

# A STUDY ON ONTOLOGY-DRIVEN GEOSPATIAL-INFORMATION RETRIEVAL SYSTEMS IN THE SEMANTIC WEB

Xu Quan-Li<sup>a</sup>, Yang Kun<sup>\*b</sup>, Wang Jun<sup>a</sup>, Peng Shuang-Yun<sup>b</sup>, Yi Jun-Hua<sup>c</sup>

<sup>a</sup>School of Power Engineering, Kun Ming University of Science and Technology, Kunming, NO.50, East Ring Road, 650031, China;

<sup>b</sup>Faculty of Tourism and Geographic Science, Yunnan Normal University, NO.298, Dec. 1st Street, Kunming, 650092, China;

<sup>c</sup>The Institute of Potogrammetry and Remote Sense of Yun Nan Province, Kunming, 650034, China

**KEY WORDS:** Semantic Web, Geo-Ontolog, Intelligent Search Engine, Semantic Web Services, Geography Cognize, Processing platform of Knowledge

## ABSTRACT:

Firstly, this paper expounded that the traditional searching based on keywords has a severe deficiency on Recall and Precision because the computer can't understand what information mean efficiently. In Succession, through extending from ordinary information to Geo-spatial information, on the support of geo-ontology theory and semantic web services technology, this paper discussed the principle, methods and application of ontology-driven geo-spatial information retrieving mechanism in semantic web. In other word, How to structure the framework for running this mechanism, how to build the ontology information models and how to establish the ontology-driven services models are discussed in detail in context. A design of systems for implementing the mechanism and experiment are given at the end of the text

## 1. INTRODUCTION

As the time going to Today, which is a speedy developing society for Web Information, the traditional searching tools based on Keywords can't provide satisfying service for people's searching in Internet more and more because of those limitation in Recall and Precision[1]. And what leads to the result? There are two main reasons. Firstly, Web Information is deficient in semantic description and expression. In the other words, information is not knowledge. And the second is there are no enough intelligent tools to cognize the existing knowledge, as all known, the searching engines for Web information have no ability to learn. All these lead a popular researching direction in field of searching engine for Web information, that is how to make more knowledge and make more intelligent searching tools for Web information.

The Web information bottleneck of limitation in semantic description and intelligent searching is also existing in spatial information, that is it is difficult to share and intelligently query the heterogeneous spatial information. Since the conception of Semantic Web has been defined as well as it's hiberarchy has been explained by Tim Berners-Lee – the father of WWW, the Semantic Web and it's relative technologies have been developed in progress[2]. Especially, the deep research in Semantic Web Services and Geo-Ontology has provided a new method and technology for sharing, semantic matching and intelligent searching for heterogeneous spatial information[3]. As a result, studying a retrieval mechanism for intelligent searching in spatial information has become a new hotspot in the field of geo-information processing. Under this background of study, this paper has discussed the principle of geo-information retrieving mechanism, as well as method and application. In one word, an exploring in theory and a design in

archetype system have been discussed in this paper, which are establishing a framework for geo-information's retrieving based on geo-ontology technology, building an ontology model for spatial information, building a service model for retrieving spatial information on ontology model, and constructing a geo-information searching system.

## 2. ONTOLOGY AND GEO-ONTOLOGY

In earlier, an ontology is a conception in philosophy, which is equal the view of being here. Ontology has depicted that what is the essence of objects in real world, and it means knowledge and knowing[4]. Later, ontology had been integrated with computer, which finally had made the computer-ontology come into being. In 1998, Studer gave an explicit explanation for ontologies in computer based on Gruber's research, that is "an explicit and formal specification of a conceptualization"[5, 6]. More directly, we may say that an ontology is an explicit conceptualising model, comprised of objects, their definitions, and relationships among these objects. Here, cerpualization is obtained by extracting a series of conceptions from phenomena in objective world.. Explicit means that conceptions used in ontologies and it's application are clear and assured. Formal implies that an ontology could be understand by computer. And share indicated that an ontology represents the common and basic acknowledge in one special domain, as well as a public set of conceptions abstracted from relative fields and applications. In other words, an ontology is designed for all users in one domain, but not for some special individuals. In computer science, an ontology is the product of an attempt to formulate an exhaustive and rigorous conceptual schema about a domain. An Ontology is typically a hierarchical data structure containing all the relevant entities and their relationships and

rules within that domain (eg. a domain ontology). The computer science usage of the term ontology is derived from the much older usage of the term Ontology in philosophy. An Ontology which is not tied to a particular problem domain but attempts to describe general entities is known as a foundation ontology or upper ontology. Typically, more specialized domain specific schemata must be created to make the data useful for real world decisions.

Based on the definition, explanation and hierarchy of an ontology in computer science, a Geo-ontology could be regarded as a domain ontology, as well as upper ontology. Conception models in an geo-ontology is a spatial conceptualization, which is formed by cognizing and extracting spatial phenomena. Of course, those spatial conceptions will also be formalized for machine understanding through being limited by domain knowledge. "an Ontology in GIS is a kind of application intercrossed multiply sciences" Mark(1999) indicated, "which has a strong relationship with knowing, description, interoperation and uncertainty in geographic information. And more important is it is a study on spatial semantics, or in other words, it is a research about human thinking for information systems and geographic world."

### 3. THE FRAMEWORK FOR ONTOLOGY-DRIVEN GEO-SPATIAL INFORMATION RETRIEVING

#### 3.1 Application for semantic web services [7, 8]

There are two critical problems to be solved while establishing the knowledge database and service system for geo-spatial information retrieving. One is discovery and usage for services, which requires those spatial services in web could be described, published and registered. And another is those functions in services could select, compose and explain themselves automatically, further more, the suitable services in web which offer some special geo-applications should be located intelligently and automatically. Semantic Web Services maybe implement this work properly. Semantic Web Services has inherited all the characters and advanced applications of the common Web Services. At first, the spatial services are registered and could be found in the Register Center built on UDDI (Universal Description, Discovery, and Integration), and at the same time, those services are described through WSDL(Web Service Definition Language). And then, all those services should be matched and located by users, and be sent to clients through SOAP protocol, which calls Simple Object Access Protocol. Distinguishing to the traditional Web Services, however, Semantic Web Services are beyond the universal structures as well as the basic functions, that is Semantic Web Services support the implement of semantic web. As a result, geo-spatial information would have machine-processable and machine-understandable semantics intelligently and automatically based on Semantic Web Services, which is what we are looking for all the time in this paper. The system structure of Semantic Web Services maybe display in the figure 1

The benefits that Semantic Web Services is used in geo-spatial information retrieving may be expounded as follows.

- define exhaustive description frameworks for describing Web Services and related aspects (Web Service Description Ontologies)
- support ontologies as underlying data model to allow machine supported data interpretation (Semantic Web

aspect)

- define semantically driven technologies for automation of the Web Service usage process (Web Service aspect)

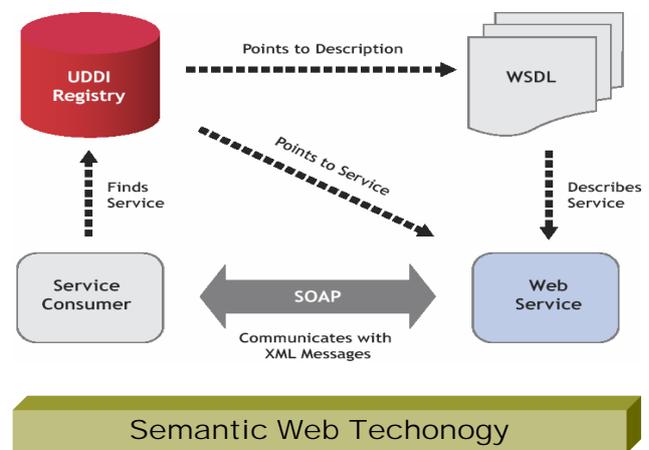


Figure 1: The General Framework Based on the Semantic Web Services (by Michael Stollberg ect., ICWE2005)

#### 3.2 General Structure of Basic Framework

The framework for geo-information retrieving mechanism has been discussed, which includes five parts as follows, and its structure is shown in figure 2.

- Client: It also is called user. And the users are usually individuals or some official groups who need the service of geo-information. This is a beginning for retrieving.
- Registry Center: This is a component in the semantic web that allows many WebGIS services to be registered with name or alias in order that the users can find what they need quickly.
- Locating Service: It is a core components in this framework, which includes a lot of geo-ontologies and spatial services. And its main function are reasoning, describing and locating the geo-ontology resources in the web based on WSDL protocol.
- Service Providers: In brief, they are a group of all kinds of WebGIS providers who distribute in the web.
- Spatial Databases: It is a collection for all kinds of geo-information, geo-ontologies and geo-metadatas.

### 4. THE BUILDING AND APPLICATION OF INFORMATION MODEL BASED ON ONTOLOGY

As the foundation of the framework, the knowledge database of geo-information plays an important role in this case. Considering specialities of geo-ontology and characters of geo-information, and according to the general steps and methods of building an ontology, an domain geo-ontology model for AIDS/HIV geo-spatial information (as an example) has been established in this paper. The continuous steps for a full geo-

ontology are the same as the figures' order listed in below pages.

At the beginning, the first step for building a geo-ontology of AIDS/HIV includes getting the important conceptions, glossaries and their relations of AIDS domain as Table1 shown. [9]

Secondly, draw a graph for AIDS geo-ontology based on Table1. The logical structure is given in figure 3.

At the end of those operations, realize the above conceptual model in protégé 2000. figure 4 and figure 5 are the capturing pictures for this case.

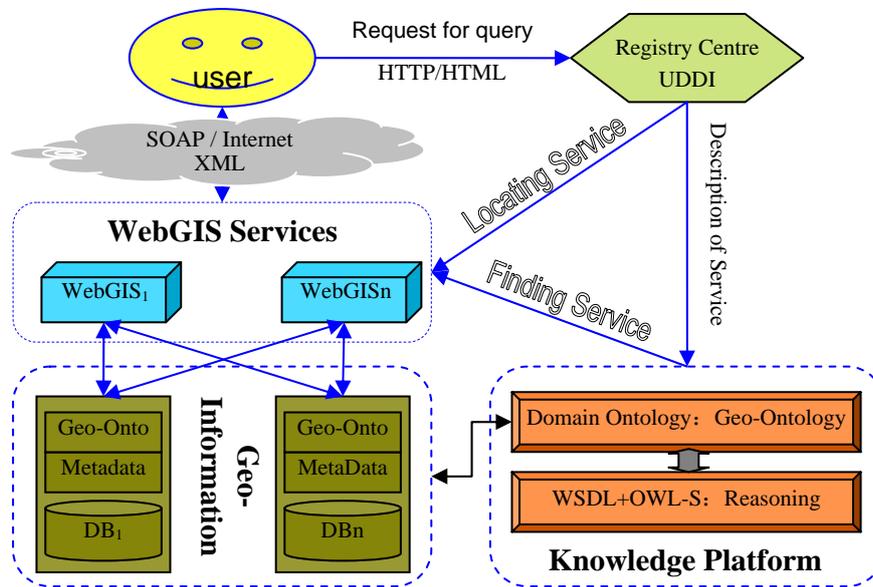


Figure 2: The General Framework Based on the Semantic Web Services

Important Concepts and Glossaries	Data Model	Spatial Characters	
		Geometry	Attribute
AIDS Zone	Popedom	polygon	Name,Count ,Area,...
People with AIDS	Individual_AIDS	point	Sex,Age,Address,Degree....
Medical Treatment	Hospital	point	Name,Address,Degree,Phone,...
Sex Composing	Sex in Individual_AIDS		.
Age Composing	Age in Individual_AIDS		Definition,Keywords,...
Epidemic			Same as Individual_AIDS
High Dangerous Crowd	Special in Individual_AIDS	point	Blood,sex,drug, childbearing
Spread path	Individual_AIDS	point	
Spatial service	Query,Geostatistics,.....		
.....	.....	.....	.....

Table 1: The List for Concepts, Glossaries and Relations of AIDS Spatial Information(in part)

### 5. THE BUILDING OF SERVICE MODEL BASED ON ONTOLOGY

Next step as the key of that framework is to construct a services model for the intelligent researching engine based on the knowledge database. Why is this step the most important study because only a service model of processing knowledge could get and know what knowledge are. On the supporting of Semantic Web Services technology, an ontology service model for geo-information retrieving based on Jena-a free middleware of knowledge processing which can reason RDF models, has been built in this case. It's logical structure is shown in figure 6. [10]

As depicted in above picture, the information source, model building, reasoning and output of geo-spatial ontology information model for AIDS have been indicated in detail and clearly. The whole service system is classified three part—

knowledge service, service description and publication for information. Knowledge service is supported by above information model of ontology, which is the base for reasoning. And service operation mainly implements query on intelligent searching engine. Restricted by the geo-spatial information model of AIDS, Jena has a powerful capacity for reasoning and producing reasoning machine. It is very convenient for Jena to gain some available subsystems of reasoning, which include DL(Description Logical), custom-build rules and tools from third parties. And then, Jena could transform the result of reasoning into information being compatible to XML format by using the RDF model. Finally, those information will be sent to user that is located in clients of WebGIS through some given search engine such as Google API. Of course, the clients are very easy tools which support XML protocol like IE and could implement the spatial services on internet

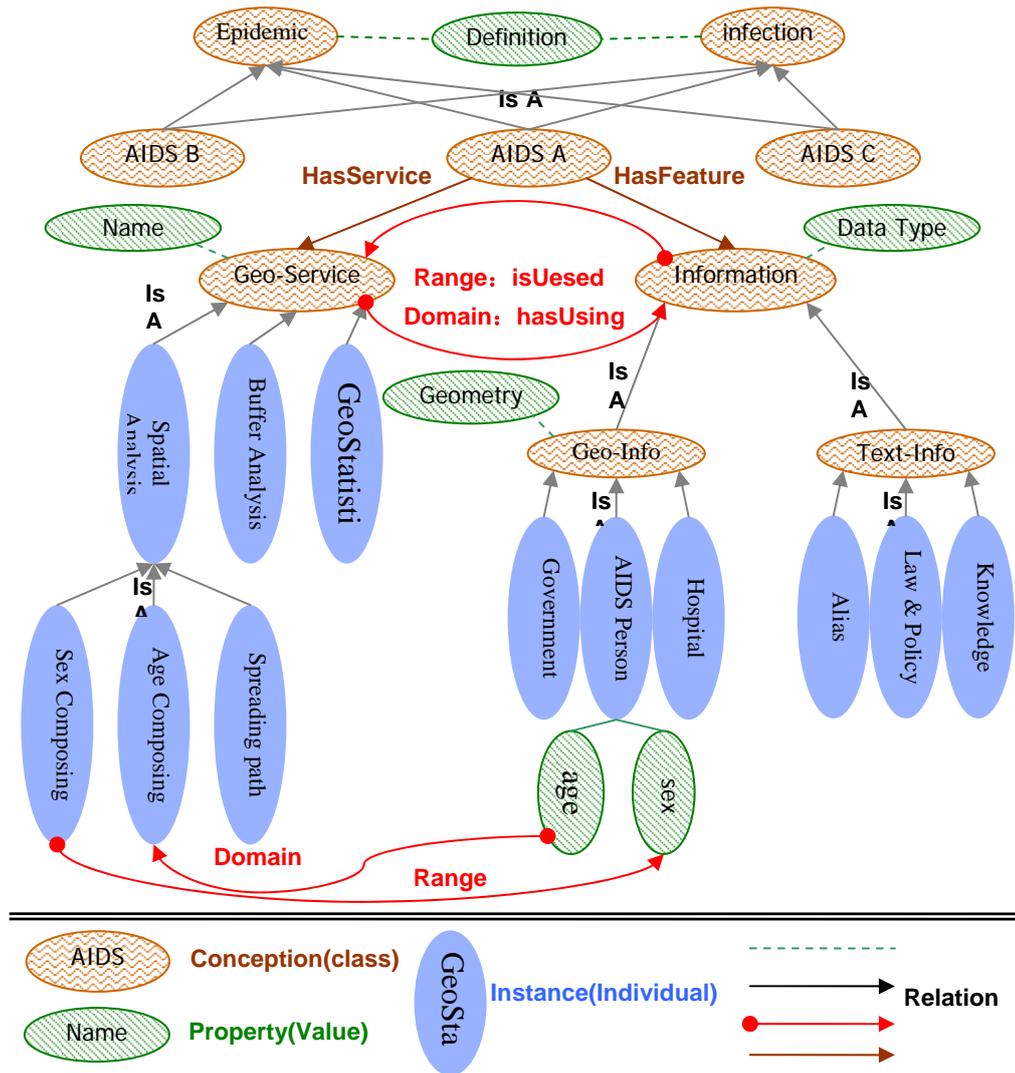


Figure 3: The Geo-Ontology Model of AIDS Spatial Information

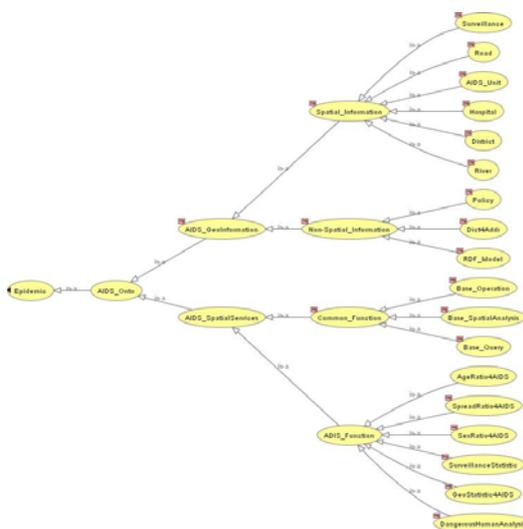


Figure 4: The Class Hierarchy of Geo-Ontology for AIDS

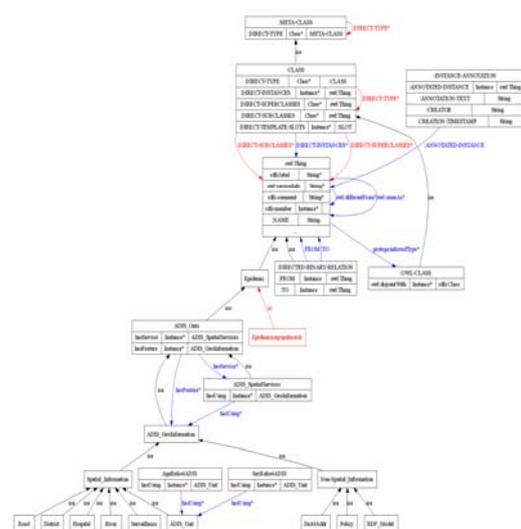


Figure 5: The OMD of Geo-Ontology for AIDS

6. APPLICATION AND CONCLUSION

All those models should be realized by a computer system in this case. In present, Java is the most popular and efficient programming language for a web application. At the same time, ArcGIS Server fully supports development for Java, which

inherits almost all abilities of ArcInfo in data description, information processing and spatial analysis. All those lead a logical structure for archetype systems, which is shown in figure 7 as follows. [11, 12]

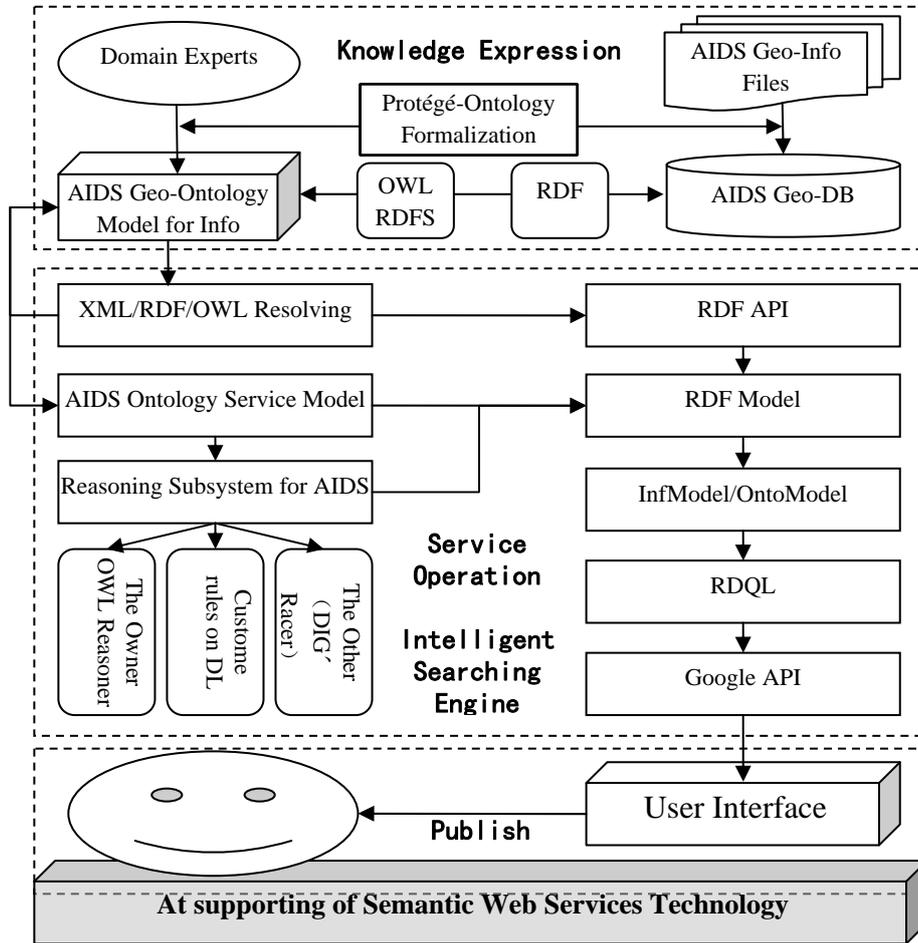


Figure 6: The Model of AIDS Geo-Ontology Based on Jena

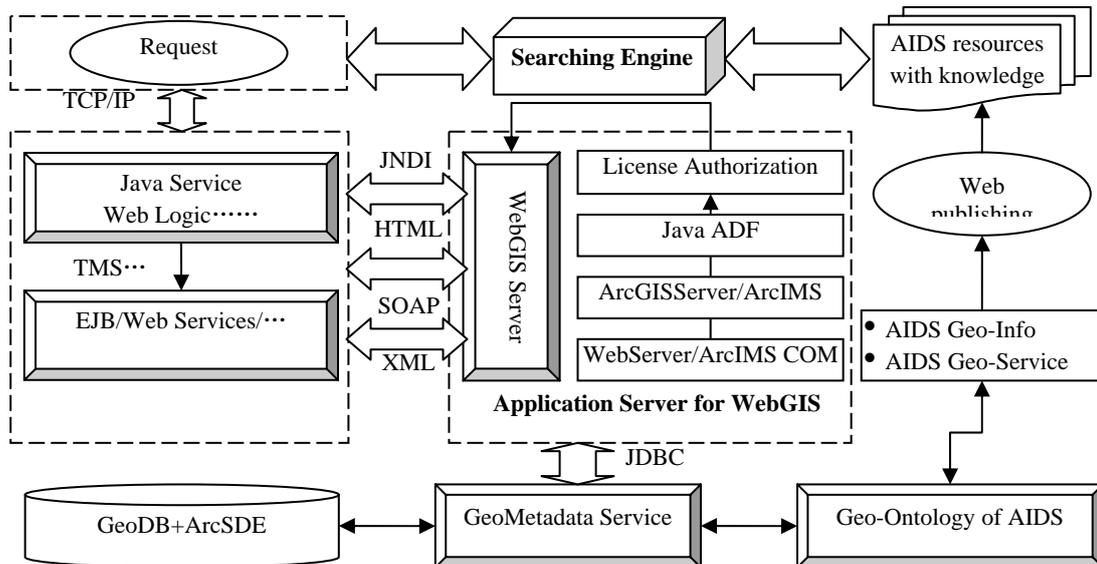


Figure 7: The System Structure of Prototype System Based on J2EE

