연안정보의 지오-온톨로지 적용에 관한 연구

강전영* · 황철수**

A Study on Geo-Ontological Application of Coastal Information

Jeon-Young Kang* · Chulsue Hwang**

요약 : 현재 우리나라의 연안정보는 단순한 질의 정도의 초보적 이용에 그치고 있다. 따라서 실질적 업무나 정책 활용에 적용하기 위해서는 구축된 정보를 재가공해야 하는 불편함을 초래한다. 관련 정보들을 의미적으로 연결하여 효율적인 활용을 가능토록 구현할 수 있는 온톨로지 기법을 이런 의미에서 주목할 필요가 있다. 본 연구는 Geo-Ontology를 이용하여 연안정보를 관리하기 위한 공간적 온톨로지 모형을 제시하고, 연안정보 온톨로지의 활용 방안에 대해서 모색하였다. 첫째, 지리 정보와 관련한 온톨로지에 대한 연구가 국내외에서 미약하기 때문에 관련 선행연구를 고찰하였다. 둘째, 연안의 지리적 사상과 관계를 정의한 Geo-Ontology를 구축하였다. 셋째, Geo-Ontology를 이용하여 사례지역인 마산만 연안의 유역권에 해당하는 연안정보 온톨로지를 구축하였다. 넷째, 의미 기반의 질의와 추론을 하고 연안정보 온톨로지를 평가하였다. 이러한 링크드 데이터를 대비하는 온톨로지 기반 접근은 장차 공간정보의 통합적 활용과 공간데이터의 공개를 통한 사회·경제적 활용의 대홍수시대를 대비하는 기초적 연구로 의미를 가진다.

주요어 : 온톨로지, 의미기반 질의, 추론, 지오-온톨로지

Abstract : It is unsuitable for Korean coastal information to work specific tasks because the coastal information of the current provides simple information, and thus the coastal information is required to reprocess. Therefore, this paper intends to present the ontology model for managing the coastal information using Geo-Ontology and seek application of ontology. The contents of this paper follow as; First of all, I considered the base theories for ontology and related researches. Second, I built Geo-Ontology which defines taxonomy of geographical features and their relationships. Third, I designed and implemented the coastal information ontology about basin of coast, Masan, using Geo-Ontology. Fourth, I carried out semantic queries and reasoning, assessment of the coastal information ontology. This paper will be a base study for many projects which are currently being conducted to integrate spatial information for more effective administrative works and easier maintenance and management of data. Also, this paper is significant in the sense that it is the study preparing for linked data.

Key Words : Ontology, Semantic Query, Reasoning, Geo-Ontology

---

* 경희대학교 지리학과 대학원생(Graduate Student, Department of Geography, Kyung Hee University), jykang22@gmail.com
** 경희대학교 지리학과 교수(Professor, Department of Geography, Kyung Hee University), hcs@khu.ac.kr

1. 서론

1) 연구배경 및 목적

정보 기술의 발달과 웹 기술의 발전은 다양한 정보 공유와 생산을 가능하게 하였다. 특히 지리정보는 지도, 이미지, 테이블 등의 다양한 형태로 웹에 존재하게 되었다. 이러한 다양한 정보들을 국가적 차원에서 관리하기 위해 그간 많은 시도들이 있었다. 우리나라 정부는 1996년부터 행정업무의 효율성 증진과 데이터의 유지 및 관리의 목적으로 그간 수조원의 예산을 투입하여 약 36개 기관(중앙부처 20개, 광역지방자치단체 16개, 총 728개 사업)에서 공간정보 사업을 추진하였다. 또한 관련된 사업과 이에 따른 대규모 예산 투입은 지속적으로 발생하고 있는 실정이다. 2000년 1월에는「국가지리정보체계의 구축 및 활용 등에 관한 법률」을 제정함으로써 5년 단위의 기본계획과 연차별 시행계획을 수립하여 명실공히 국가적 사업으로 확립되었다.

그런데 대규모 예상을 투입하여 구축되고 있는 방대한 공간정보에 대한 기존 웹 기술의 접근은 근본적으로 일종의 키워드를 통해서만 가능하기 때문에 사용자들이 이러한 정보들을 활용하기에는 어려움이 있다. 한마디로 정보의 활용에 시간과 비용, 그리고 결국 여러 형태의 자원이 낭비된다. 이 밖에도 정보 통합에 따른 공유의 문제, 기 구축된 정보들의 상호 운용성에 대한 문제, 의미적으로 다르지만 동일한 용어를 사용함에 따라 상이한 결과가 나타나는 문제 등을 내포하고 있다. 이와 같은 문제의 근본적인 원인은 컴퓨터로 대표되는 정보통신기기가 정보자원의 의미를 이해하지 못함에 있다.

정보의 의미적 이질성을 해결하고 정보의 재사용을 위해 등장한 것이 시맨틱 웹(Semantic Web)이다. 시맨틱 웹은 기존의 웹을 확장하여 컴퓨터가 이해할 수 있도록 잘 정의된 의미를 기반으로 의미적 상호운용성(Interoperability)을 실현하는데 목적이 있다(Lee *et al.*, 2001). 그리고 온톨로지는 이러한 시맨틱 웹을 구현하기 위해 실세계의 지식을 머신이 처리 가능한 형태로 표현하는 것이다. 또한 온톨로지는 추론규칙을 포함할 수 있기 때문에 새로운 사실을 자동으로 추출하거나 제약 조건에 맞지 않는 오류를 찾아낼 수 있다. 따라서 온톨로지는 지식을 수집하고 표현하는데 사용될 뿐 아니라 지식의 모델링 및 검색 시에 유용한 수단으로 이용될 수 있다(Pretorius, 2004).

본 연구는 지오-온톨로지(Geo-Ontology)를 이용한 연안정보의 지식 정보화에 목표를 두고 있다. 즉 추론과 의미에 기반한 검색을 통하여 연안정보 온톨로지를 설계하고 이를 구축하여 그 활용 방안을 모색해보고자 한다. 현재 우리 나라는 연안 관리의 중요성에도 불구하고 연안 관리에 적합한 연안정보 데이터베이스가 구축되어 있지 않다. 또한 공개하고 있는 연안정보들은 단순한 정보 제공의 수준에 그치기 때문에 특정한 업무를 하기에 적합하지 않다.

2) 연구내용 및 방법

본 연구의 내용은 다음과 같이 나뉜다. 첫째, 온톨로지의 근간이 되는 이론들과 온톨로지를 통해 지리 정보를 설계 및 구축한 선행 연구를 고찰한다. 둘째, 지리적 사상에 대한 분류체계 및 관계를 정의하는, Geo-Ontology를 구축한다. 셋째, 구축한 Geo-Ontology를 이용하여 연안정보 온톨로지를 설계 및 구현한다. 넷째, 의미기반의 질의와 추론을 실시한다. 질의의 결과와 실제 사실의 비교를 통해 구축된 연안정보 온톨로지를 평가한다.

본 연구를 진행하기 위해 필요한 지리적 지식을 설계하고 구축하기 위해서 Protégé를 이용하였다. Protégé는 1987년 의료 전문가 시스템을 위한 지식 획득을 지원하기 위해 스탠포드 대학교의 의료 정보과 학팀에 의해 개발되었다. 지식을 표현하는 언어로는 W3C에서 권고하는 OWL(Ontology Web Language)를 사용하였으며, 이는 표현력이 가장 뛰어난 시맨틱 웹 온톨로지 언어로 평가 받고 있다.

구축된 연안정보 온톨로지의 추론을 위해서 이용된 추론 엔진은 FaCT++이다. FaCT++는 기술 논리(Description Logic) 추론을 위한 시스템이며, 온톨로지의 논리적 일치성(Logical Consistency)를 검사하

며, 의미적 관계들을 통한 계층을 재구성할 수 있다(Tsarkov and Horrocks, 2006).

2. 온톨로지와 Geo-Ontology

1) 온톨로지의 개념

온톨로지에 관한 연구는 기원전 4세기 고대 그리스의 철학자 플라톤과 아리스토텔레스의 질문, "무엇이 존재하는가?"와 연관된다. 그들은 존재하는 것이 무엇인지, 존재하는 것들이 가지고 있는 공통된 특징이 무엇인지 파악하기 위해 노력하였다. 이러한 질문에 대한 해답으로 고대 그리스 철학자들은 현실 세계를 단순화하기 위해 노력했었다. 이에 온톨로지는 존재하는 것과 존재하는 것들의 기본적인 범주에 대한 연구라고 할 수 있다(Calero *et al.*, 2006).

온톨로지의 정의 중 가장 많이 인용되고 있는 것은 Gruber(1993)의 정의로,"온톨로지는 개념화의 명세이다(An Ontology is a specification of a conceptualization)." Borst(1997)가 이후 다음과 같이 의미를 확장시켰다."온톨로지는 공유된 개념의 형식적 명세서이다(An Ontology is formal specification of a shared conceptualization)."[1)]

2) 지리적 사상과 Geo-Ontology

Geo-Ontology의 대상은 바로 지리적 사상이다. 지리적 사상을 지식화하고 Geo-Ontology로 구축하는 과정은 다음과 같이 정리할 수 있다(Figure 1). 지리적 사상은 개념화의 과정에서부터 시작된다. 개념화라는 것은 지리적 대상이 무엇인가에 대해서 분류하는 과정이며 지리적 사상과 사상들 간의 관계를 파악하는 것이다. 하지만 개념화의 과정은 지리적 사상을 어떻게 바라보아야 하는지에 대한 관점이 필요하다. 왜냐하면 지리적 사상이 이산적인 것인지 연속적인 것인지 확인해야만 지리적 사상을 명확하게 정의할 수 있기 때문이다. 개념화 과정을 거친 지리적 사상은 온톨로지의 구축 방법론의 순서에 따라 Geo-Ontology로 구축하게 된다.

온톨로지 연구는 상위 온톨로지에 대한 연구와 특정 업무 및 주제 혹은 도메인의 온톨로지에 대한 연구로 구분할 수 있다. 상위 온톨로지는 일반적인 개념들을 기술하는 것으로, 지리적 연구와 관련한 대상은 공간과 시간이다. 지리적 사상을 지식으로 변환시키기 위한 지리적 사상에 대한 정의와 더불어 이들의 관계에 대한 정의가 이에 해당한다. 이때, 공간을 다루는 온톨로지를 Geo-Ontology로 칭하고 있으며 Geo-Ontology는 위상관계, 부분-전체 관계가 지리적 도메인에서의 주요 역할 이라는 점에서 다른 온톨로지

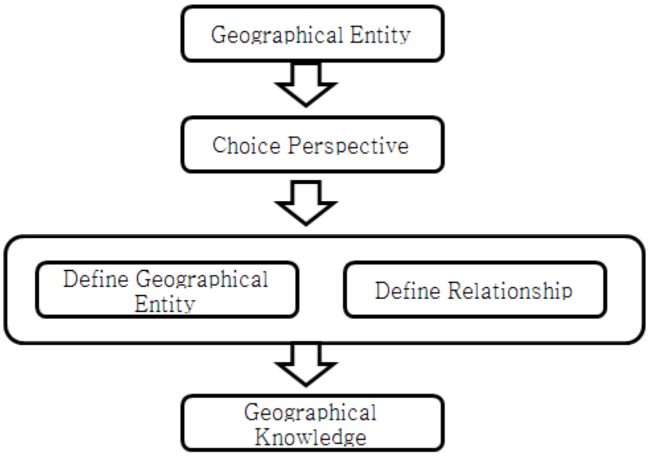

Figure 1. Transforming geographic entity into geographic knowledge. 지리적 사상을 지리적 지식으로 변환하는 과정

와 차이를 갖는다.

상위 온톨로지 연구로서, BudakArpinar *et al.*(2006)는 지리적 관계에 대한 유형을 위상적 관계, 방위적 관계, 근접성 관계로 정의하여 지리적 관계의 유형을 보다 확장시켰다. Wang *et al.*(2007)은 공간 정보를 대상으로 복잡한 공간적 위치 관계를 해결하기 위해 온톨로지의 논리적 추론을 활용하였다.

도메인 온톨로지 연구로서, Jones *et al.*(2001)는 영국의 유적 관리 시스템인 OASIS(Ontologically-Augmented Spatial Information System)을 설명하였다. 지리적 개념을 설명하기 위해 위치, 면적 등의 속성을 구성하였다. 고은정 외(2005)는 지역 정보 온톨로지 정보를 이용한 여행 정보 제공시스템을 설계하였다. 여행정보 제공 시스템에 지역 온톨로지 정보를 추가하여, 의미적으로 정확한 사례표현과 검색 결과를 생성하도록 하였다. Hong(2006)은 상이한 지리적 지식을 가진 집단 간의 랜드마크와 관련된 지식을 구조화하고 이를 통해 집단간의 지식의 공유를 제안하였다. Lee(2007)은 모바일 환경에서 사용자의 행태와 부합한 관광정보 서비스를 제공하기 위한 방법으로서 시맨틱 웹을 활용하였다. 이를 위해 관광지, 관광자원, 관광객의 활동 등을 체계적으로 개념화 및 조직화된 PARA(Place-Attrcation-Resource-Activity) 온톨로지를 구축하였다. 박지만(2009)은 공간적 의사결정 과정에 있어서 시민참여가 가능한 공공참여형태에 초점을 맞추어 개인의 심리적 인식과 지리적 지식을 온톨로지를 통해 구축하였다.

온톨로지의 기본 이론과 선행연구를 통해 도출된 연구의 시사점은 다음과 같다. 첫째, 기존의 상위 온톨로지의 연구 결과를 국내의 지리적 사상에 적용시키기는 어려움이 있다. 왜냐하면 지리적 사상들은 동일한 대상이라 하더라도 지리적 특이성에 의해 다른 구조와 형태를 갖기 때문이다. 특히 행정 구역 및 용어 등에서 국내와 국외 간의 차이가 있다. 따라서 국내의 실정이 반영된Geo-Ontology에 대해 연구가 진행되어야 한다. 둘째, 구축된 Geo-Ontology에 대한 검증의 과정이 필요하다. Geo-Ontology의 목적은 지리적 사상에 대한 분류와 관계들의 정의를 통해 지리적 지식을 확장시키기 위함이다. 그렇기 때문에 특정 주제를 Geo-Ontology에 적용하는 등의 검증과정이 반드시 필요하다. 셋째, 구축된 온톨로지는 사상들의 의미적 관계를 통한 추론 혹은 검색이 가능하여야 한다. 온톨로지는 지식과 관련한 데이터라는 의미에서 지식기반이라 할 수 있으며, 지식 내용과 절차적 추론 과정을 포함한다. 따라서 구축된 지식에 대한 추론이 가능하다. 위의 논의들을 기반으로 본 연구에서는 Geo-Ontology를 구축하고 이를 연안정보에 적용하여 연안정보 온톨로지를 구축하였다.

3. 지오-온톨로지의 구축

온톨로지를 구축하기 위한 방법론으로는 Lopez *et al.*(1999)에 의해 연구된, METHONTOLOGY 온톨로지 방법론을 수정 · 적용하였다. 왜냐하면 본 연구는 Geo-Ontolgoy의 구축 후에 연안정보를 특정 도메인으로 하는 도메인 온톨로지를 구축하기 때문에 특정 주제의 온톨로지 구축을 위한 방법론이 적절할 것으로 판단되었기 때문이다. 따라서 Geo-Ontology의 구축은 명세화(Specification), 개념화(Conceptualization), 형식화(Formalization), 구현(Implementation)의 순서로 진행하였다.

1) 명세화(Specification)

명세화 단계는 온톨로지를 개발하기 위한 사전단계이다. 이 단계에서는 온톨로지의 목적이 무엇인지, 온톨로지 설계를 위한 기초적인 질의는 어떻게 구성할 것인지, 온톨로지가 구축되었을 경우에 검증을 어떻게 할 것인지를 정하는 과정이다.

명세화 단계는 또한 지식 획득 활동과 병행으로 진행되는데 다양한 지식 출처로부터 온톨로지 개발을 위한 지식을 획득한다. 포함되는 명세서의 내용은 온톨로지의 사용 목적, 온톨로지의 구축 범위, 온톨로지의 정형성(Formality) 단계 등이다. 본 Geo-Ontology의 명세화는 Table 1과 같다.

Table 1. Specification of Geo-Ontology. 지오-온톨로지의 명세화 단계

Domain of Ontology	Geographical features and Time
Purpose	According to semantic connecting of geographic features, semantic query is possible and data can improve interoperability with others.
Level of Formality	Rigorously formal
Range of Building	• Class – Coordinate System: Latitude, Longitude Space: Region(Boundary etc.), Geographical Features(Natural Thing, Human Thing) Time: date, week, time • Properties – Spatial Properties: Direction, Distance, Topology Temporal Properties: Pre-time, Post-time
Present Language	OWL DL
Building Tool	Protégé
Builder	Jeon-Young Kang

2) 개념화(Conceptualization)

하나의 온톨로지는 담론의 도메인에서 사용되는 개념들에 대한, 정형화되고 명시적인 기술이라고 할 수 있다. 즉 클래스 또는 개념과 각 개념들의 특성과 속성, 그리고 특성과 속성에 대한 제약규정 등으로 정의된다(Noy and McGuinness, 2001).

(1) 클래스의 정의

본 연구의 최상위클래스는 Thing이다. Thing은 OWL에 내장되어 있는 클래스로서, OWL로 기술하는 모든 사물 및 현상을 담고 있는 클래스이다. 모든 클래스들은 Thing 클래스의 하위 클래스이다. Geo-Ontology에서 Thing의 하위 클래스로는 지리적 사상의 점, 선, 면과 같은 기하학적 특징을 표현한 GeoMetricThing 클래스, 클래스는 특정 지리적 사상의 좌표 체계 및 투영법을 표현한 GeoRefThing 클래스, 인공물, 자연물과 같은 지리적 사상의 특징을 표현한 SpatialThing 클래스, 날짜, 시간, 주를 표현한 TemporalThing으로 구분된다. 본 연구의 Geo-Ontology의 구조는 다음 Figure 2와 같다.

(2) 속성의 정의

본 연구에서는 topObjectProperty가 공간, 시간과 관련한 속성을 포함하는 최상위 속성이다. 공간과 관련한 속성은 hasSpatialRealationship이며 시간과 관련한 속성은 hasTemporalRealationship, 그리고 공간 대 공간 혹은 시간 대 시간 간의 관계를 표현하는 속성은 hasRelationship 이다(Figure 3). hasSpatialRelationship의 하위 속성에는 기하학적 속성을 의미하는 hasTopology 속성과, 좌표체계를 의미하는 hasReference 속성, 거리 정도를 의미하는 hasDistance 속성, 방위적 속성을 의미하는 hasDirection 속성으로 구성되어 있다. 이들 속성은 각각 위상관계, 동서남북의 방위적 관계 등을 의미한다. hasTemporalRealationship 속성은 시간적 전후 관계를 의미하는 속성이다.

속성은 Domain과 Range를 갖는다. 속성의 역할은 Domain의 객체와 Range의 객체 간의 관계를 맺어주는 것이라 할 수 있다. 쉽게 설명하면 Domain은 주어의 집합, 속성은 술어, 목적어는 Range의 집합이라 할 수 있다.

(3) 제약규정의 정의

제약 규정은 Figure 4에서 클래스를 생략한 '속성+제약규정 키워드+클래스'의 구조로 이루어져 있다. 제약 규정의 구조는 시맨틱 웹 언어의 기본 형식과 동일하게 '객체+술어+객체'에서 제약규정 키워드가 추가된 구조이다. 제약 규정 키워드는 '하나 이상의'라

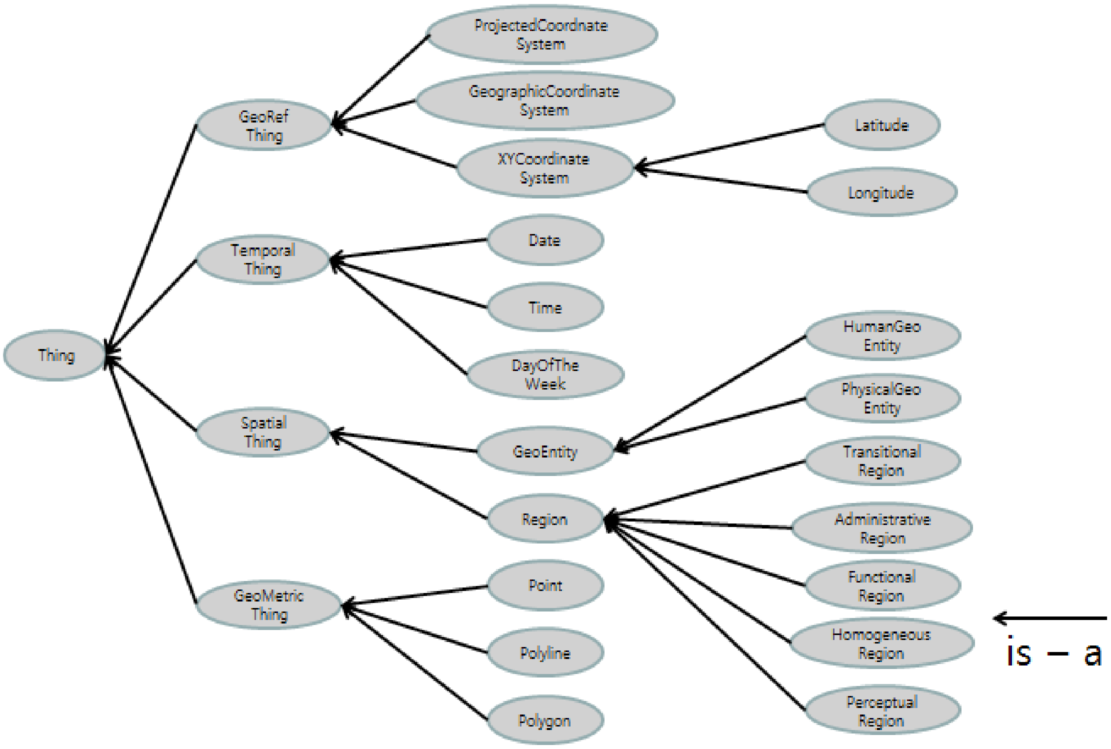

Figure 2. Structure of Geo-Ontology in this study. 지오–온톨로지의 구조

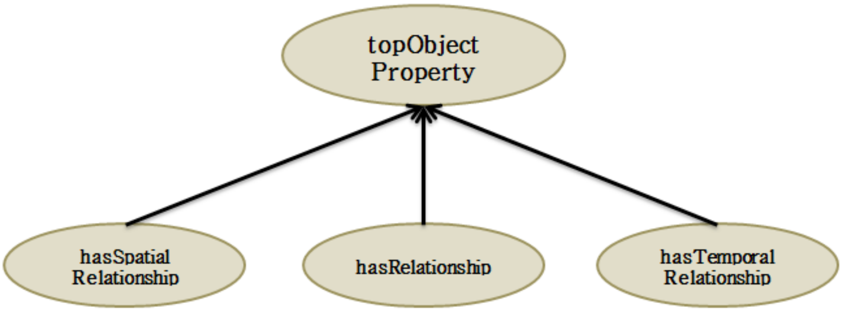

Figure 3. Upper-Property of Geo-Ontology(modified from Wang, Y. *et al.*,(2007)). 지오–온톨로지의 상위 속성

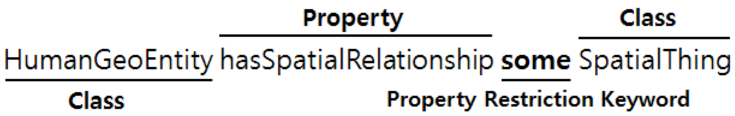

Figure 4. Example of Property Restrictions. 제약규정의 정의

는 의미를 담고 있는 some과 '단 하나의'라는 의미를 담고 있는 'only'가 있다.

(4) 형식화 및 구현

이 단계에서는 개념화 단계에서 정의한 클래스와 속성, 그리고 제약 규정들을 온톨로지 언어로 표현하였다. Geo-Ontology는 Protégé를 이용하여 OWL DL로 생성하였다. Protégé는 지식 수준에서 작성한 개념화 모델을 쉽게 온톨로지 언어로 구현할 수 있다.

4. 연안정보 온톨로지 구축

1) 사례 지역

연안은 바다와 육지가 만나는 완충지대로서 기초 생산력이 풍부하고 생물 자원의 산란장 및 서식지 일 뿐만 아니라 관광 및 사회 경제적인 활동이 이루어지는 지역이다. 마산만 연안은 2000년부터 울산연안, 부산연안, 광양만 연안, 시화호 연안, 인천 연안과 더불어 특별관리해역으로 지정되어 2012년까지 수질 2등급을 만들기 위해 노력하고 있다. 특히, 2008년부터 국토해양부와 지자체에서 인근해역의 환경개선을 위한 협약을 체결하고 연안오염총량관리제[2]의 도입 및 시행을 계획하는 등의 해역 또는 해양환경 및 생태계를 보전하기 위해 노력하고 있다. 우리나라의「연안관리법」[3]에서 연안은 연안해역과 연안육역으로 구분한다. 연안해역은 해안선으로부터 지적공부에 등록된 지역까지의 사이 혹은 해안선으로부터 영해의 외측한계까지이며, 연안육역은 무인도서를 포함한 연안해역의 육지쪽 경계선으로부터 500미터, 국가어항 또는 산업단지의 경우에는 1000미터 이내의 육지지역으로 정의하고 있다. 이러한 연안들의 수온 및 오염정도와 관련한 연안정보들은 우리나라의 연안정보의 관리를 위한 연안관리정보시스템(http://www.coast.kr/) 등에서 확인할 수 있다.

본 연구의 사례지역은 마산만 연안에 영향을 미치는 마산만 연안 유역권이다(Figure 5).

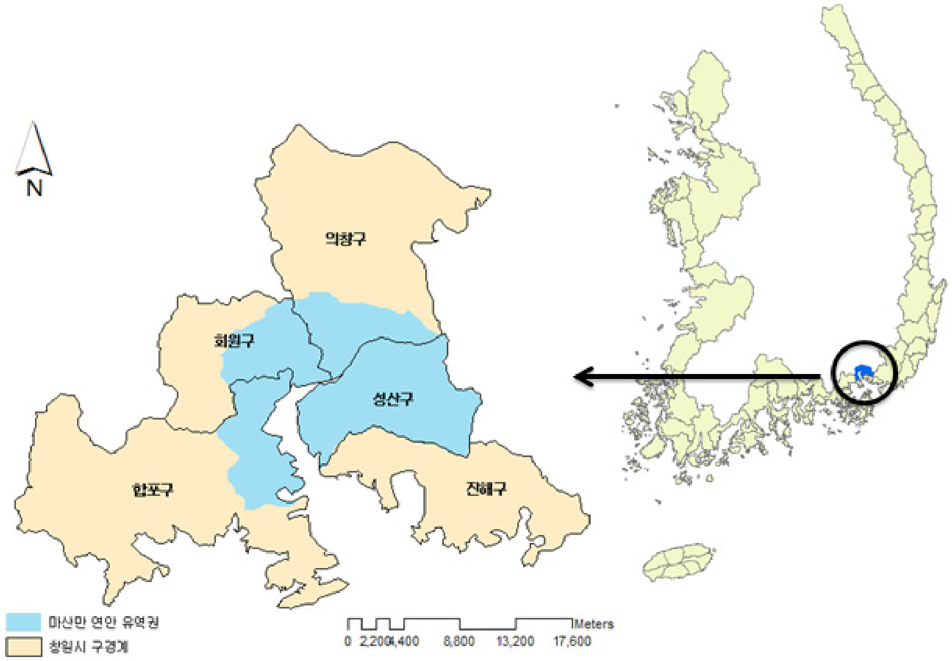

Figure 5. Basin of Coast, Masan. 사례지역인 마산만 연안

2) 연안정보 온톨로지 구축

(1) 연안정보 온톨로지의 구성 및 지리적 객체

연안정보 온톨로지는 의미기반 질의와 효율적인 정보 관리를 지원한다. 이를 위하여 연안정보들과 마산만 연안의 유역권에 속한 지리적 사상들을 온톨로지로 구축하였다. 이때, 연안정보 온톨로지는 Geo-Ontology의 클래스와 속성정보, 제약조건을 재사용하였다. 연안정보 온톨로지를 구축하기 위해서는 Geo-Ontology의 구성을 기반으로 하였고 이에 추가적으로 Geo-Ontology에서 구성한 클래스의 하위클래스와 속성의 하위 속성 등을 구축하였다.

연안정보 온톨로지는 도메인 온톨로지로, 사례지역인 마산만 연안의 유역권 내부의 하천, 행정구역, 마산만 연안의 오염 정보 등의 자료에 한정하여 구축하였다. 각 지리적 객체는 연안정보 온톨로지의 클래스의 인스턴스로 구성된다(Table 2).

연안정보 온톨로지 구축과정은 Geo-Ontology의 구축과정과 유사하게 명세화·개념화·형식화 및 구현단계로 구성된다. 그러나 본 연구에서 연안정보 온톨로지는 단순한 용어의 구조화가 아닌 연안정보라는 특정 도메인에 대한 정의와 속성의 집합체에서 지식정보화를 목적으로 하기 때문에 추론 및 평가를 수행할 수 있는 절차를 추가하여 설계하였다.

(2) 연안정보 온톨로지의 설계 및 구현

① 명세화

연안정보 온톨로지의 영역, 개발 목적, 정형성 수준, 개발 범위, 표현 언어, 개발 도구, 개발자를 표현하고 있는 온톨로지 명세서는 다음 Table 3과 같다.

② 개념화

연안정보 온톨로지를 위하여, Geo-Ontology의 클래스와 속성, 그리고 제약조건을 확장하였다. 온톨로지는 지식의 재사용이 가능하기 때문에 Geo-Ontolo-

Table 2. Examples of Geographic Instance in Coastal Information Ontology. 연안정보 온톨로지의 지리적 사상의 예

Category		Geographic instance
Rivers		Sogye-river, Naedong-river etc.
Administrative boundaries	Do	Gyeongsangnam-do
	Si, Gun, Gu	Changwon-si, Masanhoiwon-gu
	Eup, Myeon, Dong	Palyong-dong etc.
Coast	Coast, Masan	Annual Pollution information(PH, COD, TN, TP) Annual water temperature information
Building Information		Factory, house etc.

Table 3. Specification of Coastal Information Ontology. 연안정보 온톨로지의 명세화 단계

Domain of Ontology	Instance of Basin of Coast, Masan
Purpose	Semantic query and improvement management of coastal information
Level of Formality	Rigorously formal
Range of Building	• Class – Geographic Features in Basin of Coast, Masan • Instance – Information of basin of Coast, Masan • Property – Spatial properties and temporal properties of Instance
Present Language	OWL DL
Building Tool	Protégé
Builder	Jeon-Young Kang

gy의 도메인 온톨로지인, 연안정보 온톨로지의 확장이 가능하다. 즉, 구축한 Geo-Ontology의 클래스와 속성을 활용하여 연안정보 온톨로지를 구축할 수 있다.

Geo-Ontology의 PhysicalGeoEntity 클래스의 하위 클래스로 River 클래스를 추가하였으며, Human-GeoEntity 클래스의 하위 클래스로 Building 클래스와 PriorityManagement 클래스를 추가하였다. AdministrativeRegion 클래스의 하위 클래스로 Do, Gu, Dong 클래스를 추가하였으며, TransitionalRegion 클래스에는 WaterQualityClassification 클래스와 WaterTemperatureClassification 클래스가 하위 클래스로 추가되었다.

Geo-Ontology에서 추가적으로 구성된 클래스와 지리적 객체는 Table 4와 같다.

지리적 객체는 통상적으로 인스턴스라고 하며 최하위 개념이라 할 수 있다. River 클래스의 인스턴스는 소계천, 내동천 등이며 Building 클래스의 인스턴스는 건물_A, 건물_B 등이다.

연안정보 온톨로지를 위해 추가적으로 정의한 속성은 다음과 같다. 데이터 유형 속성으로는 오염의 정도를 나타내는 hasDegreeOfContamination 속성

Table 4. Classes and Instances. 연안정보 온톨로지의 클래스와 인스턴스

Class		Instance
Upper Class	Lower Class	
PhysicalGeoEntity	River	SogyeRiver, NaedongRiver etc.
HumanGeoEntity	Building	Building_A etc.
	PriorityManagement	Building_A etc.
AdministrativeRegion	Do	Gyungsangnam-do
	Gu	Changwon-si, Masanhoiwon-gu
	Dong	Palyong-dong etc.
TransitionalRegion	WaterQualityClassification	Annual Pollution information(PH, COD, TN, TP)
	WaterTemperatureClassification	Annual water temperature information

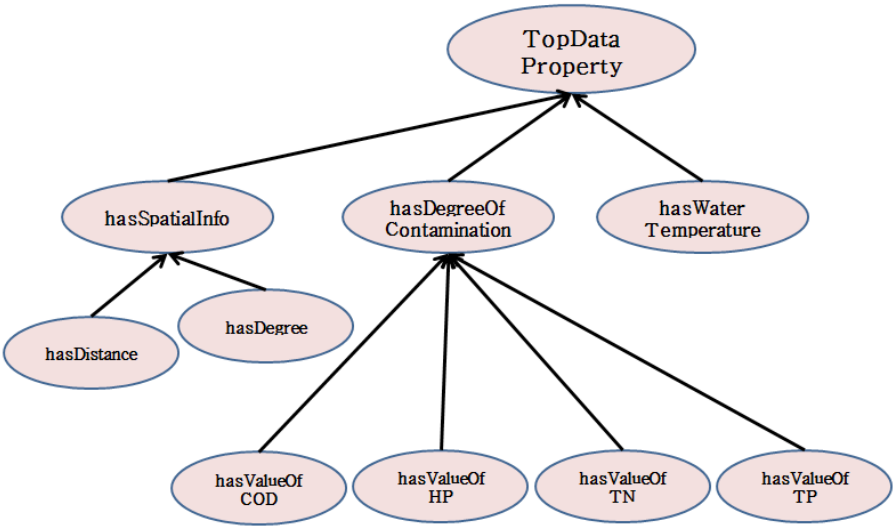

Figure 6. Properties. 연안정보 온톨로지의 속성

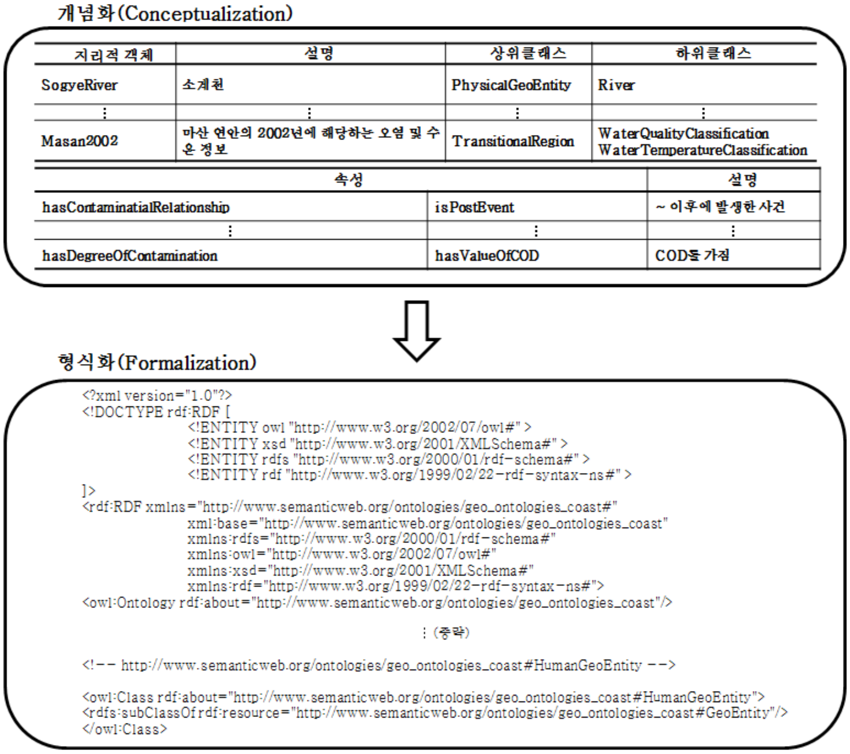

Figure 7. Formalization of Coastal Information Ontology. 연안정보 온톨로지의 형식화 단계

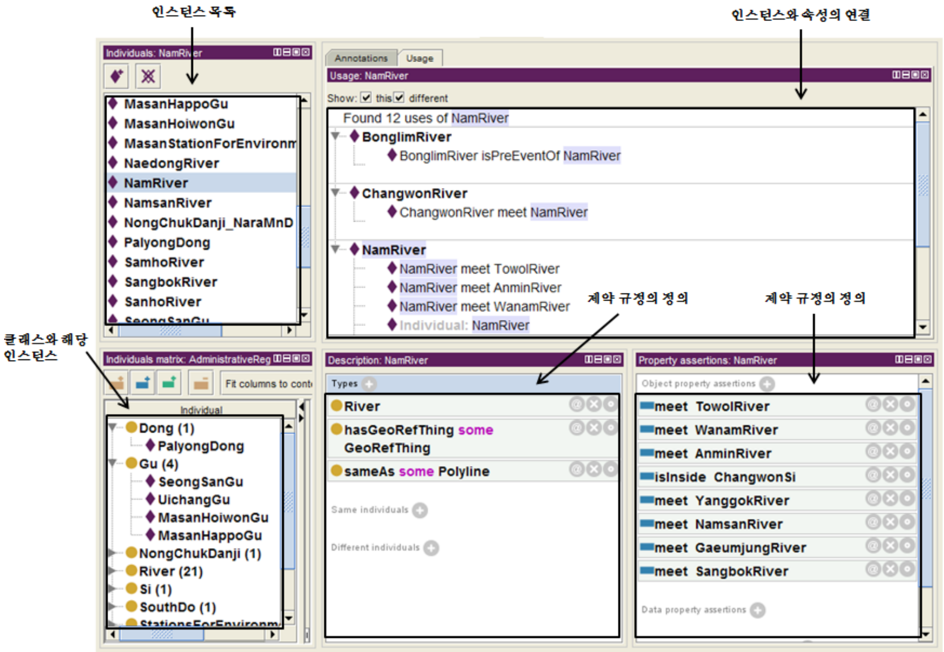

Figure 8. Implementation of Coastal Information Ontology. 연안정보 온톨로지의 구현 단계

을 추가하였다. 이의 하위 속성으로 COD 값을 표현하는 hasValueOfCOD, HP 값을 표현하는 hasValueOfHP, TN 값을 표현하는 hasValueOfTN, TP 값을 표현하는 hasValueOfTP 속성을 추가하였으며 수온을 표현하는 hasWaterTemperature 속성을 추가하였다. 이에 지리적 사상과 관련한 hasSpatialInfo 속성을 추가하였으며 하위 속성으로는 거리를 표현하는 hasDistance, 경사도를 표현하는 hasDegree 속성을 추가하였다.

③ 형식화 및 구현

연안정보 온톨로지의 형식화 단계는 Geo-Ontology에서 이용한 방법과 동일한 온톨로지 언어 OWL DL로 생성하였다. 형식화 단계는 Figure 7과 같고 연안정보 온톨로지를 구현한 것은 Figure 8과 같다.

5. 추론 및 평가

1) 추론의 정의

온톨로지로 구축된 데이터베이스는 각각의 지식의 관계에 따라서 추론이 가능하다. 본 연안정보 온톨로지는 기술 논리의 모델을 채택하고 있는 OWL DL을 이용하여 구축을 하였으므로 기술 논리 추론이 가능하다.

일반적으로 우리가 기술 논리를 통해 가지고 있는 지식을 표현하는 경우 Abox(Assertional Box, 선언적 박스), Tbox(Terminological Box, 술어적 박스)의 구조를 활용하여 지식을 표현한다. Abox는 선언적 지식이며, 인스턴스로 구성되어 있다. Tbox는 술어적 지식이며 온톨로지의 스키마를 포함하고 있다.

기술 논리 추론은 기술 논리를 통해 지식을 표현하는 방법에 따라 크게Abox와Tbox 추론으로 나뉜다. Abox 추론은 인스턴스를 반영한 추론으로, 인스턴스화(Instantiation), 현실화(Realization), 검색(Retrieval)추론으로 나뉜다. 인스턴스화 추론은 특정 인스턴스가 어떠한 클래스에 포함되는 지를 확인하는 것이며, 현실화 추론은 특정 인스턴스가 속하는 가장 특정한 클래스를 알아내는 것이다. 그리고 검색추론은 클래스에 속하는 인스턴스를 모두 알아내는 추론이다. Tbox추론은 개념간의 관계를 명시하는 추론으로, 포섭(Subsuption) 추론, 분류(Classification) 추론, 일관성(Consistency) 추론으로 나뉜다. 포섭 추론은 임의의 개념이 주어졌을 때 상·하위 관계를 구축하는 추론이며, 분류추론은 주어진 임의의 개념에 대한 포섭 혹은 포함 여부를 결정하는 추론이다. 그리고 일관성 추론은 개념 정의에 대한 일관성을 체크하는 것이다(Choi, J.H. and Park, Y.T, 2006).

하지만 기술 논리를 통해 구축된 온톨로지에는 단점이 존재한다. 이러한 단점은 기술 논리에 의한 문제점으로 다음과 같다. 첫째, 수식이나 함수 호출을 표현할 수 없다는 점이다. 이러한 이유로 지식들 간의 관계는 명확한 논리를 통해서만 정의할 수 있다. 하지만 범위의 설정과 같은 규칙을 이용하면 단점을 어느 정도 극복할 수 있다. 둘째, 불규칙한 예측 혹은 불가능한 요인에 의해 발생하는 불규칙성 비단조 추론(Nonmonotonic Reasoning)이 불가능하다. 이는 결국 개념에 대한 새로운 관계의 정의가 기존에 참으로 정의한 진술을 거짓으로 바꿀 수 없다. 따라서 복잡한 사회 현상을 밝혀내기 위한 추론의 방법으로는 부적합할 수 있다.

추론의 과정은 앞서 연안정보 온톨로지의 구축과정에서 설명했듯이, 기술논리의 충돌의 유무를 확인하는 일치성(Consistency)의 검사를 선행하고 일치성의 오류가 없다면 추론을 진행한다. 본 연구에서는 FaCT++ 를 이용하여 T-box 추론을 실시하였다.

온톨로지의 추론은 계층이 정의되지 않은 인스턴스 및 클래스를 속성을 통해 계층의 분류가 가능하다. 연도별 마산만 연안의 수질 오염 정보와 수온 정보, 마산만 연안 유역권에 해당하는 건물 정보를 이용하였다.

Table 5는 연도별 마산만 연안의 수질 오염 정보와 수온 정보를 나타낸 표이다. 수질 조사 시에 고려되는 오염물질 항목은 수온, 수소이온농도(PH), 전기전도도, 총부유물질(TSS), 화학적산소요구량(COD), 총질소(T-N), 총인(T-P) 등이다. 이 중에서 본 연구에서

Table 5. Water Pollution and Temperature Information of Coast, Masan. 마산만 연안의 수질 오염, 수온 정보

Unit(mg/L)

Year	HP	COD	TN	TP	HP
2002	7.99	2.515	1.066	0.1345	14.95
2003	7.82	2.685	0.78	0.146	15.1
2004	8.095	2.325	0.7375	0.084	15.5
2005	8.02	2.54	1.0935	0.1205	15.2
2006	8.135	2.5	0.639	0.72	15.4
2007	7.965	2.435	0.514	0.605	15.4
2008	8.11	2.26	0.5135	0.0685	15.25

Source: National Fisheries Research and Development Institute(2009), 출처: 국립수산과학원(2009)

Table 6. Classification standard for Coast, Masan. 마산만 연안의 수질 분류 기준

Unit(mg/L)

Quality of Water Class	HP	COD	TN	TP
1 Class	7.8~8.3	Less than1	Less than 0.3	Less than 0.03
2 Class	6.5~8.5	Less than2	Less than 0.6	Less than 0.05
3 Class	6.5~8.5	Less than 4	Less than 1	Less than0.09
4 Class	Except	Less than	Except	Except

Source: Basic Act on Environmental Policy No.3 Ra(Standard of Waters) to Modify,
출처: 환경정책기본법 시행령 별표1 제3호 라 목(해역기준)을 수정

Table 7. Classification standard for Coast, Masan. 마산만 연안의 수온 분류 기준

Unit(℃)

Classification Waters	Water Temperature
High Temperature Waters	Over 15.0
Low Temperature Waters	Under 15.0

는 수소이온농도, 화학적산소요구량, 총질소, 총인, 수온 정보를 이용하였다.

Table 6과 Table 7은 마산만 연안의 수질과 수온에 따라 분류하기 위한 분류 기준표이다. Table 8은 실제 해역 수질 등급을 분류하기 위한 기준표를 본 연구에 맞게 수정하였다.

연도별 마산만 연안의 수질 오염 정보와 수온 정보를 분류하기 위해서 점이지대를 의미하는 TransitionalRegion 클래스에서 연안을 의미하는 Coast 클래스를 구성하였으며 Coast 클래스의 하위 클래스로 수질 오염 정도를 의미하는 WaterQualityClassification의 클래스와 수온을 의미하는 WaterTemperatureClassification을 구성하였다. 추가적으로 수질 등급이 1등급은 CoastClass1로, 2등급은 CoastClass2로, 3등급은 CoastClass3으로, 마지막으로 4등급은 CoastClass4로 구성하였으며, 고온 해역은 HighTemperature 로, 저온 해역은 LowTemperature로 구성하였다.

본 연구의 추론은 대표적인 T-box 추론이라고 할 수 있다. 마산만 연안들의 수질 분류 클래스와 수온 분류 클래스에 속하는 마산만 연안정보 인스턴스와 우선 관리 대상 클래스에 속하는 건물 인스턴스가 어떠한 클래스에 포함되는지 알 수 있는 T-box 추론이다. Table 8은 본 연구에서 구현된 T-box 추론식이다. 연안을 의미하는 Coast 클래스는 점이지대를 의미

Table 8. T-box Reasoning. T−box 추론식

TransitionalRegion ≡ SpatialThing∧Region
Coast ≡ Transitional Region ∧ ≥1hasCostalInformation
CoastGrade1≡Coast∧∃hasCostalInformation∧hasPHof∧hasCODof∧hasTNof∧hasTPof∧hasTemperatureof

Table 9. Result of Reasoning. 추론 결과

Class	Explanation	Instance
CoastClass3	3 Class	Masan2004, Masan2008
CoastClass4	4 Class	Masan2002, Masan2003, Masan2005, Masan2006, Masan2007
HighTemperature	High Temperature Waters	Masan2002
LowTemperature	Low Temperature Waters	Masan2003, Masan2004, Masan2005, Masan2006, Masan2007, Masan2008

하는 TransitionalRegion 클래스에 포함되며 하나 이상의 연안정보를 포함한다. 또한 TransitionalRegion 클래스는 지역을 의미하는 Region 클래스와 지리적 사상을 의미하는 SpatialThing 클래스에 포함된다. Coast 클래스는 Masan2002, Masan2003과 같은 인스턴스를 가진다. CoastGrade1 클래스는 Coast 클래스이며 hasPHof, hasCODof, hasTNof, hasTPof, 그리고 hasTemperatureof의 속성을 가져야만 한다.

T-box 추론의 결과는 Table 9와 같다.

2) 정확성 평가

온톨로지에서의 의미 기반 질의는 명시적 지식(Explicit Knowledge)과 암시적 지식(Implicit Knowledge)를 발견하기 위함이지만, 구축된 온톨로지의 정확성을 판단하기 위해 이용되기도 한다. 이에 본 연구에서는 질의 결과를 통해서 연안정보 온톨로지의 정확성 유무를 판단하고자 한다. 정확성은 실제사실에서 참인 값을 갖는 추론 결과비율이며, 정확성을 구하기 위한 공식은 다음 Figure 9와 같다.

$$AC(\%)=\frac{RF \sqcap QR}{RF}\times 100$$

Figure 9. Formation of Accuracy Calculation. 정확성 측정 공식

AC는 Accuracy의 약어로 정확성을 의미하고, RF는 RealFact의 약어로 실제 사실을 의미한다. QR은 QueryResult의 약어로 질의 결과를 의미한다. 의미 기반의 질의를 통한 결과와 실제 사실과의 정확성을 비교하는 것이다. 즉, 위의 공식은 실제 사실에서 실제 사실과 질의 결과가 얼마나 부합하는가를 나타내는 지표이다.

다음 Table 9는 본 연구에서 정확성을 평가하기 위한 질의와 이의 결과, 그리고 정확성을 표현한 것이다. 의미 기반 질의는 '속성+제약규정 키워드+인스턴스' 의 형태로 실시한다. 'meet value NamRiver'를 예를 들어 설명하면, 'NamRiver'와 접하고 있는 하천은 무엇인가에 대한 질의이며, 질의의 결과로 'SanbokRiver', 'GaeumjungRiver','WanamRiver' 등이 도출된다.

정확성 평가 결과는 본 연구에서 구축한 온톨로지에 대해서는 100%의 정확성을 갖는다. 이러한 이유는 다음과 같다. 첫째, 실제 온톨로지를 구축하는 과정에서는 생명주기모형을 따르기 때문에 지속적인 온톨로지의 수정 과정을 거치기 때문이다. 둘째, 구축된 연안정보 온톨로지의 사례 개수가 적기 때문이다. 사례 개수가 적다는 것은 인스턴스 간의 관계가 비교적 단순할 수 있다는 것을 의미한다.

Table 9. Query and Reasoning Result, Accuracy. 질의와 추론 결과 및 정확성

Unit(%)

Query	Result	Accuracy
meet value NamRiver	SanbokRiver, GaeumjungRiver, WanamRiver, NamanRiver, ChangwonRiver, AnminRiver, YanggokRiver, TowolRiver	100
isOutside value Building_D	Palyong-dong, Uichang-gu, ChangWon-si, Gyungsangnam-do	100
isPreTime value Masan2008	Masan2002, Masan2003, Masan2004, Masan 2005, Masan2006, Masan2007,	100
-	-	-
isInside value Gyungsang-SouthDo and AdministrativeReigion	Seongsan-gu, Uichang-gu, Changwon-si, Masanhappo-gu, Masanhoiwon-gu, Palyong-dong	100

6. 결론

본 연구는 연안정보의 관리를 위한 데이터베이스의 구축방법으로서 온톨로지를 제시함이 목적이다. 목적을 달성하기 위하여 지리 정보들의 분류체계 및 관계를 정의하는 Geo-Ontology를 구축하였으며, 구축된 Geo-Ontology를 연안정보 데이터베이스에 적용하고, 이에 대한 정확성 평가를 실시하였다.

현재 지리적 대상과 관련한 온톨로지의 연구는 크게 상위 온톨로지에 관한 연구인 Geo-Ontology연구와 특정 주제를 온톨로지로 구축하는 도메인 온톨로지 연구로 구분된다. Geo-Ontology연구는 지리적 대상의 상위 개념에 대한 연구임에도 불구하고 현재 표준이 없기 때문에 연구자에 따라 개념화와 지리적 대상 간의 관계가 각기 다르게 정의하고 있다. 그리고 도메인 온톨로지 연구는 특정 주제에 대해 추론을 통한 의미적 검색, 혹은 사용자의 특징에 적합한 정보를 제공하는 방안으로 온톨로지를 제시하고 있다.

연구의 결과물인 Geo-Ontology는 지리적 대상을 크게 지리적 대상의 기하학적 형태를 의미하는 GeoMetricThing 클래스와 지리적 대상의 좌표체계를 의미하는 GeoRefThing 클래스와 지리적 대상을 의미하는 SpatialThing 클래스로 구분하여 구축하였다. 이외에 시간적 대상을 의미하는 TemporalThing 클래스를 구축하였다. 이에 대한 속성으로는 같은 의미임을 표현하는 hasRelationship, 공간적 관계를 의미하는 hasSpatialRelationship, 그리고 시간적 관계를 의미하는 hasTemporalRelationship으로 정의하고, 각각의 클래스에 대한 제약조건을 정의하였다.

Geo-Ontology에 대한 사례 연구로는 마산만 연안의 유역권에 대하여 연안정보 온톨로지를 구축하였다. 기존에 구축한 Geo-Ontology의 클래스와 속성, 제약조건 등을 재사용하였으며, 추론을 통한 연안정보 활용을 모색하기 위하여 클래스와 속성, 인스턴스 등을 추가하였다. 연안정보 온톨로지는 기술 논리를 통해 지리적 대상을 정의하였기 때문에 기술 논리 추론이 적합하다. 그리고 연안의 정보는 인스턴스로 구성하여, 이에 적합한 추론을 실시하였다. 마지막으로 의미 기반의 결과와 실제 사실의 비교를 통해 연안정보 온톨로지의 정확성을 평가하였다.

본 연구는 연안정보의 효율적 관리를 위해 온톨로지를 제안하였고, 구축된 온톨로지를 통한 활용방안에 대해서 모색하였다. 구축된 연안정보 온톨로지는 연안의 수질 등급 및 수온의 변화를 보다 쉽게 파악할 수 있기 때문에 연안정보 및 연안을 관리하는 데 있어 보다 효과적이다. 즉, 본 연구는 연안정보와 관련된 모니터링 업무에 효율성을 증가시킬 수 있다는 점으로 미루어보아, 현재 행정업무의 효율성 증진과 데이터의 유지 및 관리의 목적으로 진행하고 있는 국가공간정보 통합 사업과 관련한 기초 연구가 될 수 있을 것으로 판단된다. 또한, 본 연구에서 방법론으로 이용한 온톨로지는 온톨로지를 구성하는 인스턴스에 대해 URI을 부여하고, 공유한다는 점이 링크드 데이

터와 유사하다고 볼 수 있다. 따라서 데이터의 공개를 통한 사회·경제적 가치를 증진시키기 위해 영국과 미국 등에서 실시하고 있는 링크드 데이터에 대비하는 연구가 될 수 있을 것으로 기대된다.

주

1) '공유'는 온톨로지가 '합의된 지식(Consensual Knowledge)'을 표현해야 한다는 것이다. '개념화'란 특정 영역 혹은 분야의 현실 세계와 관련된 개념을 나타내는 추상 모델을 일컫는다. '형식적'이란 온톨로지의 내용을 컴퓨터가 읽을 수 있고 처리가 가능한 형태로 표현해야 한다는 뜻이다. '명확한'은 특정 영역을 모델링 할 때 사용하는 개념들과 이러한 개념들을 사용할 때 적용되는 제약 조건들을 명시적으로 정의해야 한다는 것이다(Rho, S.K and Park, J. S., 2008).
2) 연안오염총량관리제는 육지에서 바다로 흘러 들어가는 오염물질을 바다의 자정능력 범위 내에서 관리하는 제도로 국내에는 2008년 마산만에 최초로 도입되어 시행되고 있다. 2011년 4월에는 부산시 연안해역에 시행하고 있으며 2013년에는 시화호에서도 시행할 예정이다.
3) 연안관리법은 연안의 효율적인 보전·이용 및 개발에 필요한 사항을 규정함으로써 연안환경을 보전하고 연안의 지속가능한 개발을 도모하여 연안을 쾌적하고 풍요로운 삶의 터전으로 조성하는 것을 목적으로 한다.

참고문헌

Borst, W. N., 1997, Construction of Engineering Ontologies for Knowledge Sharing and Reuse, *Ph.D. Dissertation, University of Twente.*

Budak Arpinar, I. *et al*., 2006, Geospatial Ontology Development and Semantic Analytics, Transactions in GIS, Vol.10, 551-575.

Calero, C. *et al*., 2006, *Ontologies for Software Engineering and Software Technology*. Springer-Verlag New York Inc.

Choi, J. H. and Park, J. T., 2006, Ontology Reasoning Technology Trends, *The Journal of Korean Institute of Information Technology*, Vol. 24, No. 12(in Korean)

Gruber, T. R., 1993, A Translation Approach to Portavel Ontology Specifications, *Knowledge Acquisition*, Vol.5, 199-220.

Henriksson, R. *et al*., 2008, Core Geographical Concepts: Case Finnish Geo-ontology, 57-60. ACM.

Hong, I, Y., 2006, Communal Ontology of Landmarks for Urban Regional Navigation, *Journal of Korean Geographical Society*, Vol.41, No. 5.

Jones, C. *et al*., 2001, Geographical information retrieval with ontologies of place, *Spatial Information Theory*, 322-335.

Ko, E. J. *et al*., 2004, An Intelligent Travel Agent System using Region Ontology, Proceedings of the KIISE Korea Computer Congress 2004, 31(1), 610-612 (in Korean).

Lee, Y. W., 2007, A Semantic Web Service for Tourism Information over the Mobile Web, *Journal of Korean Geographical Society*, Vol.42, 788-807.

Lee, T. B. *et al*., 2001, The semantic web, *Scientific American*, Vol.284, 34-43.

Lopez, M. F. *et al*., 1999, Building a chemical ontology using methontology and the ontology design environment, *Intelligent Systems and their Applications, IEEE*, Vol.14, pp.37-46.

National Fisheries Research and Development Institute, 2009, Annual Report of Korean Marine Environment Monitoring 2008, 13, Busan (국립수산과학원, 2009, 한국해양환경 조사연보 2008, 부산).

Noy, N. F. and D. L. McGuinness, 2001, Ontology development 101: A guide to creating your first ontology, Technical Report SMI-2001-0880, Stanford Medical Informatics.

Park, J. M., 2009, A Desingn and Implementation on Ontology for Public Participation GIS, *Journal of Korean Geographical Society*, Vol.44, No.3 (In Korean).

Pretorius, A. J., 2004, Ontologies-Introduction and Overview, *Semantic technology and applications research laboratory.*

Rho, S. K. and Park, J, S, 2008, *The Key for Internet Evolution, Web 2.0 to Web 3.0*, God's toy(노상규 · 박진

수, 2008, 인터넷 진화의 열쇠 웹2.0에서 웹3.0으로, 가즈토이).

Tsarkov, D. and I. Horrocks, 2006, FaCT++ description logic reasoner: System description, *Automated Reasoning*, 292-297.

Wang, Y. *et al.*, 2007, Geo-ontology design and its logic reasoning. In *Geoinformatics 2007: Geospatial Information Science*, 675309-675301. Nanjing, China.

Yoo, D. H. and Suh, Y. M., 2008, An Ontology-based Hotel Search System Using Semantic Web Technologies, *The Journal of Korean Institute of CALS/EC*, 13(4), 1-232 (in Korean).

교신: 황철수, 서울시 동대문구 회기동 1번지 경희대학교 지리학과(이메일: hcs@khu.ac.kr, 전화: 02-961-9313)

Correspondence: Chulsue Hwang, Department of Geography, Kyung Hee University, #1 Hoegi-dong, Dongdaemun-gu, Seoul, Republic of Korea (e-mail: hcs@khu.ac.kr, phone: +82-2-961-9313)

최초투고일 2012.12. 8
수정일 2013. 2. 8
최종접수일 2013. 2.14