

## 시맨틱웹을 위한 온톨로지 주석 시스템의 설계와 구현

류영현<sup>○</sup>, 왕용, 한성국  
 원광대학교 컴퓨터공학과  
 {spirra<sup>○</sup>, crowtf, skhan}@wonkwang.ac.kr

### Design and Implementation of Ontology-Annotation System for Semantic Web

YeongHyeon Ryu<sup>○</sup>, Wang Yong, SungKook Han  
 Dept of Computer Engineering, Wonkwang University

#### 요 약

현재 웹의 발전과 더불어 시맨틱웹의 응용에 대한 연구가 활발히 진행되고 있다. 그러나 아직까지 그것에 대한 결과물이 나오지 않는 것은 정보자원을 의미적으로 정확히 분석하고, 관리할 수 없는 시스템의 부재라고 할 수 있다. 본 논문에서는 기존의 Annotator들을 분석하고, 시맨틱웹의 응용에 필요한 온톨로지 annotation system을 구현하여, 사용자가 원하는 정확한 정보를 검색하고, 사용자가 편리하게 관리·저장할 수 있는 방법을 제시하였다.

#### 1. 서론

최근 들어 인터넷의 발전으로 엄청나게 늘어나고 있는 정보의 양은 사용자들에게 많은 지식과 다양한 서비스를 제공하고 있는 반면에 정보 과다(information overload)라는 새로운 문제점을 가지게 되었다. 검색엔진이 점차 발전되어 이러한 문제점을 해결하려고 시도하고 있지만 대부분의 검색엔진이 웹 문서의 내용보다는 단어나 구문 등 단편적인 방법으로 관련성을 검사하므로 사용자가 의미적으로 원하는 정보의 검색이 어려운 실정이다. 그래서 현재 온톨로지를 이용하여 의미기반으로 정보를 검색할 수 있도록 하는 많은 연구가 진행 중이고, 시맨틱 웹과 관련된 많은 논문들이 배출되고 있다.

본 연구에서는 이러한 온톨로지 연구를 기반으로 시맨틱웹으로 발전하기 위한 온톨로지 annotator를 개발하였다. 온톨로지 annotator를 이용해 온톨로지 문서나 웹 정보들에 주석을 두고 또 하나의 의미기반 RDF문서를 생성함으로써, 사용자에게 더욱 정확한 정보를 제공할 수 있게 하였다[1].

본 논문의 구성은 다음과 같다. 2장에서는 관련연구로 기존의 온톨로지 주석시스템에 대한 비교분석을 하고, 본 논문에서 구현 될 주석시스템에 대한 요구사항을 기술한다. 3장에서는 이 논문에서 구현한 온톨로지 기반 주석 시스템의 전체적인 설계와 구성, 각 모듈에 대한 기능을 설명하고, 마지막으로 4장에서는 연구 내용을 요약하고 향후 과제와 함께 결론을 내린다.

#### 2. 관련연구

##### 2.1 온톨로지 주석 시스템의 개요

우리는 온톨로지를 수많은 틀을 이용하여 손쉽게 구축할 수가 있다. 하지만 정작 그 온톨로지 만으로는 아무것도 할 수가 없는 게 사실이다. 우리는 그 온톨로지를 가지고 시맨틱웹을 구현할 수가 있다. 수많은 정보들에 꼬리표를 붙여 컴퓨터가 의미처리를 할 수 있게 지식베이스를 구축하는 게 주석부여의

목적이다. 그렇게 주석처리를 하게 되면 컴퓨터는 사용자가 원하는 정확한 정보를 찾아낼 수가 있고, 온톨로지 Crawler를 사용하여 그 주석처리 한 결과들을 모아서 검색한다면 검색결과로 그냥 웹페이지를 보여주는 게 아닌 필요한 정보만을 직접 찾아주는 것이 된다. 그것이 시맨틱웹이다[2].

##### 2.2 기존 온톨로지 주석 시스템의 분석

현재 온톨로지 주석 시스템으로 가장 많이 사용되는 도구는 MnM, SMORE, OntoMat 등이 있다.

MnM은 웹온톨로지와 RDF, DAML+OIL 문서들을 지원하고, 드래그 앤 드롭 형식과 그래픽한 요소를 통한 주석처리를 제공하고 있다. 또한, 주석처리 시 웹페이지의 HTML 소스자체에 컨셉의 이름으로 태그를 만들어 사용자에게 보여주고, KB(Knowledge Base)를 생성하여 주석의 출처와 주석처리 된 부분만을 새로운 RDF파일로 생성하여 새로운 지식을 따로 보관할 수 있게 한다[3].

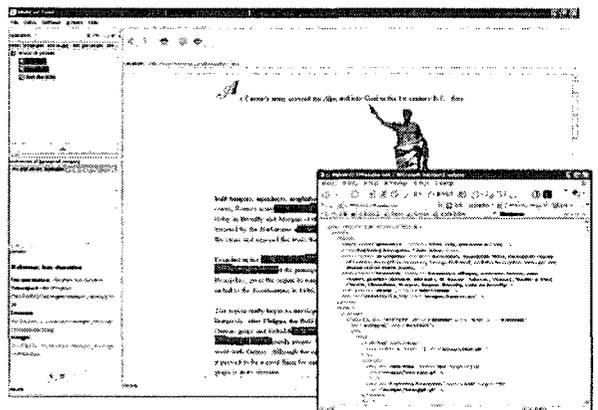


그림 1. MnM 온톨로지 주석 시스템 (예 : wine.owl)

SMORE는 온톨로지 에디터 수준의 온톨로지 문서 파싱능력을 갖추고 있고, 웹상의 온톨로지나 RDF, DAML+OIL, OWL 문서를 모두 사용가능하며, HTML 에디터를 제공함으로써 직접 만들고 수정한 웹페이지에서 온톨로지에 대한 주석을 명시할 수 있다. 주석처리시 주석이 명시된 새로운 온톨로지 문서를 만들어 보여주고, 그 구조를 테이블로 간단히 보여주는 창을 도입하여 쉽게 주석 처리된 문서를 파악할 수 있도록 하였다 [4].

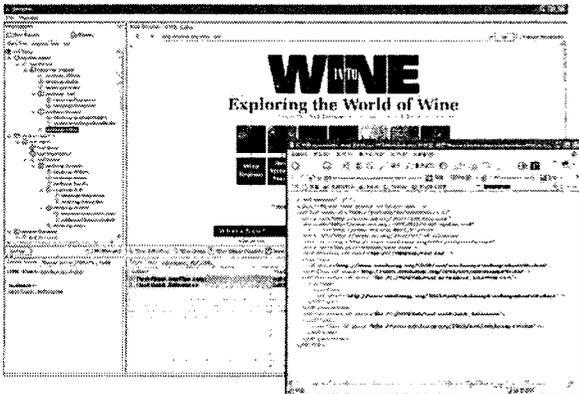


그림 2. SMORE 온톨로지 주석 시스템 (예 : wine.owl)

OntoMat는 간단한 비와 쉬운 사용법으로 가장 인기있는 주석처리 도구로써 온톨로지 문서내에 출처를 포함한 주석처리한 내용을 삽입하는 특징이 있다. N3와 OWL 파일을 지원하고, 별다른 설치없이 웹에서 바로 사용 가능하다는 게 장점이다.

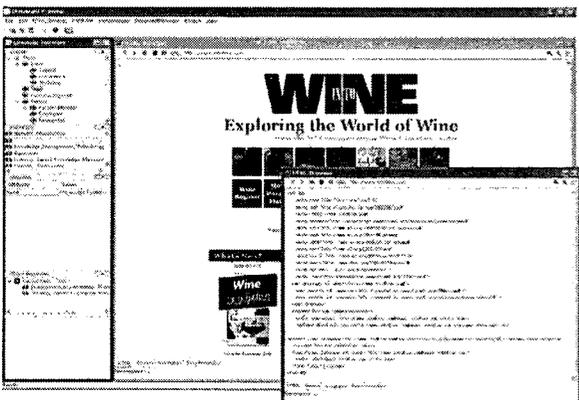


그림 3. OntoMat 온톨로지 주석 시스템 (예 : wine.owl)

### 2.3 온톨로지 주석 시스템의 요구사항

앞서 언급한 바와 같이 기존의 도구들을 분석해 본 결과 온톨로지에 주석을 부여하는 방식에는 크게 3가지가 있다는 것을 알 수 있다.

- ① 웹페이지의 HTML 파일 자체에 주석처리 한 결과를 소스

자체에 일체 쓰는 방식.

- ② 다른 하나의 XML 문서를 만들어서 주석처리 한 결과만을 따로 저장하는 방식.

- ③ 직접 온톨로지 문서에 주석을 부여 하여 저장하는 방식.

비교 결과, SMORE처럼 다른 RDF 문서를 만들어서 주석처리 한 결과만을 저장하는 방식을 사용한다면, 보다 가벼운 문서들로 인해 빠른 검색과 수정과 관리가 용이할 거라 판단된다. 그리고 온톨로지 주석 시스템은 사용자가 직접 온톨로지 문서에 대한 주석을 생성하기 때문에 쉬운 사용법(드래그 앤드롭)과 필요한 기능만 추가된 간편한 인터페이스가 필요하다. 또한 온톨로지 표준 언어인 OWL과 RDF 문서를 동시에 import 할 수 있어야 하며, 웹페이지 뷰어 뿐 아니라 Text로 직접 주석을 기술할 수 있는 공간도 필요하다.

### 3. 온톨로지 주석 시스템의 구현

#### 3.1 온톨로지 주석 시스템의 설계

다음 그림은 시맨틱웹 응용을 위한 온톨로지 주석 시스템 설계 모델의 구체적인 모습을 Database 구현을 기준으로 두 부분으로 나누어 자세히 표현하고 있다.

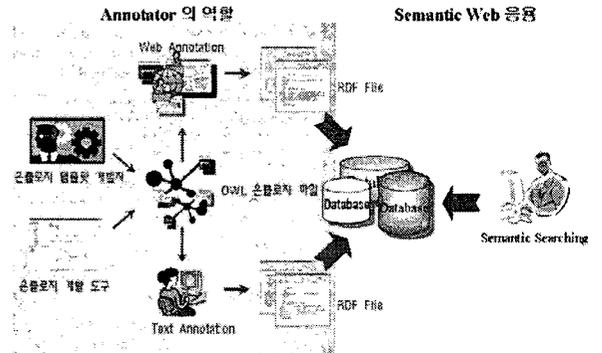


그림 4. 온톨로지 주석 시스템의 설계 모델

이 설계 모델에서는 온톨로지 개발도구 등을 이용해 OWL 온톨로지 파일에 사용자가 웹정보나 Text로 주석처리 하여 다른 RDF파일을 작성하고, 이것을 관계형 구조의 DB에 저장하여 사용자가 의미 검색을 할 수 있도록 하는 것을 보여주고 있다. 여기에서 주석 시스템의 역할 뿐 아니라 시맨틱웹을 위한 DB구축도 보여주고 있지만, DB는 주석 시스템 구현 후 다음 개발단계인 사용자 검색환경 개발에서 다루어지도록 한다[5].

#### 3.2 온톨로지 주석 시스템의 구현

본 논문에서 구현한 Onto-Annotator는 온톨로지 문서의 컨셉들을 분류하여 용어들 간의 관계를 설정해서 tree 구조로 보여주고, text로 직접 쓰거나 웹상의 정보를 그대로 가져와서 annotation함으로써 사용자들에게 보다 정확한 컨텐츠를 제공할 수 있게 하는 annotation-tool로서, Semantic Web을 위한 응용 개발에 매우 유용한 기술이며, 웹을 기반으로 하는 멀티

미디어 정보 서비스 등 다양한 분야에 활용될 수 있다[6].

Onto-Annotator은 앞서 말한 다른 툴들의 기본 컨셉을 기반으로 하여 Java 환경으로 개발되었고, 사용자에게 조금 더 쉽고, 편리한 Annotation과 관리를 제공한다. 먼저, 온톨로지 개발자나 일반 사용자가 누구나 쉽게 사용할 수 있는 온톨로지 개발 도구 등을 이용하여 OWL 온톨로지 파일을 설계·구축하고, 그 파일들의 컨셉들을 추출하여 Web이나 모든 참고자료에서 발췌한 정보들을 드래그 앤 드롭이나 Text 타이핑을 통해 Annotation 할 수 있다. 그리고나서 Onto-Annotator의 RDF SAVE 기능을 통해 유효한 RDF File로 작성되게 된다. 작성된 RDF File들은 테이블로 변환되어 지정된 관계형 DB에 저장되게 되고, crawler를 통해 원하는 정보를 검색, 이용할 수 있다.

```
<rdf:Description rdf:about="file://C:/s/w/s/ow#(063)850-6749">
  <rdf:type>
    <owl:Class rdf:about="http://www.w3.org/WKLab/SWS#Tel">
    </owl:Class>
  </rdf:type>
</rdf:Description>
```

그림 5. Annotation 후 생성된 RDF file source

다음 그림은 본 논문에서 구현한 Onto-Annotator의 인터페이스와 Annotation 후 생성된 RDF file을 보여주고 있다.

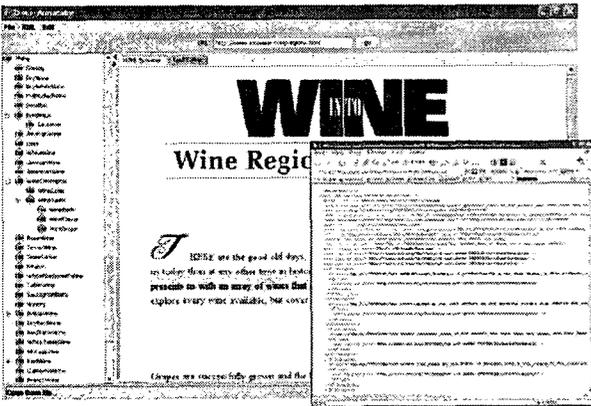


그림 6. Onto-Annotator의 사용 예(wine.owl)

그림의 좌측 프레임에서는 OWL문서를 읽어들이어 Tree 구조로 컨셉들을 나열시켜주고, 상단의 URL을 통하여 인터넷에 접근하여 원하는 정보를 드래그 앤 드롭 방식으로 컨셉의 Annotation을 부여하게 된다. 웹문서가 아닌 일상의 정보들을 Annotation하기 위해 Text 버튼을 두고 사용자 임의의 Annotation이 가능하도록 설계되었다. 그림과 같이 Annotation한 결과는 유효한 RDF file로 그대로 저장되고, 이 파일을 DB에 관계형 테이블을 통해 저장할 수 있도록 하였다. 위의 그림은 wine.owl 을 이용한 Onto-Annotator의 사용 예이다. 한 와인인에 대해 사용자가 웹페이지로부터 주석을 생성하고 저장한 뒤, 시스템에서 사용자가 그 와인을 요청하면 주석 처리된 부

분을 사용하여 더 정확한 의미를 보여줄 수 있고, 이를 저장하여 지식베이스(Knowledge Base)화 한다.

4. 결론

이 논문에서 구현한 Onto-Annotator는 온톨로지 기반의 Annotation System을 이용해 온톨로지 문서나 웹 정보들에 주석을 두고 또 하나의 의미기반 RDF 문서를 생성함으로써, 더욱 정확한 정보를 검색하고 사용자가 편리하게 관리·저장할 수 있도록 하였다. 앞으로의 과제는 이 시스템에 crawler 기능과 merge 기능을 첨가시켜 검색에 따른 자세하고 정확한 정보를 보여주고, 주석처리 된 정보들의 병합과 분석을 통한 중복 및 결핍요소에 대한 처리를 진행할 수 있도록 하며, 온톨로지에 부가한 주석을 다른 사용자가 재사용할 수 있도록 공유성을 극대화 시키는 것이다. 그리하여 주석처리 된 결과물을 가지고 Semantic Web에 응용할 수 있는 시스템을 만드는 게 이 연구의 최종 목표이다[7].

감사의 글

이 논문은 2006년 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임 (전북대학교 헬스케어기술개발사업단).

참고문헌

- [1] Ilia A. Ovsianikov, "Annotation Technology", International Journal of Human-Computer Studies v.50 n.4, 1999.
- [2] Johan Hjelm, Creating the Semantic Web with RDF : Professional Developer's Guide, WILEY Press 2001.
- [3] Siegfried Handschuh, Steffen Staab, Authoring and Annotation of Web Pages in CREAM, In Proceedings of the 11th International World Wide Web Conference, WWW 2002, Honolulu, Hawaii, May 7-11, 2002.
- [4] M. Erdmann, A. Maedche, H.-P. Schnurr, and Steffen Staab. From manual to semi-automatic semantic annotation: About ontology-based text annotation tools. In P. Buitelaar & K. Hasida (eds). Proceedings of the COLING 2000 Workshop on Semantic, An notation and Intelligent Content, Luxembourg, August 2000.
- [5] Jeff Heflin, James Hendler, Semantic Interoperability on the Web, In Proceedings of Extreme Markup Languages 2000, pp.111-120. 2000.
- [6] Stefan Decker, Sergey Melnik, Frank Van Harmelen, Dieter Fensel, Michel Klein, Jeen Broekstra, Michael Erdmann, Ian Horrocks, Teh Semantic Web: The Roles of XML and RDF, IEEE Internet Computing, Vol. 4(5), pp. 63-74, Sept/Oct 2000.
- [7] Noy, Natalya Fridman and Carole D. Hafner. 1997. "The State of the Art in Ontology Design: A Survey and Comparative Review". AI Magazine, Fall 1997.