Design and Development of Job Ontology using Semantic Web for Job and Skill Search

[1]Prof. Hina J. Chokshi, [2]Dr. Dhaval Vyas, [3]Dr. Priya Swaminarayan
[1]Assistant Professor,PICA [2]Dean(MPhil), C U Shah University, [3] Dean(FITCS), Parul University
[1]Parul Institute of Computer Application,
[1]Parul University, Vadodara, India

Abstract : Semantic search has become the next level searching concept for promoting web searching. Semantic web based search promotes semantic-based (meaning) search and gives exact match. In the proposed paper we have developed a Job Search Ontology for Effective Job Searching in the domain of IT in Protégé 5.1.1. The first part of paper shows the job search ontology consisting of various classes, the Dataset and axioms set on it and SQL Queries executed on it. Multiple SPARQL queries are executed on the created Ontology to get the desired output. Second part of paper shows the use of Jena Framework, for creating a Java interface to accept the input data from the Job Search OWL file and same queries are executed to get the required output to provide match between Job Seeker (Applicant) and Job Provider (Company) using Semantic Web. Future work can be corresponding to executing the SPARQL query in a de-centralized database, where multiple ontologies belonging to different URL are combined, to improve the job search results.

IndexTerms **- Web, Ontology, Job Search, Jena Framework, SPARQL, OWL**

I. INTRODUCTION

Ontology is a formal specification of shared conceptualization10. It is description of concepts in a domain of classes, properties and attributes. It also refers to crafting hierarchies of categories so that things can be organized and efficiently accessed. Ontology can be defined through following steps:

- Determine the scope of the Ontology.
- Define concepts (classes) to be modeled in Ontology and relationship among them.
- Reuse concepts which are available in similar existing Ontology, if any.
- Arrange the concepts in a hierarchy (sub class-super class hierarchy).
- Determine attributes and properties (slots) for each class and constraints on their values.
- Define instance and fill the slot values for each class.

II. ONTOLOGY FRAMEWORK

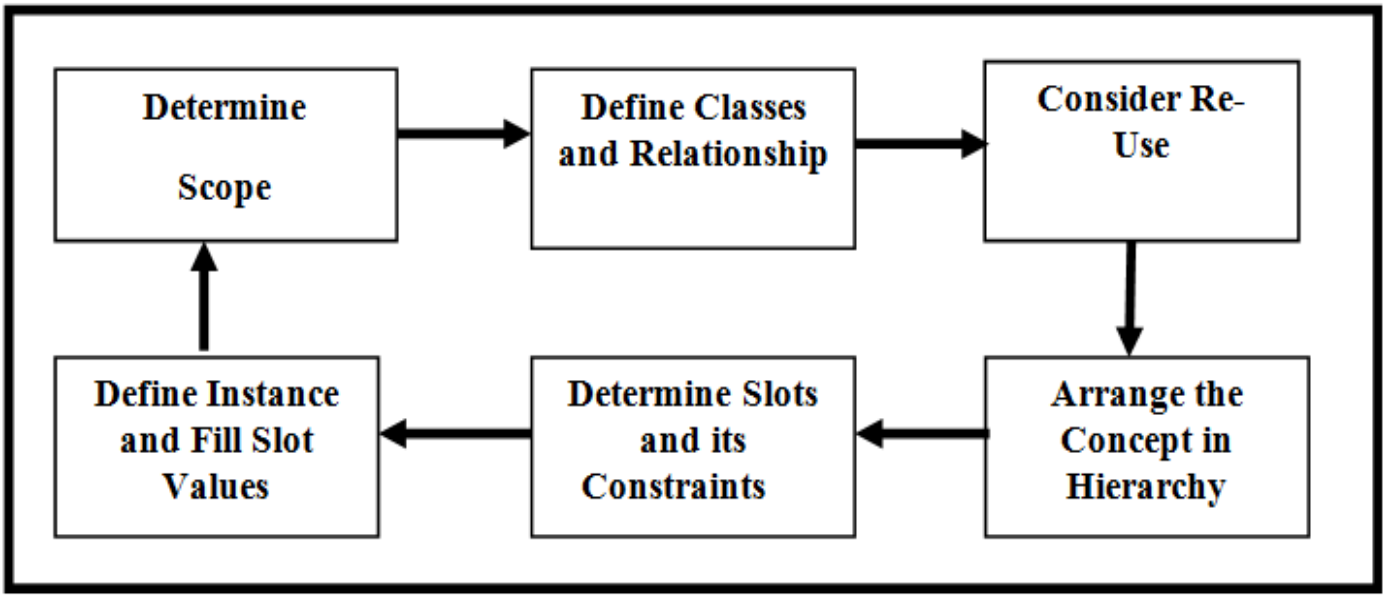

Figure 1 Ontology Framework

Importance of Ontology

A. To share common understanding of the structure of information among people or software agents.
B. To enable reuse of domain knowledge.
C. To make domain assumptions explicit.
D. To separate domain knowledge from the operational knowledge.

III. SEMANTIC WEB

The semantic web facilitates publishing the data in different languages designed for data namely

1. **Resource Description Framework (RDF)-** Resource Description Framework is a framework that is used to express data in a meaningful way. Data is expressed in form of triples, which can be easily expressed in form of a graph.[42]
2. **Web Ontology Language (OWL)** – It is a fomal representation of a new Ontology in semantic web.[43]
3. **Extensible markup Language (XML)** – It is a markup language used for exchanging data over the Internet.[44]

In contrast to HTML that describes data and provides links between them, here RDF, OWL and XML are combined to add description and meaning to the terms, which enables machine to process data as human being with reasoning and inference and result with meaningful data gathering and research.
In non-semantic web page a tag is represented as

<item>blog</item>

Using semantic web page the same is represented as

<item rdf:about="http://example.org/semantic-web/">Semantic web</item>

IV. METHODOLOGY

4.1 Design of the Job Search Ontology

The Job search ontology is represented in a hierarchical manner of a chart. The hierarchy represents the data requirement flow in an IT domain.

4.2 Hierarchical representation of Ontology for Job Search

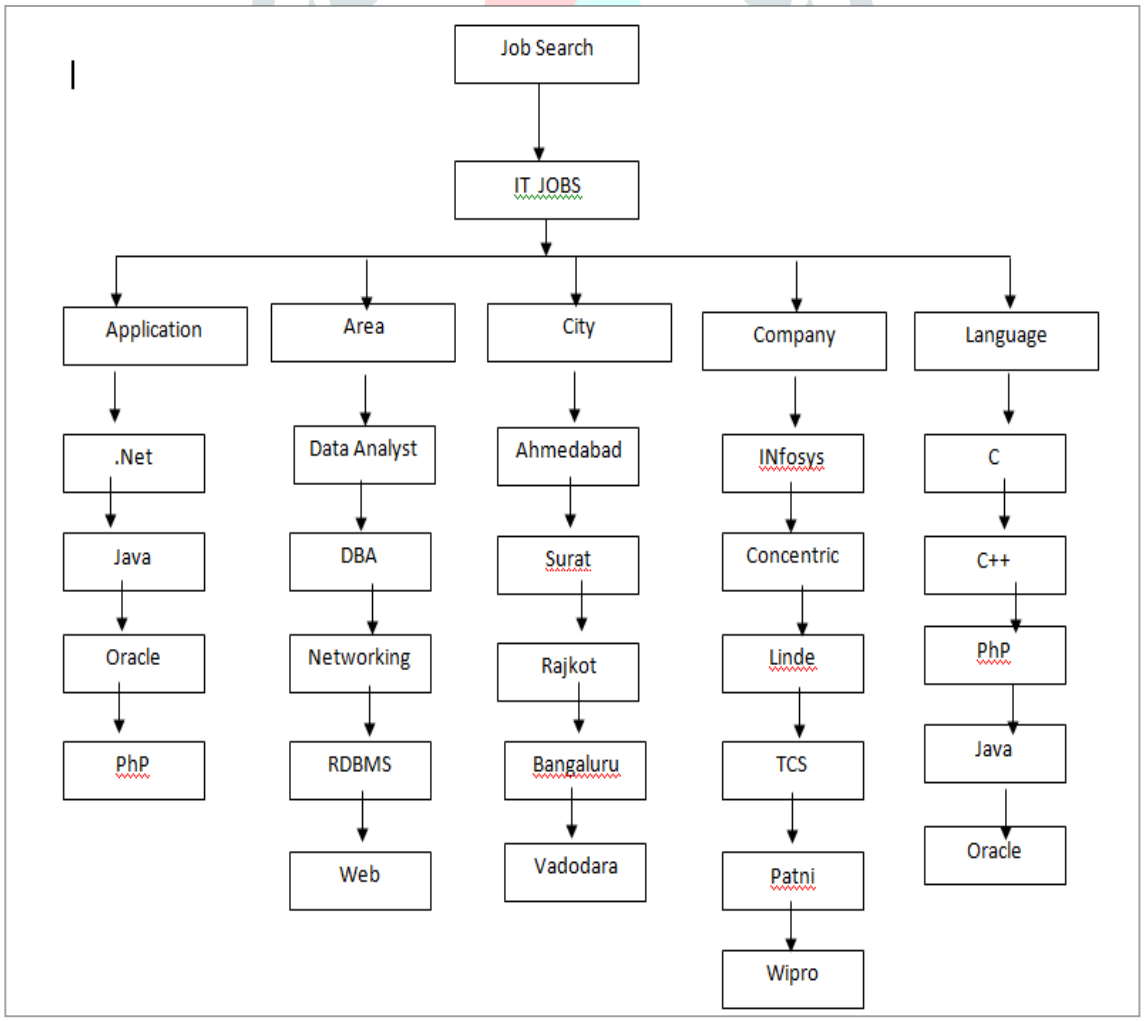

Figure 2. JobOntology Hierarchical view

4.3 Structure of JobOntology

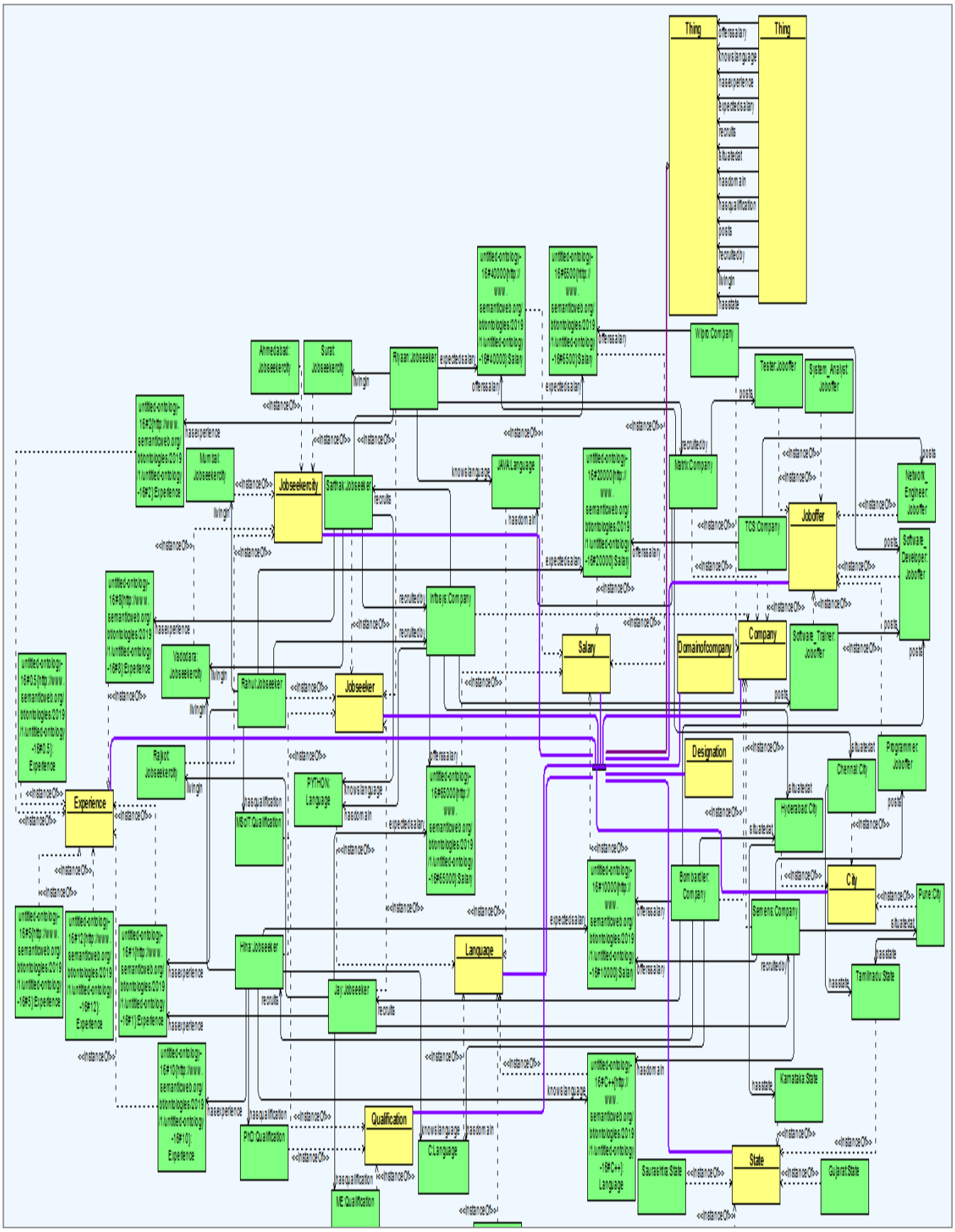

Figure 3. Job Search Ontology Design structure through OWLGrEd

4.4 Formation of class hierarchy in protégé

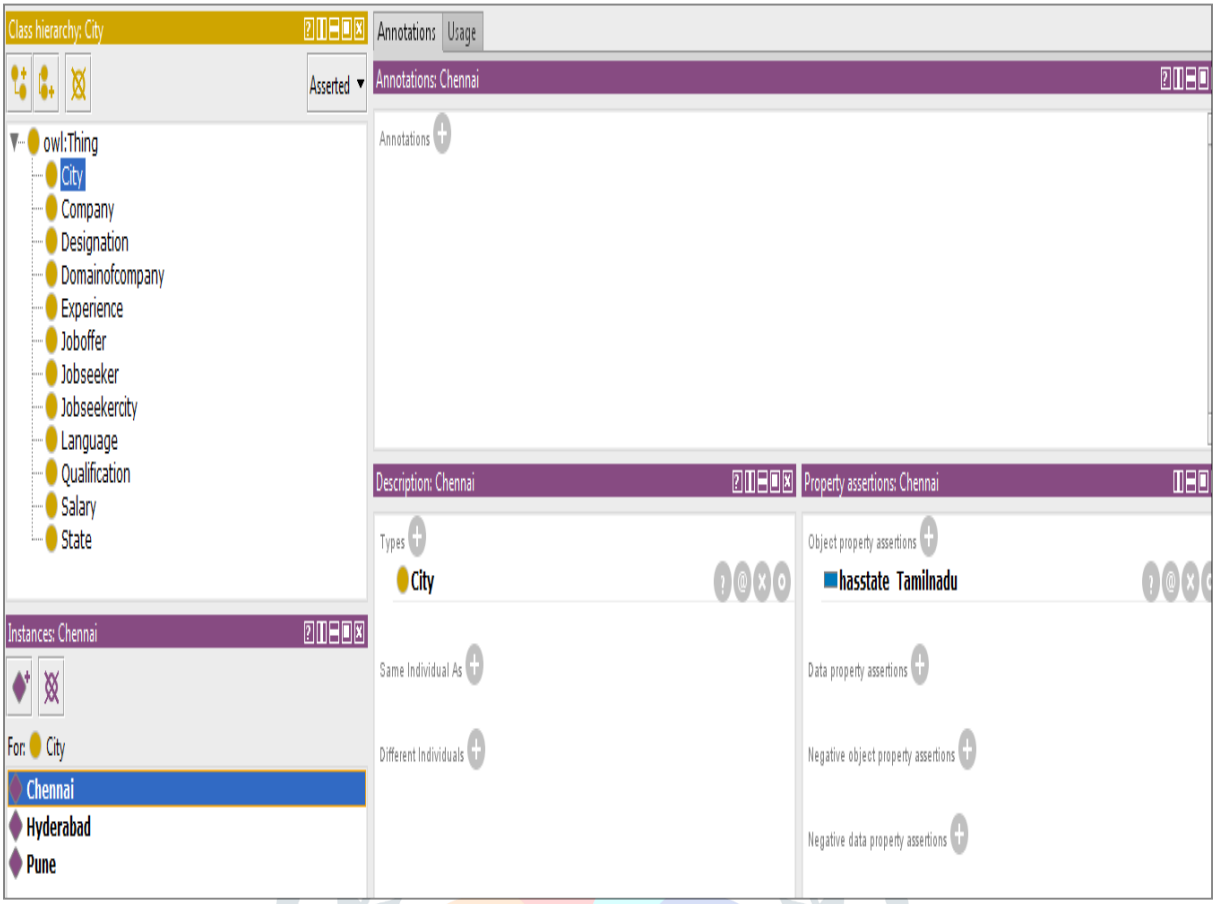

Figure 4 Class/Entity description of JobOntology in Protege

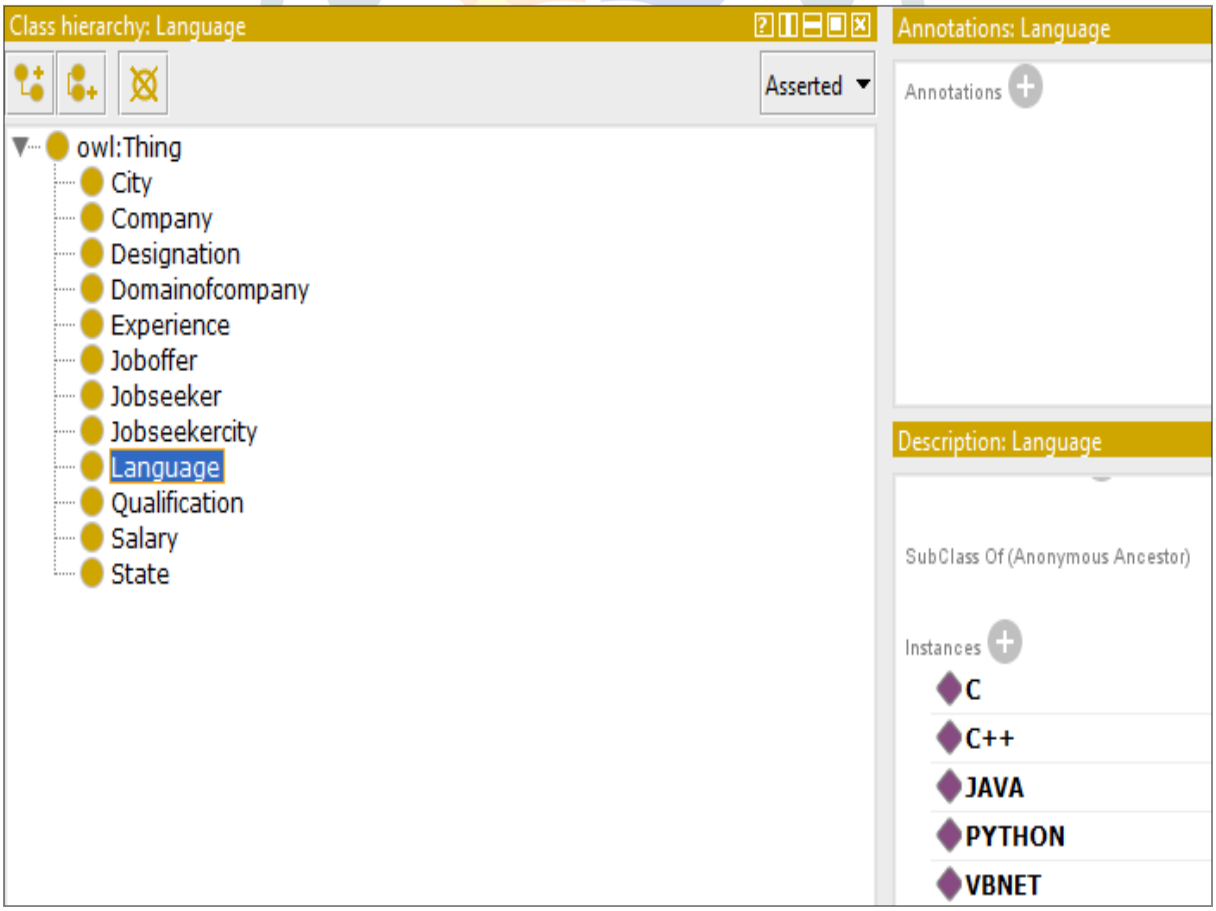

Figure 5 JobOntology instance representation through Protégé

V. SPARQL QUERY

SPARQL stands for **SPARQL Protocol and RDF Query Language.** Information in the Web is represented using SPARQL which is a semantic based database query language that fetches and updates data stored in the RDF(Resource Description Framework) format.
For the proposed JobOntology queries are created using SPARQL for the purpose of creating and retrieving data from the OWL file using the Java interface.

5.1 Structure of a SPARQL query

The SPARQL query consists of four main components namely Prefix, Select, From and Where Clause.

SPARQL query comprises of the following, in order:

- Prefix declarations: used for abbreviating URIs.
- Dataset definition: which states what RDF graphs are being queried for.
- Result clause: Identifies what data has to be returned from the query.
- Query pattern: Specifies what the data set has to be queried for
- Query modifiers: Used for data slicing, ordering, and rearranging query results

 # prefix declarations using the
PREFIX foo: <http://example.com/resources/>
...
 # dataset definition using FROM Keyword
FROM ...

 # result clause using the SELECT clause
SELECT ...

 # query pattern using the WHERE clause
WHERE {
...
}
 # query modifiers like
ORDER BY ...

5.2 Executing SPARQL Query for JobOntology in Protégé

Different SPARQL queries are created and executed on JobOntology, for extracting information related to Company, Jobseeker and Domain.

❖ **Displaying all Company Names offering Jobs**

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 : <http://www.owl-ontologies.com/untitled-ontology-424#>

SELECT DISTINCT ?companyname
WHERE {
  ?companyname rdf:type :Company.
}
```

companyname
Wipro
Tcs
Infosys

Figure 6 Query to retrieve Company offering Jobs

- **Displaying all JobSeekers name with their Citynames**

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 : <http://www.semanticweb.org/bt/ontologies/2019/1/untitled-ontology-16#>
SELECT ?jobseekername  ?jobseekercityname
        WHERE { ?jobseekername :livingin  ?jobseekercityname}
```

jobseekername	jobseekercityname
Rahul	Mumbai
Hina	Vadodara
Sarthak	Vadodara
Riyaan	Surat
Jay	Rajkot

Figure 7 Query to retrieve JobSeeker information

- **To find all company name which are in Gujarat State working in Dotnet Domain, and which is paying minimum salary 45000.**

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 : <http://www.owl-ontologies.com/untitled-ontology-424#>

SELECT ?companyname
WHERE {
?companyname :situatedAt  ?cityname.
?cityname :hasState :Gujarat.
?companyname :hasDomain :Dotnet.
?companyname :payingsalary ?salary.
            FILTER(?salary>=45000).
}
```

companyname
Tcs

Figure 8 Query to retrieve Company name

VI. JAVA FRAMEWORK

To develop Semantic Web Applications Apache Jena which is a semantic web Java framework is used. Apache Jena provides Java libraries for creating code that handles RDF, RDFs and OWL, SPARQL. It provides an API to extract data from and write to RDF graphs. It includes a rule-based inference engine to perform reasoning based on OWL and RDFS ontologies, and different types of storage strategies to store RDF triples in memory or disk.

6.1 Java Interface

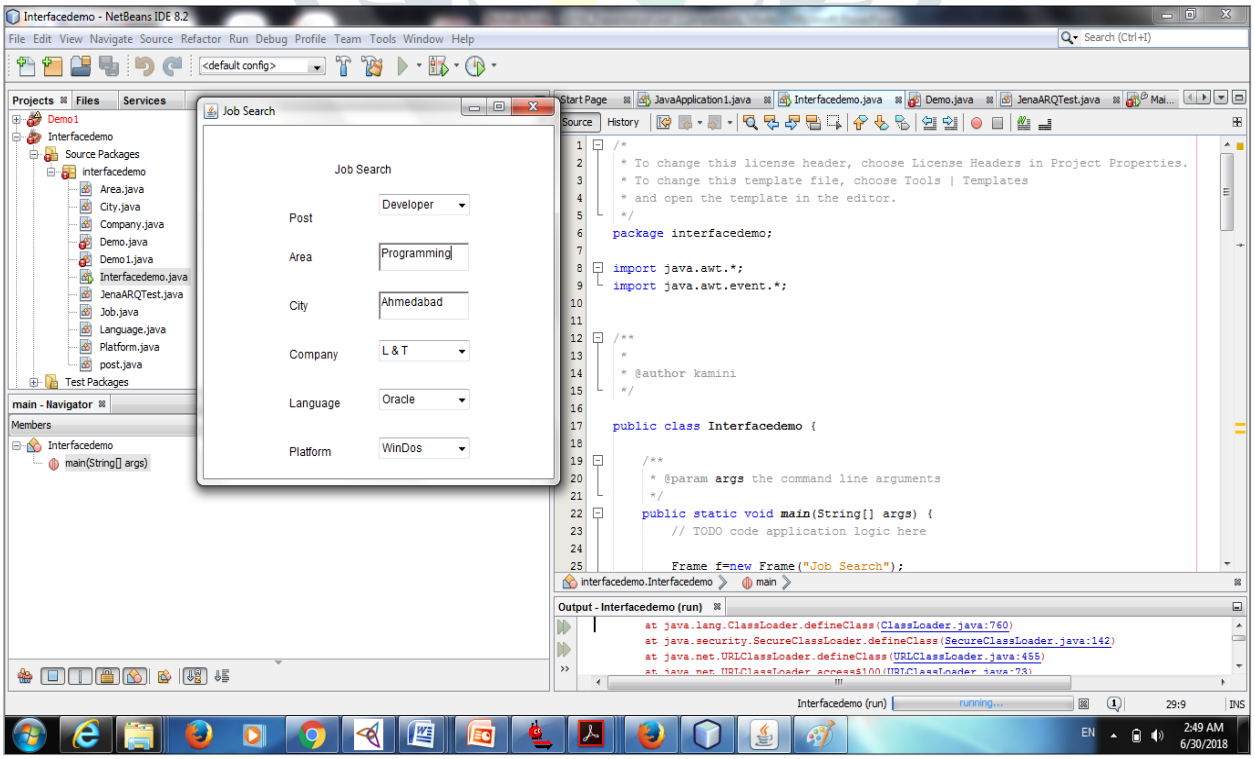

Figure 9 Snapshot of JavaInterface

6.2 SPARQL Query implemented on Java Interface

```
package JobOntology;
public class JavaApplication1 {
   public static void main(String[] args) {
        String fname="//JobOntology.owl ";
        Model model1=ModelFactory.createDefaultModel();
       model1=ModelFactory.createOntologyModel(OntModelSpec.OWL_DL_MEM);
        try
        {
           File file=new File(fname);
           FileReader reader= new FileReader(file);
            model1.read(reader,null);

                    String query1= "PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>\n" + "PREFIX owl:
      <http://www.w3.org/2002/07/owl#>\n" + "PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>\n" + "PREFIX
       xsd: <http://www.w3.org/2001/XMLSchema#>\n" + "SELECT ?subject ?object\n" + "          WHERE { ?subject
                                                rdfs:subClassOf ?object }";

           org.apache.jena.query.Query query= QueryFactory.create(query1);
           QueryExecution exe=QueryExecutionFactory.create(query, model1);
           ResultSet results =exe.execSelect();
           ResultSetFormatter.out(System.out, results, query);
        }
        catch(Exception e) {
           e.printStackTrace(); }}}
```

The above code when executed on Java Platform shall return subjects and predicates.

VII. Conclusion and Future Work

In this paper we proposed an Ontology of Job Search in IT domain. The proposed Ontology design includes matching the job seekers requirement exactly with the company's requirement. Different queries formations are shown depicting their usage. Job seeker can get more relevant opportunities for selecting the Job. Our future work shall correspond to executing the SPARQL query in a de-centralized database, where multiple ontologies belonging to different URL are combined together and executed, to improve the job search results.

VIII. Acknowledgment

I would like to express my sincere gratitude to my Guide Dr. Dhaval Vyas and my Co-Guide Dr. Priya Swaminarayan for the continuous support of my Ph.D study and related research, for their patience, motivation, and immense knowledge. Their guidance helped me in all the time of research and writing of this thesis. I am also thankful to my Colleague Dr. Kamini Solanki who provided insight and expertise that greatly assisted the research. I thank everyone who helped me directly or indirectly for the research work.

References

[1] https://www.researchgate.net/publication/259145288_Web_to_Semantic_Web_and_Role_of_Ontology_in_its_Development

[2] Fensel, D. (2001), Ontologies – A Silver Bullet for Knowledge Management and Electroni Commerce, Springer, Berlin.

[3] Stuckenschmidt, H. and Harmelen, F. van (2005), Information Sharing on the Semantic Web,Springer, Berlin

[4] Empfohlene Zitierung / Suggested Citation:Krause, J. (2008). Semantic heterogeneity: comparing new semantic web approaches with those of digital libraries.Library Review, 57(3), 235-248. https://doi.org/10.1108/00242530810865501

[5] P. R. Agarwal, "Semantic Web in Comparison to Web 2.0," *2012 Third International Conference on Intelligent Systems Modelling and Simulation*, Kota Kinabalu, 2012, pp. 558-563.doi: 10.1109/ISMS.2012.49, http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6169763&isnumber=6169663

[6] Hartig, O., Heese, R.: The SPARQL Query Graph Model for Query Optimization. In: Franconi, E., Kifer, M., May, W. (eds.) ESWC 2007. LNCS, vol. 4519, pp. 564–578. Springer, Heidelberg (2007)

[7] Stocker, M., Seaborne, A., Bernstein, A., Kiefer, C., Reynolds, D.: SPARQL basic graph pattern optimization using selectivity estimation. In: Proceeding of the 17th International Conference on World Wide Web, pp. 595–604. ACM (2008).

[8] Ladwig, G., Tran, T.: Linked Data Query Processing Strategies. In: Patel-Schneider, P.F., Pan, Y., Hitzler, P., Mika, P., Zhang, L., Pan, J.Z., Horrocks, I., Glimm, B. (eds.) ISWC 2010, Part I. LNCS, vol. 6496, pp. 453–469. Springer, Heidelberg (2010)

[9] Reddy, BR Kuldeep, and P. Sreenivasa Kumar. "Optimizing SPARQL queries over the Web of Linked Data." Proceedings of the VLDB Workshop on Semantic Data Management, SemData. 2010.

[10] Schwarte, A., Haase, P., Hose, K., Schenkel, R., Schmidt, M.: FedX: Optimization Techniques for Federated Query Processing on Linked Data. In: Aroyo, L., et al. (eds.) ISWC 2011, Part I. LNCS, vol. 7031, pp. 601–616. Springer, Heidelberg (2011)

[11] X. Wang, T. Tiropanis, and H. C. Davis. Evaluating Graph Traversal Algorithms for Distributed SPARQL Query Optimization. In: J.Z. Pan et al. (Eds.): JIST 2011, LNCS 7185, pp. 210–225. Springer, Heidelberg (2012)

[12] Vandervalk, B.P., McCarthy, E.L., Wilkinson, M.D.: Optimization of Distributed SPARQL Queries Using Edmonds' Algorithm and Prim's Algorithm. In:

[13] Hartig, O., Bizer, C., Freytag, J.: Executing SPARQL queries over the web of linked data. In: Bernstein, A., Karger, D.R., Heath, T., Feigenbaum, L. Maynard, D., Motta, E., Thirunarayan, K. (eds.) ISWC 2009. LNCS, vol. 5823, pp. 293–309, Springer, Heidelberg (2009)

[14] E. Prud'hommeaux and A. Seaborne. SPARQL query language for RDF. W3C recommendation, World Wide Web Consortium, 2008.

[15] Raval Shruti , Prof. Richa Sinha and Lajja Vyas. SPARQL Query Optimization categories. In. International Journal of Application or Innovation in Engineering & Management (IJAIEM), Volume 2, Issue 12, December 2013

[16] Drashty R. Dadhaniya, Ashwin Makwana. A survey paper on different SPARQL query optimization techniques. In, A Multidisciplinary Journal of Scientific Research & Education, 2(8), 2016.

[17] Tanvi Chawla, Girdhari Singh, Emmanuel S. Pilli. A Shortest Path Approach to SPARQL Chain Query Optimisation.In, 978-1-5090-6367-3/17/$31.00 ©2017 IEEE 2017.

[18] Marcello Leida, Andrej Chu. Distributd SPARQL query answering over RDF data stream. In, 2013 IEEE International Congress on Big Data

[19] Johannes Busse1, Bernhard Humm, Christoph Lubbert, Frank Moelter, Anatol Reibold5, Matthias Rewald, Veronika Schluter , Bernhard Seiler, Erwin Tegtmeier, Thomas Zeh, "Journal of Computing and Information Technology" - CIT 23, 2015, 1, 29–41doi:10.2498/cit.1002508

[20] Bhaskar Kapoor and Savita Sharma, "A Comparative Study Ontology Building Tools for Semantic Web Applications" : International journal of Web & Semantic Technology (IJWesT) Vol.1, Num.3, July 2010

[21] Sisay Chala, Scott Harrison, Madjid Fathi. Knowledge Extraction from Online Vacancies for Effective Job Matching. In: 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)

[22] Bastian Quilitz and Ulf Leser, Querying Distributed RDF Data Sources with SPARQL,In. S. Bechhofer et al.(Eds.): ESWC 2008, LNCS 5021, pp. 524–538, 2008. Springer-Verlag Berlin Heidelberg 2008

[23] S. Chala and M. Fathi, "Job seeker to vacancy matching using social network analysis," in Industrial Electronics Society (IES), 18th Annual International Conference on Industrial Technology, IEEE, 2017.

[24] S. Chala, F. Ansari, and M. Fathi, "A framework for enriching job vacancies and job descriptions through bidirectional matching," in 12th International Confer-ence on Web Information Systems and Technologies, SCITEPRESS Digital Library (Science and Technology Publications, Lda), vol. 2, 2016, pp. 219–226.

[25] E. Prud'hommeaux and A. Seaborne, "Sparql query language for rdf," W3C, Recommendation, 2008. [Online]. Available: http: //www.w3.org/TR/rdf-sparql-query

[26] Sheng-Wei Huang1, Chia-Ho Yu1, Ce-Kuen Shieh1, Ming-Fong Tsai2, "Efficient and Scalable SPARQL Query Processing with Transformed Table": IEEE Wireless Communications and Networking Conference (WCNC) - Workshop - Next Generation WiFi Technology, 2015

[27] Vijay Rana, Gurdev Singh, "MBSOM: An Agent Based Semantic Ontology Matching Technique": 1st International Conference on Futuristic trend in Computational Analysis and Knowledge Management (ABLAZE-2015)

[28] Micha¨el Guedj, "Levelized Taxonomy Approach for The Job Seeking/Recruitment Problem": IEEE International Conference on Computational Science and Engineering, IEEE International Conference on Embedded, 2016

[29] Weidong Hao1, Yang Yang1, Chuang Lin2, Zhengli Zhai1, "QoS-aware Scheduling Algorithm Based on Complete Matching of User Jobs and Grid Services " : Proceedings of the 2006 IEEE Asia-Pacific Conference on Services Computing (APSCC'06)0-7695-2751-5/06 $20.00 © 2006

[30] Duygu Çelik, 2Aşkın Karakaş, 3Gülşen Bal, 4Cem Gültunca, 5Atilla Elçi, 6Başak uluz, Murat Can Alevli, "Towards an Information Extraction System based on Ontology to Match Résumés and Jobs":IEEE 37th Annual Computer Software and Applications Conference Workshops,2013

[31] Soumaya Amdouni Wahiba Ben abdessalem Karaa, "Web-based recruiting Framework for CV structuring": Conference Paper · June 2010 DOI: 10.1109/AICCSA.2010.5587018 · Source: IEEE Xplore https://www.researchgate.net/publication/224178254

[32] Kumar Sharma, Ujjal Marjit, Utpal Biswas, "Efficient Provenance Storage for RDF Dataset in Semantic Web Environment": 15 14th International Conference on Information Technology, 2014

[33] Vasudha Sarda Ȧ, Prasham SakariaȦ and Sindhu NairȦ, "Relevance Ranking Algorithm for Job Portals" : International Journal of Current Engineering and Technology E-ISSN 2277 – 4106, P-ISSN 2347 – 5161

[34] S.K.Aparnaa, K.Kousalya, An Enhanced Adaptive Scoring Job Scheduling Algorithm for Minimizing Job Failure in Heterogeneous Grid Network. In 2014 International Conference on Recent Trends in Information Technology

[35] Vijay Rana, Gurdev Singh, MBSOM: An Agent Based Semantic Ontology Matching Technique. In, 2015 1st International Conference on Futuristic trend in Computational Analysis and Knowledge Management (ABLAZE-2015)

[36] Duygu Çelik, Aşkın Karakaş, Gülşen Bal, Cem Gültunca, Atilla Elçi, Başak Buluz, Murat Can Alevli , Towards an Information Extraction System based on Ontology to Match Résumés and Jobs. In, 2013 IEEE 37th Annual Computer Software and Applications Conference Workshops

[37] S.Amdouni, Wahiba Abdessalem, Web-based recruiting. In, IEEE Xplore,2010.

[37] Elena Simperi, Practical Guidelines for Building Semantic eRecruitment Applications. In, www.researchgate.net/publication/228610902

[38] Hexin Lv, Bin Zhu, Skill-Ontology based Semantic Model and its Matching Algorithm , 2006 7th International Conference on Computer-Aided Industrial Design and Conceptual Design, IEEE Explore

[39] Paolo Montuschi, Valentina Gatteschi, Fabrizia Lamberti, Job Recruitment and Job Seeking Processes: How Technology Can Help, In. https://ieeexplore.ieee.org/document/

[40] Johannes Busse1, Bernhard Humm, Christoph Lubbert, Frank Moelter, Anatol Reibold5, Matthias Rewald, Veronika Schluter , Bernhard Seiler, Erwin Tegtmeier, Thomas Zeh, "Journal of Computing and Information Technology" - CIT 23, 2015, 1, 29–41doi:10.2498/cit.1002508

[41] V. A. Chakkarwar, Amruta A. Joshi, "International Journal of Computer Applications" (0975 – 8887) Volume 133 – No.10, January 2016

[42] IanHorrocksaPeter F.Patel-SchneiderbFrankvan Harmelen, "Journal of Web Semantics",Volume 1, Issue 1, December 2003, Pages 7-26

[43] Tim Bray, Jean Paoli, C M Sperberg-McQueen, "Extensible markup language (XML) 1.0", Indian Journal of Science and Technology, Vol 9(9), DOI: 0.17485/ijst/2016/v9i9/72130, March 2016