each category. Each image can consist of any number of properties. The project can also be implemented by considering only the features of the images as a dataset rather than considering the images directly where we can save the memory space, so that the system efficiency will be increased.

IX. REFERENCES

- [1]. Ruofei Zhang, Zhongfei (Mark) Zhang, Mingjing Li, Wei-Ying Ma, Hong-Jiang Zhang, "A Probabilistic Semantic Model for Image Annotation and Multi-Modal Image Retrieval", <http://www.springerlink.com/content/u1t220x838372257/>, 27 April 2006
- [2]. Peter Mika- "Social Networks and the Semantic Web" <http://dare.ubvu.vu.nl/bitstream/1871/13263/5/7915.pdf>
- [3]. "The Semantic Web: An Introduction" <http://infomesh.net/2001/swintro/>
- [4]. Graham Klyne, "Semantic Web Applications", <http://www.ninebynine.net/papers/SemanticWebApplications.ppt>, 26 February 2004
- [5]. E. Motta and M. Sabou, "Next Generation Semantic Web Applications", http://books.google.co.in/books?id=-YeUh7-PCkgC&pg=PA259&dq=E.+Motta+and+M.Sabou+%2B+Next+Generation+Semantic+Web+Applications&hl=en&ei=eulmTm0CcbOrQfx492CCw&sa=X&oi=book_result&ct=result&resnum=2&ved=0CDwQ6AEwAQ#v=onepage&q&f=false
- [6]. Jena – A Semantic Web Framework for Java, <http://jena.sourceforge.net>
- [7]. RDF Working Group, <http://www.w3.org/rdf>, 2004-02-10
- [8]. "An Introduction to RDF and the Jena RDF API" <http://jena.sourceforge.net/downloads.html>
- [9]. Tim Berners-Lee: Why RDF model is different from the XML model, <http://www.w3.org/DesignIssues/RDF-XML>, October 1998
- [10]. The SPARQL specification from W3C: Read the specification of the SPARQL query language. <http://www.w3.org/TR/rdf-sparql-query/>, 15 January 2008
- [11]. SPARQL NEW FEATURES AND RATIONALE W3C WORKING DRAFT [HTTP://WWW.W3.ORG/TR-SPARQL](http://www.w3.org/TR-SPARQL), 2 JULY 2009
- [12]. Lee Feigenbaum, "Why SPARQL" http://www.thefigtrees.net/lee/blog/2008/01/why_sparql.html
- [13]. D. Blei and M. Jordan, "Modeling annotated data", <http://www.cs.princeton.edu/~blei/papers/BleiJordan2003.pdf>, (SIGIR), 2003.
- [14]. K. Barnard, P. Duygulu, N. de Freitas, D. Forsyth, D. Blei, and M. I. Jordan, "Matching words and pictures Journal of Machine Learning Research", www.cs.cmu.edu/~efros/courses/AP06/Papers/barnard-jmlr-03.pdf, 3:1107–1135, 2003.
- [15]. Charles Petrie, "Semantic Web Premier", <http://books.google.co.in/books?id=hXK125b2MRAC&lpg=PA1&dq=Semantic%20Web%20Premier&pg=PA1#v=onepage&q&f=false>
- [16]. W.R. Dillon and M. Goldstein, "Multivariate Analysis, Methods and Applications", http://148.201.94.3:8991/F?func=direct¤t_base=ITE01&doc_number=000034796, 1984.