韓國地形空間情報學會誌 第17卷 第1號 2009年 3月 pp. 121-129

연구논문

시멘틱 공유를 위한 MDA기반 지하공간정보 온톨로지 모델 개발

Development of MDA-based Subsurface Spatial Ontology Model for Semantic Sharing

이상훈* · 장병욱**

Lee, Sang Hoon · Chang, Pyoung Wuck

要 旨

최근 이질적인 정보의 폭발적인 증가와 다양한 관리주체별로 생산, 축적되는 공간정보의 특성으로 인하여 공간정보의 재사용과 공유가 어려운 실정이다. 국가공간정보체계의 하나인 지하공간정보도 공간분석을 위해서는 지형도, 지질도, 지하시설물도 등 여타 공간정보와 공유가 필수적이다. 그러나, 기존 표준 혹은 데이터웨어하우스에 의한 공유방법은 시멘틱 이질성을 고려할 수 없다. 본 연구는 지하공간정보의 시멘틱 공유를 위해 일반개념, 측정스케일, 공간모델을 포함한 온톨로지 레이어 모델을 개발하였다. 또한, 기존의 수작업에 의존하는 온톨로지 개발방법론이 아닌 MDA기반 방법론을 적용하여 직관적이며 환경변화에 쉽게 대응할 수 있는 메타모델(UML Profile)을 개발하였다. 개발된 온톨로지 모델의 시멘틱 품질은 Pellet 추론엔진을 통해 검증하였다. 본 연구를 통해 시멘틱 공유를 증진시키고, 온톨로지의 지식표현 능력을 이용하여 GIS 전문가시스템 개발이 가능할 것으로 기대된다.

핵심용어 : 시멘틱공유, 온톨로지개발, MDA, 지하공간정보, 시멘틱GIS

Abstract

Today, it is difficult to re-use and share spatial information, because of the explosive growth of heterogeneous information and specific characters of spatial information accumulated by diverse local agency. A spatial analysis of subsurface spatial informa-tion, one of the National Spatial Data Infrastructure, needs related spatial information such as, topographical map, geologic map, underground facility map, etc. However, current methods using standard format or spatial datawarehouse cannot consider a se-mantic hetergenity. In this paper, the layered ontology model which consists of generic concept, measuremnt scale, spatial model, and subsurface spatial information has developed. Also, the current ontology building method pertained to human experts is a expensive and time-consuming process. We have developed the MDA-based metamodel(UML Profile) of ontology that can be a easy under-standing and flexiblity of environment change. The semantic quality of devleoped ontology model has evaluated by reasoning engine, Pellet. We expect to improve a semantic sharing, and strengthen capacities for developing GIS experts system using knowledge representation ability of ontology.

Keywords : Semantic Sharing, Ontology Development, Model-Driven Architecture, Subsurface Spatial Information, Semantic GIS

1. 서 론

1.1 연구 배경 및 목적

인터넷의 폭발적인 발전으로 웹을 통해 문서, 이미지, 동영상 및 지도와 같이 이질적인 정보가 다량 배포되고 있다. 상이한 공급자에 의해 생산된 이질적인 정보는 사용자가 자신에게 필요한 정보를 찾아 공유하기 어렵게 하는 주요 원인이다. 특히, GIS시스템의 경우에는 중앙정부, 지자체 및 각 관리주체 별로 특정한 목적에 따라 공간정보를 수집하고, 많은 양의 데이터를 상이한 시스템과 다양한 포맷으로 축적하고 있어, 재사용과 공유가 매우 어렵다(Devogele, 1998).

현재 GIS분야에서 이질적인 정보를 공유하기 위한 방법으로는 표준 개발과 데이터웨어하우스 구축이 있다.

2009년 2월 23일 접수, 2009년 3월 19일 채택

* 교신저자·정회원·한국건설기술연구원 U-국토연구실, 연구원 (sanghoon@kict.re.kr)
** 서울대학교 지역시스템공학과 교수 (pwchang@snu.ac.kr)

먼저, 표준에 의한 방법은 SAIF, SDTS등 교환포맷의 개발, ISO/TC211의 표준활동, OGC의 GML, WMS, WFS 개발 등이 있다. 또한, 공간데이터웨어하우스 구축은 국내의 국가공간정보체계, 미국의 NSDI(National Spatial Data Infrastructure), 캐나다의 CGDI(Canadian Geospatial Data In-frastructure) 등이 있다. 공간데이터웨어하우스에는 '데이터의 데이터'로 정의되는 메타데이터(FGDC, 1998)를 이용하여 공간정보를 분류하고 관련 지식을 제공한다. 이를 통해 사용자는 원하는 정보를 찾고 공유할 수 있다.

정보의 이질성은 문법(syntactic), 구조(structure), 시멘틱(semantic)의 세 요소로 구분할 수 있다(Stuckenschmidt, 2000). 앞서 언급한 표준과 데이터웨어하우스에 의한 방법은 정보의 문법적, 구조적 이질성은 해결할 수 있지만, 사용자 요구에 적합하도록 시멘틱 즉, 내재된 의미까지 검색하여 공유할 수는 없다. 본 연구는 국가공간정보체계의 하나인 지하공간정보의 시멘틱 이질성을 극복하기 위해 온톨로지(ontology) 방법론을 적용하였다. 온톨로지는 시멘틱을 공유하는 주요한 방법론의 하나로, 정보공유 뿐만 아니라, 사용자, 어플리케이션 및 컴퓨터와 자유로운 통신을 지원하여 시스템의 활용폭을 확장시킬 수 있다. 이를 통해 기존 특정업무를 위해 개발된 공간정보와 GIS시스템을 다양한 시나리오에서도 활용할 수 있다.

GIS 분야에서 시멘틱 공유를 위한 온톨로지 개발에 대한 여러 연구가 있었지만, 그 활용은 제한적이었다. 이는 GIS가 다양한 도메인이 통합되어 있는 시스템이기 때문에 모든 사용자를 만족시키는 온톨로지 개발이 어렵기 때문이다.

따라서, 본 연구에서는 기존 GIS온톨로지를 상위개념으로 공유하면서, 지하공간정보에 적합하도록 도메인 온톨로지를 개발하였다.

또한, 기존의 온톨로지 개발방법론은 전문가의 수작업에 의존하고 있다(Nicola, 2009). 이는 많은 시간과 비용이 소요될 뿐만 아니라, 도메인 전문가의 경험과 직관에 의존하는 휴리스틱(heuristic)한 방법론을 적용하여 전문가에 따라 객관성이 결여될 수 있다. 본 연구는 MDA (Model-Driven Architecture)의 메타모델을 적용하여, 현재 소프트웨어 개발방법론의 표준으로 인식되는 UML (Unified Modeling Language)모델을 이용하여 온톨로지를 개발하였다. 이를 통해 누구나 쉽게 재사용가능한 온톨로지를 설계할 수 있다. 또한, MDA가 기술, 상황, 인프라 등 환경변화에 메타모델 변경만으로 쉽게 대응할 수 있기 때문에, 적은 시간과 비용으로 신뢰성 높은 온톨로지를 재생산할 수 있다.

2. 배경 이론 및 기존 연구

2.1 온톨로지 및 개발 방법론

온톨로지는 물질의 존재와 그 관계를 연구하는 철학의 한 분야에서 출발하여 정보시스템에 적용되어 많은 연구가 이뤄지고 있다. 현재 다양한 온톨로지 정의가 있지만, 대표적인 정의는 "An ontology is a formal, explicit specification of a shared conceptualization of a domain of interest"이다(Gruber, 1992). 온톨로지는 관심 도메인에 대한 개념화라는 측면에서 기존의 데이터모델링과 유사하다. 그러나, 데이터 모델링이 데이터의 무결성, 분류체계를 정의하여 데이터의 품질 확인을 목적으로 한다면, 온톨로지는 자연어 처리 혹은 인터넷 환경에서 데이터의 의미 교환과 공유를 그 목적으로 한다.(Meersman, 2002).

또한, 온톨로지는 해당 도메인에서 용어와 개념(혹은 의미)를 묘사함으로써, 도메인의 지식을 표현할 수 있다.

그림 1과 같이 W3C에서 차세대 지능형 웹으로 제안하고 있는 시멘틱웹에서는 온톨로지 언어로 OWL(Web Ontology Language)을 제안하고 있다(W3C,2004). OWL은 웹에서 온톨로지를 기술하는 표준 마크업 언어로 현재 널리 쓰이고 있는 RDF 혹은 RDFS 보다 표현성(expressiveness)이 뛰어나다. 또한, 단순한 정보표현이외에 서술논리(description logic)를 지원하여 추론(reasoning)도 가능하다. OWL은 표현능력에 따라 OWL-Lite, OWL-DL, OWL-Full의 세가지 버전이 있으며, 본 연구에서는 향후 정보공유시 시멘틱 추론이 가능한 OWL-DL을 기본 언어로 채택하였다.

온톨로지 개발에는 Uschold & King Method, Toronto Virtual Enterprise Method, Methontology, KACTUS

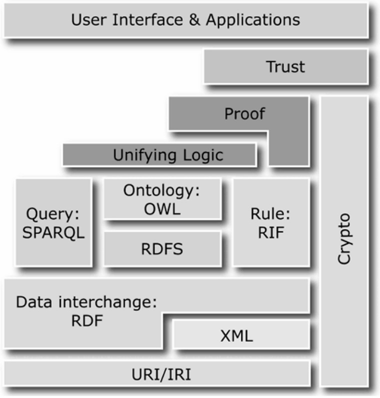

그림 1. 시멘틱웹 레이어(W3C, 2007)

Project Method, Lexicon-Based Method 등 다양한 방법론이 있다(Breitman, 2007). 이러한 방법론은 일반적으로 전문가가 참여한 수작업으로 이뤄진다. 먼저, 설문 (competence question)을 통해 온톨로지의 요구사항을 추출하고, 도메인의 데이터 수집을 통해 주요 용어에 대한 용어집 (glossary)을 작성한다. 객체에 대한 클래스 택소너미와 내외부 속성을 정의하고, 개념 사이의 관계를 설정한 후 공리를 포함하여 인스턴스를 생성, 테스트한다. 설문의 요구사항을 만족하도록 위 과정을 반복하여 최종적으로 무결성과 정확성을 확보한 온톨로지를 개발한다.

또한, 온톨로지 구축을 상향식과 하향식으로 구분할 수 있다. 상향식은 현재 주로 이용되는 방법으로 먼저 용어집이나 전문가 인터뷰에서 개념을 추출하여 온톨로지를 작성한다. 이 방법은 상세한 개념을 공급할 수 있지만, 온톨로지의 무결성과 일관성에 문제가 발생할 수 있으며, 다양한 전문가가 참여하지 못하면 객관성이 결여되는 문제가 발생한다(Teller, 2007). 하향식 방법은 기 구축된 DB 혹은 모델에서 직접적으로 온톨로지를 추출하는 방법이다. 하향식 방법은 무결성을 가진 온톨로지 생성이 쉽고, 새로운 사용자도 적응이 쉽다는 장점이 있지만, 관련된 시스템 혹은 모델이 이미 구축되어야 한다.

본 연구는 기 개발된 국토지반정보DB구축 사업의 지하공간정보시스템을 기반으로 온톨로지를 도출하는 하향식 방법을 적용하였다. 이를 통해 휴러스틱하고, 반복적인 수작업으로 인해 발생하는 시간과 비용을 최소화하고, 무결성 체크를 최소화한 온톨로지 개발이 가능하다.

2.2 MDA기반 온톨로지 개발방법론

UML 2.0에 해당하는 MDA는 소프트웨어 개발에서 코드보다 모델 생성에 중점을 둔 개발방법론으로, 소프트웨어의 구현과 명세의 관계를 구조적으로 분리하여 상호운영성, 재사용성, 휴대성의 장점을 가진다.(Miller 2003). 특히, MDA기반 방법론 중 UML모델은 요구분석, 시스템 설계, 구현 등의 객체지향 개발과정에서 그래픽한 방법을 제공하고 설계 결과물을 공유할 수 있는 방법론을 제공한다. 현재 산업계의 표준으로 인식된 UML을 적용한 시스템은 신뢰성 있는 시스템으로 평가받고 있다 (OMG, 2001).

MDA는 그림 2와 같이 네 가지 레이어의 메타모델링 구조를 가지며, MOF(Meta-Object Facility), UML, XMI (XML Metadata Interchange) 등으로 구성된다. MDA는 모델링언어와 구현언어, 그리고 메타모델간의 매핑을 통해 상호변환함으로써 플랫폼 독립적이며, 대용량 어플리케이션과 기술변화에 효과적으로 대응할 수 있다.

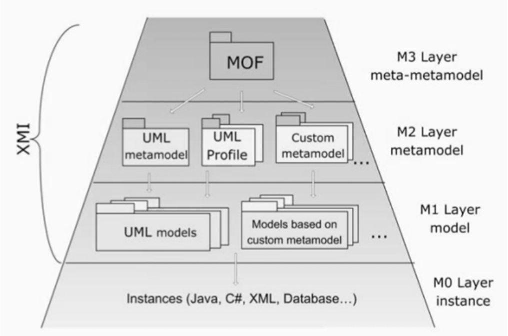

그림 2. MDA의 구조(OMG, 2003)

UML은 효율적이며, 정량화가 가능한 개념 모델링 방법으로 온톨로지 모델링 언어로 이용할 수 있다(Cranefield et al., 1999). 기존 연구에서 UML모델을 ODM(Ontology Definition Model)의 XMI를 이용하여 OWL 모델로 변환하였으며(Duric et al., 2003), UML모델을 DAML 온톨로지 모델로 변환하였다(Baclawski, et al., 2001). 이러한 방법은 UML의 표준적인 그래픽 도구를 이용하기 때문에 사용자가 직접 온톨로지를 표현할 수 있으며, 변환은 XMI 교환포맷과 MOF에서 정의된 인터페이스로 자동적으로 이뤄진다. 또한, 객체지향 모델인 UML의 모듈적인 성격으로 인해 특정 객체의 변화가 다른 객체에 영향을 주지 않는 장점이 있다.

2.3 GIS 온톨로지 개발

이질적인 GIS시스템 간에 정보교환을 위하여 지리정보 온톨로지의 필요성이 제기되었다(Smith & Mark, 1998). 그러나, 다른 정보시스템 온톨로지와 달리 공간정보는 위상관계와 부분-전체관계 등이 존재하여 차별화가 필요하다(Smith, 1995). GIS 온톨로지의 연구분야는 GIS 온톨로지 개발 및 표현, GIS서비스의 발견 및 검색, GIS 정보의 중재와 변환 등이 있다. 현재까지 개발되거나, 운영되고 있는 GIS 온톨로지 개발그룹에는 MUSIL, OntoGeo, OntoSpace, SEEK, SWEET, USGS-ONT, W3C Geospatial Incubator Group 등 다수 존재한다.

그러나, 현재의 GIS 온톨로지 개발은 제한적인 성공이었다. GIS와 같이 다양한 도메인이 통합되어 있는 시스템의 온톨로지를 만드는 것은 매우 어려운 작업이기 때문이다(Torress, 2005). 현재 개발 중인 GIS 온톨로지는 가능한 많은 시스템에서 공유할 수 있도록 공간정보를 묘사하기 때문에 상위 온톨로지로 적용하는 것이 바람직하다. 본 연구와 같이 특정한 시스템에서는 그에 적합한 도메인과 업무 온톨로지 개발이 이뤄져야 한다.

2.4 국내의 지하공간정보 시스템

국토해양부에서는 지하국토의 개발 및 보전을 위하여 제2차 국가지리정보체계 기본계획에서 ‘지반・지질・자원의 시추정보, 지질도, 시험자료, 지하수, 광물자원, 석재 등의 정보를 지하지리정보”로 규정하였다(건설교통부, 2000). 또한, 국토지반정보DB구축 사업을 추진하고, 건설공사시 필수적으로 지반조사자료를 등록하도록 하여 현재 전국적으로 10만여 시추공을 구축하고 지속적으로 증가될 예정이다(국토해양부, 2008). 지하공간정보는 지층의 형상 및 위치, 지하구조물 설계에 필요한 강도 등의 공학적 특성을 수치지도와 함께 GIS-DB로 구축된다. 이러한 정보는 웹GIS를 통해 지도상에서 시추공의 위치검색과 텍스트 기반의 속성검색으로 정보를 열람할 수 있다(지반정보DB포털, 2009).

최근 지형도, 지질도, 토지이용현황도와 같은 여타의 GIS자료와 공유하거나, 건설공사에서 기초 및 지하구조물 설계시 활용에 대한 요구가 증가되었다. 그러나, 사용자가 원하는 지하공간정보를 얻기 위해서는 DB구조 및 시스템에 대한 이해가 필요할 뿐만 아니라, GIS 전문가, 지반공학 전문가 등의 정보해석이 필요하다. 즉, 사용자가 정보의 시멘틱을 쉽게 파악할 수 없기 때문에 별도의 전문가의 정보해석이 필요한 것이다. 따라서, 지하공간정보 시스템 온톨로지 개발은 사용자가 별도의 정보해석을 하지 않고도 시멘틱 파악이 가능하여 원하는 정보의 검색과 공유를 도와준다.

3.지하공간정보 온톨로지 레이어 설계

3.1 온톨로지 레이어 구성

온톨로지는 특정한 업무 혹은 관점에서 의존성 수준(level of dependence)을 달리하며, 이에 따라 상위(top-level), 도메인(domain), 업무(task), 응용(application) 온톨로지로 구분할 수 있다(Guarino, 1997). 본 연구는 지하공간정보 온톨로지와 시멘틱 공유 어플리케이션을 매핑시키기 위하여 온톨로지를 세분화하여 레이어로 구성하였다. 특히 지하공간정보 도메인에서 담기 어려운 일반개념, 단어사전, 측정스케일, 공간모델에 관한 온톨로지를 그림 3과 같이 레이어로 구성하였다.

3.1.1 상위온톨로지

상위온톨로지는 사람의 인식에 기초하여 온톨로지의 개념과 관계를 분류한 것으로 철학, 인식학, 심리학 등에 의해 기초 분류가 이뤄진다(Mika, 2004). 여러 도메인에서 동일하게 취급될 수 있는 일반적인 개념을 묘사하며, 가장 넓은 범위의 시멘틱 공유를 목표로 한다. 현재 개발

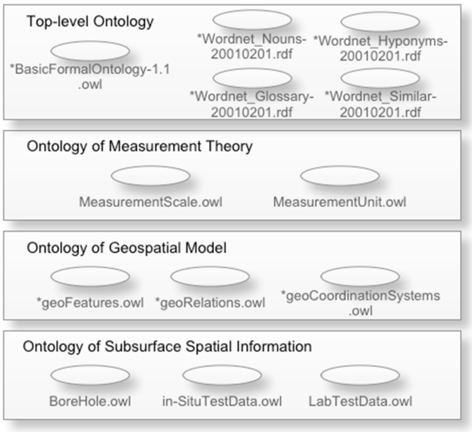

그림 3. 온톨로지 레이어 모델

된 상위온톨로지에는 Cyc, BFO(Basic Formal Ontology), DOLCE, GFO, IDEAS, Wordnet, SUMO 등이 있다. 본 연구에서는 지하공간을 묘사하기 쉽고 지진, 산사태 등 지질현상을 표현가능한 BFO와 자연어 사전으로 이용가능한 Wordnet을 적용하였다.

BFO는 SNAP과 SPAN의 두 가지 다른 레벨의 서브 온톨로지로 구성된다. SNAP은 현재의 객체를 이해하기 위해 1, 2, 3차원으로 객체를 묘사하며, SPAN은 주어진 시간 간격에서 발생하는 프로세스를 묘사한다. BFO는 두 서브 온톨로지 사이에 상호관계를 정의하여 정적 혹은 동적, 공간적 혹은 시간적 객체를 모두 나타낼 수 있다. Wordnet은 인식관계학에 기초한 시멘틱네트워크로 출발하여 현재는 사전으로 그 영역을 확장하였다. 대부분의 일반적인 개념을 포함하는 상위온톨로지로서 “isKindOf (hypernym/hyponym)”, “isPartOf (holonym/meronym)” 등 다양한 시멘틱 관계를 포함하여 현재 자연어 연구에 널리 이용되고 있다.

3.1.2 측정온톨로지

지하공간정보는 시추공 등 관측조사 및 실험결과를 묘사한다. 이러한 정보의 측정은 이름(nominal), 순서(ordinal), 간격(interval), 비율(ratio)로 구분할 수 있다(Chrisman, 1995). 지반분류, USCS코드와 같이 이름과 순서 단위는 정성적인 속성으로 인해 수학적 연산할 할 수 없지만, 심도, N값과 같은 간격과 비율은 정량적 속성으로 연산이 가능하다. 비율의 경우에는 나누거나 곱할 수 있으나, 간격은 더하거나 뺄 수 밖에 없다. 정량적 속성은 g/cm^2와 같이 측정단위에 대한 정량적인 묘사가 가능하며, 본 연구에서는 지하공간정보모델에 필요한 측정단위를 “MeasurementUnit.owl”에서 표현하였다.

이렇게 측정 단위는 측정값에 대한 품질의 시멘틱을 묘사한다. 그러나, 상위온톨로지로 채택한 BFO가 측정단위를 정량적, 정성적으로 구분하고 있지 않다. 따라서, 본 연구는 BFO의 “Quality”개념을 확장하여 측정가능한 정량자료(measurable quantity)와 측정가능한 정성자료(measurable quality)를 추가하여 “MeasurementScale.owl”을 작성하였다.

3.1.3 공간온톨로지

공간 온톨로지를 구성하기 위해서는 객체의 정의를 포함하여 위상관계, 거리, 방향 및 전체-부분 관계를 묘사해야 한다. 먼저, 객체의 정의는 OGC의 공간참조시스템인 OGM(Open Geodata Model)과 GML을 적용하였다. GML은 공간객체를 XML로 표현하면서 객체의 Geometry, Properties, Reference System을 포함한다. 공간 관계는 객체의 상세사항을 제공한다. 잘 알려진 관계로 피쳐타입(e.g. face, line, point, location, dimension)과 위상 관계(e.g. disjoint, equal, contains, overlaps, touches)가 있다.

본 연구에서 공간모델 온톨로지는 GML에 기초한 OWL 온톨로지로 작성된 GeoOntologies(2004)를 참조하여 사용하였다. GeoOntologies는 포인트와 멀티폴리곤과 같은 공간 디스크립터 (“geoCoordinateSystems.owl”)와, 도시, 빌딩과 같은 공간피쳐 (geoFeatures.owl”), 공간디크립터 사이에 관계 (“geoRealtions.owl”) 세 가지로 구성된다.

3.2 지하공간정보의 구성

지하공간을 이루는 토사와 암반 등은 복잡하고 불확실한 성격을 가지기 때문에 지하공간 정보는 불완전하고 이질적이다. 이러한 정보를 좀 더 정확하고 자세히 표현하기 위해 시추주상도, 지층단면도, 지질도, 구조지질도, DEM 등 다양한 자료가 이용되고 있다.

지하공간정보 중 건설공사 및 환경조사에서 그 이용도가 높은 정보로는 시추조사, 실내시험, 현장시험의 정보순으로 높은 비중을 차지하며, 이외에 지질조사, 물리탐사, 현장수리시험 등이 중요도가 높다고 조사되었다(PEER, 2004). 이는 본 연구에서 개념모델로 이용한 국가공간정보체계의 “국토지반정보시스템”과 유사하다.

본 연구의 지하공간정보는 그림 4의 시추공 관련정보로 구성하였다. 대상지(site)는 시추공 혹은 샘플채취를 위한 지역의 항공사진, 지형도 등의 지도에 시추위치를 표시한다. 시추공(hole)은 지하공간 조사를 위한 시추(boring)의 결과이며, 지층(layer)은 지반의 특정 깊이 간격에서 물질의 텍스쳐(texture)와 물리적 성격을 묘사한다. 지층구성(component)은 지반의 특정 깊이의 지점 혹은 범위에서 관찰되는 물리적인 형태이다. 지층구성은 물리적, 화학적, 생물학적, 광물적 특성과 시간에 따른 지질작용을 나타내며, 지층내에 존재하며 다수의 지층을 포함할 수 있다. 코어(core)는 시추공으로부터 추출되는 특정 범위의 정보 샘플이며, 시료(specimen)은 해당 지점에 대한 묘사와 실내시험을 목적으로 코어에서 분리된 것이다. 일반적으로 시료는 기본적인 토사의 물성으로 함수비, 입도크기, 액소성한계을 측정한다. 이외에도 현장시험과 실내시험으로 지하공간정보의 특성을 나타낸다.

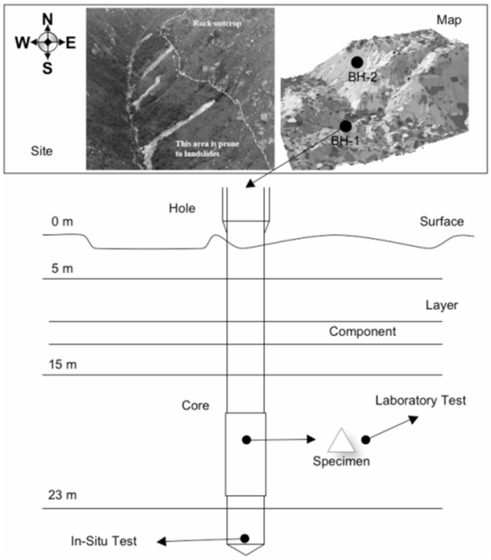

그림 4. 시추공 관련정보의 구성

3.3 MDA기반 지하공간정보 OWL모델 구축

MDA기반의 온톨로지 설계를 위해서는 UML의 클래스 다이어그램을 메타모델인 UML Profile로 확장해야 한다. 본 연구에서는 UML의 의미를 확장하여 새로운 종류의 구성요소를 생성하는데 사용되는 스테레오 타입을 이용하였다. 스테레오 타입을 통해 UML 메타클래스를 가상의 서브클래스로 정의하고, 클래스, 연관성, 의존성과 같은 부가적인 시멘틱을 할당한다.

UML Profile모델은 XMI문서로 변환되고, XML파서로 XSLT(eXtensible Stylesheet Language Transformations)를 거쳐 OWL모델로 변환하였다. 본 연구에서 UML 모델링 툴로는 UML XMI 1.4를 지원하는 Poseidon for UML v4.0을 적용하였고, XSLT 프로세서로는 Xalan Java v2.7.1를 이용하여 XMI문서로 UML모델을 변환시켰다. 또한, 구현된 온톨로지 모델의 편집은 Protege v3.4를 이용하였다.

클래스 다이어그램의 패키지는 스테레오타입 "<<ontology>>"으로 표시하였으며, 각 클래스는 "<<OntClass>>"로 표시하였다. 이외에 온톨로지에서 존재하는 Enumeration, Union, Intersection, Complement, Restriction, AllDifferent 등의 클래스도 스테레오타입으로 정의하였다. 간단한 데이터 타입의 클래스 속성은 "<<DatatypeProperty>>"으로 표시하고, 클래스 사이에 복잡한 데이터 타입 혹은 관계와 관련한 속성은 "<<ObjectProperty>>"로 표시하였다. 이러한 규칙에 따라 Poseidon for UML을 이용하여 그림 5와 같이UML Profile 모델을 작성하였다. 작성된 UML모델은 Xalan XSLT 프로세서를 통해 변환작업이 이뤄지며, 앞서 설계한 온톨로지 레이어와 함께 최종적으로 OWL-DL모델을 개발하였다. 그림 6은 Protege-OntoViz를 이용하여 작성된 OWL모델의 클래스를 표현한 것이다.

4.온톨로지의 유효성 검증 및 활용

4.1 추론엔진을 통한 온톨로지 검증

다차원적 성격을 가지는 온톨로지 품질은 문법적, 시멘틱, 실용적(pragmatic), 사회적 품질로 나눌 수 있다 (Burton-Jones, 2005). 문법적 품질은 정형화된 문법으로 작성되었는지를 평가하고, 시멘틱 품질은 모순된 개념이 존재하는지를 측정하여 온톨로지의 무결성을 검토한다. 실용적 품질은 사용자를 위한 온톨로지의 유용성을 평가하고, 사회적 품질은 얼마나 일반성을 반영하여 다른 온톨로지와 연결시킬 수 있느냐를 검토한다.

본 연구에서는 이러한 온톨로지 품질 중에서 시멘틱 품질을 우선 고려하였다. 문법적 유효성은 UML Profiles에서 XMI를 통해 OWL모델로 변환하는 과정에서 이미 확인되며, 사회적 및 실용적 품질은 온톨로지 레이어의 상위 온톨로지의 범용성으로 인해 달성될 수 있다. 시멘틱 품질의 평가는 추론엔진(reasoner)를 통해 무결성 검증이 이뤄진다. 본 연구는 그림 7과 같이 Mindswap에서 제공하는 OWL 추론엔진인 Pellet v1.5.1을 적용하여 지하공간정보 온톨로지 모델의 개념과 클래스 계층이 오류없이 무결성을 가짐을 확인하였다(Consistent : Yes).

4.2 시멘틱 공유를 위한 온톨로지 모델 활용

지하공간정보는 대규모 토목공사의 타당성 검토와 사업비 산정, 지반해석을 통한 설계지원, 그리고 지하시설

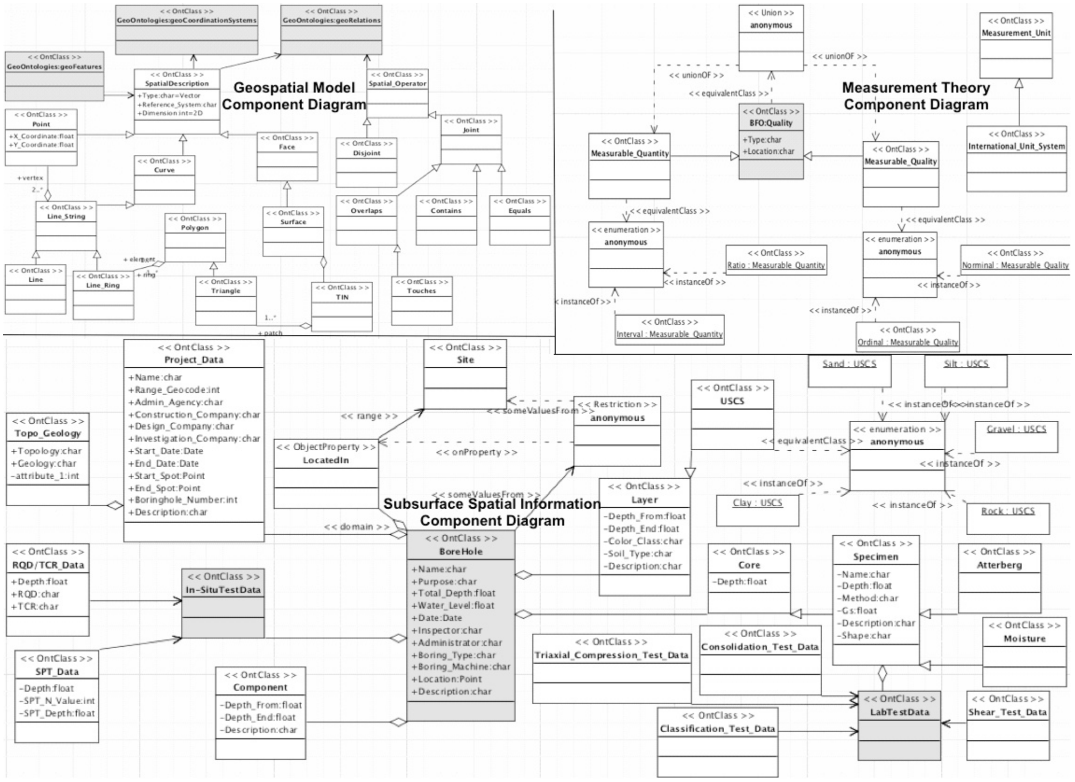

그림 5. 지하공간정보 레이어 UML Profile 모델 일부발췌 (Poseidon for UML 4.0로 작성)

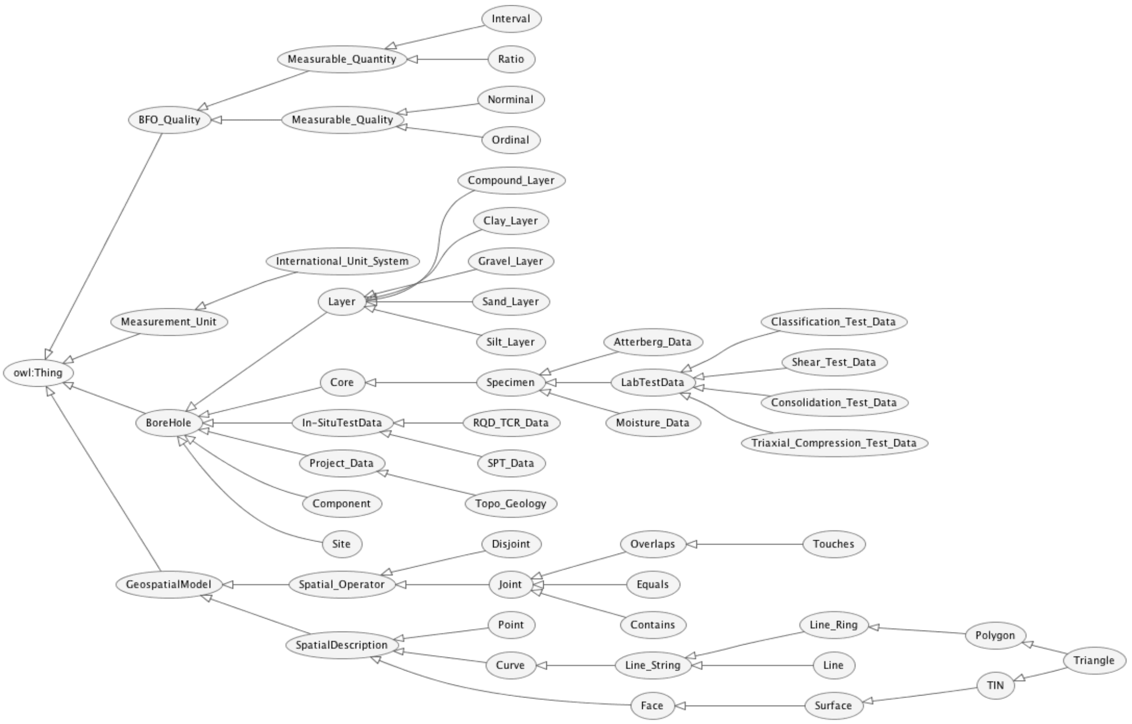

그림 6. 지하공간정보 레이어 OWL모델(클래스) 일부발췌 (Protégé-OWLViz로 작성)

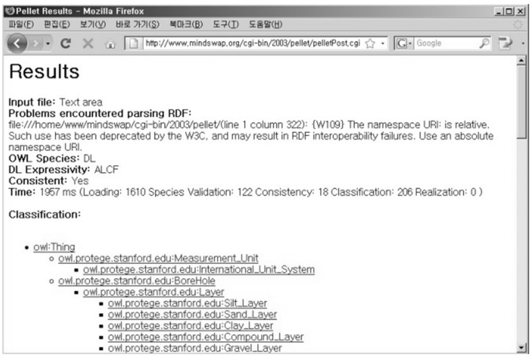

그림 7. Pellet을 이용한 온톨로지 검증

물 최적위치 및 터파기 비용산정 등 지하굴착을 포함하는 모든 건설행위에서 반드시 필요한 정보이다. 그러나, 현재 구축된 지하공간정보 만으로는 대규모 건설공사에 적합한 공간적 분석에는 한계가 있기 때문에, 지형도, 지질도, 지하시설물도 등 관련 공간정보의 공유가 필수적이다.

공간정보의 시멘틱은 개념과 관계로 나뉠 수 있다. 시멘틱 개념은 “body of water”를 “lake” 혹은 “sea”로 인식하는 것처럼 동일한 기호에 대해 서로 다른 개념을 가리키는 것을 나타낸다. 또한, 개념표현과정에서도 “lake”를 포인트 혹은 폴리곤으로 다르게 표현하는 시멘틱 세분화(semantic granularity) 문제를 고려해야 한다. 본 연구에서는 상위 온톨로지의 일반개념과 어휘사전, 측정온톨로지의 측정단위, 그리고 지하공간정보 온톨로지를 통해 이러한 시멘틱 개념의 문제를 고려하였다. 시멘틱 관계의 경우에는 정량적인 관계를 주로 다루는 현행 GIS시스템에서 “near”, “far” 등의 자연어를 통한 관계표현이 이뤄져야 한다. 본 연구에서는 시멘틱 관계를 위상관계(e.g. “contains”), 방향관계(e.g. “isSouthOf”), 거리관계(e.g. “isCloseTo”), 부분-전체관계(e.g. “isPartOf”)로 나누어 모델에 적용하였다.

기존 정보공유에서 이러한 시멘틱 이질성을 극복하기 위해서는 전문가의 해석이 수반되어야 했다. 본 연구에서 개발한 온톨로지 레이어 모델은 그림 8과 같이 시멘틱을 고려한 GIS 공유를 가능케하고, 추론을 통해 자동 공유가 가능한 기반이 된다.

또한, 온톨로지는 사용자와 컴퓨터간의 자유로운 통신을 가능케하고, 관련 지식을 표현, 저장할 수 있어 시스템의 활용 폭을 확장시킬 수 있다. 일반적인 전문가시스템은 문제를 해결하는 PSM(Problem Solving Mehtod)과 해당 도메인의 지식모듈로 나뉜다. 온톨로지는 도메인의

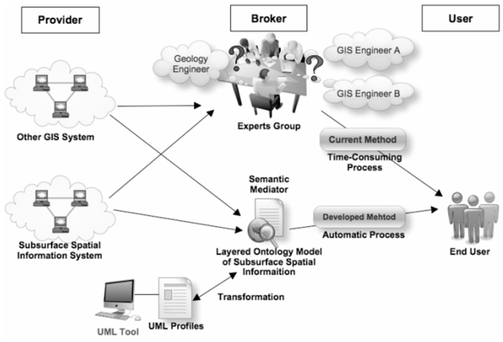

그림 8. 온톨로지를 이용한 시멘틱 공유 활용

지식을 표현, 정리하기 때문에 본 연구성과로 GIS 전문가시스템을 구현하여 기존 특정업무에 맞춰진 GIS시스템을 다양한 시나리오에서 활용가능하게 할 수 있다.

5. 결론 및 향후연구

현재의 GIS시스템은 각 관리주체별로 서로 다른 목적에 따라 공간정보를 생산하고 있어 근본적으로 이질성을 지닌다. 정보의 문법적, 구조적 이질성은 변환포맷 혹은 표준화를 통해 해결 가능하지만, 시멘틱 이질성은 해결하기 어렵다.

본 연구는 국가공간정보체계의 하나인 지하공간정보의 시멘틱 공유를 위하여 온톨로지 모델을 개발하였다. 개발된 온톨로지 모델은 GIS의 다양한 도메인 특성을 고려하여 상위온톨로지, 측정, 공간, 그리고 지하공간정보 온톨로지가 레이어 형태로 구성된다. 이를 통해 시멘틱을 고려한 지하공간정보 시스템의 정보공유 기반을 마련하였다.

또한, 본 연구에서는 도메인 전문가의 경험과 직관에 의존하는 휴리스틱한 온톨로지 개발방법론이 아닌, 표준 소프트웨어 개발방법론인 MDA를 적용하였다. UML 2.0인 MDA는 메타모델을 이용하여 누구나 쉽게 재사용 가능한 온톨로지 개발 환경을 구현한다. UML은 직관적이며 다양한 시각화 도구가 존재할 뿐만 아니라, 환경변화에도 메타모델 변경만으로 쉽게 온톨로지를 재생산할 수 있다. 따라서, 본 연구를 통해 개발된 지하공간정보 UML Profile 모델 수정만으로 손쉽게 지하공간정보 온톨로지를 수정/보완할 수 있다.

향후에는 개발된 지하공간정보 온톨로지인 OWL모델을 기초로 OWL-S(Web Ontology Language for Services)로 확장하여, 공간정보 웹서비스에서 시멘틱 정보공유와 공간쿼리가 가능한 시멘틱GIS를 구현할 계획이다.

감사의 글

본 연구는 국토해양부 첨단도시기술개발사업 - 지능형국토정보기술혁신사업과제의 연구비지원 (06국토정보C01)에 의해 수행되었습니다.

참고문헌

1. 건설교통부, 2000, “제2차 국가지리정보체계 기본계획”, pp. 12-13.
2. 국토해양부, 2008, “국토지반정보DB구축 사업 최종보고서”, pp. 7-10.
3. 지반정보DB포털, http://www.geoinfo.or.kr.
4. Baclawski, K., et al., 2001, “Extending UML to Support Ontology Engineering for the Semantic Web”, 4th Intl. Conf. on UML, Toronto.
5. Breitman K. K., Casanova M. A., Truszkowski, W., 2007, “Methods for Ontology Development in Semantic Web : Concepts, Technologies and Applications”, NASA Mono-graphs in Systems and Software Engineering, Springer London, pp.155-173.
6. Burton-Jones, A., et al., 2005, “A semiotic metrics suite for assessing the quality of ontologies”, Data Knowledge Engineering.
7. 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.
8. Cranefield, S., Purvis, M., 1999, “UML as an Ontology Modeling Language”, In Proc. of the Workshop on Intelligent Information Integration.
9. Devogele, T., Parent C., Spaccapietra S., 1998, “On Spatial Database Integration”, Intl. Journal of Geographic Information Science, pp.335-352.
10. Duric, D., Gasevic, D., Devdzic, V., 2003, “A MDA-based Approach to the Ontology Definition Metamodel”, 6th Int. Conf. on Information Technology, pp.193-196.
11. FGDC, 1998, “Content Standard for Digital Geospatial Metadata”.
12. Geoontologies, http://www.mindswap.org/2004/geo/geoOntologies.shtml.
13. 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.
14. Gruber, T., 1992, “A Translation Approach to Portable Ontology Specification”, Knowledge Systems Laboratory, Stanford University, Stanford, CA, Technical Report KSL 92-71.
15. Meersman, R., Spyns, P., Jarrar, M., 2002, “Data modelling versus ontology engineering”, ACM SIGMO Vol.31, Issue4, pp. 12-17.

16. Mika, P., et al., 2004, "Foundations for Service Ontologies: Aligning OWL-S to DOLCE", WWW2004, New York, NY.
17. Miller, J., Mukerji, J., 2003, "MDA Guide Version 1.0", OMG.
18. MUSIL, http://musil.uni-muenster.de.
19. Nicola A. D., Missikoff M., Navigli R., 2009, "A Software engineering approach to ontology building", Information Systems, pp. 258-275.
20. OMG(Object Management Group), http://www.omg.org.
21. OntoGeo, http://ontogeo.ntua.gr.
22. OntoSpace, http://www.ontospace.uni-bremen.de.
23. PEER, 2004, "Archiving and Web Dissemination of Geo-technical Data: Development of a Pilot Geotechnical Virtual Data Center", Consortium of Organizations for Strong-Motion Observation Systems.
24. Pellet, http://www.mindswap.org/2003/pellet/demo.shtml.
25. Poseidon for UML, http://www.gentleware.com.
26. Protégé, http://protege.stanford.edu.
27. SEEK, http://seek.ecoinformatics.org.
28. SWEET, http://sweet.jpl.nasa.gov.
29. Smith, B., 1995, "On Drawing Lines on a Map", Spatial Information Theory, COSIT 95, pp. 475-484.
30. Smith, B., Mark, D.M., 1998, "Ontology and Geographic Kinds", Intl. Symposium on Spatial Data Handling, Vancouver, pp. 308-320.
31. Stuckenschmidt, H., Visser, U., 2000, "Semantic translation based on approximate re-classification", 7th Intl. Conf. on Principles of Knowledge Representation and Resoning.
32. Teller, J., 2007, "Ontologies for an Improved Communication in Urban Development Projects", Ontologies for Urban Development, Springer, pp. 1-14.
33. Torres, M., et al., 2005, "Ontology-Driven Description of Spatial Data for Their Semantic", 5th Intl. Conference, GeoS, pp. 242-249.
34. USGS-ONT, http://www.geog.buffalo.edu/ncgia/i21/i21ontology.html.
35. W3C(World Wide Web Consortium), http://www.w3c.org.
36. W3C Geospatial Incubator Group, http://www.w3.org/2005/incubator/geo.
37. Xalan, http://xml.apache.org/xalan-j.