SEMANTIC WEB ANALYSIS

K.C.Gouda[1], K. Deepan siddarthan[2] , Nagaraj Bhat[3]

[1]CSIR-Centre for Mathematical Modelling and Computer Simulation,
Wind tunnel Road, Bangalore-37 Karnataka, India
kcgouda@cmmacs.ernet.in

Dept of ISE,Jeppiar Engineering College,
Anna University, Chennai, 600 025
deepansiddarthan.k@gmail.com

[1]CSIR-Centre for Mathematical Modelling and Computer Simulation,
Wind tunnel Road, Bangalore-37 Karnataka, India
nagakallare@gmail.com

Abstract **The ability to search for climate data has proven to be essential for many predictions on weather and prevention of natural disasters. However the available data search engines are typically limited to lexical searches and do not consider the underlying semantics of the query typed for searching data. This work addresses the analysis of climate data from diverse resources through semantic web, here we present Semantic Web-based approach climate data search that uses ontologies to model and integrate the climate data segments from varied sources on the Internet. This approach allows us to search for the desired data avoiding ambiguities and in-complete knowledge. The semantic web allows the exchange of the queries along with their semantics without being bound to programming language or any other factors by extracting the essentials for climate data retrieval, from the ontologies being created and defined. The main intention is to create a set of ontologies that can be mapped with the search query selected through the GUI and recover substantial data as desired**

Keywords— **Delay, Queuing theory, arrival and service.**

I. INTRODUCTION

The awakening of a new technology happens only to reduce human efforts. A natural language is the easier way of communication because they are effortless in conveying and straight to point. As a simple example consider the following two sentences. Both are of the form “Subject-verb-object,” one of the simplest possible grammatical structures:

1. John enjoys horror films.
2. Horror films scare Jamie.

Each of these sentences represents a piece of information. The words “Jamie” and “John” refer to specific people, the word “horror films” refers to a class of films, and the words enjoys” and “scare” tell you the relationship between the person and the film. Because we know from previous experience what the verbs “enjoy” and “Scare” mean, and we’ve probably seen a horror films before, we’re able to understand the two sentences. And now that we’ve read them, we’re equipped with new knowledge of the world. This is an example of semantics.Symbols can refer to things or concepts, and sequences of symbols convey meaning. Semantics [1] is the process of communicating enough meaning to result in an action. A sequence of symbols can be used to communicate meanings. These sentences or queries have relationship within one another or between each other. Semantic Web[2] helps the machines or the computers to extract meanings. In the current web(Web 2.0) the computer sees the web page as a bunch of characters and images, videos etc. In others words it cannot see as the way as we do. They can only parse and process the data without understanding the underlying semantics. We understand semantics of information by representing and classifying them in a domain. In Semantic Web we create several domains for knowledge representations. These domains for the denotation and depiction of knowledge are called as Ontologies [3].Semantic web works by mapping these ontologies based on the relationships between the query selected.

II. SEMANTIC WEB

The next generation of web is Semantic Web or Web 3.0. The Semantic Web initiative of the World Wide Web Consortium (W3C)has been, of great interest and has attracted innovation and also skepticism in equal measure. This creative approach was inspired by the vision of its founder, Tim Berners-Lee, of a more flexible, integrated, and automatic and self-adapting

Fig 1: Interaction between the machines through semantic web

providing a richer and more interactive experience for the users. The W3C has developed a set of standards and tools to support this vision. This paved ways for the blooming of a new era of Web 3.0.

Semantic web can be accomplished by varied methods and mechanisms. Each platform has its own framework for supporting semantic web. This paper represents a certain approach for the implementation of the web 3.0. The Semantic Web and ontologies integrate heterogeneous data and enable interoperability among disparate systems by the use of ontology mapping. Ontologies are created according to the data dependencies and the relationships between heterogeneous datasets, by integrating the data with the ontologies. The Ontologies are created by a wide variety of application programs available in the Internet. Among those we utilize the software provided by Stanford, which is Protégé 4.2(for educational purpose only)

2.1SEMANTIC WEB TECHNIQUES:

The semantic web can be implemented in different ways. Each of these mechanisms serve the purpose of creating a Semantic Web

2.1.1Resource Description Framework (RDF):

RDF[4] is a standardized model by W3C for the data interchange on the Web. RDF has features that facilitate the data merging even when the underneath schemas differ, and it specifically supports the evolution of the schemas over a period without all data consumers to be changed. These utilize .Triples are database normalization taken to a logical extreme. They have the advantage that we can load them form heterogeneous sources into one database without configuration. RDFdefines a predicate called “rdf:type”.This is used to depict or denote that entities are of certain types.

Eg:

```
<?xml version="1.0"?>
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:cd="http://www.recshop.fake/cd#">
<rdf:Description
rdf:about="http://www.recshop.fake/cd/Empire Burlesque">
<cd:artist>John urbano</cd:artist>
<cd:country>USA</cd:country>
<cd:company>Columbia</cd:company>
<cd:price>10.90</cd:price>
<cd:year>2013</cd:year>
</rdf:Description>
<rdf:Description
rdf:about="http://www.recshop.fake/cd/Hide your heart">
<cd:artist>Bonnie Tyler</cd:artist>
<cd:country>UK</cd:country>
<cd:company>CBS Records</cd:company>
<cd:price>9.90</cd:price>
<cd:year>1988</cd:year>
</rdf:Description>
</rdf:RDF>
```

2.1.2Web Ontology Language (OWL):

Web ontology language add semantics to the schema. The term ontology has a complex history both in and out of computer science. An ontology is a set of precise descriptive statements about some part of the world (usually referred to as domain of interest or the subject matter of the ontology). Precise descriptions serve several purposes, especially they prevent misunderstandings in human communication and they ensure that software behaves in a uniform, predictable way and works well in other software.

In order to precisely describe a domain of interest, it is helpful to come up with a set of central terms –often called as vocabulary.

Are generally represented as Subject-Predicate-Object. The object is either another or literal such as number or string .For example, a triple might describe the fact that Charles is John’s Father

<http://example.com/person/john><http://familyontology.net/1.0#hasFather><http://example.com/person/Charles>

OWL[5] is not a programming language. It is declarative and it describes a state affairs in a logical way. Appropriate tools, also called as [11], can be used to infer further information about the state of affairs in a logical way. It is a knowledge representation, which is built as extension for RDF.OWL is the current W3C standard for defining semantic web schemas, and tools and API support for OWL are rapidly expanding.OWL is a very large language with a lot of complicated parts. OWL itself is broken intothree sub-languages of increasing complexity and expressiveness called OWL-Lite (thesimplest), OWL DL, and OWL Full.

OWL Full: The entire language is called OWL Full, and uses all the OWLlanguages primitives. It alsoallows to combine these primitives in arbitrary ways with RDF and RDFSchema. This includes thepossibility (also present in RDF) to change themeaning of the pre-defined (RDF or OWL) primitives, by applying thelanguage primitives to each other. For example, in OWL Full we could impose a cardinality constraint on the class of all classes, essentially limitingthe number of classes that can be described in any ontology. The advantage of OWL Full is that it is fully upward compatible withRDF, both syntactically and semantically: any legal RDF document is alsoa legal OWL Full document, and any valid RDF/RDF Schema conclusionis also a valid OWL Full conclusion.The disadvantage of OWL Full is the language has become so powerfulas to be undecidable, dashing any hope of complete (let alone efficient)reasoning support.

OWL DL: In order to regain computational efficiency, OWL DL (shortfor: Description Logic) is a sublanguage of OWL Full which restricts theway in which the constructors from OWL and RDF can be used. The advantage of this is that it permits efficient reasoning support.Theadvantage is that we lose full compatibility with RDF: an RDFdocument will in general have to beextended in some ways and restrictedin others before it is a legal OWL DL document. Conversely, every legalOWL DL document is still a legal RDF document.

OWL Lite: An ever further restriction limits OWL DL to a subset ofthe language constructors. For example, OWL Lite excludes enumeratedclasses, disjointness statements and arbitrary cardinality (among others).The advantage of this is a language that is both easier to grasp (for users)and easier to implement (for tool builders).The disadvantage is of course a restricted expressivity.

Example:

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:owl="http://www.w3.org/2002/07/owl#"
xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns="http://sweet.jpl.nasa.gov/ontology/service.owl#">
<owl:Ontology rdf:about="">
 <!--
<dc:title>
Service</dc:title>
<dc:date>
1/20/2004 7:59:03 PM</dc:date>
<dc:creator>
SWEET project</dc:creator>
<dc:description>
</dc:description>
<dc:subject>
</dc:subject>
-->
<owl:versionInfo>
1.0</owl:versionInfo>
</owl:Ontology>
<owl:Class
rdf:about="http://www.w3.org/2002/07/owl#Thing"/>
<owl:Class rdf:ID="Web-based">
<rdfs:subClassOf rdf:resource="#InteractiveProgram"/>
</owl:Class>
<owl:Class rdf:ID="InteractiveProgram">
<rdfs:subClassOf rdf:resource="#Education-outreach"/>
</owl:Class>
<owl:Class rdf:ID="Ask-aBiologist">
<rdfs:subClassOf
rdf:resource="#Digital-VirtualReferenceDesk"/>
</owl:Class>
<owl:Class rdf:ID="Digital-VirtualReferenceDesk"/>
<owl:Class rdf:ID="Frost-FreezeWarning">
<rdfs:subClassOf rdf:resource="#Weather-
ClimateAdvisory"/>
</owl:Class>
```

2.2Semantic Web Architecture:

A layered architecture for the Semantic Web that adheres tosoftware engineering principles and the fundamental aspects of layeredarchitectures will assist in the development ofSemantic Web specificationsand applications. The most well-known versions of the layered architecture that exist within literature have been proposed by Berners-Lee. It is possible to indicate inconsistencies and discrepancies in thedifferent versions of the architecture, leading to confusion, as well asconflicting proposals and adoptions by the Semantic Web community.

The proposed CFL[6] architecture (Figure 1) for the Semantic Web varies fromthe previously suggested versions of the architectures mainly becauseit adheres to the evaluation criteria for layered architectures that are based onestablished software engineering principles. In addition, it is noticeablethat the CFL architecture abstracts and depicts related functionalities ratherthan the W3C technologies used to instantiate these functionalities

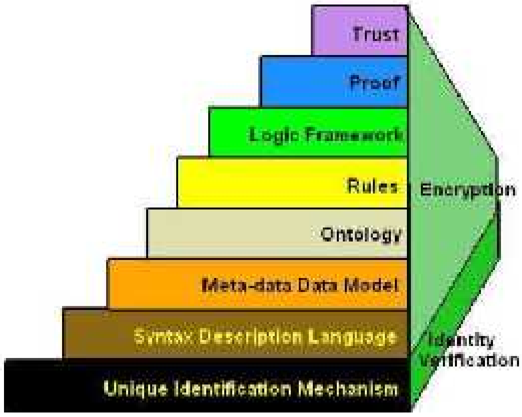


Fig 2: The proposed CLF architecture for semantic web

III. INTERFACE

Interface design is most often associated with the development of Web pages. In this scenario we use the J2EE[7] Technology for the development of user interfaces .Java platform, Enterprise Edition or Java EE is the Oracle's enterprise java computing platform. The platform paves way for the design of API and runtime environment for the developing the enterprise software applications. In java EE a UI can be built using Servlet, Java ServerPages (JSP)[8] or using Html and css styles .Java ServerPages (JSP) technology provides a simplified, fast way to create a dynamic content. This enables the rapid development of Web- based applications that are both server and platform independent. Further it, separates the user Interface from the content generation, enabling the designers to change the overall page layout without altering the underlying dynamic content. The Java Server pages uses XML-like tags that encapsulates the logic that creates or generates the contents for the contents of the page. The application logic can reside in server-based resources (such as JavaBeans component architecture) that the page accesses with these tags. Any and all formatting (HTML or XML) tags are passed directly back to the response page. By separating the page logic from its design and display and supporting a reusable component-based design, JSP technology makes it faster and easier than ever to build Web-based applications.

Java Server Pages technology is an extension of the Java Servlet technology. Servlets are platform-independent, server-side modules that fit seamlessly into a Web server framework and can be used to extend the capabilities of a Web server with minimal overhead, maintenance, and support. Unlike other scripting languages, servlets involve no platform-specific consideration or modifications; they are application components that are downloaded, on demand, to the part of the system that needs them. Together, JSP technology and servlets provide an attractive alternative to other types of dynamic Web scripting/programming by offering: platform independence; enhanced performance; separation of logic from display; ease of administration; extensibility into the enterprise; and, most importantly, ease of use. Third-party servlet containers are available for Apache Web Server, Microsoft IIS, and others. Servlet containers are usually a component of Web and application servers, such as BEA Web Logic Application Server, IBM Web Sphere, Sun Java System Web Server, Sun Java System Application Server, and others.

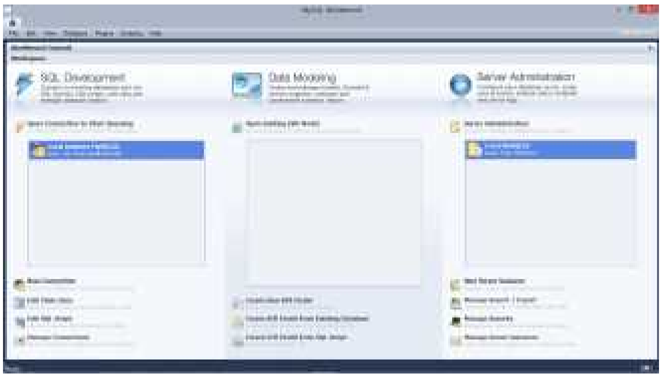
IV. DATABASE

A database is a collection of information that is organized so that it can easily be accessed, managed, and updated. In one view, databases can be classified according to types of content: bibliographic, full-text, numeric, and images. In computing, databases are sometimes classified according to their organizational approach. The most prevalent approach is the relational database, a tabular database in which data is defined so that it can be reorganized and accessed in a number of different ways. A distributed database is one that can be dispersed or replicated among different points in a network. An object-oriented programming database is one that is congruent with the data defined in object classes and subclasses.

Computer databases typically contain aggregations of data records or files, such as sales transactions, product catalogs and inventories, and customer profiles. Typically, a database manager provides users the capabilities of controlling read/write access, specifying report generation, and analyzing usage. Databases and database managers are prevalent in large mainframe systems, but are also present in smaller distributed workstation and mid-range systems such as the AS/400 and on personal computers. SQL (Structured Query Language) is a standard language for making interactive queries from and updating a database such as IBM's DB2, MYSQL, Microsoft's SQL Server, and database products from Oracle, Sybase, and Computer Associates.

5.1 MYSQL:

We use MYSQL[9] database to manage organize all data including the climate data[10]. The MSQL workbench provides us user friendly working environment for the development and maintenance for databases.



MYSQL is one of the top databases available in the market. MYSQL is a relational database with many advanced features and options. Over time, MYSQL has proved itself to be a fast, reliable and cost effective competitor to the other more expensive databases like MS SQL Server and Oracle. Here are a few of the advantages of using MYSQL in database development.

Open Source. MYSQL is an open source database system which means that anyone can use it for free. Developers can amend its code to suit their requirements which means that MYSQL is highly customizable as well. Another edge of using MYSQL over other database systems is that; it is available widely in the market with no ownership cost.

Fast Development. A lot of people around the globe are continuously developing new modules for integration with MYSQL. This means that it has a wider and faster development circle. Patches, upgrades and fixes are developed fast and become available in forums, blogs and developer sites on the internet.

Better for Small Businesses

This relational database system is free so it reduces the cost of overall database solution for small businesses and companies. This database is relatively easy to learn and operate, so operational cost is reduced substantially which is again an important factor in classifying MYSQL as an applicable and practical tool for small businesses.

Cross Platform Operability

MYSQL is easily installable and operable on different platforms including Windows, Linux, OS2 and Solaris. Cross platform operability makes it a favorable choice for development companies. MYSQL database system also contains APIs for integration with C, C++, PHP, Java, Perl, Python, Tcl, and Ruby etc. You can connect it easily with different development platforms so you can actually integrate applications developed in different OS and development platforms.

Security

MYSQL as a relational database is secure as all access passwords are stored in an encrypted format restricting any unauthorized access to the system. It also encrypts the transactions so eavesdroppers and data harvest tools cannot replicate or regenerate the database transactions once they are processed.

Connectivity

MYSQL clients can access this relational database through standard TCP/IP sockets, named pipes, UNIX sockets and many more. Standard ODBC 2.5 and above functions and commands are also supported in MYSQL.

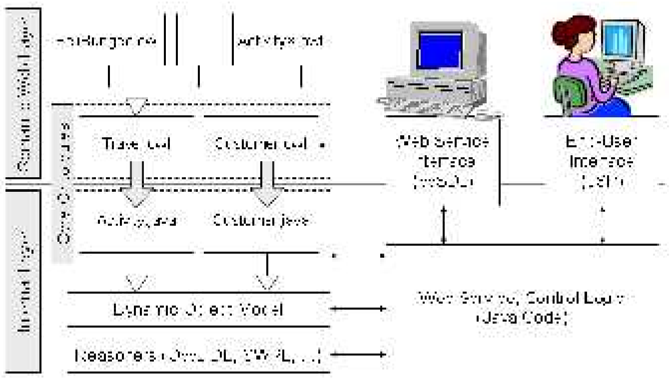
V. ALGORITHM DESIGN AND METHODOLGY:

The basic approach is to select the query as per the requirements of the user through the GUIand retrieve the information. The query is sent to the database and it matches with the request and performs the task specified. Finally it retrieves the data from the database server and displays. The figure below describes the user Interface and its purpose

User Interface containing the required links

VI. IMPLEMENTATION

The implementation of the semantic Web requires the proper functioning of all the necessary modules. These include segment of modules from ontologies to the user Interface. The owl editor Protégé provides wide opportunities of creating ontologies and mapping with the search query being searched.



REFERENCES

[1]Kalashnikov, D.V.; Mehrotra, S.; Jie Xu; Venkatasubramanian, N., "A Semantics-Based Approach for Speech Annotation of Images," Knowledge and Data Engineering, IEEE Transactions on , vol.23, no.9, pp.1373,1387, Sept. 2011
doi: 10.1109/TKDE.2010.185

[2] S. Alexaki, V. Christophides, G. Karvounarakis, D. Plexousakis, K. Tolle, The ICS-FORTH RDFSuite: Managing Voluminous RDF Description Bases, 2nd Intl Workshop on the Semantic Web (SemWeb'01, with WWW10), pp. 1-13, Hongkong, May 1, 2001.

[3]Jellouli, I.; Mohajir, M.E., "An ontology-based approach for web information extraction," Information Science and Technology (CIST), 2011 Colloquium in , vol., no., pp.5,5, 11-12 May 2011 doi: 10.1109/CIST.2011.6148583

[4]Gang Wu; Juanzi Li; Jianqiang Hu; Kehong Wang, "System II: A Native RDF Repository Based on the Hypergraph Representation for RDF Data Model," Web-Age Information Management, 2008. WAIM '08. The Ninth International Conference on , vol., no., pp.62,69, 20-22 July 2008
doi: 10.1109/WAIM.2008.90

[5]Zuo zhihong; Zhou mingtian, "Web Ontology Language OWL and its description logic foundation," Parallel and Distributed Computing, Applications and Technologies, 2003. PDCAT'2003. Proceedings of the Fourth International Conference on , vol., no., pp.157,160, 27-29 Aug. 2003
doi: 10.1109/PDCAT.2003.1236278

[6]'A Functional Semantic Web Architecture' Aurona Gerber, Alta van der Merwe, and Andries BarnardMeraka Institute and University of South Africa (Unisa), Pretoria, South Africa

[7]Perin, F., "Enabling the Evolution of J2EE Applications through Reverse Engineering and Quality Assurance," Reverse Engineering, 2009. WCRE '09. 16th Working Conference on , vol., no., pp.291,294, 13-16 Oct. 2009
doi: 10.1109/WCRE.2009.45

[8]Hu Yong; Song Jingan; Li Yanli; Zhu Lida; Zhang Jun, "Research and development of JSP based online testing system for bilingual teaching of fundamentals of manufacturing technology," Multimedia Technology (ICMT), 2011 International Conference on , vol., no., pp.671,674, 26-28 July 2011
doi: 10.1109/ICMT.2011.6001935

[9]Naim, N.F.; Yassin, A.; Zamri, W.M.A.W.; Sarnin, S.S., "MySQL Database for Storage of Fingerprint Data," Computer Modelling and Simulation (UKSim), 2011 UkSim 13th International Conference on , vol., no., pp.293,298, March 30 2011-April 1 2011,doi: 10.1109/UKSIM.2011.62

[10]"Improving the Weather Information Available to Simulation Programs"

by D. Crawley, J. Hand, and L. Lawrie.
[11]Meditskos, Georgios; Bassiliades, N., "A Rule-Based Object-Oriented OWL Reasoner," Knowledge and Data Engineering, IEEE Transactions on , vol.20, no.3, pp.397,410, March 2008 doi: 10.1109/TKDE.2007.190699

K C Gouda is currently working as a Scientist at CSIR Centre for Mathematical Modeling and Computer Simulation (CSIR C-MMACS). His research and professional career spans about twelve years of research and capacity building in modeling and computer simulation, satellite data processing, numerical modeling, Data mining, Data assimilation, cloud computing knowledge engineering and related subjects. His expertise is primarily in the domains of Software development for modeling and simulation. He is presently involved in several international and national projects related to HPC enabled modeling for weather and climate forecasting and analysis. He has published about 75 peer-reviewed papers as journal articles, book chapters, Technical reports and contributions to conference proceedings. He is a member of IMS, IEEE, AGU, AOGS and EGU. He is also a member in the board of studies of Department of Computer Science in the Jain University and Dayananda Sagar Institutes, Bangalore. He obtained his M.Sc., M.Phil, from Berhampur University, MCA from IGNOU, New Delhi and completed PhD from Mangalore University. He has Guided 15 M.Tech., 55 Masters and 30 B.E students for their academic project.

K.Deepan siddarthan presently pursuing his B.Tech (Information Science) at the Dept of ISE,Jeppiar Engineering College, Anna University, Chennai, 600 025 His Research interests includes Semantic web, Image Processing, Big data analysis, Simulation and Computer Networks. He is working for his B.Tech final semester technical project at CSIR C-MMACS under SPARK programme under the guidance of K C Gouda . He has presented two papers in the International conferences.

Nagaraj Bhat is currently working as a a Project Scientist at CSIR Centre for Mathematical Modeling and Computer Simulation (CSIR C-MMACS), Bangalore, India. His research interest includes satellite remote sensing, Cyclone studies, software development for remote sensing data analysis and Geographical Information System modeling..He obtained B.E (Computer Sc.) during 2009 and M.Tech degree in Geoinformatics in 2011 from Visvesvaraya Technological University (VTU), Karnataka. He has published seven technical papers in international and national conference proceedings.