SPATIAL-DATA REASONING FRAMEWORK FOR PREVENTING NATURAL DISASTERS

SangHoon, Lee [a]

[a] Ubiquitous Land Research Department, Korea Institute of Construction Technology, 2311, Daehwa-dong, Goyang-shi, Geyounggi-do, 411-712, Republic of Korea - sanghoon@kict.re.kr

Commission VI, WG VI/4

KEY WORDS: Semantic Reasoning Framework, Semantic Query, Subsurface Ontology, Open-source GIS, Natural Disasters

ABSTRACT:

Many heterogeneous data are being distributed via the Web. The heterogeneous data that different suppliers produce make it difficult for users to find and share the data they need. Especially in the GIS(Geospatial Information System), reuse and sharing are very difficult. In this study, the ontological methodology was used to overcome the semantic heterogeneity in the subsurface spatial database system, which is one of Korea's national spatial information systems and essential for the preventing the diverse subsurface disasters. The subsurface ontology model, which consists of a generic concept, a measurement unit, a spatial model, and subsurface information, was developed using OWL-DL. Also, spatial-data reasoning framework, which need to query(or discovery), interpret(by Jena reasoning engine), and integrate thematic information in the interoperable repositories, open-source geospatial services(WMS, WFS by GeoServer and PostGIS), and external open-map services(Naver map), was implemented to raise a natural language-style sample query for the retrieving boring-hole or subsurface area in high liquefaction potential. In the future, this study will be expanded to establish domain ontology related fog hazard and semantic reasoning system for preventing natural disasters by which semantic sharing could be made available in the national spatial information system.

1. INTRODUCTION

In the geospatial information domain, there has been challenges to share, integrate, inter-operate distributed geospatial information produced by different organization. Accordingly, many researchers suggest data standards to interchange, such as SAIF(Spatial Archive Interchange Format), SDTS(Spatial Data Transfer Standard), and OGC standard (GML, WMS, WFS, etc), and spatial data warehouse which offers relevant information or knowledge using metadata, for instance, Korea's National Spatial Data System, NSDI (National Spatial Data Infrastructure, US), and CGDI (Canadian Geospatial Data Infrastructure, Canada). The heterogeneity of data consists of three factors: syntax, structure, and semantics (Stuckenschmidt and Visser, 2000). The aforementioned sharing methods that involve interchange standards and data warehouse construction can solve the problems of syntactic and structural heterogeneity, but cannot search for the semantic or innate meanings in compliance with the user's demand. In this study, the ontological methodology was used to overcome the semantic heterogeneity in subsurface spatial information system developed for the preventing subsurface disasters, for example, landslide, land subsidence, and liquefaction.

The earth's subsurface space is essential for the survival and development of humna beings. The lack of land due to global urbanization is accelerating the use of the earth's subsurface space. The land subsidence or liquefaction due to an earthquake brings much damage to life and property. With the expansion of cities and the development of rural areas, the incidences of landslides are increasing and more damges to life and property are expected. Accordingly, it is very important to analyze the spatial distribution by calculating the liquefaction potential in the estimation of the risk of land liquefaction occurrence.

The subsurface spatial information is essential for all construction projects as it is used for feasibility studies and cost estimation of large-scale engineering projects, design support via geotechnical analysis, and calculation of the optimal location of subsurface structures considering subsurface disasters such as those mentioned. The present subsurface spatial information is not enough to spatially analyze land liquefaction potential, however, and relevant spatial information, including those from topographical maps, geological maps, and subsurface structure maps, must be shared.

To overcome this semantic heterogeneity in the existing information sharing system, experts' analysis was essential. The ontology model and semantic reasoning framework that was designed in this study facilitates GIS sharing, considering semantics, as the basis for making automatic sharing available through reasoning.

2. SPATIAL INFORMATION SYSTEM IN KOREA

As a large quantity of subsurface spatial data are generated according to geotechnical or geological survey in construction projects, the management of existing survey documents has reached its limit, and the subsurface data are now required to be stored and shared through computerization. Accordingly, the Association of Geo-technical and Geo-environmental Specialists (AGS, 1992) announced its Electronic Transfer of Geo-technical and Geo-environmental Data Project, which has become the standard for computerized transfer of subsurface data. Petroleum development enterprises that conduct numerous boring surveys also developed and are using the Public Petroleum Data Model (PPDM, 1991).

In Japan, the applications of subsurface information such as the standardization of soil and geological survey information and the preparation of a simplified soil diagnosis tool and a geological map based on three-dimensional (3D) geological model construction are actively done at the Construction Center, the Regional Soil Environment Institute, the Osaka Soil Test Laboratory, etc (GeoDAS, 1998). The U.S. is promoting ease of use by individual of GIS-based subsurface information search, mapping, and analysis, through the development of the GeoLibrary system. In Australia, the use of subsurface information is maximized through the development of a soil analysis support system, 3D analysis, and data sharing based on distributed computing technology (MLTM, 2008).

Korea's Ministry of Land, Transport, and Maritime Affairs defined subsurface geospatial information as "information on drilling for soil, geological features and resources, geological maps, test data, groundwater, mineral resources, stones, etc." in the 2nd National Geographic Information System Master Plan for Developing and Preserving Subsurface Land (MLTM, 2000). In addition, it constructed a database for over 100,000 boreholes nationwide under its National Subsurface Information Database System Project shown in Figure 1, and made it mandatory to register geotechnical survey data to expand the database. The subsurface spatial information, including technical features such as the shapes and locations of layers and the strength required for the subsurface structure design, is inputted in the geospatial database along with the digital map. The results can be viewed on the map using the borehole location search or the text-based property search via the Web GIS (Geoinfo DB, 2000).

Recently, there is increasing demand for the use of geotechnical information for sharing with other geospatial data including topographical maps, geological maps, and land use maps, or for its use in the design of foundations and subsurface structures for analyzing subsurface disasters. To obtain the subsurface spatial information that a user requires, however, the analysis of GIS experts and geotechnical engineers, as well as understanding of the database architecture and system, is needed. The additional information analysis by experts is needed because users cannot easily understand the semantics of the information. Accordingly, the development of the subsurface ontology and its reasoning system helps users search and share the information, as they can understand the semantics without the need for additional information analysis.

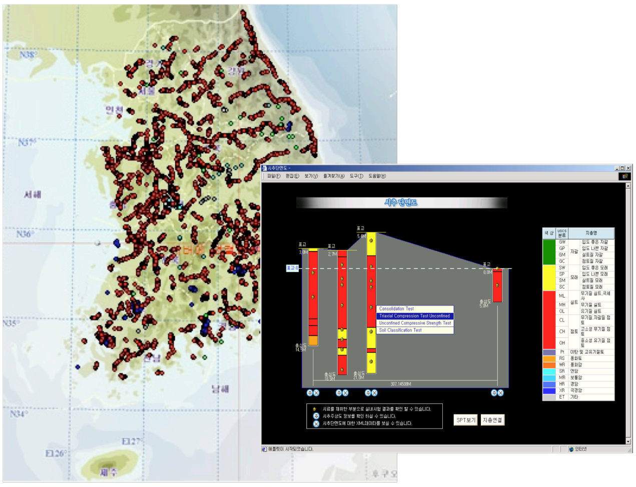


Figure 1. Web-GIS-based Subsurface System in Korea

3. SUBSURFACE ONTOLOGY MODEL

Ontology originated from a sector of philosophy that studies the existence of materials and the relationships between them. It is applied to information systems. There are many definitions of ontology, one of which is that it is "a formal, explicit specification of a shared conceptualization of a domain of interest" (Gruber, 1992).

The semantic Web, which has been presented as the next-generation intelligent Web, proposes OWL (Web Ontology Language) as the ontology language (W3C, 2004). OWL is more expressive than RDF or RDFS, which is now widely used as the standard markup language for ontology description on the Web. Besides, it makes reasoning possible by supporting the description logic other than mere simple data expression. There are three OWL versions (OWL-Lite, OWL-DL, and OWL-Full) that differ in their expression capability. In this study, OWL-DL was chosen as the basic language with which semantic reasoning is possible when the information is shared in the future.

Ontologies have different levels of dependence according to specific tasks or views, and can be divided into the top-level, domain, task, and application ontologies (Guarino, 1997). In this study, the ontology was divided into layers for the mapping of the subsurface ontology and for the semantic sharing application. Especially, the ontology of the general concept, term dictionary, measurement scale, and spatial model, which are difficult to include in the subsurface spatial information domain, was organized in layers.

3.1 Top-Level, Measurement, Geospatial Ontology

In the top-level ontology, the concept and relationships of the ontology are classified based on human recognition. It describes the general concepts that can be identically treated in many domains, and aims to produce the widest range of semantic sharing. The top-level ontology that was developed includes Cyc, BFO (basic formal ontology), DOLCE, GFO, IDEAS, Wordnet, and SUMO. BFO, which can express geological phenomena such as earthquakes and landslides, and Wordnet, which can be used as a natural language dictionary, were employed in this study.

Subsurface spatial information describes survey and test results, including boreholes. These data are measured according to their name, order, interval, and ratio (Chrisman, 1995). Nominal and ordinal units such as soil classification and USCS(Unified Soil Classification System) codes cannot be mathematically calculated due to their qualitative properties, but the interval and the ratio, such as the depth and the N value, are quantitative properties that can be calculated. The ratio can be divided or multiplied, but the interval can only be added or subtracted. The quantitative properties can be quantitatively described using measurement units (e.g., g/cm2), and the measurement units that were required for the subsurface spatial information model were expressed in "MeasurementUnit.owl".

To organize the geospatial ontology, the topology relationship, distance, direction, and whole-part relationship, as well as the object definition, must be described. In this study, the geospatial model ontology was based on GeoOntologies, which was made with the GML-based OWL ontology. GeoOntologies consists of spatial descriptors ("geoCoordinateSystems.owl") such as points and multi-polygons, spatial features (geoFeatures.owl") such as cities and buildings, and relationships between spatial descriptors ("geoRealtions.owl").

3.2 Subsurface Information Ontology

Subsurface spatial information is imperfect and heterogeneous because soil and rocks in the subsurface space have complex and uncertain properties. Accurate and detailed expression of these data requires diverse data, including data on the drilling log, layer profile, geological map, structural geological map, and DEM. The most frequently used subsurface spatial information test in construction and environmental surveys is the drilling test, followed by the lab test and the in-situ test.

The subsurface spatial information in this study consisted of the boring hole-related information. The site indicates the drilling locations on the aerial photographs or topographic maps of the area for the hole or sample collection. The hole is the result of the boring for the subsurface space survey, and the layer describes the textures and physical properties of materials at specific depth intervals in the soil. The component is a physical form that is observed at a specific depth or scope in the soil. It expresses physical, chemical, biological, and mineral characteristics and geological behavior with the lapse of time, and exists in a layer or includes several layers. The core is an information sample with a specific scope, which is extracted from holes, and the specimen is a sample that is separated from the core for the purposes of description and lab testing of the subject point. The water content, grading, and liquid/plastic limit are measured from the specimens as basic soil properties. Besides, the subsurface spatial information is expressed via the in-situ test and the lab test.

Finally, we developed OWL model according to aforementioned design of subsurface ontology layer, as shown in Figure 2.

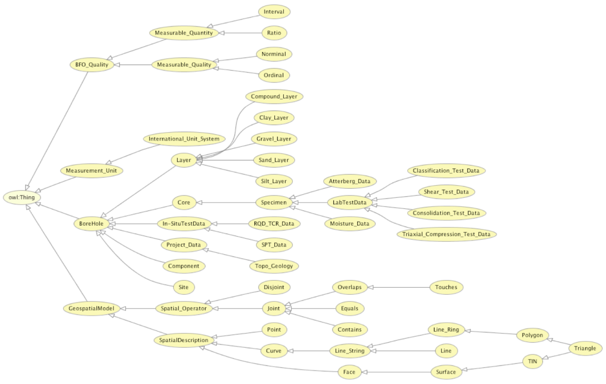


Figure 2. An excerpt of the subsurface ontology OWL model(class) using Protégé 4.0

4. SPATIAL-DATA REASONING FRAMEWORK

4.1 Overall Architecture

The proposed framework for the creation of semantic-based subsurface information system has two major goals : to support different target domains, such as top-level, measurement, geospatial and subsurface in a single ontology model, and to set up a reasoning platform to be able to automatically overcome semantic heterogeneity in related subsurface information systems. Ontology is able to manage and operate domain model in a consistent and uniform way. While being able to manipulate the concepts at the instance level of the ontology, the inference mechanisms may take both levels in consideration and the result may improve and alter either the model or metamodel of the particular targe domain. We focus on the semantic geospatial web service that agents need to query(or discovery), interpret, and integrate thematic information in the interoperable repositories and open web map services. These semantic geospatial webservice comprise 1)obtaining datasets and values for a selected theme, 2)interpreting a dataset or a dataset value to different vocabulary, and 3) intergrating different datasets into a new one depicting a particular theme.

In our design we take advantage of MVC-based frameworks, component-based web development and XML processing, which are based on the pipes and filters architectural pattern, what makes them specifically suitable for OWL processing by ontology engine(Jena) and SPARQL query enginer(joseki). One such framework is the open-source geospatial service framework(Geoserver and PostGIS), and Ajax-based web service framework which is combine with external web map service(Korea's Naver map or Google map). Figure 3 depicts an overview of the reasoning architecture that extends the basic functionality of Jena and Geoserver with additional software components in order to fulfill the aforementioned requirements.

General query method only provide a very simple support for semantics through keywords. The only "semantics" service that catalogues offer is a keyword-based service for retrieving datasets. However, this does not addresses the problem of semantic heterogeneity. We can observe that our semantic framework based on an ontology of the repository including DL definitions for themes enables us to define new functionalities that are of special importance for the integration of thematic information from various data sources

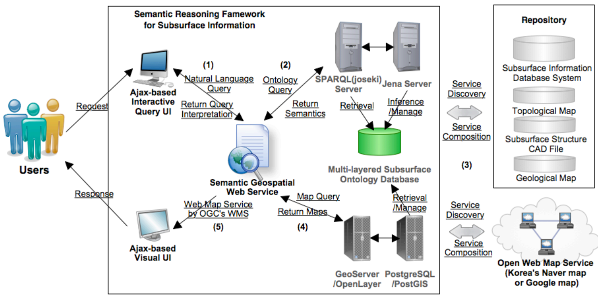


Figure 3. Overall Architecture

4.2 Semantic Query for Retrieving Liqueifaction Potential

It is neccssary for semantic query to integrate with various data sources for corresponding with user's semantic. For example, there is sample query , "Which area within 1km of Meyoung-Gi Boulevard is to be expected in the high liquefaction potential ?". This query includes various concepts to optimize decision making. These concepts have the location of "Meyoung-Gi Boulevard" indicated by user, the definition of "within" intended by user, the earhquake engineering definition of "liquefaction potential", and so on. This simple query can be attained diverse results because of being contained user's semantics. Figure 4 depicts UML sequence diagram for semantic reasoning system. From step 1 to step 13, user can obtain semi-results to interpret first query through developed ontology model and GIS database, and choose second query suitable to user's semantics. From setp 14 to step 25, user can discovery final results.

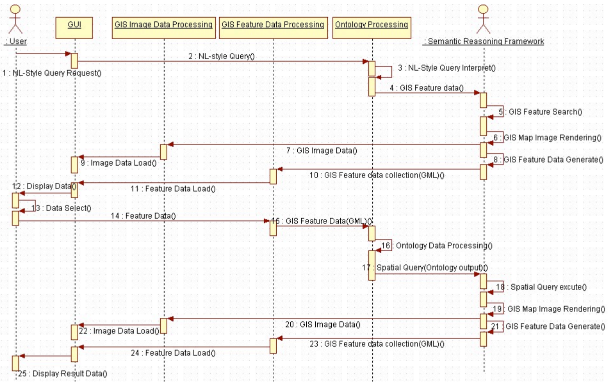
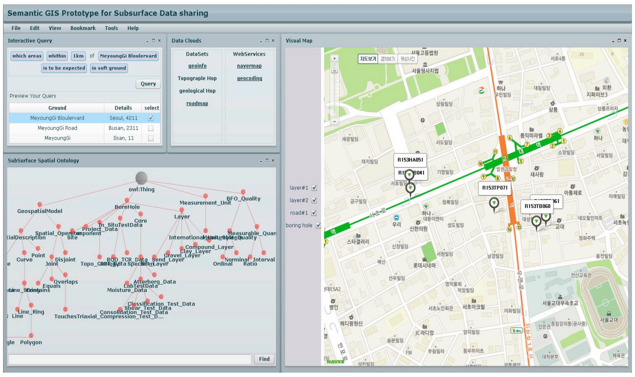


Figure 4. Sequence Diagram for Reasoning Framework

As shown in Figure 5, Interactive query window shows results list with additional information through geocoding service of "Meyoung-GI Boulevard", then user can choose more specific query. In the case of engineering judgement, definition of "high liquefaction(if silt or silty-sand, N values <= 20 and under the ground water level) (Seed R.B., et al., 2003)" described in subsurface ontology model is refered to interpreting query, related concepts and relation of ontology is shown in Subsurface ontology window. Here, N value is quantitiative data that are represented as numerical values in SPT(Standard Penetration Test). The purpose of SPT is to measure the relative density of sand or gravel. It is used to estimate the strength parameter as a good guide for the determination of geotechnical conditions, specially liquefaction. These restriction properties in liquefaction condition is described in subsurface ontology, as shown in Figure 6.


Semantic GIS Prototype for Subsurface Data sharing

which areas whithin 1km of MeyoungGi Bloulervard is to be expected in soft ground

Preview Your Query

Ground	Details	select
MeyoungGi Bloulervard	Seoul, 4211	☑
MeyoungGi Road	Busan, 2311	☐
MeyoungGi	Ilsan, 11	☐

Figure 5. Protype Screenshot

bore:LiquefactionGround =
(Sand and SPT_N some integer(<="20"integer])
and DEPTH some integer(<=GWL integer])
or (Silt and SPT_N some integer(<="20"integer])
and DEPTH some integer(<=GWL integer])
or (LooseSand or VeryLooseSand)
bore:LiquefactionGround = bore:hasUSCS and bore:hasSPT

Figure 6. An excerpt of restriction tag in the subsurface ontology

As shown in Figure 5, spatial-data reasoning web service has interactive query window, data clouds window, subsurface ontology window, and visual map window. Each window will be implemented by development environment in Table 1.

Section	Item	Product
Base	OS	Mac OS/X 10.5.8
	Development Environment	Eclipse 3.4
	Development Language	Adobe FLEX 3.0 SDK(Action Script)
	Web Client	Mozilla, Safari, Explorer, etc
Interactive Query Window	Inference Engine	Jena 2.5
	Query Engine (SPARQL Server)	Joseki 3.3
Subsurface Ontology Window	Visualization Component	RaVis(Relational Analysis Component)
Visual Map Window	GIS Database	PostgreSQL 8.3 + PostGIS
	Map Rendering Server	GeoServer 1.7
	Open API	Naver Map Open API

Table 1. Development Environment in Reasoning System

5. CONCLUSION

The present GIS system is basically heterogeneous, as it manages entities to produce spatial information according to diverse purposes. The syntactic and structural heterogeneity of the information can be ensured through format conversion or standardization, but it is difficult to address the semantic heterogeneity.

In this study, the ontological methodology was used to overcome the semantic heterogeneity in the subsurface spatial database system, which is essential for the analyzing the diverse subsurface disasters. We developed ontology model consists of the top-level, measurement, space, and subsurface information ontologies in the form of layers. Thus, a basis was established on which the subsurface information system can share information by considering semantics. Also, the spatial-data reasoning framework was designed to raise a natural language-style sample query for the retrieving high liquefaction area. The query sequence scenarios, user interface design, and development environment is developed for semantic framework. In the future, this study will be expanded to establish the semantic reasoning system for preventing fog disaster by which semantic sharing and spatial queries could be made available in the spatial database system.

6. REFERENCES

Stuckenschmidt, H., Visser, U., 2000. Semantic translation based on approximate re-classification. In: 7th Intl. Conf. on Principles of Knowledge Representation and Resoning

AGS(Association of Geotechnical and Geoenvironmental Specialists), 1992. http://www.ags.org.uk (accessed 10 Nov. 2009)

PPDM(Professional Petroleum Data Management Association), 1991. http://www.ppdm.org (accessed 10 Nov. 2009)

GEODAS, 1998. http://www.jiban.co.jp (accessed 10 Nov. 2009)

MLTM(The Ministry of Land, Transport, and Maritime), 2008. The 2008 National Geotechnical Information Database Project Report, pp. 7-10 (in Korean)

MLTM(The Ministry of Land, Transport, and Maritime), 2000. The 2nd National Geographical Information System Master Plan, pp. 12-13 (in Korean)

Geoinfo DB, 2000. http://www.geoinfo.or.kr (access 10 Nov. 2009)

Gruber, T. A., 1992. Translation Approach to Portable Ontology Specification. Knowledge Systems Laboratory. Stanford University. Stanford. CA. Technical Report KSL, pp. 92-71.

W3C(World Wide Web Consortium), 2004. http://www.w3c.org (accessed 10 Nov. 2009)

Guarino, N., 1997. Semantic matching: formal ontological distinctions for information organization, extraction, and integration. Information Extraction: A Multidisciplinary Approach to an Emerging Information Technology. Springer Verlag, pp. 139-170

Chrisman, N.R., 1995. Beyond Stevens: A Revised Approach to Measurement for Geographic Information. In Proc. of AUTO-CARTO 12 Charlotte 4, pp. 271-280

Geoontologies. http://www.mindswap.org/2004/geo/geoOntologies.shtml (access 10 Nov. 2009)

Seed R.B., et al., 2003. Recent advances in soil liquefaction engineering : a unified and consistent framework. 26th Annual ASCE Los Angeles Geotechnical Spring Seminar, Keynote Presentation, H.M.S. Queen Mary, Log Beach, California, pp. 2-7

RaVis(Relational Analysis Components).http://code.google.com/p/birdeye/wiki/RaVis (access 10 Nov. 2009)

Naver Map Open API. http://dev.naver.com/openapi/apis/map/flash/reference (access 10 Nov. 2009)

7. ACKNOWLEDGMENTS

This research was supported by the grant (09CCTI-A052531-02-000000) from the Ministry of Land, Transport and Maritime of Korean government through the Core Research Institute at Seoul National University for Core Engineering Technology Development of Super Long Span Bridge R&D Center.