웹 자원을 이용한 지리정보 온톨로지 확장
(Populating Geo-ontology with Web resources)

송 원 용 † 백 두 권 †† 정 동 원 †††
(Wonyong Song) (Doo-Kwon Baik) (Dongwon Jeong)

요 약 시맨틱 웹과 지리정보가 접목된 시맨틱 지리정보 시스템에 대한 연구가 활발히 진행되어 왔다. 그러나 현재의 지리정보 시스템은 사용자의 비공간 정보 입력 및 검색이 특정 시스템에 종속적이라는 문제점을 지닌다. 이는 또한 풍부한 서비스 제공을 어렵게 한다. 따라서 이 논문에서는 비 구조화된 웹 자원으로부터 지리정보 온톨로지(Geo-ontology)를 확장할 수 있는 구현 모델을 제안한다. 제안 모델은 웹 온톨로지 인스턴스를 시스템에 독립적으로 확장할 수 있으며, 따라서 보다 풍부한 웹 서비스 개발이 가능하다. 마지막으로 대학교를 도메인으로 선정하여, 지리정보 온톨로지 구축 사례와 함께 인스턴스를 확장하는 프로토타입을 보인다.

키워드 : 시맨틱 웹, 지리정보 시스템, 지리정보 온톨로지, 온톨로지 확장

Abstract Much research on semantic geographic information systems that incorporating the Semantic Web and geographic information has been actively studied. However, the existing geographic information systems have a system dependency problem that users can input and retrieve non-spatial information only in a specific system. It also causes difficulty in providing rich services. Therefore, this paper proposes an implementation model for population of Geo-ontology from non-structured Web resources. The proposed model can populate instances for Web ontology independently of systems, and thus it enables a richer Web service development. Finally, this paper shows the prototype that populates instances including a Geo-ontology building example for a University selected as an application domain.

Key words : Semantic Web, GIS, Geo-ontology, Population

1. 서 론

최근 웹의 확산에 의하여 일반인들의 인터넷 이용이 급증하고 있다. 따라서 사용자들이 일방적으로 정보를 제공받던 웹 1.0에서 사용자들이 생산자의 역할을 수행하여 스스로 정보 및 네트워크를 창조하고 공유하는 웹 2.0으로 변화하고 있다[1].

웹 2.0 환경에서는 사용자들의 상호작용이 중요한 키워드로 부각되면서, 다양한 서비스를 제공하기 위해 근간이 되는 데이터가 가장 중요시 되었다. 또한 사용자들의 적극적인 참여를 기반으로 하여 Wikipedia, 블로그 등과 같이 스스로 진화해 나가는 데이터를 제공할 수 있는 환경들이 웹 2.0 환경의 중요한 요소로 대두되었다. 이러한 환경 속에서 웹상의 수많은 정보 가운데 사용자에게 알맞은 정보만을 제공하기 위해 시맨틱 웹이 등장하게 되었으며, 시맨틱 웹에서의 핵심 기술인 온톨로지가 필요하게 되었다. 온톨로지란 의미 정보를 표현하기 위한 개념(concept)과 관계(relationship)를 정의한 정형적인 명세(formal specification)를 의미한다[2-4].

웹 환경의 진화로 웹에 존재하는 데이터 및 서비스가 방대해지고 참여와 공유 개방을 모토로 하는 웹 2.0이 지리정보 시스템(GIS, Geographic Information System)

· 2008년 정부(교육과학기술부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구(KRF-2008-314-D00485)

† 학생회원 : 군산대학교 정보통계학과
swy0111@kunsan.ac.kr
†† 종신회원 : 고려대학교 컴퓨터학과 교수
baikdk@korea.ac.kr
††† 종신회원 : 군산대학교 정보통계학과 교수
djeong@kunsan.ac.kr
(Corresponding author임)

논문접수 : 2009년 7월 6일
심사완료 : 2009년 8월 14일

Copyright©2009 한국정보과학회 : 개인 목적이나 교육 목적인 경우, 이 저작물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야 합니다.

기술과 접목되면서 지리정보 시스템 분야에 Where 2.0이 제시되었다[5,6].

기존의 지리정보 시스템에서는 전문가가 비공간 정보들을 지리정보 시스템에 저장하여 국한된 정보만을 사용자에게 제공해왔다. 그러나 최근 연구·개발되고 있는 지리정보 시스템은 일반 사용자 중심으로 시맨틱 웹과 융합되어 시맨틱 지리정보 시스템이라는 새로운 서비스가 개발되고 있다. 시맨틱 지리정보 시스템을 실현하기 위해 현재 시맨틱 웹 서비스의 선두 기업인 구글과 마이크로소프트의 경우, 웹 기반의 지리정보 시스템을 바탕으로 Wikipedia나 YouTube에 있는 콘텐츠를 제공함으로써 시맨틱 지리정보 시스템을 구현하고 있다. 또한 매쉬업(Mash-up)을 통하여 웹 콘텐츠의 활용을 강화할 수 있는 환경을 제공하고 있다[7].

국내외 지리정보 시스템의 대표 사이트인 구글의 Google Maps나 다음(Daum)의 지도 서비스 같은 경우, 사용자가 직접 서비스 되고 있는 시스템에 접속하여 정보를 입력해야 하는 특정 시스템에 종속적이라는 구조를 지닌다. 따라서 웹상에서 생성되는 방대한 정보, 즉 비공간 정보와의 연계를 통한 서비스 개발이 어렵다는 문제점을 지닌다.

이 논문에서는 비 구조화된 웹 자원을 정의된 지리정보 온톨로지의 인스턴스로 추가하여 지리정보 시스템의 비공간 정보를 확장함으로써 특정 시스템에 종속적이지 않고 사용자에게 보다 풍부한 서비스를 제공하기 위한 지리정보 온톨로지 확장 모델을 제안한다. 또한 제안 확장 모델을 위한 프로토타입 시스템을 개발하고 그 구현 결과를 보인다. 프로토타입 구현을 위한 대상 응용 도메인은 대학교를 대상으로 한다. 특히 대학교와 관련된 이미지, 평언과 주변의 맛집에 대한 정보를 확장하도록 구현한다.

이 논문의 구성은 다음과 같다. 제2장에서는 지리정보 온톨로지 시스템을 위해 필요한 기본 개념에 대해 소개하며, 제3장에서는 제안 모델의 전반적인 구조와 프로세스에 대하여 서술한다. 제4장에서는 프로토타입 구현 및 평가에 대하여 기술하고 제5장에서는 결론 및 향후 연구에 대하여 기술한다.

2. 관련 연구

이 장에서는 시맨틱 지리정보 시스템의 필요성과 현재 제공되고 있는 국내외 웹 지리정보 시스템의 문제점을 정의한다. 추가적으로, 프로토타입 구현을 위해 이용한 Open API에 대하여 간략하게 소개한다.

2.1 시맨틱 지리정보 시스템

시맨틱 지리정보 시스템은 시맨틱 웹과 지리정보 시스템이 접목된 지능적인 지리정보 시스템을 말한다. 웹 2.0 환경에서 제공하는 수많은 정보들을 이용하여 지도상에 정보를 표현한다[8-10]. 예를 들어, 특정 건물 상가나 음식점에 대한 정보를 건물의 위치와 이름을 통해 Wikipedia, 블로그, 카페에서 비공간 정보들을 검색해 사용자에게 보여줌으로써 사용자는 특정 건물에 대한 위치와 특성, 속성 등을 쉽게 이해할 수 있다.

2.2 국내외 지리정보 시스템

국내외에서 지도 서비스가 널리 사용되고 있으나, 이 논문에서는 다른 지리정보 시스템에 비하여 다양한 서비스를 제공하고 있는 다음(Daum)의 지리정보 시스템과 구글의 지리정보 시스템에 대하여 기술한다.

그림 1은 국내의 지리정보 시스템인 다음(Daum)에서

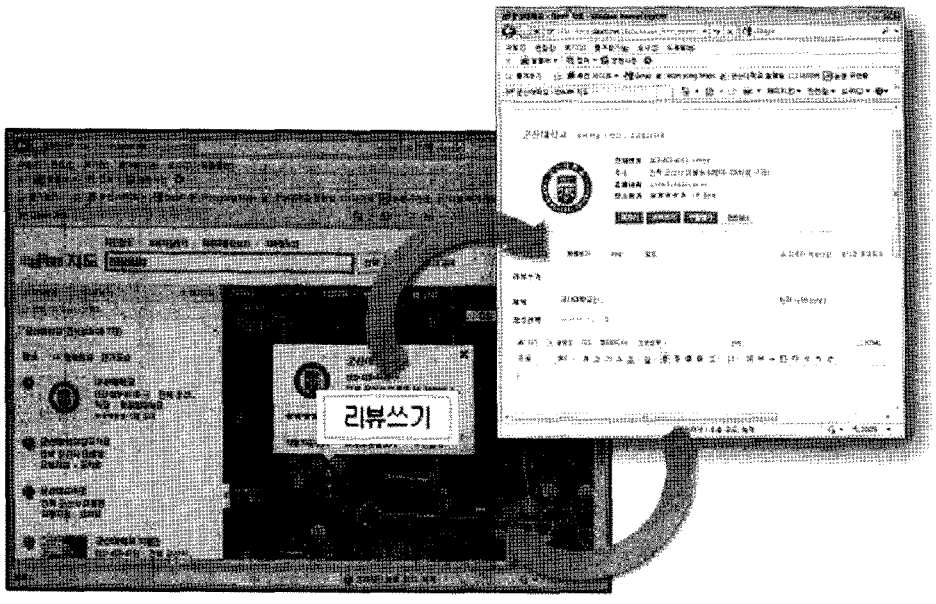


그림 1 다음(Daum) 지도 비공간 정보 등록과 가시화 화면

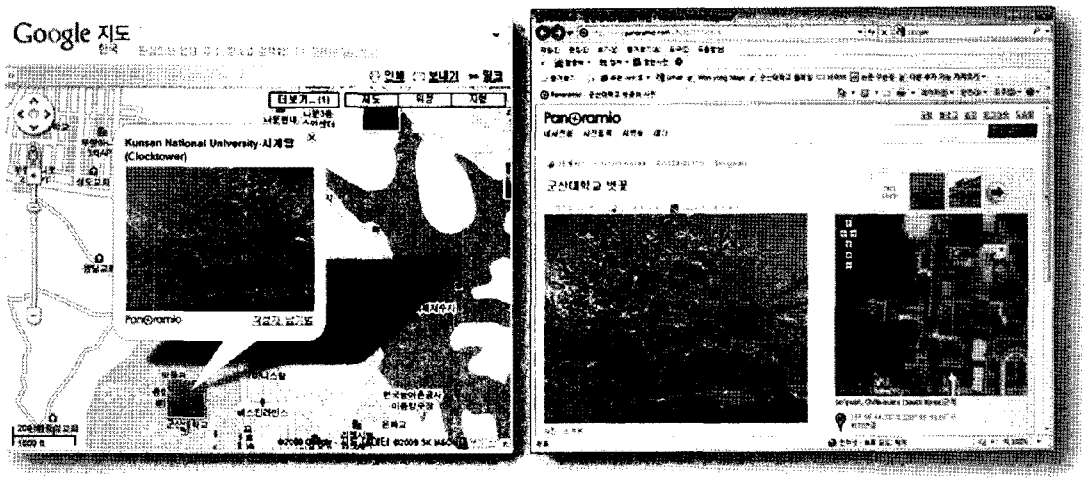


그림 2 Google Maps의 비공간 정보 등록과 가시화 화면

제공하는 지도 서비스로 지도에 비공간 정보를 입력·검색할 수 있는 서비스를 제공하고 있다.

그림 2는 국외의 지리정보 시스템인 구글에서 제공하는 지도 서비스로 다음(Daum)에서 제공하는 지도 서비스와 같이 지도에 비공간 정보를 입력·검색할 수 있는 서비스를 제공하고 있다. 두 지리정보 시스템 모두 비공간 정보의 확장을 위한 지리정보 시스템이라는 공통점을 지닌다. 그러나 그림 1과 그림 2의 예에서 알 수 있듯이, 두 지리정보 시스템은 사용자가 직접 웹 지리정보 시스템에 접속하여 자신이 알고 있는 정보를 등록함으로써 비공간 정보를 확장하는 시스템이다[11,12]. 그림 2에서 왼쪽에 있는 그림은 Google Maps에서 이미지가 표현된 화면이고, 오른쪽 그림은 왼쪽의 이미지를 등록하는 Panoramio의 화면으로 구글과 연계되어 이미지를 등록하는 서비스이다[13].

따라서 사용자가 직접 특정 시스템에 종속되어 있는 지리정보 시스템에 접속하여 비공간 정보를 입력·검색해야 하는 불편함이 있다. 또한, 현재 서비스되고 있는 지리정보 시스템의 비공간 정보들이 온톨로지화 되어있지 않기 때문에 정보 간의 관계를 알 수 없다는 단점을 지닌다. 특히 기존 웹 지리정보 서비스가 특정 시스템에 종속적인 이유는 지리정보의 관리 접근 방법에서 찾을 수 있다. 이 논문에서 제안하는 접근 방법의 경우, 지리정보에 대한 개념화를 통해 온톨로지로 구축한다. 즉, 지리정보를 위한 클래스를 정의하고 클래스 간 명시적인 관계성을 정의한다. 그러나 기존 접근 방법의 경우, 지리정보 온톨로지가 정의되지 않고 단순히 고전적인 접근 방법과 같이 지리정보를 데이터베이스에 저장한다. 따라서 데이터베이스 검색과 같은 기능 수행만으로 서비스를 사용자에게 제공한다. 이로 인하여 각각의 내부 저장 구조를 개념화하여 사용자에게 제공할 수 없기 때문에 자체 사이트에 한정하고 이미 정의해 놓은 규칙에 따라 사용자가 단순한 정보를 입력하는 구조를 지닌다. 이와 같이, 현재 웹에서 제공하는 지리정보 시스템은 여러 가지 문제점을 지니고 있으므로 이 논문에서는 이러한 문제점들을 보완할 수 있는 모델을 제안한다.

2.3 Open API

API(Application Program Interface)의 사전적 의미는 개발자들이 프로그램 개발을 위해 사용하는 함수나 메소드의 집합을 의미한다. 다양한 어플리케이션 개발에 이용할 수 있도록 개방하여 공유할 수 있도록 제공하는 API를 Open API라고 한다[14]. 현재 국내에도 Open API를 제공하는 웹 서비스가 증가하고 있을 뿐만 아니라, 점차 그 범위 또한 확대되고 있다. 이러한 Open API를 활용하여 개발자의 필요에 따라 자신만의 콘텐츠를 생성할 수 있고, 공유할 수 있다.

3. 제안 모델

이 장에서는 앞에서 언급했듯이 현재 웹에서 제공하는 지리정보 시스템의 문제점을 보완하기 위하여 제안한 접근 방법에 대하여 기술한다. 또한 지리정보 온톨로지 확장의 필요성에 대하여 서술한다.

3.1 개념 모델

이 논문에서 제안하는 개념 모델은 그림 3과 같이 웹 2.0의 철학인 사용자들의 참여, 공유, 협업, 개방을 통해 생성된 데이터를 웹에 등록하고, 지리정보 시스템과 결합시켜 기존의 지리정보 시스템의 단점인 특정 시스템에 종속성 문제를 보완하고 웹에 등록된 비공간 정보를 추출하여 온톨로지화 함으로써 사용자에게 기존의 지리정보 시스템보다 정확하고 풍부한 정보를 제공할 수 있다.

3.2 접근 방법

이 논문에서 제안하는 모델은 앞서 언급했던 현재 웹

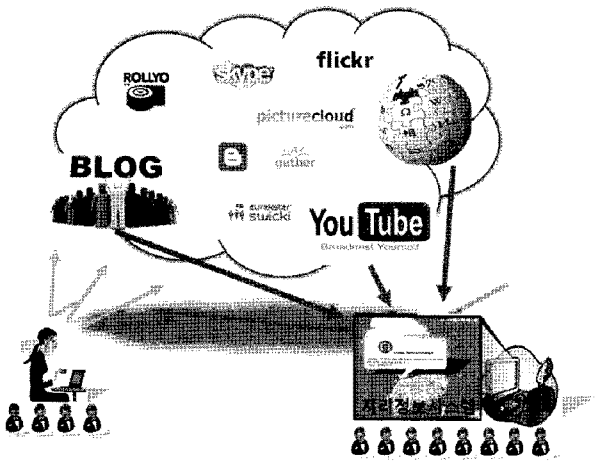


그림 3 제안 모델 개념도

에서 제공하는 지리정보 시스템의 문제점을 보완하기 위하여 다음과 같은 두 가지 작업이 요구된다.

3.2.1 지리정보 온톨로지 정의

변화하는 웹 환경하에서, 현재 제공되고 있는 국내외 지리정보 시스템의 문제점은 비공간 정보들 간의 관계가 정의되어 있지 않아 컴퓨터가 이해할 수 있는 환경이 제공되지 않는다는 점이다. 이는 보다 다양하고 정확한 서비스 제공을 어렵게 한다. 이 논문에서는 이러한 문제를 해결하기 위하여 지리정보 온톨로지를 정의함으로써 정확한 정보를 제공할 수 있다. 또한, 프로토타입 개발을 위해 대학을 응용 도메인으로 선정하고 이를 위한 지리정보 온톨로지를 정의한다.

3.2.2 웹 자원을 이용한 비공간 정보 추출

웹 자원을 이용하여 지리정보 온톨로지의 인스턴스를 확장함으로써 웹 2.0 환경에서 다양하고 점점 더 방대해지는 비공간 정보를 사용자에게 제공할 수 있다. 지금까지 웹 자원을 통해 온톨로지를 확장하고자 하는 다양한 연구들이 진행되어 왔으며, 기본적으로 온톨로지 공학자의 경험 지식에 의존하여 클래스를 정의하고 인스턴스를 추가하는 방식이 전통적으로 널리 이용되어 왔다. 그러나 이러한 접근 방법은 온톨로지 클래스에 대한 정의가 정확하다는 장점을 지니는 반면, 다양한 웹 자원을 사용자에게 서비스로 제공될 인스턴스로 확장하여 활용하지 못한다는 문제점을 지닌다. 이는 또한 웹 2.0을 넘어 웹 3.0을 향해 발전하고 있는 현재의 패러다임에 적합하지 않다. 따라서 이러한 접근 방법은 이 논문에서 제안하는 접근 방법에 적용될 수 없다. 경험 지식에 의한 온톨로지 확장의 문제점을 해결하기 위한 대안으로서, 자연어 처리 기법을 이용한 접근 방법이 대두되었으며, TextOntoEx, ONTOSTRUCT 등과 같은 다양한 방법이 제안되었다[15,16]. 이러한 접근 방법은 자연어로 서술된 텍스트로부터 클래스 혹은 인스턴스를 생성하는 방법이다. 이 논문에서 제안한 접근 방법을 위해서는 웹 상에 존재하는 자연어 형태로 표현된 웹 자원으로부터 인스턴스 추출하는 기법이 요구된다. 따라서 이 논문에서는 웹 문서에 대한 자연어처리를 통해 인스턴스 후보를 추출하는 기법을 이용한다.

그림 4는 제안하는 모델의 접근 방법을 개념적으로 보여준다. 그림 4에서, GOS(Geo-ontology Storage)는 지리정보 온톨로지 저장소로서, 온톨로지 스키마와 인스턴스를 저장한다. RE(Resource Extractor)는 웹 자원으로부터 공간 정보를 검색하여 후보 인스턴스를 제공하는 역할을 수행한다. RE의 결과는 IS(Instance Selector)를 통해 확인 연산을 수행한 후, 중복된 정보 제거 및 URI 검증을 위해 IV(Information Validator)에 전달된다. IV에 의한 처리가 완료되면 추출된 인스턴스는 지리정보 온톨로지 클래스의 인스턴스로 추가, 저장된다. 이와 같은 연산을 반복적으로 수행함으로써 지리정보 온톨로지

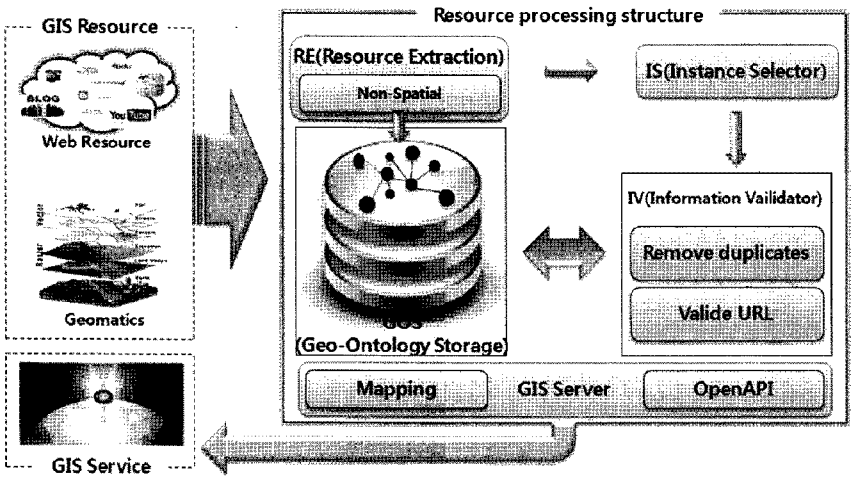


그림 4 제안 모델의 접근 방법

의 확장이 이루어지며, 사용자에게 보다 풍부한 지리정보 서비스의 제공이 가능하다.

3.3 제안 모델의 구조 및 프로세스

그림 5는 제안 모델의 구조를 보여준다. 제안 모델에서는 지리정보 온톨로지의 확장을 위하여 정의된 지리정보 온톨로지에 웹으로부터 추출한 정보를 인스턴스로 추가하는 방법을 사용한다.

그림 5의 지도 정보 계층(Map data Layer)에서는 구글의 Open API를 이용하여 지도를 표현하고, 공간 정보 계층(Spatial Layer)에서는 지도로부터 좌표를 추출하여 지리정보 온톨로지에 저장한다. 또한, 비공간 정보 계층(Non-Spatial data Layer)에서는 웹 2.0 환경으로부터 비공간 정보를 추출하기 위하여 키워드 입력 모듈, Open API를 통한 키워드 검색과 URL 추출 모듈, 추출된 URL의 유효성 검증 모듈, 그리고 URL 저장 모듈로 구성된다. 이렇게 추출된 비공간 정보와 공간 정보를 서버 계층(Server Layer)에서 사상시켜 지리정보 서버를 통해 사용자에게 제공한다.

그림 6은 제안 모델의 프로세스를 도식화한 것으로서 크게 세 개의 주요 프로세스로 분류할 수 있다. 첫 번째 단계에서는 최초 지리정보 온톨로지를 생성하고, 지도로부터 공간정보를 추출하여 지리정보 온톨로지에 저장한다. 두 번째 단계에서는 비공간 정보를 추출하기 위하여 지리정보 온톨로지의 클래스와 프로퍼티를 선택하고 키워드 검색을 통하여 웹으로부터 비공간 정보를 검색하게 된다. 두 번째 단계를 반복하여 계속적으로 비공간 정보를 확장할 수 있다. 마지막으로 세 번째 단계에서는 첫 번째 단계와 두 번째 단계에서 추출된 공간 정보와 비공간 정보를 맵핑시켜 지리정보 온톨로지에 저장하게 된다.

3.4 지리정보 온톨로지 정의

프로토타입 구현을 위한 응용 도메인을 "대학교"로 선정하고 그림 7과 같은 지리정보 온톨로지를 설계한다. 또한 설계된 지리정보 온톨로지를 그림 8과 같이 Jena를 활용하여 관계형 데이터베이스에 저장한다.

지리정보 시스템을 온톨로지화 함으로써 클래스와 클래스, 클래스와 인스턴스, 프로퍼티와 프로퍼티 등 자원과 자원 사이의 관계를 정의할 수 있다. 따라서 데이터를 검색할 경우 보다 빠르고 정확한 정보를 사용자에게 제공할 수 있다.

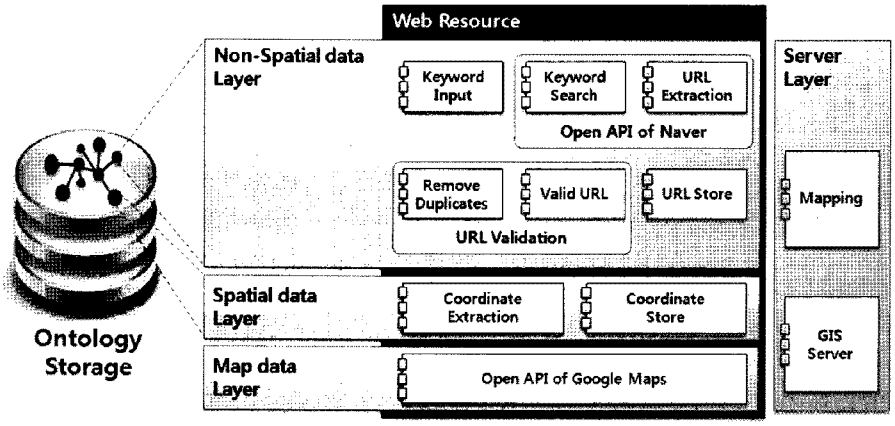


그림 5 제안 모델의 구조

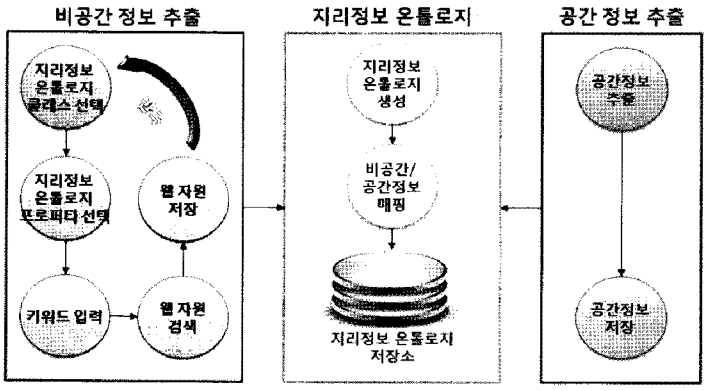


그림 6 제안 모델의 주요 프로세스

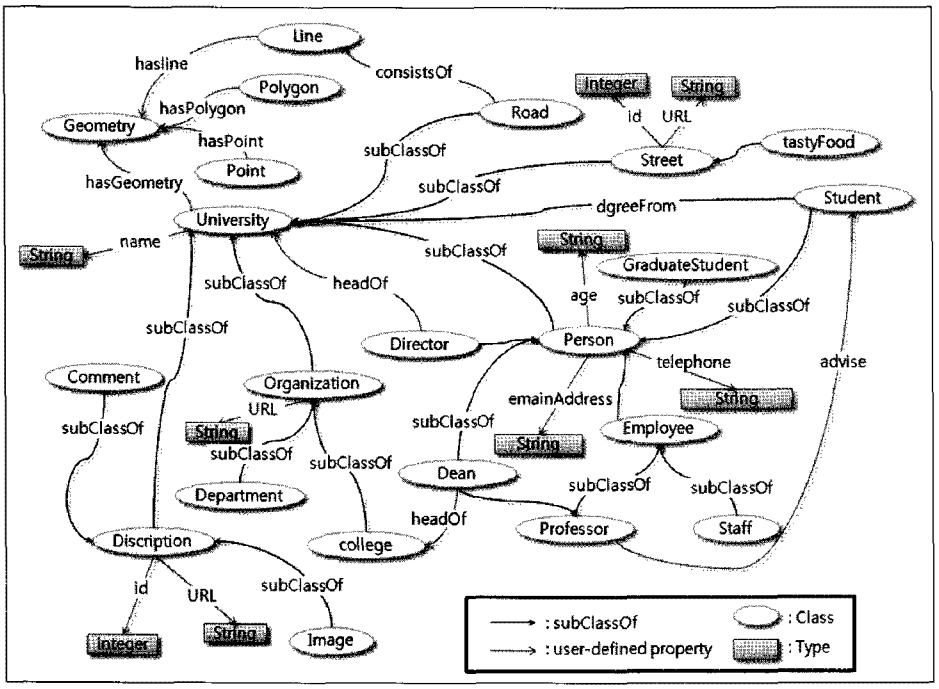


그림 7 대학교 도메인을 위한 지리정보 온톨로지 예제

SUBJ	PROP	OBJ
Uv::http://www.owl-ontologies.com/ontology.owl#name	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#type	Uv::http://www.w3.org/2002/07/owl#DatatypeProperty
Uv::http://www.owl-ontologies.com/ontology.owl#name	Uv::http://www.w3.org/2000/01/rdf-schema#range	Uv::http://www.w3.org/2001/XMLSchema#string
Uv::http://www.owl-ontologies.com/ontology.owl#telephone	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#type	Uv::http://www.w3.org/2002/07/owl#DatatypeProperty
Bv::12afdc33:120af111943:-7db	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#type	Uv::http://www.w3.org/2002/07/owl#Class
Uv::http://www.owl-ontologies.com/ontology.owl#telephone	Uv::http://www.w3.org/2000/01/rdf-schema#domain	Bv::12afdc33:120af111943:-7db
Bv::12afdc33:120af111943:-7db	Uv::http://www.w3.org/2002/07/owl#unionOf	Bv::12afdc33:120af111943:-7da
Bv::12afdc33:120af111943:-7da	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#first	Uv::http://www.owl-ontologies.com/ontology.owl#Organization
Bv::12afdc33:120af111943:-7da	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#rest	Bv::12afdc33:120af111943:-7d9
Bv::12afdc33:120af111943:-7d9	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#first	Uv::http://www.owl-ontologies.com/ontology.owl#Person
Bv::12afdc33:120af111943:-7d9	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#rest	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#nil
Uv::http://www.owl-ontologies.com/ontology.owl#telephone	Uv::http://www.w3.org/2000/01/rdf-schema#range	Uv::http://www.w3.org/2001/XMLSchema#string
Uv::http://www.owl-ontologies.com/ontology.owl#emailAddress	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#type	Uv::http://www.w3.org/2002/07/owl#DatatypeProperty
Uv::http://www.owl-ontologies.com/ontology.owl#emailAddress	Uv::http://www.w3.org/2000/01/rdf-schema#range	Uv::http://www.w3.org/2001/XMLSchema#string
Uv::http://www.owl-ontologies.com/ontology.owl#emailAddress	Uv::http://www.w3.org/2000/01/rdf-schema#domain	Uv::http://www.owl-ontologies.com/ontology.owl#Person
Uv::http://www.owl-ontologies.com/ontology.owl#id	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#type	Uv::http://www.w3.org/2002/07/owl#DatatypeProperty
Uv::http://www.owl-ontologies.com/ontology.owl#id	Uv::http://www.w3.org/2000/01/rdf-schema#range	Uv::http://www.w3.org/2001/XMLSchema#int
Bv::12afdc33:120af111943:-7d8	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#type	Uv::http://www.w3.org/2002/07/owl#Class
Uv::http://www.owl-ontologies.com/ontology.owl#id	Uv::http://www.w3.org/2000/01/rdf-schema#domain	Bv::12afdc33:120af111943:-7d8
Bv::12afdc33:120af111943:-7d8	Uv::http://www.w3.org/2002/07/owl#unionOf	Bv::12afdc33:120af111943:-7d7
Bv::12afdc33:120af111943:-7d7	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#first	Uv::http://www.owl-ontologies.com/ontology.owl#Discription
Bv::12afdc33:120af111943:-7d7	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#rest	Bv::12afdc33:120af111943:-7d6
Bv::12afdc33:120af111943:-7d6	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#first	Uv::http://www.owl-ontologies.com/ontology.owl#Street
Bv::12afdc33:120af111943:-7d6	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#rest	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#nil
Uv::http://www.owl-ontologies.com/ontology.owl#age	Uv::http://www.w3.org/1999/02/22-rdf-syntax-ns#type	Uv::http://www.w3.org/2002/07/owl#DatatypeProperty
Uv::http://www.owl-ontologies.com/ontology.owl#age	Uv::http://www.w3.org/2000/01/rdf-schema#domain	Uv::http://www.owl-ontologies.com/ontology.owl#Person

그림 8 관계형 데이터베이스에 저장된 지리정보 온톨로지

그림 7은 대학교에 관련된 비공간 정보와 공간 정보를 포함하는 지리정보 온톨로지의 예로서, 이 논문에서 제안한 방법을 이용하여 프로토타입을 개발하고 그 결과를 보이기 위해 이용된다. 특히 프로토타입 개발을 통해 사용자에게 제공하는 서비스를 대학교에 대한 서평 및 주변의 맛집 등을 대상으로 한정하였기 때문에 이에 필요한 일부 비공간 정보를 포함할 수 있도록 정의되었다. 따라서 실제 다양한 서비스를 제공하기 위해서는 보다 많은 비공간 정보를 포함할 수 있도록 지리정보 온톨로지를 정의해야 한다. 그림 7에 정의된 지리정보 온톨로지에서 볼 수 있듯이, 지리정보 온톨로지를 정의함으로써 클래스와 클래스, 클래스와 프로퍼티간의 관계를 표현한다. 그림 7의 정의된 지리정보 온톨로지는 웹에서 추출한 비공간 정보를 인스턴스로 추가하여 확장된다.

그림 8은 그림 7에서 표현된 지리정보 온톨로지를 Jena를 통해 트리플(Triple) 구조 즉, subject, predicate (property), object 형태로 관계형 데이터베이스에 저장한 결과를 보여준다.

3.5 지리정보 온톨로지 확장

웹 자원 즉, 블로그나 게시판, Wikipedia 등과 같은

시맨틱 웹 서비스에서 비공간 정보를 추출하여 온톨로지의 인스턴스를 확장함으로써 풍부한 비공간 정보를 제공할 수 있다. 표 1은 웹 자원을 이용한 지리정보 온톨로지 확장 알고리즘에 대한 의사코드(pseudocode)이다. 먼저 지리정보 온톨로지를 생성하고 키워드를 입력하여 웹으로부터 자원을 추출한다. 추출된 자원은 개발자가 확인하여 최종적으로 지리정보 온톨로지의 인스턴스로 추가하게 된다. 이러한 방법을 반복하여 지리정보 온톨로지를 확장할 수 있다.

표 1 웹 자원으로부터 비공간 정보를 추출하기 위한 의사코드

```
1. Create Geo-ontology;
2. List K; //입력된 키워드 리스트
3. while(K.next() is NULL)
4.        Get a keyword;
5.        Search a keyword from Web;
6.        Extract URL from searched web resources;
7.        Check the URL by developer;
8.        Store the URL to Geo-ontology instances;
9. end while
```

표 2 추출된 URL 유효성 검증을 위한 의사코드

```
1.  Connect Oracle DBMS;
2.  Select field;
3.  if(duplication instances)
4.     create temp table from selected distinct instances ;
5.  end if;
6.  URL = instance;
7.  Connect URL.getRespondCode();
8.  if(URL == invalid)
9.     Remove a invalid URL from temp table;
10. end if;
11. Disconnect Oracle DBMS;
```

위의 표 1을 통해 확장된 지리정보 온톨로지 인스턴스의 유효성을 검증하기 위하여 표 2에서처럼 지리정보 온톨로지의 인스턴스인 URL에 접속하여 중복된 URL과 시간이 흐르면서 변경되거나 삭제된 URL을 추출하여 유효한 URL을 사용자에게 제공할 수 있다.

4. 구현 및 평가

이 장에서는 대학교를 대상으로 한 지리정보 온톨로지의 확장을 위한 프로토타입을 구현하고 제안 모델로 인한 장점과 프로토토입의 기능적 제약성에 대하여 논의한다.

4.1 구현

이 절에서는 지리정보 온톨로지 확장에 대한 프로토타입을 기술한다. 구현에 앞서, 지도는 구글의 Google Maps를 사용하고, 비공간 정보 검색 API는 네이버의 Open API를 사용한다.

그림 9는 구현된 시스템의 메인 화면으로 어떤 대학교의 비공간 정보를 확장할 것인지 선택하는 화면이다. 현재 9개 대학교에 기본 인스턴스를 저장해 놓고 대학교를 선택하면 왼쪽의 지도에 위치와 기본 비공간 정보가 디스플레이 된다.

그림 10은 그림 9에서 선택한 대학교의 프로퍼티를 선택하는 화면이다. 이 프로토타입에서는 대학교의 평언과, 이미지, 맛집에 대한 비공간 정보를 확장할 수 있도록 프로퍼티가 정의되어 있다.

그림 11은 그림 9와 그림 10에서 선택된 대학 및 프로퍼티에 추가할 인스턴스를 선택하는 화면이다. 키워드를 입력하여 대학교의 이미지나 평언, 맛집에 관한 비공간 정보를 Open API를 통해 블로그나 카페, 뉴스 등에서 검색하고, 이렇게 검색된 비공간 정보를 개발자의 검

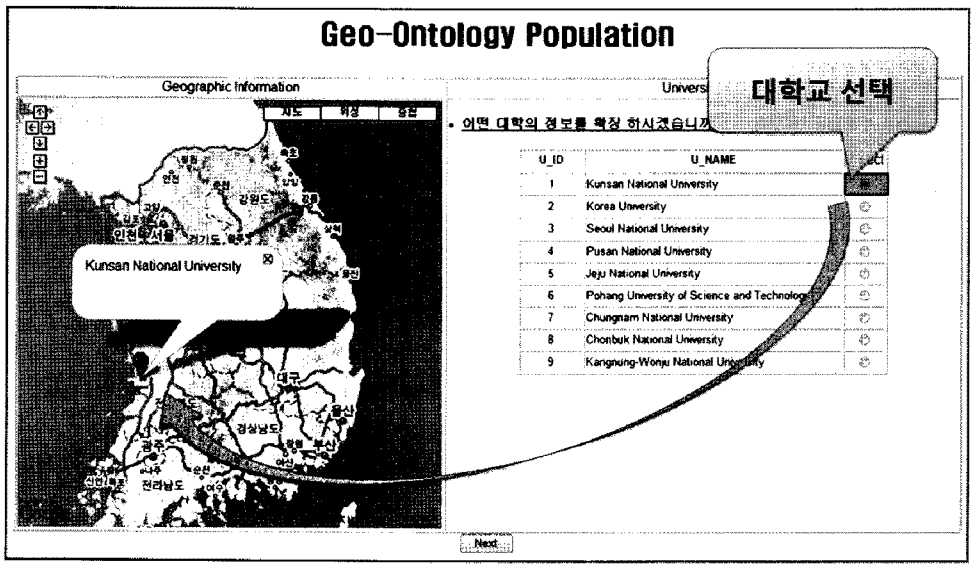


그림 9 대학교 선택 화면

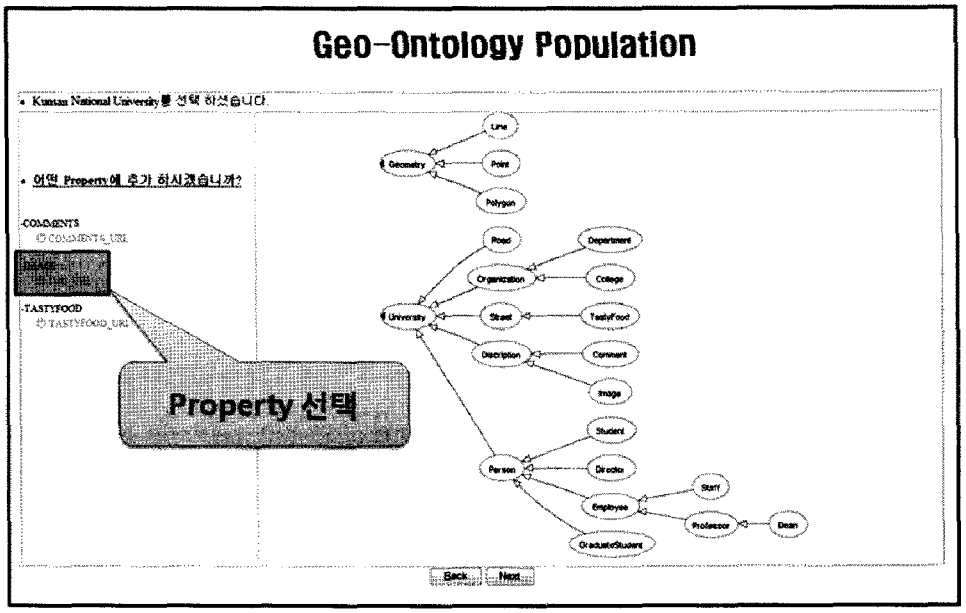


그림 10 대학교 프로퍼티 선택 화면

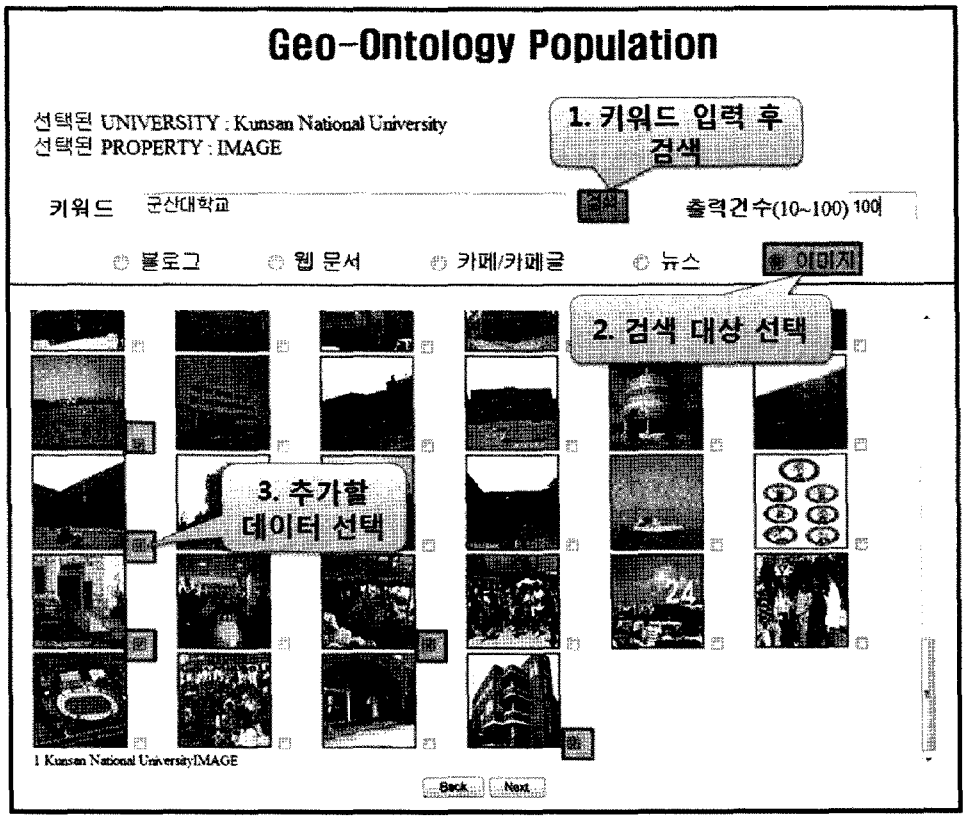


그림 11 대학교 인스턴스 검색, 선택 후 저장 화면

증 작업을 통하여 선택하고 지리정보 온톨로지에 저장한다.

그림 12는 그림 11에서 저장된 대학교 지리정보 온톨로지의 확장된 인스턴스를 보여주는 화면이다. 그림 12와 같이 대학교를 선택하고 프로퍼티(이미지, 평언, 맛집)를 선택한 후 로드를 클릭하면 아래에 이미지나, URL들이 디스플레이 된다. 예를 들어, 대학교에 군산대학교를 클릭하고 프로퍼티에 이미지 선택 후 로드를 하게 되면 그림 12와 같은 화면이 디스플레이 된다. 이렇게 확장된 URL의 유효성을 검증하기 위하여 그림 13과 같은 도구를 사용한다.

그림 13은 URL의 유효성 여부를 판단하기 위한 도구

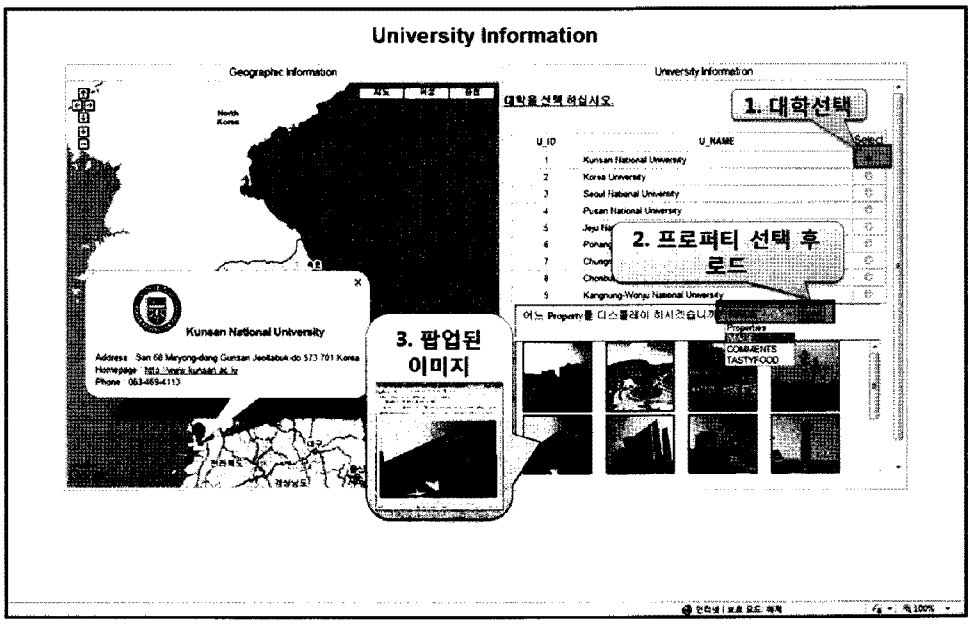


그림 12 지리정보 온톨로지의 확장된 인스턴스 화면

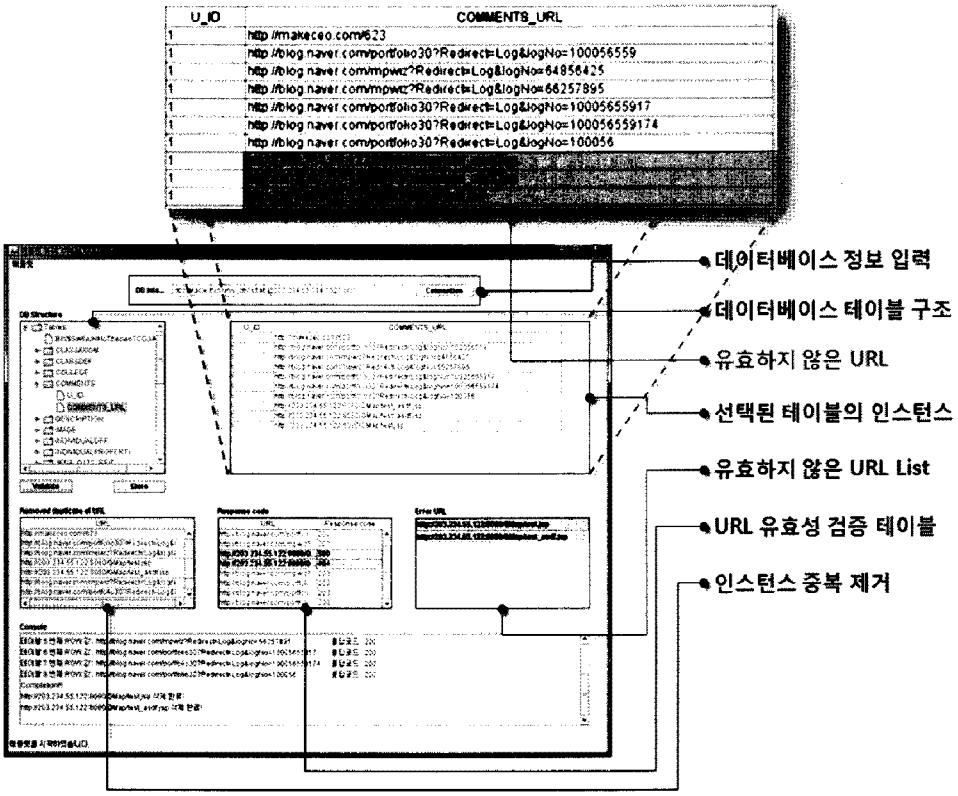


그림 13 URL 유효성 검증을 위한 도구 화면

를 자바를 이용하여 개발한 것으로서 최상단의 "데이터베이스 정보 입력" 부분에 데이터베이스 정보를 입력하고 "Connection"을 클릭한다. 데이터베이스에 연결이 되면 왼쪽 "데이터베이스 테이블 구조" 부분에 트리형태로 테이블명과 컬럼명이 출력된다. 출력된 트리 구조에서 테이블명을 클릭하면 트리구조 오른쪽의 "선택된 테이블의 인스턴스" 부분에 선택된 테이블의 컬럼명과 인스턴스가 출력된다. 이 출력된 인스턴스의 유효성을 검증하기 위하여 "Validate" 버튼을 클릭한다. "Validate" 버튼을 클릭하게 되면 인스턴스들을 검색하여 중복된 URL을 하나로 통합하고, 각각의 URL에 접속하여 리턴되는 응답코드로 연결 가능 여부를 확인한다. "Respond

duplicate of URL" 부분에는 중복이 제거된 URL 리스트가 출력되고 "Respond code" 부분에는 URL에 접속하여 리턴된 응답코드들이 출력된다. 그리고 응답코드를 확인하여 200코드(접속 가능코드)를 제외하고 404코드(URL 접속 불가) 또는 500코드(페이지 오류)와 같은 오류가 발생한 URL 리스트가 "Error URL" 부분에 출력된다. 예를 들어, 그림 13에서 "유효하지 않은 URL" 부분에 선택된 3개의 URL이 "Error URL"에 출력되었다. 그리고 그림 14와 같이 데이터베이스에서도 삭제된 것을 확인할 수 있다.

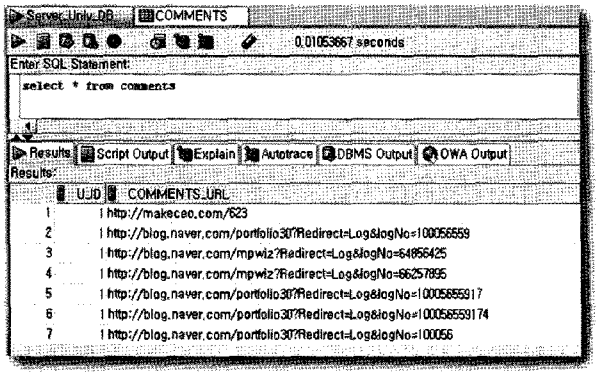


그림 14 유효성이 검증된 URL

표 3과 표 4는 그림 13의 유효성 검증에 필요한 모듈을 의사결정 코드로 나타낸 표이다. 표 3의 중복된 URL 제거 모듈은 먼저 데이터베이스에 접속하여 중복된 URL을 하나로 출력해주는 "DISTINCT" 명령어를 사용하여 그 결과 값을 임시 테이블을 생성하여 저장한다. 그리고 중복된 URL을 가지는 원래 테이블을 삭제하고 임시테이블의 이름을 원래 테이블의 이름으로 바꿔주면 중복된 URL을 하나로 통합할 수 있다.

표 4 URL 유효성 검증 모듈

```
1. public int chkUrl(chk_url){
2.     response_code = 404;
3.     try {
4.         new URL(chk_url);
5.         URL.setConnectTimeout(time);
6.         URL.openConnection();
7.         response_code = URL.getResponseCode();
8.         if(response_code of URL != 200)
9.             Remove the URL;
10.        end if;
11.    end try;
12.    catch (Exception e) {
13.        return response_code; // URL 접속 불가, 404 코드 리턴
14.    end catch;
15.    finally{
16.      return response_code;
17.    end finally;
18. }
```

표 4는 중복이 제거된 URL에 접속하여 응답코드 분석 후 접속 가능 코드를 제외한 나머지를 삭제하는 부분이다. 유효성 검증 모듈에 URL이 인자 값으로 들어오면 URL에 일정 시간동안 접속을 시도하여 getRespondCode() 메소드를 통해 응답코드를 리턴 받는다. 접속되지 않는다면 URL이 존재하지 않는 것이므로 404코드로 미리 설정한다. 리턴된 응답 코드가 404코드나 500코드이면 데이터베이스에서 삭제한다.

이러한 과정을 정기적으로 실행하여 지리정보 온톨로지의 비공간 정보인 URL 자원들을 효율적으로 관리할 수 있을 뿐만 아니라 유효한 자원을 관리할 수 있다.

4.2 평가

이 절에서는 웹 자원을 이용하여 지리정보 온톨로지를 확장할 때 얻을 수 있는 장점과 프로토타입의 기능적 제약성에 대해 논의한다.

표 3 중복된 URL 제거 모듈

```
1. Oracle DB 연동;
2. public DefaultTableModel getTableModel(테이블명, 필드명) {
3.     SQL1 = "CREATE TABLE temporary_table AS SELECT DISTINCT * FROM " +
   table_name;
4.     SQL2 = "DROP TABLE " + table_name;
5.     SQL3 = "RENAME temporary_table TO " + table_name;
6.     SQL4 = "SELECT "+ 필드명 + " FROM" + table_name;
7.     Execute query(SQL1 to SQL3);
8.     resultSet = Execute query(SQL4);
9.
10.    while(resultSet.next())
11.        TableModel.addRow(resultSet.getString(1)); // 테이블 모델에 추가
12.    end while;
13.    return TableModel;
14.}
```

표 5 제안 모델의 정성적 평가표

평가 항목	다음(Daum) 지도	Google Maps	제안 모델
시스템 종속성	종속적	부분 종속	독립적
정보의 획득 범위	제한적	부분 제한	제한 없음
정보 간 의미 관계성	관계성 없음	관계성 없음	관계성 유지
질의 결과의 정확성	낮음	낮음	높음

먼저, 온톨로지를 확장할 때 얻을 수 있는 장점은 표 5의 제안 모델의 3가지 요인에 대한 정성적 평가표에서 나타나듯이 첫째는 시스템에 독립적이다. 현재 국내에서 지리정보 시스템을 제공하는 다음(Daum) 지도는 다양한 비공간 정보를 제공하지만 다음(Daum)의 시스템에 사용자가 직접 등록하고 접속하여 정보를 제공받는 것이므로 시스템에 종속적이다. 또한 국외의 Google Maps에서는 좀 더 다양한 비공간 정보를 제공하기 위하여 YouTube, Panoramio, Wikipedia와 연계하여 동영상, 사진, 지식 등을 제공하고 있다. 이는 일부 웹 2.0 환경의 서비스로부터 정보를 검색하여 제공하고 있기 때문에 특정 시스템에 부분 종속적이라고 할 수 있다. 그러나 제안 모델은 웹 전체로부터 비공간 정보를 검색하기 때문에 특정 시스템으로부터 독립적이다.

둘째는 정보의 획득 범위이다. 정보의 획득 범위는 첫 번째 장점에서 언급했듯이 다음(Daum) 지도 서비스와 Google Maps는 각각 시스템에 종속적, 부분 종속이라는 특성을 지닌다. 다음(Daum) 서비스의 경우, 특정 시스템에 종속적이므로 정보를 획득하는 범위는 제한적이라고 할 수 있다. 또한 Google Maps의 경우는 일부 웹 2.0 환경의 서비스와 연계되어 있기 때문에 정보의 획득 범위가 부분 제한적이라고 할 수 있다. 그러나 제안 모델의 경우, 다양한 웹 서비스, 혹은 시스템을 통해 사용자가 자유롭게 정보 생성에 참여할 수 있으며, 이러한 웹 자원을 통해 정보를 획득하여 확장하기 때문에 정보의 획득 범위는 제한이 없다.

마지막으로, 제안 모델의 접근 방법인 지리정보를 온톨로지화 함으로써 정보 간의 의미 관계성을 정의할 수 있다. 또한 지리정보 온톨로지는 이러한 의미 관계를 이용한 질의 처리가 가능하며 따라서 질의 결과에 대한 정확도를 향상시킬 수 있다. 즉 기존의 정보 검색 방법의 하나인 키워드 맵핑을 통한 검색은 많은 불필요한 정보를 포함하여 사용자에게 제공하는데 반해, 제안 모델의 경우, 온톨로지를 이용한 보다 정확한 정보 제공이 가능하므로 질의 결과의 정확성이 높다.

한편, 제안 모델을 적용하여 실제 응용 서비스를 제공하기 위해서는 보다 정확하게 후보군을 추출할 수 있는 알고리즘 개발이 요구되며 무엇보다 논문에서 보인 프로토타입의 기능적 제약성을 해결하여야 한다. 현재 개발한 프로토타입의 문제점 보완을 위해서는 우선적으로 다양한 도구와의 연계가 요구된다. 특히 다양한 지리정보 온톨로지를 이용하는 서비스에 적용할 수 있도록 하기 위해 편집 도구나 보다 다양한 형태로 정의된 온톨로지를 확인할 수 있는 가시화 도구와의 연계 및 통합이 요구된다. 또한 그림 13과 같은 도구에 의한 작업이 자동적으로 처리될 수 있는 최적화 노력이 요구된다. 이는 높은 성능의 후보 인스턴스 추출 알고리즘과 함께 전체적인 시스템의 성능을 향상시킬 수 있다. 마지막으로, 현재 개발되어 서비스되고 있는 Google Maps 등과 같은 시스템과의 연동 및 이를 통한 활용성을 향상시키기 위해 플러그인 형태로의 개발이 요구된다.

5. 결론 및 향후 연구

이 논문에서는 현재 웹에서 제공하는 지리정보 시스템의 문제점을 정의하고, 비 구조화된 웹 자원을 통한 지리정보 온톨로지의 확장모델을 제안하였다. 또한 제안한 방법을 통해 지리정보 온톨로지를 확장한 예를 보였으며, 제안 모델의 장점을 중심으로 평가에 대하여 기술하였다. 제안 모델의 평가 결과를 통하여 알 수 있듯이 제안 모델은 특정 시스템에 종속적이지 않고 사용자에게 보다 풍부한 서비스를 제공할 수 있으며, 온톨로지를 이용한 보다 정확한 정보 제공을 가능하게 한다.

향후 연구로는 프로토타입을 웹 브라우저의 플러그인 제공을 위한 연구와, 제안 모델의 활용성 향상을 위한 보다 자동화된 기능 개발이 요구된다. 또한 추론 엔진과의 결합을 통하여 보다 정확하고 다양한 정보 제공을 위한 연구가 요구된다.

참 고 문 헌

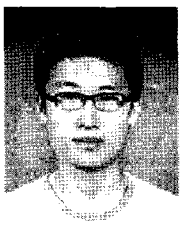
[1] T. O'Reilly, "What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software," *journal of Communications & Strategies*, no.65, 1st quarter, pp.17-37, Jan. 2007.

[2] T. Berners-Lee, J. Hendler, O. Lassila, "The Semantic Web," *journal of Scientific American*, vol.284, no.5, pp.33-43, May 2001.

[3] T. R. Gruber, "A Translation Approach to Portable Ontology Specifications," *journal of Knowledge Acquisition*, vol.5, no.2, pp.199-220, 1993.

[4] A. Maedche, *Ontology Learning for the Semantic Web*, pp.11-27, Kluwer Academic Publisher, 2002.

[5] Where 2.0, http://radar.oreilly.com/2005/05/what-is-where-20.html.
[6] Y.-W. Lee, "A Semantic Web Service for Tourism Information over the Mobile Web," *Journal of the Korean Geographical Society*, vol.42, no.5, 2007.
[7] E. Kim, "A Study on the Service Model for Custom-tailored Geospatial Services," *journal of KSII*, vol.9, no.1, pp.51-58, Mar. 2008. (in Korean)
[8] A. Mikroyannidis, "Toward a Social Semantic Web," *journal of IEEE Computer*, vol.40, no.11, pp.113-115, Nov. 2007.
[9] L. Stoimenov and S. Djordjevic-Kajan "Framework for Semantic GIS Interoperability," Series Mathematic and Informatics, 2002.
[10] V. Tanasescu and J. Domingue, "Toward User Oriented Semantic Geographical Information Systems," 2nd AKT Doctoral Symposium, 2006.
[11] Daum Map, http://local.daum.net/map/index.jsp.
[12] Google Maps, http://maps.google.com/maps.
[13] Panoramio, http://www.panoramio.com.
[14] Open API, http://en.wikipedia.org/wiki/Open_API.
[15] M. Y. Dahab, H. A. Hassan, A. Rafea, "TextOntoEx: Automatic Ontology Construction from Natural English Text," *Conference of AIML 06*, Jun. 2006.
[16] M. Degeratu, V. Hatzivassiloglou, "Building Automatically a Business Registration Ontology," *Proc. of the ACM International Conference*, vol.129, pp.1-7, 2002.

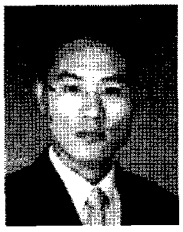


송 원 용

2009년 군산대학교 수학정보통계학부 정보통계학전공(이학사). 2009년~현재 군산대학교 정보통계학과 석사과정. 관심분야는 시맨틱 웹, 시맨틱 GIS, 데이터베이스

백 두 권

1974년 고려대학교 수학과 이학사. 1977년 고려대학교 대학원 산업공학과 석사 1983년 Wayne State Univ. 전산학과 석사. 1985년 Wayne State Univ. 전산학과 박사. 1989년~2007년 (사)한국정보과학회 이사/평의원/부회장. 1986년~현재 고려대학교 컴퓨터전파통신공학과 교수. 1991년~현재 (사)한국시뮬레이션학회 이사/부회장/감사/회장/고문. 1991년~현재 ISO/IEC JTC 1/SC 32 전문위원회 위원장. 2001년~현재 (사)도산아카데미 원장. 2002년~2004년 고려대학교 정보통신대학 초대학장. 2004년~2005년 (사)한국정보처리학회 부회장. 2009년 고려대학교 정보통신대학 학장. 관심분야는 메타데이타, 소프트웨어공학, 데이타공학, 컴포넌트 기반 시스템, 메타데이타 레지스트리, 프로젝트 매니지먼트 등



정 동 원

1997년 군산대학교 컴퓨터과학과 이학사 1999년 충북대학교 전산학과 이학석사 2004년 고려대학교 컴퓨터학과 이학박사 1998년 전자통신연구원 위촉연구원. 1999년~2000년 ICU 부설 한국정보통신교육원 GIS 분원 전임강사. 2000년~2001년 (주)지구넷 부설 연구소 선임연구원. 2002년~2005년 라임디미디어 테크놀로지 부설 연구소 연구원. 2004년~2005년 고려대학교 정보통신기술연구소 연구조교수. 2005년 Pennsylvania State University PostDoc. 2002년~2004년 TTA 표준화위원회-데이터연구회(SG08.02) 특별위원. 2004년~현재 TTA 표준화위원회-메타데이타표준화 프로젝트그룹(PG406) 위원. 2005년~현재 군산대학교 정보통계학과 교수. 2006년~현재 데이터관리서비스 전문위원회 (ISO/IEC JTC1/SC32 국내위원회) 위원. 2008년~현재 지리정보 전문위원회(ISO/TC211 국내위원회) 위원. 2009년~현재 TTA 표준화위원회-NGIS 프로젝트그룹(PG409). 관심분야는 데이타베이스, 시맨틱 웹, 시맨틱 GIS, 유비쿼터스 컴퓨팅, 정보보안