

INFORMATION RETRIEVAL FROM GOOGLE MAP USING SEMANTIC WEB TECHNOLOGY

Lutfun Nahar¹, Rahima Afrose², Sanjida Sharmin³

¹Lecturer, Department of CSE, International Islamic University Chittagong, Chittagong, Bangladesh

²Lecturer, Department of CSE, International Islamic University Chittagong, Chittagong, Bangladesh

³Lecturer, Department of CSE, International Islamic University Chittagong, Chittagong, Bangladesh

Abstract

Web resource such as information and service for providers and user are given in a more efficient way by today's web. This service is more enhancing by Semantic Web Technology. Various technologies have been developed to deploy the Semantic Web, such as machine understandable description languages, resource composition algorithms, RDF etc. It is known by us that a web based service is Google Map that provides full information about geographical region and sites around the world. But we cannot retrieve any information about unavoidable circumstances from map. Such as a strike is called in any district which roads are affected and which are safe this type of information are not given by map. For this reason we suggest a model to retrieve information from Google Map by using Semantic Web Technology. To develop a model where we can use a useful Semantic Web and the applications, different types of the technologies must be analyzed. In our thesis we analyze different aspects for service, which is not easy in the context of the Semantic Web. Theoretically and experimental data is used in our proposed system. Our evaluation shows the effectiveness for extracting information or content to get the regional blocked area and details information.

Keywords: Semantic Web; Ontology; RDF; Linked Data.

1. INTRODUCTION

People are providing the information globally from the large database storehouse, by using World Wide Web (WWW). Many databases are being diluted with the volume of information. The information need to be searched by people with the help of specialized tools known as search engines. Now a days lots of search engines are available. So, recovering meaningful information has become tough. Semantic web technologies keep a major contribution in order to recuperate significant information intellectually and to prevail the problem of search engines.

The signification of information are unerringly described in terms of clear lexicon by using the Semantic Web technology at the current Web [1], which is well understood by common people and computers. Semantic Web information is described with the new W3C standard known as Resource Description Framework (RDF). Semantic Web (web of data) mainly use Semantic Web Search as a search engine. Both people and computers use current web site to exactly allocate and gather information which are publicized on the Semantic Web. The ontology [2] is mostly used in the semantic web infrastructure which is most important concepts. Two W3C recommended data representation patterns used to convey ontologies are RDF(S) [18] (Resource Description Framework/Schema) and OWL [19] (Web Ontology Languages).

"An ontology is an manifest specification of a conceptualization" is a leading statement by T.R. Gruber in 1995 (Gruber 1995). The definition was then extended by R. Studer et al., in 1998 as "An ontology is an plain, formal

identification of a imparted conceptualization of a domain of interest". Current web technologies has some limitation such as integration and reuse of data, the Semantic Web will resolve this problem. It also resort more effective invention, automation, and provide facility for interoperability problem. The research on semantic web search engine are in the preliminary stage. The present markets of search engines are seized by usual search engines such as Bing (MSN), Google and Yahoo so forth.

Most of the search engines [20][21] is used to search the keywords. Web pages are searched by search engine to find required information. They use advanced algorithms [22] to filter the pages by searching unnecessary pages. Usually these kind of search engines responses successfully by manifesting state-of art algorithms. Search engine usually failed in replying intelligent questions from the user because their outcome depends on information provided in web pages. The primary goal of the search engines is to liquefy these queries quickly and with exact results using researched [25][26] methods. Accurate but unreliable results and improper outputs are provided with this approach. Due to lack of confidence on blogs the operator cannot have a gratification. Semantic web technology plays an important role in order to conquer this problem in search engines [3] to reclaim incidental and meaningful information intelligently. Search engines are used to gives the closer to intended results to the user with the help of semantic technology.

In our thesis we use the semantic web technology to retrieve information from the Google Map. We know Google Map is a web based service that provides detail information about

geographical region and sites around the world. Google map offers satellite views of many place. But it did not give any information about block area(place) or restricted zone due an unavoidable circumstance. The main focus of our thesis is to find out the blocked road due to the strike. District data is a domain that needs to create machine understandable. Our target is to Integrate the district data and a machine understandable Linked Data[16][17] using semantic knowledge.

The paper is designed as follows: Section one gives the introduction and section two ,three, four and five provides research motivation, literature review, proposed system and results and discussion respectively and finally section six provides the conclusion and future work..

2. RESEARCH MOTIVATION

Strike is a common phenomenon for Bangladesh. When strike or hartal are called by a party in a district the road within the district and also road that are directly connected to this district are blocked. This information can not be retrieve from google map(which road will be used and which not).

The main motivation of this thesis is to show the blocked road accurately and then forecasting blocked area more accurately than other methods. The key challenges of this research are to avoid the block road and to use the alternate solution. However, this paper proposes a technology to find out the block area.

3. LITERATURE REVIEW

By comparing with existing technology we find there are many challenges in searching information on the web for information retrieval. Due to the variation in search process and indexing different search results can be found from different search engine. After manipulating the keywords Google, Yahoo, and Bing are used to deal the searching. Search results are provided on the web page. Recently some researches are found that gives the results by semantics based search engines.

We know that a vast storehouses of content is current web . That have some limitation such as data are not semantically structure ,so it is very tough to understand the data by the machine that is provided to user. There are some research problems in search engine when the information was distributed in web i.e.

1.An intelligent and meaning full information does not retrieve even all information is available. What is the procedure that a search engine can convert query to file?

2.The searching results are distributed created by search engine by various file that are linked with hyperlink. How such a distributed results could be identified efficiently?

First make an intelligent and meaningful information by using searching technology and ontology, first problem[4][5] will be solve.

The linked search models [6] is used to solve the second problem. Semantic web would require to solve the hard problem like natural sentence understand. The figure that given below are the general semantic web framework.

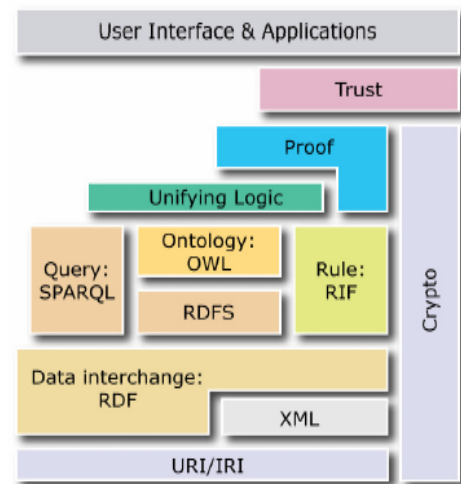


Fig -1: Semantic web framework

3.1 Present Web & Limitations

The vast global database [15] World Wide Web has some shortcoming and therefore it grows very hard for the machine to understand the data given to the user in the structure of search strings. As a results, ambiguous or partially suspicious result data set are given by the search engines.

The google map also don't have any Semantic district data. Semantic web is being used to integrate the following difficulties of present web.

- The web data has some infrastructure to represent the information.
- Reason for not having the interconnection of information gives the obscurity of information.
- Automatic information transfer lack age[24].
- For the weakness of a universal format machines is unable to understand the information that is provided.

The aim of semantic search[23] engine is to searched formed text a like Wikipedia. To give search output based on interpretation match a "meaning-based (semantic) search engine"[7] [8] are take an action, rather than by the vogue of search terms.

By using the semantic search engine first we have to understand the machine the meaningful location or place .In the figure we show many information. That are given in the google map using present web technology such as HTML or XML. But there is no semantic location based data. Our approach is to solve this by using semantic search engine

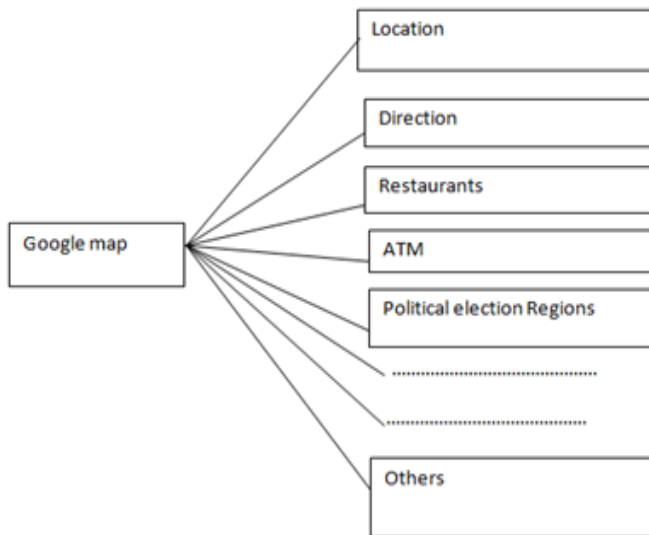


Fig -2: Layer of Information Retrieve from Google Map

4. OVERVIEW OF PROPOSED SYSTEM

Nowadays, various type of search engines are use and implemented for various work and also in many place and the approach that understand these search technique is individual. Our research focus on the combination of inference system and ontology to perform as an intelligent search engine [10]. Presenting demands can be solved by using the technology of search engine and that increase the efficiency of search engine.

Only text is use on the web page to search any content .Our methodology not only use for text but also gives the road information and street view .After finding the street view it also provides the safe and blocked area for any circumstance.

In the google map all information are given in different layer. That are shown in figure 2.Our research area are to choose one layer and retrieve all meaningful information from map to describe the layer semantically. Suppose we choose location. Here the location is different district data or divisional data. First of all find the latitude and longitude of any division to show it in the map as a marker that will a point notation. Then find out all point of this division to show the polygon.

To show the division as a marker and a polygon all point, latitude and longitude are to be trained to the machine. The whole information is given by RDF for machine understandable. RDF (Resource Description Framework) is represented as a number of triples consisting subject, predicate, and object. A format for storing and transmitting data is called triple.

After making the triple [13[14]]it is feed to the machine and the meaningful information are retrieve from map. In our research we show how from the divisional or district data (block road) area information can be retrieve. Figure 3 depict an overall structure of our research methodology.

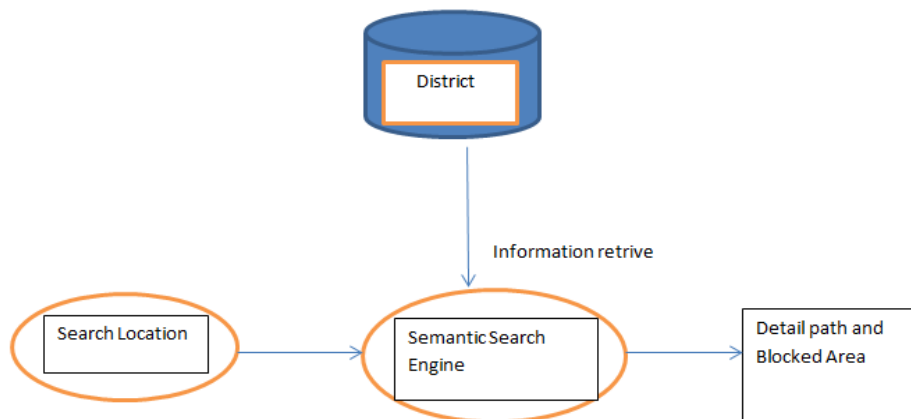


Fig -3: Proposed method

4.1 Proposed Methodology

The working procedure are summarize here for a divisional data.

1. Finding all point for a division to show it as a polygon.
2. Finding the latitude and longitude to show it as a marker shown in figure 4.

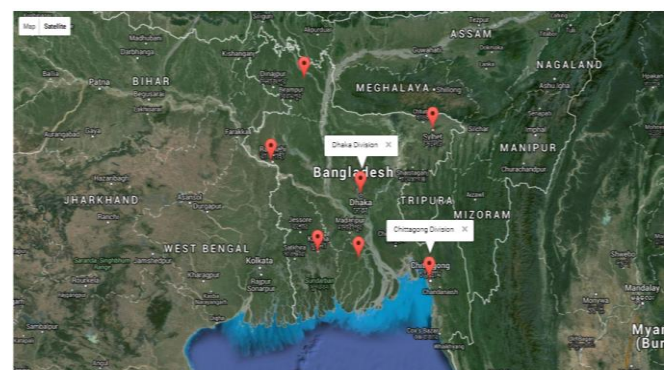


Fig-4: Location Data

- Findout all district for a division. If we consider for Dhaka division show all 17 district by latitude and longitude.
- Collect all information about a division which district are directly connected with which district in a division .It is most important because our research area focus on which district or district road are affected due to the strike on another district. We know if one district called strike the another district that are directly connected via street road

obviously affected. All transport will be stopped on this road and the adjacent area or road will be blocked .Our researches focus to show the blocked area due to strike or any unavoidable circumstance.

- Find an adjacent matrices to find out the region that are blocked. Some text type data are given to find out the adjacent matrices.

| আসুঞ্জজেলা ও দূরপাল্লা রুটের ভাড়ার তালিকা (গাবতলী)- জানুয়ারি/২০১৩ | | | | | | | | | | |
|---|-----------|--|---|---|-------------------|--|--|----------------------|--|---------|
| ক্রঃ নং | রুট নং | রুটের বর্ণনা | স ও জ হতে প্রাপ্ত দূরত্ব (কিমিঃ) | বাস ভাড়া | | | মিনিবাস ভাড়া | | | মন্তব্য |
| | | | | প্রতি কিঃমিঃ ১.৪৫ টাকা হিসাবে (যাত্রী আসন, চালক ব্যতীত ৫১ বিবেচনায়) | টোল | টোলসহ ভাড়া (প্রযোজ্য ক্ষেত্রে) | প্রতি কিঃমিঃ ১.৪৫ টাকা হিসাবে (যাত্রী আসন, চালক ব্যতীত ৩০ বিবেচনায়) | টোল | টোলসহ ভাড়া (প্রযোজ্য ক্ষেত্রে) | |
| ১ | ৬২ | ঢাকা (গাবতলী)- ফরিদপুর, ভায়া- পাটুরিয়া । | ১০৯.৭০ | ১৫৯.০৬ | ১৩৫০+৫১= ২৬.৪৭ | ১৮৫.৫৩ | ১৫৯.০৬ | ৫১৫+৩০ =১৭.১৬ | ১৭৬.২২ | |
| ২ | ৬৩ | ঢাকা (গাবতলী)- আলফাডালা, ভায়া- পাটুরিয়া, ফরিদপুর । | ১৫৮ | ২২৯.১০ | ১৩৫০+৫১= ২৬.৪৭ | ২৫৫.৫৭ | ২২৯.১০ | ৫১৫+৩০ =১৭.১৬ | ২৪৬.২৬ | |
| ৩ | ৬৪ | ঢাকা (গাবতলী)- মোকসেদপুর, ভায়া- পাটুরিয়া, ফরিদপুর । | ২০৩ | ২৯৪.৩৫ | ১৩৫০+৫১= ২৬.৪৭ | ৩২০.৮২ | ২৯৪.৩৫ | ৫১৫+৩০ =১৭.১৬ | ৩১১.৫১ | |
| ৪ | ৬৫ | ঢাকা (গাবতলী)- বরিশাল, ভায়া- পাটুরিয়া, ফরিদপুর, মাদারীপুর । | ২৪২ | ৩৫০.৯০ | ১৫৫০+৫১= ৩০.৩৯ | ৩৮১.২৯ | ৩৫০.৯০ | ৬১৫+৩০ = ২০.৫০ | ৩৭১.৪৪ | |
| ৫ | ৬৬ | ঢাকা (গাবতলী)- পটুয়াখালী, ভায়া- পাটুরিয়া, ফরিদপুর, মাদারীপুর, বরিশাল । | ২৭৮ | ৪০৩.১০ | ১৮৬৫+৫১= ৩৬.৫৬ | ৪৩৯.৬৬ | ৪০৩.১০ | ৭৫১+৩০ = ২৫.০৩ | ৪২৮.১৩ | |
| ৬ | ৬৭ | ঢাকা (গাবতলী)- মাদারীপুর, ভায়া- পাটুরিয়া, ফরিদপুর । | ১৭৭ | ২৫৬.৬৫ | ১৩৫০+৫১= ২৬.৪৭ | ২৮৩.১২ | ২৫৬.৬৫ | ৫১৫+৩০ = ১৭.১৬ | ২৭৩.৮৩ | |
| ৭ | ৬৯ | ঢাকা (গাবতলী)- রাজবাড়ী, ভায়া- পাটুরিয়া, গোয়ালন্দ । | ১০৮ | ১৫৬.৬০ | ১৩৫০+৫১= ২৬.৪৭ | ১৮৩.০৭ | ১৫৬.৬০ | ৫১৫+৩০ = ১৭.১৬ | ১৭৩.৭৬ | |
| ৮ | ৭০ | ঢাকা (গাবতলী)- যশোর, ভায়া- পাটুরিয়া, ফরিদপুর, ঝিনাইদহ । | ২৩২ | ৩৩৬.৪০ | ১৪৩০+৫১= ২৮.০৩ | ৩৬৪.৪৩ | ৩৩৬.৪০ | ৫৫৫+৩০ = ১৮.৫০ | ৩৫৪.৯০ | |

Fig-5: Real Data Set

- After finding all district information ,all adjacent matrices for a division create a RDF by triple.

If we want to visit Dhaka to Netrokona then the direction road will be like that.Dhaka-
→Joydebpur→Baluka→Mymensing→Netrokona

Here we will consider only Mymensing and Netrokona because this are district and part of Dhaka division.

So we can say the adjacent matrix Dhaka →Mymensing =1
Dhaka →Netrokona = 0 , Mymensing →Netrokona = 1

And by using this method we can create the adjacentmatrices for the 17 district of Dhaka division.

- A triple example are given here for a division.
- Now feed the triple to the machine to understand and find out the adjacent area.
- Finally if you click on the marker of a district on a division the blocked area can be shown.

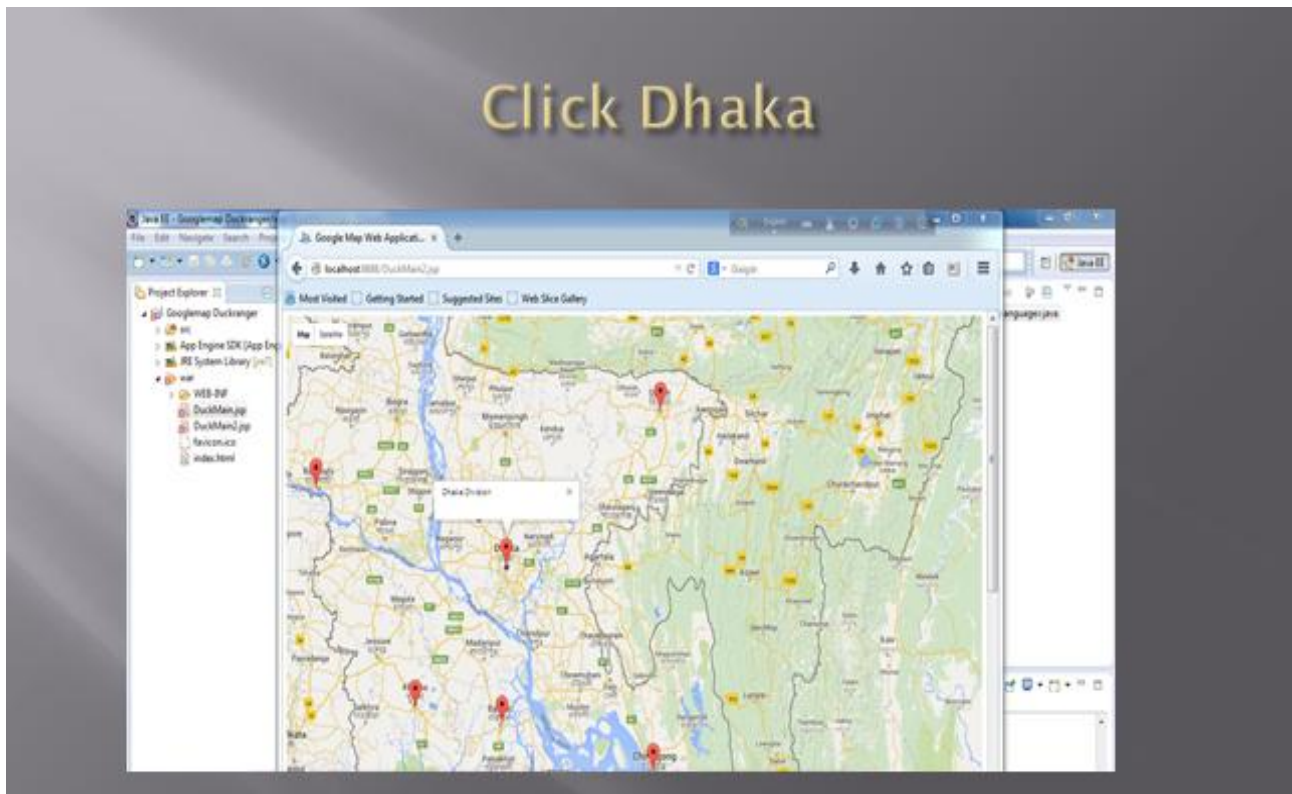


Fig-6: Divisional data

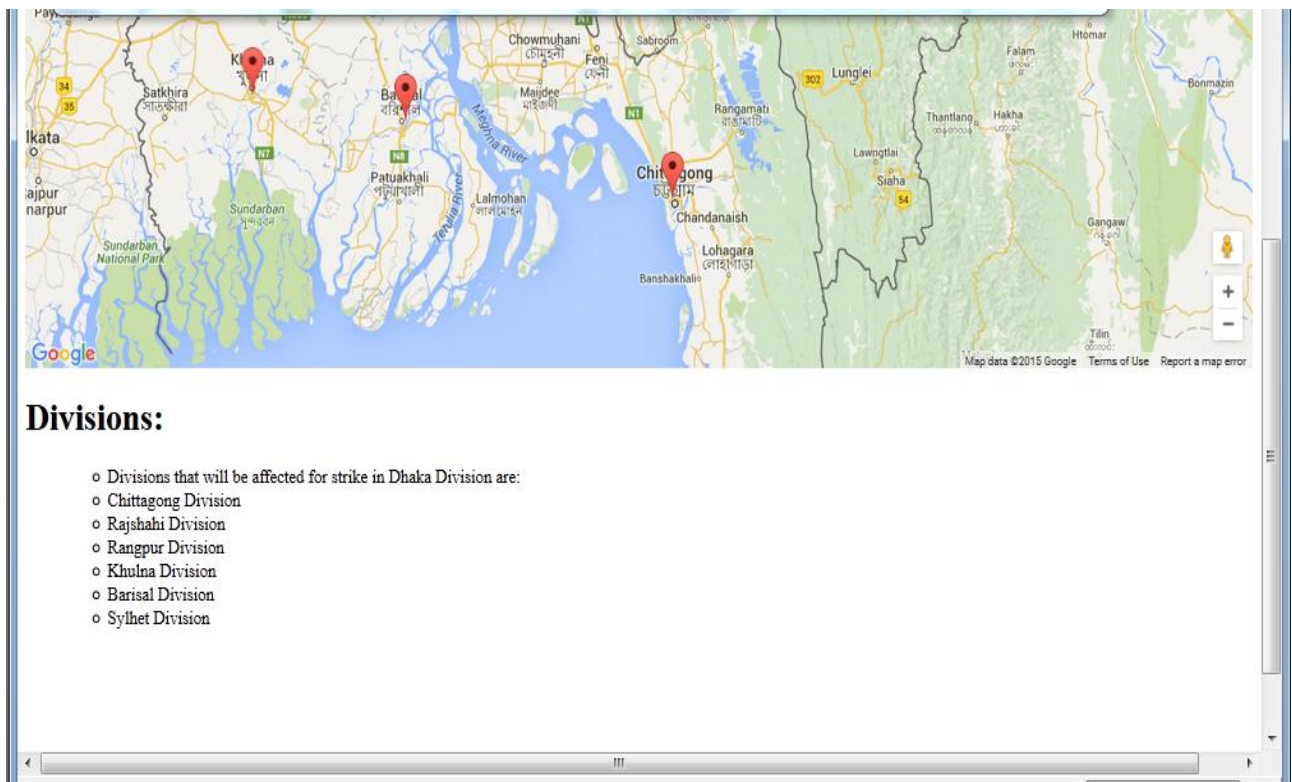


Fig-7: Blocked Area

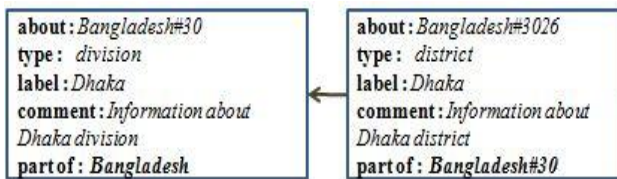
5. RESULTS AND DISCUSSION

In our research, we used a real dataset which was collected from BRTA. This data is given real street view in Bangladesh. For making decision about blocked road, we

collected real traffic data .First of all we have to convert this data into RDF format .Then feed to the machine so that the machine can understand. Machine Understandable dataset are given in table 3.

Table-3: Rdf Representation

| Subject | Predicate | Object |
|-----------------------------|---|---|
| <http://www.skeim.org#30> | <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> | <http://www.skeim.org#division>. |
| <http://www.skeim.org#30> | <http://www.w3.org/2000/01/rdf-schema#label> | "Dhaka @en". |
| <http://www.skeim.org#30> | <http://www.w3.org/2000/01/rdf-schema#comment> | "Information about Dhaka division @en". |
| <http://www.skeim.org#30> | <http://www.w3.org/2003/01/geo/wgs84_pos#lat> | "24.16667@en". |
| <http://www.skeim.org#30> | <http://www.w3.org/2003/01/geo/wgs84_pos#long> | "90.41667@en". |
| <http://www.skeim.org#30> | <http://www.w3.org/2000/01/rdf-schema#partOf> | <http://www.skeim.org#Bangladesh>. |
| <http://www.skeim.org#3026> | <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> | <http://www.skeim.org#district>. |
| <http://www.skeim.org#3026> | <http://www.w3.org/2000/01/rdf-schema#label> | "Dhaka@en". |
| <http://www.skeim.org#3026> | <http://www.w3.org/2000/01/rdf-schema#comment> | "Information about Dhaka district @en". |
| <http://www.skeim.org#3026> | <http://www.w3.org/2003/01/geo/wgs84_pos#lat> | "23.81092@en". |
| <http://www.skeim.org#3026> | <http://www.w3.org/2003/01/geo/wgs84_pos#long> | "90.36542@en". |



| Bangladesh | Dhaka District | Dhaka Division |
|-----------------|----------------|----------------|
| Geo-code | 26 | 30 |

Fig:8. Research motivation

```

rdf representation - Notepad
File Edit Format View Help
<html xmlns = "http://www.w3.org/1999/xhtml"
xmlns:foaf="http://xmlns.com/foaf/o.1/"
xmlns:dc="http://purl.org/dc/term/"
xmlns:rdf = "http://www.w3.org/1999/02/22 -rdf-syntax-ns#"
xmlns:rdfs = "http://www.w3.org/2000/01/rdf-schema#">
<head>
<meta name = "dc.title" content="Information retrieve from map"/>
<meta name="dc.creator" content="lutfun Nahar,Shima chakbort1,Md Hanif seddiqui"/>
<title property="dc.title">Division of Bangladesh</title>
</head>
<body>
<rdf:Description rdf:about = "http://www.skeim.org/Division#3026">
<h2 property="rdfs:label">Dhaka@en</h2>
<p><b>Lacation:</b><a property =dc:lacation href=http://www.skeim.org/
Geo-BD#3026>Latitude and longitude</a></p>
</body>
</html>
    
```

Fig-9: Rdf and xml data

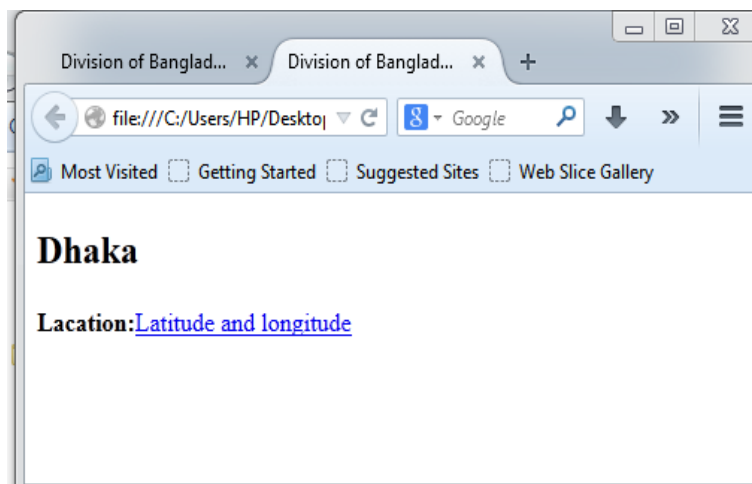


Fig-10: Research output link data representation

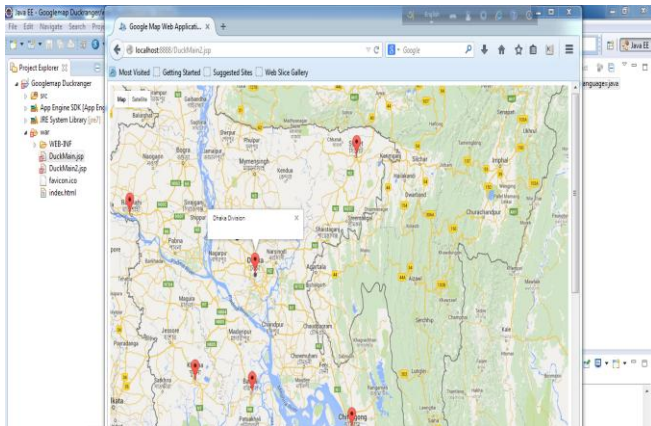


Fig-11: Research output

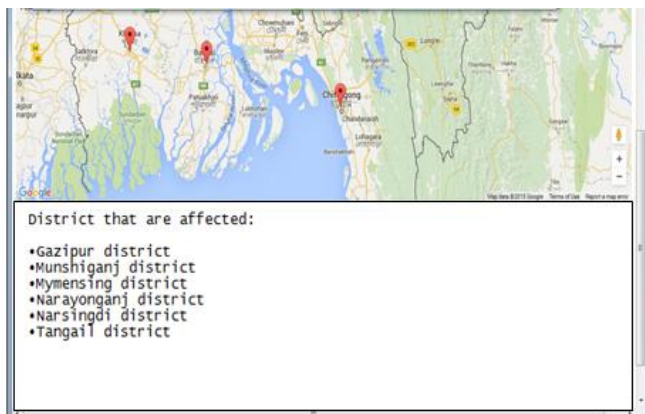


Fig.12. Research output

6. CONCLUSION AND FUTURE WORK

In these researches, we proposed a search that retrieve information from google map due to an unavoidable circumstance and solve the limitation of present search technology. We summarize the semantic search methodology for information retrieval here.. Finally we can show the block road and area due to strike. It is very effective work because such type of information is not yet added in the google map. People who want to travel along one district to another, after calling strike in one district they can find easily the safe road or area. And they can also see the blocked road or area by which they cannot travel. This make easier to take decision about traveling and also safe our life and time.

In future, the goal of our work will be more study about semantic search to find the alternative road so that we can reach in our destination. If one road is blocked then the area that are directly connected will obviously blocked. Our future work will be focus on how it will be possible to go the blocked district area by using another road .If possible than show all path and road to reach this district.

REFERENCES

- [1] Berners-Lee, T., Hendler, J. and Lassila, O. "The Semantic Web", *ScientificAmerican*, May 2001.
- [2] Deborah L. McGuinness. "Ontologies Come of Age". In Dieter Fensel, J im Hendler, Henry Lieberman, and

Wolfgang Wahlster, editors. *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*. MIT Press, 2002.

[3] Ramprakash et al "Role of Search Engines in Intelligent Information Retrieval on Web", Proceedings of the 2nd National Conference; INDIACom-2008.

[4] T. Berner-Lee and M. Fishetti, *Weaving the web "chapter Machines and the web," Chapter Machines and the web*, pp. 177-198, 1999.

[5] D. Fensal, W. Wahlster, H. Lieberman, "Spanning the semantic web: Bringing the worldwide web to its full potential," MIT Press 2003.

[6] G. Bholotia et al.: "Keyword searching and browsing in database using BANKS," 18th Intl. conf. on Data Engineering (ICDE 2002), San Jose, USA, 2002.

[7] D. Tümer, M. A. Shah, and Y. Bitirim, An Empirical Evaluation on Semantic Search Performance of Keyword-Based and Semantic Search Engines: Google, Yahoo, Msn and Hakia, *2009 4th International Conference on Internet Monitoring and Protection (ICIMP '09)* 2009.

[8] "To 5 Semantic Search Engines" <http://www.pandia.com/>

[9] H. Dietze and M. Schroeder, GoWeb: a semantic search engine for the life science web. *BMC bioinformatics*, Vol. 10, No. Suppl 10, pp. S7, 2009.

[10] S. A. Inamdar¹ and G. N. Shinde "An Agent Based Intelligent Search Engine System for Web mining" Research, Reflections and Innovations in Integrating ICT in education. 2008.

[11] Fu-Ming Huang et al. "Intelligent Search Engine with Semantic Technologies" International journal of Web & Semantic Technology (IJWeST) Vol.2, No.1, January 2011

[12] Li Zhan, Liu Zhijing, , ' Web Mining Based On Multi-Agents ', COMPUTER SOCIETY, IEEE(2003)

[13] Patrick Lambrix et al, " Dwebic :An Intelligent Search Engine based on Default Description Logics"-1997.

[14] K. Satya Sai Prakash and S. V. Raghavan "Intelligent Search Engine: Simulation to Implementation", In the proceedings of 6th International conference on Information Integration and Web-based Applications and Services (iiWAS2004), pp. 203-212, September 27 - 29, 2004, Jakarta, Indonesia, ISBN 3-85403-183-01.

[15] Xiaijiong Shen Yan Xu Junyang Yu Ke Zhang "Intelligent Search Engine Based on Formal Concept Analysis" IEEE International Conference on Granular Computing, pp 669, 2-4 Nov, 2007.

[16] Dan Meng, Xu Huang "An Interactive Intelligent Search Engine Model Research Based on User Information Preference", 9th International Conference on Computer Science and Informatics, 2006 Proceedings, ISBN 978-90-78677-01-7.

[17] Sanjib kumar, Sanjay kumar malik "TOWARDS SEMANTIC WEB BASED SEARCH ENGINES" National Conference on "Advances in Computer Networks & Information Technology (NCACNIT-09) March 24-25

[18] F. F. Ramos, H. Unger, V. Larios (Eds.): LNCS 3061, pp. 145-157, Springer-Verlag Berlin Heidelberg 2004.

[19] Cohen, S. Mamou, J. Kanza, Y. Sagiv, Y "XSearch: A Semantic Search Engine for XML" proceedings of the international conference on very large databases, pages 45-56, 2003.

- [20] D. Bhagwat and N. Polyzotis, "Searching a file system using inferred semantic links," in *Proceedings of HYPERTEXT '05* Salzburg, 2005, pp. 85-87.
- [21] H. L. Wang, S. H. Wu, I. C. Wang, C. L. Sung, W. L. Hsu, and W. K. Shih, "Semantic search on Internet tabular information extraction for answering queries," in *Proceedings of CIKM '00* McLean, 2000, pp.243-249.
- [22] A. Maedche, B. Motik, L. Stojanovic, R. Studer, and R. Volz, "An infrastructure for searching, reusing and evolving distributed ontologies," in *Proceedings of WWW '03* Budapest, 2003, pp. 439-448.
- [23] E. Kandogan, R. Krishnamurthy, S. Raghavan, S. Vaithyanathan, and H. Zhu, "Avatar semantic search: a database approach to information retrieval," in *Proceedings of SIGMOD '06* Chicago, 2006, pp. 790-792.
- [24] www.georges.gardarin.free.fr/Articles/Sewise_NLDB2003.pdf.
- [25] D. Ding, J. Yang, Q. Li, L. Wang, and W. Liu, "Towards a flash search engine based on expressive semantics," in *Proceedings of WWW Alt.'04* New York, 2004, pp. 472-473.
- [26] Chiung-Hon Leon Lee, Alan Liu, "Toward Intention Aware Semantic Web Service Systems," *scc*, vol. 1, pp.69-76, 2005 IEEE International Conference on Service Computing (SCC'05) Vol-1, 2005.

BIOGRAPHIES



Lutfun Nahar, M.S (Eng.) and B.Sc (Engg) in CSE from Chittagong University(CU). Lecturer at the Department of CSE in the International Islamic University Chittagong, Bangladesh



Rahima Afrose, B.Sc (Engg.) in CSE from International Islamic University Chittagong (IIUC). Lecturer at the Department of CSE in the International Islamic University Chittagong, Bangladesh



Sanjida Sharmin, B.Sc (Engg.) in CSE from Chittagong University of Engineering and Technology (CUET). Lecturer at the Department of CSE in the International Islamic University Chittagong, Bangladesh