

OpenAPI 검색시스템을 위한 시맨틱 어노테이션 도구 개발

A Development of Semantic Annotation Tool for OpenAPI retrieval System

최영호(Young-ho Choi)*, 차승준(Seung-Jun Cha)**, 이규철(Kyu-Chul Lee)***
iloot@cnu.ac.kr, junii@cnu.ac.kr, kcleee@cnu.ac.kr

초 록

본 논문에서는 OpenAPI 검색 시스템에 활용되는 시맨틱 어노테이션 도구를 개발하였다. OpenAPI 검색 시스템에서의 시맨틱 어노테이션이란 온톨로지를 이용하여 웹페이지에 의미 정보를 추가하는 작업이다. 하지만 시맨틱 어노테이션 수행 시 웹페이지에 의미정보를 직접 입력해야 했기 때문에 웹 구조 및 내용에 대한 지식이 있는 전문가가 아닌 경우 입력이 어렵고, 일일이 수작업으로 입력해야 하기 때문에 시간이 오래 걸리고 잘못 입력하는 경우가 많이 발생했다. 이를 해결하기 위해 본 논문에서는 우선 관련연구로 MicroWSMO modelReference 를 위한 시맨틱 어노테이션 도구인 SWEET 와 온톨로지를 활용한 시맨틱 어노테이션을 위한 OntospaceAnnotator 를 분석하여 어노테이션에 필요한 기능을 도출하였다. 이를 참조하여 어노테이션에 필요한 항목들을 추출하여 테이블화한 후 선택하는 기능과, 참조할 온톨로지 클래스의 선택을 위한 온톨로지 트리/그래프 뷰어를 개발했다. 도구는 OpenAPI 검색 시스템과의 연동을 위해 자바 스윙 기반으로 작성하여 웹페이지에서 애플릿 형태로 동작하도록 개발하였다. 개발된 시맨틱 어노테이션 도구를 통한 손쉬운 입력으로 효과적인 OpenAPI 의미검색을 위한 기반이 마련되었다.

1. 서론

특정 서비스 제공자들만이 서비스를 제공하는 기존의 웹과는 달리 최근의 웹에서는 블로그나 미니홈피로 인해 데이터를 생산하고 공유하는 서비스 제공자들이 늘어남에 따라 사용자들이 서비스의 작성과 배포가

쉬운 방식을 찾게 된 웹 2.0 시대로 돌입했다. 이에 따라 서비스의 내부 동작 과정을 모르더라도 공개된 인터페이스를 이용하여 해당 서비스의 기능을 손쉽게 사용할 수 있는 OpenAPI 가 발전하게 되었다[1][5].

OpenAPI 란 웹 2.0의 ‘플랫폼으로서의

본 연구는 국토해양부 첨단도시기술개발사업-지능형국토정보기술혁신 사업과제의 연구비지원(07 국토정보 C05)에 의해 수행되었음.

* 충남대학교 컴퓨터공학과 석사과정

** 충남대학교 컴퓨터공학과 박사과정

*** 충남대학교 컴퓨터공학과 교수, 교신저자

웹'이 기술적으로 구현된 방식으로 웹 사이트가 자신의 기능을 이용할 수 있도록 공개한 인터페이스를 뜻한다[2]. 최근 웹에서는 OpenAPI 를 이용한 서비스들에 대한 사용자들의 공유와 참여가 증가함에 따라, OpenAPI 서비스의 숫자는 빠르게 증가하고 있다. 그에 따라 사용자는 자신이 원하는 기능을 정확히 제공하는 서비스를 검색하기 힘들어졌다.

이런 문제점을 해결하기 위해 시맨틱 웹 기술을 활용한 OpenAPI 통합 검색 시스템(이하 OpenAPI 검색 시스템)[3] 이 개발되었다. OpenAPI 검색 시스템은 검색에 사용될 OpenAPI 설명정보페이지를 찾아 모델링 후 시스템에 저장하여 검색에 사용한다. 모델링 시 태깅과 시맨틱 어노테이션을 이용하여 SA-OpenAPI 라고 정의한 통합 서비스 정보모델로 변환했다.

그 중 시맨틱 어노테이션은 온톨로지를 이용하여 웹페이지에 의미정보를 추가하는 기술이다. OpenAPI 검색 시스템에서는 시맨틱 어노테이션을 위해 W3C 표준기술인 SAWSDL 의 모델 참조, 리프팅 스키마 매핑, 로어링 스키마 매핑을 활용하여 도메인 온톨로지를 바탕으로 태깅이 완료된 웹페이지의 오퍼레이션 입/출력 정보에 의미정보를 추가했다.

그러나 시맨틱 어노테이션 작업을 위해서는 의미정보를 웹페이지에 직접 입력해야 하기 때문에 웹 구조 및 내용에 대해 잘 알고 있는 전문가만 입력가능하며, 일일이 수작업을 통해 입력해야 하기 때문에 시간이 오래 걸리고 잘못 입력하는 경우가 발생했다.

본 논문에서는 이러한 문제점들을 해결하기 위해 시맨틱 어노테이션 도구를 개발했다. 개발된 도구는 SA-OpenAPI 구조 중

어노테이션이 필요한 요소를 테이블화하여 제공했고, 온톨로지 트리/그래프 뷰어를 통해 사용자가 온톨로지 클래스를 참조하기 편리하도록 했으며, 모델 참조, 리프팅 스키마 매핑, 로어링 스키마 매핑을 이용한 의미정보 입력 기능을 제공한다.

이를 통해 비전문가도 시맨틱 어노테이션이 가능해졌고, 입력에 걸리는 시간과 잘못된 입력하는 실수를 줄일 수 있었다.

본 논문은 다음과 같은 구성을 가지고 있다. 2장에서는 관련연구로 SOA4ALL 프로젝트에서 개발한 SWEET 과 원광대학교에서 개발한 OntoSpaceAnnotator 를 분석했다. 이를 바탕으로 3장에서는 개발한 도구와 기능들에 대해서 설명한다. 4장에서는 논문의 결론과 향후 연구내용에 대해서 설명한다.

2. 관련연구

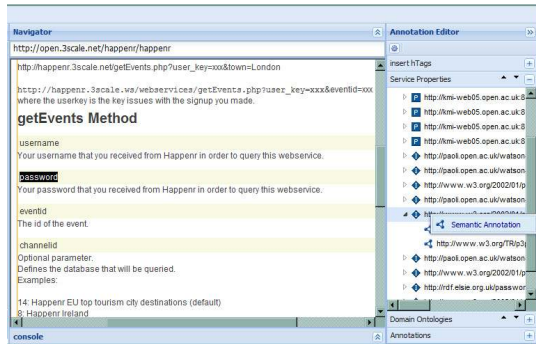
2.1 SWEET

SWEET 은 'Semantic Web sErvice Editing Tool'의 약자로 RESTful 웹서비스로 제공되는 웹페이지에 hRESTS, MicroWSMO 태그를 이용하여 시맨틱 어노테이션을 할 수 있도록 도와주는 도구[6]이다.

SWEET 의 기능은 다음과 같이 크게 5가지로 구분된다. hRESTS 의 마이크로포맷 태그를 추가할 수 있는 기능과 통합된 온톨로지를 검색하여 사용할 수 있는 기능, MicroWSMO 모델참조 태그를 추가하는 기능, HTML 에 시맨틱 어노테이션 한 결과를 저장하는 기능, RDF MicroWSMO 를 추출하는 기능이 이에 해당한다..

이 중 MicroWSMO 모델참조 태그를 추가하는 기능은 <그림 1>과 같이 개발되었다. SWEET 에서는 불러온 온톨로지를 오른

쪽에 URL 형태로 나열하고, 호출된 온톨로지 클래스 중 어노테이션에 필요한 클래스를 선택한 후 해당 클래스에서 오른쪽 마우스 버튼을 통해 시맨틱 어노테이션 버튼을 클릭한다. 그러면 선택된 온톨로지에 대한 모델참조 태그가 HTML 에 추가된다.



<그림 1> SWEET

2.2 OntoSpaceAnnotator

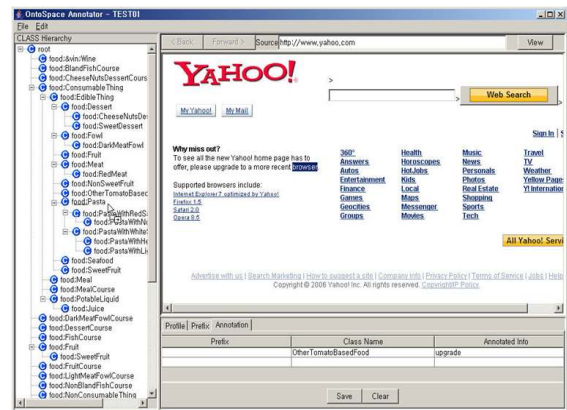
OntoSpaceAnnotator[4]는 원광대학교에서 2006년에 발표되었던 논문인 온톨로지 기반의 시맨틱 어노테이터 구현에서 제안된 웹페이지에 대한 어노테이션을 위한 도구다.

이 도구는 온톨로지를 참조한 시맨틱 어노테이션을 좀 더 간편하게 하기 위해 개발됐다. 도구에서는 어노테이션을 위해 사용될 OWL 파일의 온톨로지를 파싱하여 클래스를 추출하는 작업이 선행되어야 한다. 추출은 로컬 드라이브의 온톨로지 파일을 도구로 불러와서 파일 내용을 버퍼에 저장하고 읽어가면서 클래스만 추출해서 계층구조인 트리 형태로 구성한다.

<그림 2>에서와 같이 도구의 왼쪽은 온톨로지 파일을 파싱하여 추출해낸 클래스를 트리 구조로 표현해주고 있고 오른쪽은 URL 입력을 통해 웹사이트를 열 수 있는 기능을 가지고 있다.

메뉴에서 온톨로지 파일을 불러와서 파싱

하는 작업을 끝마치면 왼쪽의 창에 트리 형태로 각각의 클래스들이 계층구조로 'subclass' 관계를 나타낸다. 오른쪽의 URL 입력창을 통해 불러온 웹사이트의 내용 중 어노테이션이 필요한 단어나 문장을 드래그한 후 왼쪽의 온톨로지 트리의 클래스 중 참조할 클래스에 드래그 앤 드롭 했을 때 해당 항목에 대해 어노테이션이 자동으로 이루어지게 된다. 그리고 어노테이션이 완료된 내용은 오른쪽 아래의 창에서 설명정보를 보여준다.



<그림 2> OntoSpaceAnnotator

3. 시맨틱 어노테이션 도구 개발

OpenAPI 검색 시스템에서 태깅이 완료된 웹페이지의 오퍼레이션 입/출력 정보에 대한 어노테이션을 제공하는 도구와는 달리 각각 SWEET 은 MicroWSMO 모델참조, OntoSpaceAnnotator 는 일반 웹페이지의 본문에 대한 어노테이션을 제공한다

따라서 관련연구에서 제공하는 모든 기능을 수용할 수 없다. 하지만 온톨로지를 이용한 시맨틱 어노테이션을 제공하기 때문에 어노테이션 대상과 온톨로지 정보를 제공하는 창으로 구성된 도구의 구조와 온톨로지의 트리 뷰를 참고했다.

이런 기능들을 바탕으로 개발된 도구에서 필요한 서비스 설명정보에 대한 테이블화와 OWL API 를 이용한 온톨로지 뷰어, 참조된 온톨로지 클래스를 이용한 의미정보 입력 기능을 구현했다.

이렇게 개발된 도구의 전체 모습은 <그림 3>과 같다. 도구의 왼쪽 윗부분은 서비스 설명정보의 테이블화 부분으로 서비스 설명정보 중 시맨틱 어노테이션이 필요한 오퍼레이션의 입/출력값을 테이블 형식으로 구조화하여 출력한다.

왼쪽 아래 부분은 의미정보 입력부로 모델 참조를 통해 참조될 클래스 정보가 들어가는 대한 입력창, 데이터 미디어이션[5]을 위해 리프팅 스키마와 로어링 스키마를 이용하여 매핑 정보를 입력하는 입력창, 시맨틱 어노테이션으로 인해 수정된 코드를 확인하는 코드 확인창을 제공한다.

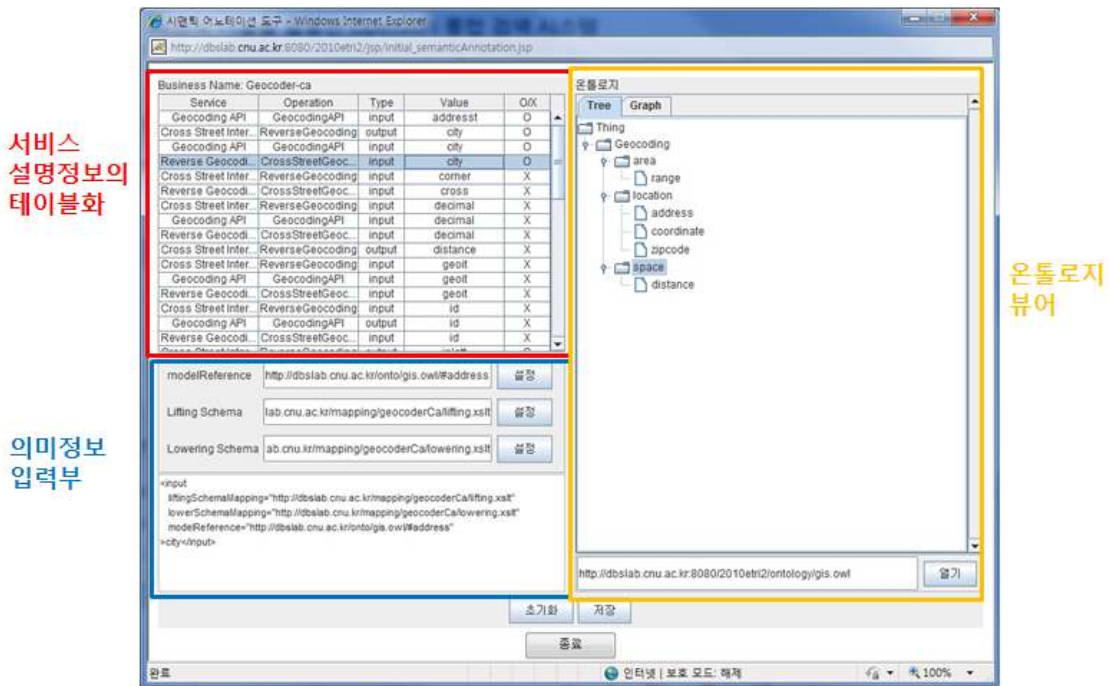
오른쪽에서는 모델 참조를 위한 온톨로지

뷰어를 출력해주며, 트리와 그래프는 탭 메뉴를 통해 선택한다. 대상 온톨로지는 아래의 입력창을 통해 입력받는다.

3.1 서비스 설명정보의 테이블화

OpenAPI 검색 시스템에서는 OpenAPI 서비스의 설명정보를 SA-OpenAPI라고 정의한 통합 모델로 변환해서 검색에 활용한다. 변환의 과정에서 시맨틱 어노테이션은 오퍼레이션의 입/출력값에 시맨틱 정보를 추가하는 것이다. 따라서 SA-OpenAPI 구조 중 시맨틱 어노테이션에 필요한 요소만 추출하여 테이블로 구성 후 제공한다.

<그림 4>는 테이블 구조로 사용자에게 제공된 서비스 설명정보의 모습이다. 테이블에서는 어노테이션이 필요한 오퍼레이션의 입/출력 인자에 대한 정보로 Business 명, 서비스명, 오퍼레이션명, 인자의 종류, 인자 값, 어노테이션 여부를 제공한다.



<그림 3> 전체 도구의 모습

Business명, 서비스명, 오퍼레이션명은 어노테이션이 필요한 입/출력값이 소속된 설명정보 페이지의 구조를 사용자가 이해할 수 있도록 테이블 항목으로 선택했다. 인자의 종류와 인자값은 어노테이션이 필요한 입/출력값의 타입에 대한 정의와 어떤 값을 가지고 있는지 사용자에게 제공하기 위해서 포함됐다. 마지막으로 어노테이션 여부의 경우 사용자가 어노테이션을 완료한 항목인지 구분해주기 위해 테이블 항목으로 선택했다.

이렇게 오퍼레이션 정보를 테이블 구조로 나타내줌에 따라 전문가가 아니더라도 시맨틱 어노테이션이 필요한 부분을 사용자가 쉽게 찾을 수 있었다.

Service	Operation	Type	Value	O/X
Geocoding API	GeocodingAPI	input	address	O
Cross Street Inters...	ReverseGeocoding	output	city	O
Geocoding API	GeocodingAPI	input	city	O
Reverse Geocodin...	CrossStreetGeoco...	input	city	O
Cross Street Inters...	ReverseGeocoding	input	corner	X
Reverse Geocodin...	CrossStreetGeoco...	input	cross	X
Cross Street Inters...	ReverseGeocoding	input	decimal	X
Geocoding API	GeocodingAPI	input	decimal	X
Reverse Geocodin...	CrossStreetGeoco...	input	decimal	X
Cross Street Inters...	ReverseGeocoding	output	distance	X
Cross Street Inters...	ReverseGeocoding	input	geoit	X
Geocoding API	GeocodingAPI	input	geoit	X
Reverse Geocodin...	CrossStreetGeoco...	input	geoit	X
Cross Street Inters...	ReverseGeocoding	input	id	X
Geocoding API	GeocodingAPI	output	id	X
Reverse Geocodin...	CrossStreetGeoco...	input	id	X

<그림 4> 서비스 설명정보 테이블화

3.2 온톨로지 뷰어 개발

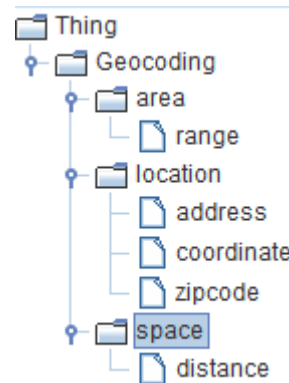
시맨틱 어노테이션을 위해서는 의미정보의 추가를 위해 의미와 관련된 온톨로지 클래스의 참조가 필요하지만 일반 사용자들이 온톨로지 구조를 이해하기 어려웠고, 참조된 내용을 입력하는 작업도 직접 해야 했기 때문에 불편했다. 이런 문제점을 해결하기 위해 온톨로지를 분석하여 트리 형태과 그래프 형태로 사용자에게 제공하는 뷰어를 개발했다.

3.2.1 온톨로지 트리 뷰어

온톨로지 트리 뷰어는 온톨로지의 최상위 클래스부터 하위 연결고리를 확인 가능하고 사용자들이 참조할 클래스를 선택할 수 있도록 만들었다.

<그림 5>는 온톨로지 트리 뷰어의 모습으로 최상위 클래스부터 하위 클래스까지 온톨로지의 클래스들을 분석하여 계층 구조로 출력했다.

온톨로지 트리 뷰어의 기능은 노드 펼치기, 노드 접기, 클래스 참조 기능으로 나뉜다. 노드 펼치기 기능은 상위 노드를 더블 클릭할 경우 하위 노드가 펼쳐지는 기능이며, 노드 접기는 하위노드를 사용자가 보고 싶지 않을 때 펼쳐진 하위 노드들의 상위 노드를 더블 클릭하면 하위 노드들을 접어 보이지 않도록 하는 기능이다. 마지막으로 클래스 참조는 모델 참조 시 필요한 클래스를 선택하여 적용하는 기능이다.



<그림 5> 온톨로지 트리 뷰어

3.2.2 온톨로지 그래프 뷰어

온톨로지 그래프 뷰어는 온톨로지의 전체 모습을 한 눈에 사용자가 볼 수 있도록 하여 각 클래스간의 관계를 쉽게 파악할 수 있으며, 트리 뷰어와 달리 사용자가 설정한

범위의 온톨로지 클래스들을 선택된 클래스를 기준으로 볼 수 있도록 만들어졌다.

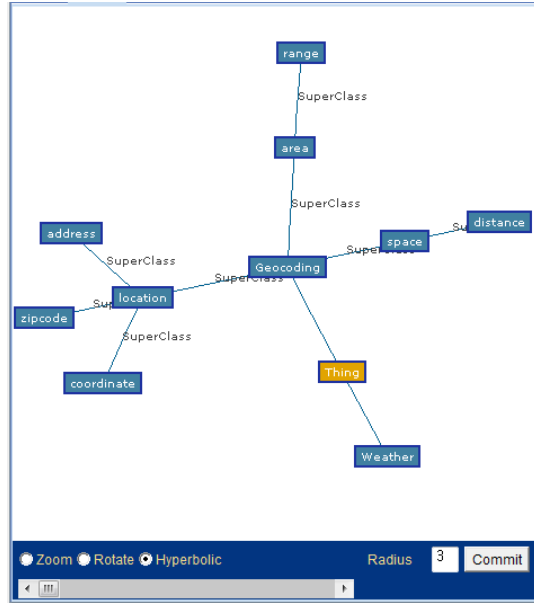
온톨로지 그래프 뷰어는 TouchGraph[7]라는 오픈소스를 활용했다. TouchGraph는 연관성 있는 정보를 그래프화 해주는 도구로 그래프화면을 구성하는 기능과 노드 제어를 위한 기능을 제공한다.

그 중 그래프 화면 구성을 위해 제공하는 기능들은 그래프의 크기를 설정해서 시각적으로 보기 편하도록 해주는 ‘Zoom’기능, 그래프가 회전이 가능하도록 해주는 ‘Rotate’기능, 그래프의 선 길이 조절을 가능하도록 해주는 ‘Hyperbolic’기능으로 이 기능들은 라디오 버튼을 통해 선택 가능하며 스크롤바를 이용하여 그래프의 크기나 회전방향, 그래프의 선 길이를 조절 가능하다.

노드 제어를 위해 제공되는 기능들은 다음과 같다. 선택된 노드의 하위 노드를 펼치는 “Expand Node”, 펼쳐진 노드를 숨기는 “Collapse Node”, 선택된 노드를 숨기는 “Hide Node”, 선택된 노드를 가운데로 이동시켜 주는 “Center Node”가 이에 해당한다.

개발된 도구에서는 OWL API를 이용하여 온톨로지를 구조화했고, 온톨로지의 관계성 중 ‘subclass’ 관계를 그래프에 출력할 수 있도록 만들었다. 그리고 ‘Radius’기능을 추가로 구현하여 펼쳐지는 온톨로지 클래스의 범위를 수치로 입력하여 사용자가 필요한 범위까지 온톨로지 구조를 볼 수 있도록 해주었다. ‘Radius’는 입력창에 수치를 입력하여 commit 버튼을 누르고 기준이 될 노드를 클릭하면 해당 노드를 기준으로 설정한 범위만큼의 온톨로지 클래스들이 화면에 출력된다.

<그림 6>은 TouchGraph를 수정하여 구현한 온톨로지 그래프 뷰어이다.



<그림 6> 온톨로지 그래프 뷰어

3.3 참조된 온톨로지 클래스 반영

온톨로지 뷰어에서 참조된 클래스를 설명 정보 페이지에 적용하려면 모델 참조, 리프팅 스키마와 로어링 스키마를 이용한 매핑 정보 입력과정을 거쳐야 한다. <그림 7>은 참조된 클래스와 매핑정보를 입력하기 위한 입력창과 입력 후 변경된 코드를 보여주는 코드 확인창이다.

모델 참조 입력창은 온톨로지 뷰어에서 선택한 클래스 정보가 자동으로 입력되고, 설정 버튼을 통해 코드 확인창에 반영된다. 또한 리프팅 스키마 매핑과 로어링 스키마 매핑 정보는 사용자에게 의해 입력되고, 모델 참조와 마찬가지로 설정 버튼을 통해 코드 확인창에 반영된다.

modelReference	<input type="text" value="http://dbslab.cnu.ac.kr/onto/gis.owl/#address"/>	<input type="button" value="설정"/>
Lifting Schema	<input type="text" value="lab.cnu.ac.kr/mapping/geocoderCa/lifting.xslt"/>	<input type="button" value="설정"/>
Lowering Schema	<input type="text" value="lab.cnu.ac.kr/mapping/geocoderCa/lowering.xslt"/>	<input type="button" value="설정"/>

```

<input
  liftingSchemaMapping="http://dbslab.cnu.ac.kr/mapping/geocoderCa/lifting.xslt"
  lowerSchemaMapping="http://dbslab.cnu.ac.kr/mapping/geocoderCa/lowering.xslt"
  modelReference="http://dbslab.cnu.ac.kr/onto/gis.owl/#address"
>city</input>
  
```

<그림 7> 참조된 온톨로지 반영

4. 결론

본 논문에서는 OpenAPI 검색시스템을 위해 개발한 시맨틱 어노테이션 도구를 제안했다.

이를 위해 관련연구로 SWEET 과 On-toSpaceAnnotator 를 분석하여 어노테이션 대상을 출력하는 창과 온톨로지 정보를 제공하는 창으로 구성된 도구 구조와 온톨로지 트리 뷰 기능을 참고했다.

참고한 내용을 바탕으로 도구에서는 SA-OpenAPI 구조에서 어노테이션이 필요한 항목들을 추출하여 테이블화했고, OWL API 와 TouchGraph 를 이용한 온톨로지 뷰어를 개발하여 사용자들의 온톨로지 구조에 대한 이해를 돕고, 온톨로지 클래스에 대한 모델 참조 기능 포함시켰다. 또한 모델 참조, 리프팅 스키마 매핑, 로어링 스키마 매핑을 통해 의미정보를 입력할 수 있는 입력부를 구현했다.

이러한 기능들로 인해 전문가가 아니더라도 시맨틱 어노테이션이 가능하며, 입력에 걸리는 시간과 입력 오류도 줄일 수 있었다.

향후 연구로 시맨틱 어노테이션 도구를 통해 메타데이터를 입력하여 검색할 수 있었던 OpenAPI 서비스들 중 조합 가능 서비스를 조합하는 방식에 대한 연구를 지속할 예정이다.

- [3] 최윤정, 차승준, 이규철, “시맨틱웹 기술을 활용한 OpenAPI 통합 검색 시스템 개발”, 데이터베이스연구, 제26권 제3호, pp45-65, 2010.
- [4] 박재훈, 유재규, 전양승, 정영식, 한성국, “온톨로지 기반의 시맨틱 어노테이터 구현” 프로그래밍언어논문지, 제20권 제2호, pp 17-22, 2006.
- [5] O'Reilly T., What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software, 2005
- [6] Maleshkova, M., Kopecky, J. and Pedrinaci, C., “Adapting SAWSDL for Semantic Annotating RESTful Services”, Beyond SAWSDL at On TheMove Federated Conferences & Workshops, 2009.
- [7] Shapiro, A., “touchgraph Project.”, <http://touchgraph.sourceforge.net>, 2001

참고문헌

- [1] 김재철, 이규철, u-GIS 국토정보 제공 시스템, 한국공간정보시스템학회, 제11권 제1호, 2009, pp1-8
- [2] 박지강, 당신은 웹 2.0 사람입니까?, 한빛미디어, 서울 2007