

## 온톨로지 인스턴스 생성 지원 도구 개발

### A development on Ontology Instance Management Tool

이미경\*, 정한민\*, 김문석\*\*, 성원경\*  
한국과학기술정보연구원\*, K4M\*\*

Mikyong Lee\*, Hanmin Jung\*, Mun Seok Kim\*\*,  
Won-Kyung Sung\*  
Information Service Research Team, KISTI\*  
Solution Business Division, K4M Inc.\*\*

#### 요약

시맨틱 웹 기술의 발전에 따라 온톨로지는 점점 복잡해지고 대용량화되고 있어서 기존의 온톨로지 저작도구를 이용하여 인스턴스를 구축하게 되면 인스턴스 관리에 많은 어려움이 따르기 때문에 사용자 입장에서 편리하게 인스턴스 생성을 지원해주는 도구가 필요하게 되었다. 본 논문에서 개발한 온톨로지 지원도구 OntoManager는 계층적 인스턴스들의 관계를 쉽게 구현하기 위해서 히스토리 기반의 인스턴스 생성 뷰를 제공하여 인스턴스 생성 작업 시 사용자의 현재 작업 뷰에서 흐트러지지 않고 새로운 클래스의 인스턴스 생성할 수 있으며 생성된 인스턴스의 자동 연결 기능을 제공한다. 그리고 웹브라우저와 개체명 인식 기능을 제공하여 웹 페이지의 텍스트에서 인스턴스 생성에 필요한 부가 정보를 자동 추천해줘서 인스턴스 속성 값을 쉽게 구축할 수 있도록 도와준다. 마지막으로 이미지 어노테이션 기능을 제공하여 이미지의 특정 영역 정보를 선택하여 속성 값으로 가지는 인스턴스 생성도 지원해준다.

#### Abstract

In this paper we present an Ontology Instance Management Tool. OntoManager is a user-friendly interactive ontology Instance management tool with webpage annotation tool and an image annotation tool. It supports the user with the task of creating and maintaining ontology-based OWL-markup, creating of OWL-instances, attributes and relationships. It include an ontology browser for the exploration of the ontology and instances and a HTML browser that will display the annotated parts of the text. And OntoManager is an image annotation tool that allows users to markup regions of an image with respect to concepts in an ontology. It provides the functionality to import images, ontologies, instance bases, perform markup, and export the resulting annotations to disk or the Web.

## I. 서론

시맨틱 웹 기술이 발달함에 따라 시맨틱 웹 관련 기술 동향도 점점 다양하게 변화하고 있다. 시맨틱 웹 등장 초기에는 많은 사람들이 온톨로지를 구축하는 것에 많은 관심을 가져서 Protégé, OntoEdit, OilEd, OntMat 등의 온톨로지 저작도구와 온톨로지 과서에 대한 연구가 활발했다. 지금은 대용량 온톨로지를 위한 추론 엔진과 자동 인스턴스 구축 또는 시맨틱 기술을 이용한 서비스와 같은 응용 분야의 연구와 개발이 활발한 실정이다. 현재 온톨로지 인스턴스 저작 도구들은 Annotation Tool이라는 이름으로 인스턴스 생성 및 인스턴스 생성 지원 기능을 가진다. 인스턴스 생성 지원 도구들은 이미지 편집 툴이나 텍스트 자동 추출 기능을 포함하여 사용자가 인스턴스를 쉽게 생성할 수 있도록 도움을 준다.

우리가 개발한 도구는 인스턴스 생성 및 지원 도구로써 기존 도구의 단점을 보완하였다. 하나의 인스턴스를 생성하기 위

해서는 그와 관련된 인스턴스들이 이미 준비된 상태에서 생성 프로세스를 마칠 수 있었다. 예를 들면 “홍길동의 소속기관은 키스티이다”라는 속성 값을 저장하기 위해서는 기관 클래스의 인스턴스로 “키스티”가 먼저 생성되어있어야 “홍길동”의 생성 작업을 완료할 수 있었다. 하지만 우리는 “키스티”가 먼저 생성되어 있지 않으면 “홍길동” 생성 작업을 중단하고 “키스티”를 생성한 후, 새롭게 “홍길동” 생성 작업을 해야하는 기존의 도구와는 다르게 “홍길동” 생성 작업 중에 히스토리 기능을 이용해 “키스티” 생성 작업을 하면 자동으로 “홍길동”의 소속 기관에 “키스티”가 입력될 수 있게 지원해 준다. 이렇게 기존에 독립적으로 생성되던 인스턴스를 온톨로지에 설계된 연결 관계를 통해 자동으로 생성할 수 있게 지원해주는 히스토리 기반의 인스턴스 생성 기능을 장점으로 가진다.

본 논문의 구성은 다음과 같다. 2장에서 관련된 기존 연구들을 살펴보고 3장에서는 온톨로지 인스턴스 생성 지원도구인 OntoManager에 대해서 자세히 살펴보고 4장 결론으로 끝을

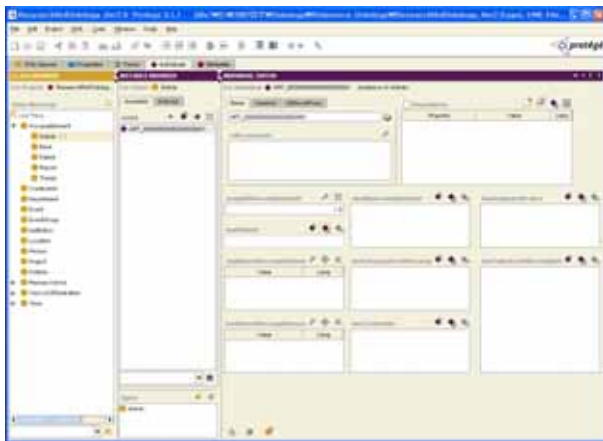
맺는다.

## II. 기존 연구

### 1. 온톨로지 저작 도구

#### 1.1 Protégé

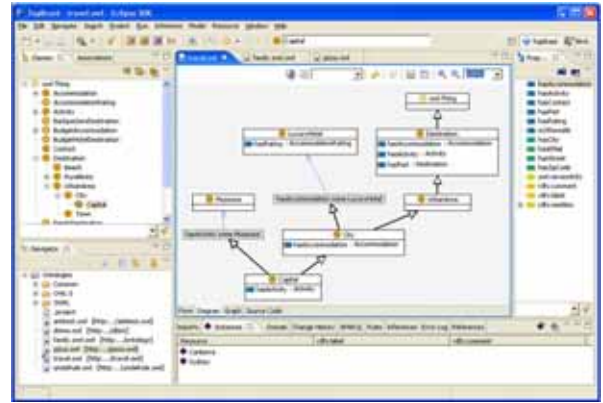
Protégé는 온톨로지 저작도구로 널리 알려져 있다[1]. 이 툴은 온톨로지 클래스와 프로퍼티를 생성하는데 주로 사용되고 있으며 그림 1과 같이 Individuals라는 탭을 통해 온톨로지 인스턴스를 생성, 편집 기능을 제공하고 있다. Protégé는 가장 기본적인 온톨로지 인스턴스 생성 및 편집 기능을 제공해주고 있다. 따라서 다른 도구에서처럼 인스턴스 편집을 편리하게 도와주는 기능은 제공되지 않으며 특정 과인스턴스를 선택할 때 URI를 이용하여 검색 및 선택을 해야하므로 URI를 알지 못하면 사용하기 어려운 단점을 가진다.



▶▶ 그림 1. Protégé

#### 1.2 TopBraid Composer

TopQuadrant에 개발된 TopBraid Composer은 온톨로지 저작도구으로써 이클립스의 플러그인 형태로 이루어져있다[2]. 그림 2에서 보는바와 같이 클래스를 정의할 때 UML과 같은 클래스 다이어그램으로 정의하기 때문에 기존의 도구들에서 다루는 트리형식보다 클래스들의 관계를 한눈에 알아볼 수 있는 장점을 가진다. 그리고 클래스들간의 무결성도 체크할 수 있으며 규칙을 생성할 수 있다. SPARQL 쿼리를 지원해주며 RDF 그래프로 온톨로지를 확인할 수 있다. DB, UML, XML 스키마와 스프레드시트를 통한 데이터 반입, 반출 지원 및 HTML 문서로 생성할 수 있는 장점을 가진다.



▶▶ 그림 2. TopiBraid

### 2. 온톨로지 주석도구

#### 2.1 OntoMat-Annotizer

OntoMat-Annotizer는 사용자에게 친숙한 웹페이지 기반의 어노테이션 도구이다[3]. 우리 시스템과 상당히 유사한 인터페이스를 갖고 있는 이 도구는 OWL 인스턴스, 프로퍼티 그리고 그들간의 관계를 생성하는 OWL 기반 온톨로지를 생성하고 관리하는 역할을 한다. 이것은 그림 3과 같이 온톨로지와 인스턴스를 탐색하는 온톨로지 브라우저, 텍스트를 어노테이션하기 위한 HTML 브라우저를 포함하고 있다. 이것은 사용자에게 웹 페이지에서 OWL-메타 데이터를 구축하기 쉽게 해준다. 텍스트를 쉽게 수정할 수 있도록 오른쪽에 나타나는 브라우저에 웹페이지 하이라이팅 기능을 제공해주며 드래그&드랍 기능을 통해 인스턴스를 쉽게 생성할 수 있게 도와준다.

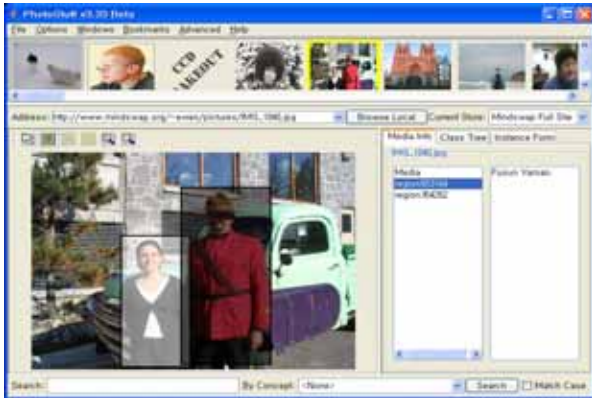


▶▶ 그림 3. OntoMat-Annotizer

#### 2.2 PhotoStuff

PhotoStuff는 시맨틱웹 상의 디지털 이미지 어노테이션 도구이다[4]. 이것은 Paperpuppy 시스템에서 사용하고 있는 인스턴스 생성 지원 도구의 하나로 온톨로지와 이미지의 특정

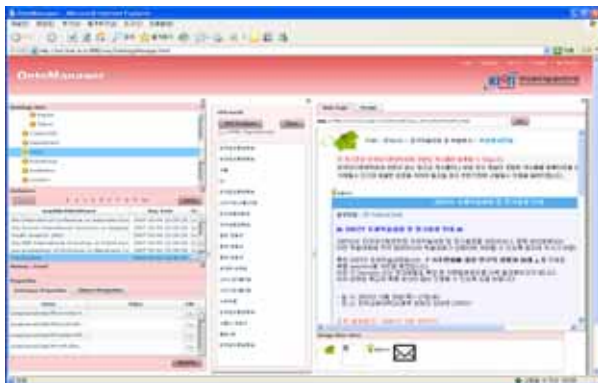
영역과의 관계를 맺어주고 어노테이션을 다는 기능을 제공하여 온톨로지 메타데이터를 자동으로 생성해준다. 이 도구는 온톨로지와 이미지 데이터의 로딩과 주석을 다는 기능과 함께 클래스의 인스턴스 생성 및 인스턴스에 대한 assertion을 추가적으로 제공해준다. 새롭게 생성된 인스턴스들은 EXIF(Exchangeable Image File Format) 메타데이터와 함께 사이트의 저장소에 관리되며 재사용이 가능하다.



▶▶ 그림 4. PhotoStuff

### III. OntoManager의 구성

OntoManger는 우리가 개발한 온톨로지 인스턴스 저작, 지원 도구이다. 현재 이 도구는 URI서버와의 웹서비스를 통해 인스턴스를 저장, 수정, 삭제 기능을 수행하지만 URI서버와의 연계 없이 독립적으로 기능을 수행할 수 있다. OntoManager의 주요 기능은 온톨로지 클래스의 인스턴스를 관리하는 생성, 수정, 삭제, 조회 기능이다. 이 도구는 그림 5와 같이 구성되어 있으며 왼쪽 영역은 온톨로지 클래스를 선택하여 인스턴스를 조회하거나 편집하는 기능을 제공하며, 오른쪽 영역은 인스턴스를 생성할 때 필요한 정보를 지원하기 위해서 웹브라우저를 통한 텍스트 추천 및 이미지 주석 기능 등을 제공하게 된다.



▶▶ 그림 5. OntoManger

## 1. 인스턴스 관리부

OntoManger의 인스턴스 관리부는 그림 6과 같으며 Ontology view, Instances view, History view, Property view로 구분할 수 있다.

### 1.1 Ontology view

OntoManger에서 온톨로지 적재는 URI 서버와의 웹서비스를 통해 이루어진다. OWL기반의 온톨로지를 선택하면 온톨로지가 파싱되어 클래스 기반의 트리 형태로 보이며, 온톨로지 뷰를 통해 선택된 클래스의 인스턴스들을 조회, 편집할 수 있게 된다.

### 1.2 Instances view

인스턴스 관리 기능은 조회, 생성, 수정, 삭제의 4가지로 구성된다. 온톨로지 뷰에서 클래스를 선택하면 인스턴스 뷰에서 선택된 클래스에 해당되는 인스턴스 목록들이 나타난다. 인스턴스를 선택하여 속성 값을 수정하거나 삭제가 가능하며 인스턴스 신규 등록을 원할 때는 프로퍼티 뷰에서 객체 속성 값과 데이터 속성 값을 등록하면 된다.

### 1.3 History view

OntoManger에서는 히스토리 기능을 통해 현재 사용자가 작업하고 있는 클래스를 기억해둔다. 사용자 편집 클래스의 히스토리를 항상 저장하고 있기 때문에 관련된 인스턴스를 생성할 때 현재 사용자의 작업 관점을 변경하지 않고 새로운 클래스의 인스턴스를 작업할 수 있고 이렇게 생성된 인스턴스가 기존에 작업 중이던 인스턴스로 자동으로 연결되기 때문에 인스턴스 생성에 많은 편리함을 제공한다[5].

### 1.4 Property view

인스턴스의 속성 값을 등록할때는 객체 속성 값과 데이터 속성 값을 탭을 이용해서 구분하고 있다. 데이터 속성 값의 경우는 String, File, URL로 구분하여 입력 형식을 지원해준다. 객체 속성 값의 입력은 기존에 존재하는 인스턴스를 검색하여 선택하는 기능과 기존에 존재하지 않는 인스턴스를 선택하기 위해 신규 인스턴스 작성 버튼을 제공한다. 신규 인스턴스 작성 버튼을 누르면 History가 계층적으로 바뀌고 작업 중이던 Property창이 신규 생성 형태로 바뀌게 된다.



▶▶ 그림 6. 온톨로지 편집부

## 2. 인스턴스 생성 지원부

온톨로지 생성 지원부는 웹브라우저와 텍스트 개체명 추출 기능과 온톨로지와 관련된 이미지 영역 추출 기능을 제공해준다. 이 기능들은 인스턴스를 생성할 때 관련 속성 값들을 입력할 때 편리하게 도움을 주고자 구현된 추가 기능이다.

### 2.1 텍스트 개체명 추출 기능

인스턴스 저장을 지원하기 위해 우리 시스템에서는 웹브라우저를 제공해준다. 웹 브라우저를 이용하여 원하는 정보를 검색해서 인스턴스 생성 및 속성 값을 구현할 때 이용할 수 있다. 그림 7에서처럼 오른쪽은 웹브라우저이고 왼쪽은 개체명 추출 영역이다. 웹페이지에서 자동으로 개체명을 인식하여 텍스트를 추출하여 개체 정보를 제공해준다.



▶▶ 그림 7. 웹브라우저와 NER 결과

만약 이벤트 클래스의 인스턴스를 작업하는 중이라면 그림 7과 같이 인스턴스로 생성할 학술대회의 홈페이지를 검색하여

학술대회명, 학술대회 기간, 학술대회 개최 위치와 같은 속성 값에 해당하는 정보들을 입력하는데 도움을 준다. 이 도구는 복사&붙여넣기와 드래그&드랍을 통해 속성 값을 입력하는 기능을 제공해 준다. 그리고 아래에 나타나는 image view area는 현재 웹페이지에 나타난 이미지들을 자동 다운로드해서 담고 있는데 원하는 이미지를 클릭하게 되면 이미지 영역 편집 탭으로 자동으로 바뀌게 된다.

### 2.2 이미지 영역 추출 기능

인스턴스 저장을 지원하기 위해 제공하는 또 다른 기능으로 이미지 영역 추출 기능이 있다. 웹 브라우저와 탭을 이용하여 구분하고 있는 이미지 브라우저는 웹 페이지의 이미지나 로컬의 이미지 파일 또는 URL을 가진 이미지를 로딩하여 이미지 파일의 인스턴스 생성 및 이미지에서의 특정 영역에 대한 인스턴스 생성을 지원해준다. 그림 8과 같이 단체 사진에서 특정 사람을 선택해서 그 정보를 Person 클래스와 관련된 속성 값으로 저장하길 원한다면, 마우스를 사용하여 원하는 영역을 선택하게 된다. 그림 왼쪽의 이미지 영역 정보 창에 선택된 이미지 영역이 임시로 메모리에 저장된다. 이것은 저장, 삭제 버튼을 제공하여 선택된 영역을 인스턴스의 속성 값으로 저장하거나 삭제하는 기능을 제공한다.



▶▶ 그림 8. 이미지 브라우저와 선택된 이미지 영역 정보

## V. 결 론

본 논문에서는 우리가 개발한 온톨로지 인스턴스 저장, 관리 도구의 개발에 대해 설명하였다. 기존에 개발된 온톨로지 저장 도구들은 독립된 인스턴스를 생성하거나 인스턴스 생성에 도움을 주기 위한 이미지 어노테이션 도구 또는 웹브라우저 텍스트 추출 기능에 초점을 맞추고 있다. 하지만 복잡한 관계를 가지는 온톨로지가 등장하고 대용량의 인스턴스를 구축, 관리

해야할 필요성이 증가함에 따라 기존에 독립적으로 생성되던 인스턴스를 좀 더 편리하게 등록할 수 있는 도구의 필요성이 대두되었다.

따라서 우리가 개발한 OntoManager는 히스토리 기반 인스턴스 저작, 관리 도구로 복잡한 온톨로지의 인스턴스를 생성할 때, 서로 연계된 인스턴스를 하나의 편집 흐름으로 구축할 수 있는 히스토리 기능을 제공함으로써 인스턴스 생성 및 편집의 효율성을 향상시키는 장점을 가진다. 기존의 도구들은 두 개의 인스턴스가 이미 존재할 경우에만 서로간의 관계를 맺을 수 있는 단점을 가지기 때문이다.

우리는 히스토리를 통해 사용자의 현재 작업 뷰에서 흐트러지지 않고 새로운 인스턴스를 생성하고 자동 연결하는 기능을 제공하여 인스턴스 생성에 편리함을 제공한다. 또한 인스턴스 생성을 지원하기 위해 제공되는 웹브라우저 개체명 추출 기능과 이미지 영역 추출 기능 및 자동 입력 기능을 통해 인스턴스를 생성할 때 편리함을 제공하였다.

#### ■ 참고 문헌 ■

- [1] <http://protege.stanford.edu/>
- [2] <http://www.topbraidcomposer.com/features.html>
- [3] <http://annotation.semanticweb.org/ontomat/index.html>
- [4] <http://www.mindswap.org/2003/PhotoStuff/>
- [5] 이미경, 정한민, 성원경 “히스토리 기반 온톨로지 인스턴스 관리 도구”, 제 19회 한글 및 한국어처리학회, 제1권 제1호, 2007.