

IMPROVEMENT IN THE USABILITY OF GIS BASED SERVICES BY APPLYING SEMANTIC ONTOLOGY

Abhijit Manepatil¹, P.R.Barapatre²

^{1,2}Department of Computer Engineering, University of Pune, SKN Sinhgad Institute of Technology & Sciences, Lonaval, Pune, Maharashtra, India

Abstract

After the invention of GIS Based services in the web and the desktop areas it becomes the most widely used services over the internet and because of its tremendous capabilities of giving accurate and most efficient information about GIS related queries of particular organizations, for users or industries these services are on the peak of popularity. Now a days there are lots of options available in the market who provides these GIS related functionalities to the users such as for e.g. ArcMap, GRASS and Google map APIs etc.

This system improves the user usability on geo services. Purpose of this system to develop a user convenient system related to GIS that help any normal user to handle it in its natural way. User do a spatial thinking in his/her mind according to that some natural question where generated from this system take that questions as it is and make that questions. Filter those questions in a way that system can understand it and produce an exact result. Semantic concept gives a common and meaningful data format by integrating this concept in GIS which will come to know as Geosemantic. OWL files provide well defined relationship between the object by mapping with real word things.

Keywords— GIS, Semantic, Geosemantic, OWL, Spatial thinking

1. INTRODUCTION

Natural language is the language that can be used by the any normal user. Language in which user can think. Here the system concerned with the spatial thinking of normal user with the natural language. According to natural thinking of the user whatever the questions comes in users mind in a general way that questions are directly handled by the system and which gives the exact answer according to the user's need, according to user's thinking. Means we develop a system in such a way that it can think according to user's point of view. So the user can easily deal with the system which is related with the geographical information.

In this system some methods and techniques are used to evaluate a user's questions. Such as stop word removal, Stemming, Tokenization where this methods helps to evaluate a user's question in a machine understandable form. OWL file provides a system to define an object with a meaningful way which provides a relationship with a real world object. Where protégé tool helps to develop OWL files. Ontology framework defines a certain rules and regulations. Finally system can map these object with the geographical information, geographical services. This geoinformation provided by certain GIS tools such as ArcMap, Google map etc.

Application should be designed in a way so that user can execute any service with minimum key strokes. User should be free from following methods, syntax and predefined rules

and regulation to finds out required geographical information. System's different modules are communicating with one another on the following scenarios

- i. From User Question layer module to semantic layer module
- ii. From semantic layer module to implementation layer module.
- iii. From implementation layer module to answer extraction module.

User Interfaces for this system interacts with user while giving Spatial Questions as input to the system, while giving determining previous reports and answers.

Hardware Interfaces for this system interacts with secondary storage memory while reading OWL Ontology Files for semantic analyzing. System also interacts with secondary storage memory while accessing GIS API or tools.

2. LITERATURE SURVEY

As the technology grows up with new system, new methods and new models. It's all main aim is that to provide an easy way to handle that thing. User should become very familiar, friendly with the developed system that's the main of every technology. So that there will be more interaction between the user the computer. Computer can understand the users need and works according to that and provide a better solution to the user.

In web technology as the beginning to the web at the starting era of the web technology the very first step was web 1.0 were data, records, things were seen over the web from the anywhere in the world that's just a staring era of web technology. Where user can only view contain, only see the paragraphs and stable images. That was known as the static era. In that there were no interaction between the user and the data presented over web. User only see contain but cannot modify, cannot comment on that. User were not allowed to puts its view against the contain of web. That the main drawback of web 1.0.

So to overcome the drawback of the web1.0 new technology were invented in web technology that is Web 2.0 which we are using in today's web. In web 2.0 data presented over the web come up with the dynamic web pages. Which provide the more interaction between the user and the computers. Where user can deal with the web of the data. The data become more interactive to the user so user can easily deal with the system. User can easily handle the web relates thing in its own way. It give a more efficient way to the user to interact, to understand the web things in an easy manner. But their need to be more required to exist more interaction user and the computer.

To make a more and well interaction between the users and the computer the new technology Web 3.0 (Semantic Web) were invented by a Tem bernals lee who was the inventor of the WWW. In with he defines a structure to the web of data. Which provides a common platform for the data which will going to upload on the web. Which come up with the RDF (Resource Description Framework) structure which describes the each sentence in a tuples that is object-predicates-subject. RDFS gives the more hierarchical structure about the object-predicates-subject. Which helps to link each word in to a real world things. So each word get certain meaning with the real world object that can help to find out it in an easy manner and we will get the exact result(exact meaningful word) according to users request. XML is used to manage the thing in to a user defined tags so that each word will de get unique identification over the document that document over the web will be uniquely identify by the URL. Ontology provides the particular meaning to the web of data. Ontology makes a hierarchical structure of the data by linking them with the class, subclass. With the certain syntax presented to define ontology.

So, we come to know that semantic web technology can defines the web of data in a well manner that can be easily understand by the user. It is also known as intelligent web. Where machine can itself understand the data and provide an exact result according to users need. That will helps user to deal with the data presented over web. Which provides the easier way for user to handle the web in an efficient way. Now, we are applying this semantic concept with the Geoinformatic relates information so which will come to know that Geosemantic technology. Which is the combination of

Geo related concepts, structures and the sematic technology concepts.

Geosemantic technology provides the concepts of spatial thinking which is the normal users thinking related to the geographical information. Spatial thinking is the structure were user's point of related to the GIS, Geoinformation etc.

Development of intelligent geographical information systems. User acceptance criteria of any GIS software in which user compares the analysis of geographical information with other production processes in which a user can only accept an end-product if it meets certain quality requirements. Defined approach of this paper suggest some acceptance criteria for GIS software's on the basis of some important points such as (1) the data requirements, sensitivity and error propagation in models. (2) Data collection, level of resolution and quality; and (3) the use of basic analytical functions of the geographical information system. Conclusion is that the intelligent geographical information system would present a range of alternative strategies—better methods, more data, different data, different models, well model calibration, or better spatial resolution—and their costs by which the user's aims could reasonably be achieved.

Design of a geographical information system for a heterogeneous scientific community. Accepted approach of this paper basically focuses on the study of user needs for a geographical information system (GIS) within the community of scientists in the Natural Environment Research Council (NERC).Which Includes a description of the user needs study is followed by a discussion of the relevant characteristics of spatial data structures and how these match the functionality identified as needed and in the form of the ideal user interface, need for GIS training facilities. At the end Based upon this analysis the author defines, a conceptual design for a NERC GIS is described, together with a review of the commercial GIS on offer and of the literature, and a consideration of both the rapid evolution of technology and those constraints imposed by-other decisions by the Council.

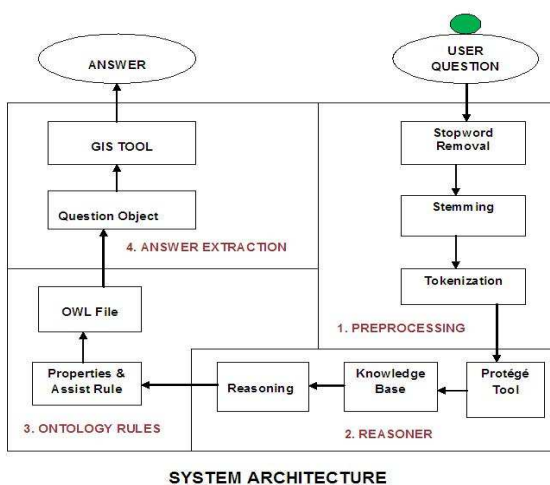
Ontology-driven geographic information integration: A survey of current approaches. In this paper the authors Agustina Buccella, Alejandra Cechich, Pablo Fillotrani analyzes and compares the geographic information integration, focusing on those using ontology's as semantic tools to represent the sources, and to facilitate the integration process. The authors describe the need of Integrating different information sources is a growing research areas particularly in GIS Application Software's. They focuses our attention towards the GIS (Geographic Information Systems) on the Web which is leading to the proliferation of different geospatial information repositories and the subsequent need to integrate information across repositories to get consistent information and to overcome them they suggest concept of ontology's to be included in the integration processes of GIS in to the web

Observation-Driven Geo-Ontology Engineering This paper the author Krzysztof Janowicz studied some research trends that surfaced over the last years such as Big Data, Linked Data, Smart Dust, Digital Earth, and e-Science etc. further he described that these trends addresses different visions and need yet they are having common problems in some areas such as How do we manage massive amounts of heterogeneous data, derive knowledge out of them instead of drowning in information, and how do we make our findings reproducible and reusable by others? For solving these problems he suggested the solution that these trends require a radical paradigm shift in ontology's. Author also describe the concept of global ontology's developed top-down, to a high number of local ontology's that are driven by application needs and developed bottom-up out of observation data. Author also focuses on point that existing ontology engineering frameworks are not well suited for this new perspective and thus he proposed an observation-driven ontology engineering framework which conclude that how its layers can be realized using specific methodologies, and relate the framework to existing work on geo-ontology's .

3. IMPLEMENTATION DETAILS

The basic ideas of designing a question-based user interface that integrates different levels of ontologies (spatial concept ontology, domain ontology and task ontology) to guide the process of extracting the core spatial concepts and translating them into a set of equivalent computational or operational GIS tasks. We also list some typical spatial questions that might be posed for spatial analysis and computation. The principle introduced in this paper could be applied not only to desktop-GIS software but also to web map services.

For convenient understanding of the system we are dividing our system is in to four categories Preprocessing, Reasoner, Ontology Rules and Answer Extraction which are explained in details.



User question:

User is the most important part of any system so as here, we developing our system according to user's point of view. We are making our system in a way that user can easily interact with that and get the appropriate result according to his/her request. We cannot predict what type of user will approaches to our system. User may be knowledgeable person or any unknowledgeable person.

In our system we are making our system for users convenient. Here user can deal with the geographical information so before user going to interact with the geographical system before start using geospatial services. User thinking on that how to interact with that in our system it is related to geographical information. So user think on geographical related information i.e. how to use? What is happening where? etc. So according to the users thinking on geo related some questions are arrives is user mind. He/she thinking on geographical information in his/her own natural language. This is a spatial thinking of a user.

So that first input or data set for our system is that user natural thinking in the form of natural/general questions. These questions are further handle by the other parts of system which described in detail below.

3.1 Preprocessing

Preprocessing is the first step in a system which takes the user input in the form of natural language and evaluate that input in to an appropriate format give to the next step of system which is Reasoner. Preprocessing perform the basic operation on users input question. Which make that input into a machine understandable form. Certain steps involved in the Preprocessing which are described below.

Stop word removal: in this whole sentence in the form of questions will be filtered out. So that only useful and meaningful words will be sorted out from a process of stop word processing. Stop word removal sorted those words and removed that from original question given from a user. Typically stop word lists contain words that don't carry as much meaning, such as determiners and prepositions Words like *the, is, at, which, and on*. Due to stop word removal process we get only main required worlds which are really useful for answering to the user's question. By using this process we are improving the answering performance of system. The word in a whole query of users question where words are preceded by the plus sign (+) defined as the stop words. We are using this stop word removal process because that words are not indexed in webpages and thus are not used in search engine queries. At the end stop word removal module gives us an only definite words which is useful for further execution process.

Stemming: A stemming is a process of linguistic normalization, where variant forms of a word are reduced to a

common normal form. We can use a stemming for increasing a performance of the system to provide an exact result from a system. Variables having a ending part or any suffix taking of it is known as a Stemming. For example -ion, -ions, -ive, -ed, -ing. We search through this suffix by using a stemming structure where we are using certain rules and regulations to find out stemming part.

Stemming is a process of removing prefixes and suffixes from words. By using stemming process we can combine word forms to avoid mismatches that may be occurs during a searching time. Which make a system more understandable according to user's point of view.

Tokenization: Firstly we are having a raw state which is our systems inputs comes from user's minds in the form of natural language. By using a tokenization process without changing its meaning whole text, it is segmented into sequential manner of words and sentences which represent a token. A *token* is an instance of a sequence of characters in some text that are grouped together as a useful semantic unit for processing.

3.2 Reasoner:

Protégé tool: Once we got the token of user's natural language spatial thinking question. We are providing that token to the protégé tool. We are having tokens in the sequential, hierarchical forms so by using that sequence protégé tool makes an OWL structure of that token. That is a Web Ontology language. OWL is an ontology language formally defined for the *Semantic Web* where it provides a meaning to the data represented on it. This OWL provides a classes, properties, individuals and data values which will be stored as a Semantic web documents. That document contain a RDF(Resource Description Framework) structure where it provides a common platform for the integration of data. With a RDFS structure each token within a sentence can be map with a real world object. In that it make a relationship between current object and existing object. Which makes a whole data presented over web in a suitable formant by linking them together with the meaningful related existing object. (XML). This protégé tools helps to load and save OWL and RDF ontology that allows us to define, edit, and visualize classes according to tokens. Which execute reasoners such as description logic classifiers.

Knowledge base: This system can solve a difficult problem in to an appropriate manner. Knowledge based system is a productive Artificial Intelligence system. It adds powers to the solution and concentrates on effective solution. It provides a knowledge based processing approach.

3.3 Ontology Rules :

Ontology share a similarities between different structures. It describes instances or objects, concepts, attributes and relationship in between classes and objects. Where rules are

the statements in the form of conditions where relationship is defined like if then etc. by defining rules we are making a relationship between classes and subclasses. Here the objects are classified by classes and is-a-subclass defines a further classification between object and relationship between them with the help of taxonomy.

3.4 Answer Extraction:

Whatever the object extracted from OWL files that are provided to a GIS tools which provides us a geographical information. According to that object, according to user's questions exact objects are filtered out and due to which exact query will be given to a GIS tool so that exact result will be generated according to user's request. That answer in the form of any GIS service which will going to show to a user. Here is the definite result comes from a normal users point of view, from the normal users natural thinking about geoinfromation. This is a form of Geosemantic concept

Algorithm 1 GIS QA SYSTEM

```

Input:
 $Q_n$  is user question;
 $D_e$  is Dictionary;
Token file containing Token, supporting token, related answer string and NLP rule;
File system are the documents containing parsed web page contents.

Output:
Short string or list

1: procedure QA( $Q_n, D_e$ )
2:                                      $\triangleright Q_n$  Question,
3:                                      $\triangleright D_e = \{d_1, \dots, d_n\}$  Dictionary words
4:    $q_w = \text{preprocess}(Q_n)$ 
5:   Set  $M_e = \{q_w, s_w, a_w\}$ 
6:                                      $\triangleright q_w =$  query words after processing of  $Q_n$ 
7:                                      $\triangleright M_e =$  Master vector,
8:                                      $\triangleright a_w =$  Related answer strings
9:                                      $\triangleright s_w$  are supporting Token or query words
10:  for  $d \leftarrow 1, n$  do
11:    for  $i \leftarrow 1, k$  do
12:      if  $M_e \in S_i$  then
13:        Tag  $S_i$  as  $S_{imp}$ 
14:        end if
15:      end for
16:    end for
17:  for  $i \leftarrow 1, n_o$  do
18:    if  $q_w \in D_e$  then
19:      Tag  $q_w$  as Token
20:    else
21:      Tag  $q_w$  as Proper noun  $P_n$ 
22:    end if
23:  end for
24:  for  $i \leftarrow 1, n_{osen}$  do
25:    if  $S_{imp} \in R_i$  then
26:      termweight(Pronoun  $P_n \in S_{imp}$ )  $\rightarrow A_n$ 
27:    end if
28:  end for
29:  Rank  $A_n$  and display
30: end procedure

```

This algorithm explain a step by step process of an execution. When user provides a question in the form of natural language by the process of stopword removal, stemming, tokenization whole question will be filtered. Token file store a sequence of words. Through which token words are relate with a particular meaning. That is a data file that contain more information related to that object e.g. Distance→Km, Height→ft etc. where ontology provides a hierarchical structure. That can be handle with the procedural program.

4. RESULTS

4.1 Data Set & Result Set

Data Set:

Our system deals with the normal user's behavior. Spatial natural thinking of the any type of user having some natural, normal questions in mind so he/she directly provided with the platform where he/she directly throw a questions, whatever comes in his/her mind by spatial thinking and system can understand this natural human behavior and evaluating that questing defining into an appropriate format so that machine can understand and which gives an exact results according to the uses need.

So we are having data set with the natural human minds questions related to the spatial thinking.

Geographical related information required to search the answer for users question that can be provided through a different GIS tools like Archmap, Google map etc.

We are putting user's spatial thinking question into machine understandable form that can be done through an Ontology tool like protégé. Which define a data in in OWL form which contain RDF (Resource description framework) and XML structure. Which help to define a data in a format that machine can compile and understand. And the appropriate result will be created in a result set.

Result Set:

Our systems main is to produce a result according to the users need and where user get exact result with its natural behavior by natural thinking. Here any normal user can easily deal with the geographical functionality in its own way. Our result will satisfies a usability functionality of system where user feels that system is easy to use. Enjoyable, our final result is user's satisfaction. Actual result in the form of exact GIS service which is requested by a user.

5. CONCLUSIONS

The contribution of this work is to provide a better platform for any normal user according to his/her spatial thinking in a natural way that can be easily handle by system which leads to generate exact result according to user's request. Normal questions that comes in user's mind that can be handle well that uses a semantic and Geosemantic framework for designing new GIS user interface.

In Future work our topic leads to generate a new concept on artificial intelligence in area of geoinformation. Our framework has different ontology structure that useful in further research and improvement on spatial reasoning in web. Which will leads to develop a more powerful data over web.

This proposed system leads to development into Intelligent GIS application.

ACKNOWLEDGEMENTS

We would like to thanks to our guide & respected teachers for their constant support and motivation for us. Our sincere thanks to Sinhgad Institute of Technology for providing a strong platform to develop our skill and capabilities.

REFERENCES

- [1] P. A. Longley, M. F. Goodchild, D. J. Maguire, and D. W. Rhind, *Geographic Information Systems and Science* (Third Edition). Hoboken, NY: Wiley, 2011.
- [2] M. F. Goodchild, "Spatial thinking and the gis user interface," *Procedia-Social and Behavioral Sciences*, vol. 21, 2011.
- [3] F. T. Fonseca, M. J. Egenhofer, P. Agouris, and G. Cmara, "Using ontologies for integrated geographic information systems," *Transactions in GIS*, vol. 6, 2002.
- [4] D. Mark, M. Egenhofer, S. Hirtle, and B. Smith, "Ontological foundations for geographic information science," *A research agenda for geographic information science*, 2004.
- [5] Y. Gao, S. Gao, R. Li, and Y. Liu, "A semantic geographical knowledge wiki system mashed up with google maps," *Science China Technological Sciences*, vol. 53, 2010.
- [6] K. Janowicz, "Observation-driven geo-ontology engineering," *Transactions in GIS*, 2012.
- [7] N. Guarino, "Semantic matching: Formal ontological distinctions for information organization, extraction, and integration," in *Information Extraction a Multidisciplinary Approach to an Emerging Information Technology*, M. Paziienza, Ed., 1997.
- [8] R. G. Golledge, "Primitives of spatial knowledge," in *Cognitive aspects of human-computer interaction for geographic information systems*. Springer, 1995. *Information Systems and Science* (Third Edition). Hoboken, NY: Wiley, 2011.