

온톨로지를 이용한 건축 도면데이터 레포지터리 구성

Construction of Construction Drawing Data Repository using Ontology

이희재(Lee Huijae)*, 유상봉(Yoo Sang Bong)*, 김인한(Kim In Han)**

초 록

W3C는 웹 어플리케이션의 온톨로지 활용을 위한 RDF 표준을 개발해 왔다. 이 논문은 도면 데이터의 온톨로지를 추출하고 적용한다. 설계 파일에 포함 된 도면의 문서정보와 관리 정보를 주 예로 삼는다. 도면 데이터의 관계를 분석함으로써 RDF 스키마를 먼저 디자인 한다. 스키마를 기반으로 온톨로지가 추출 되어 XML 파일로 저장이 된다. 온톨로지 저장을 위한 어플리케이션으로서 재구성 가능한 도면 데이터 레포지터리를 구성할 수 있다. 이 논문의 예제로 사용자는 관심을 가지는 항목의 뷰를 선택할 수 있다 (예를 들어, 프로젝트, 도면, 또는 디자이너). 도면 데이터의 다양한 관점으로 사용자는 특정 데이터에 보다 효율적으로 접근 할 수 있다.

ABSTRACT

W3C has developed the RDF standard for utilizing ontology in Web applications. This paper presents extracting, storing, and applying ontology on product data. The management and document information included in DWG files is focused as an example. By analyzing the relationship among the drawing data, the RDF schema is designed first. Based on the schema ontology is extracted and stored in XML files. As an application of the stored ontology, we can reconfigure the sitemap of drawing data repositories. In this example, the users can select the view that he or she is interested in (e.g., designer, document, project). With such various views of an drawing data repository, the users can access the specific data more effectively.

키워드 : 온톨로지, 메타데이터, 건축 도면 데이터, 사이트맵 재구성

RDF, Ontology, Metadata, Construction Drawing Data, Reconfigurable Sitemap

본 연구는 한국과학재단 목적기초연구(R01-2001-00467)의 연구비 지원에 의해 수행되었음.

* 인하대학교 컴퓨터정보공학과

** 경희대학교 건축공학과

1. 서 론

사용하고자 하는 설계 데이터에 대한 효율적인 검색과 접근은 제품데이터의 활용 위해서 매우 중요하다. 이러한 효율성은 데이터의 검색에 대한 정확성과 접근에 대한 용이성으로 평가될 수 있다.

시스템에 대한 접근 환경 또한 제한이 없어 설계자뿐만 아니라 일반 사용자도 어디에서나 쉽게 원하는 데이터를 검색하고 접근할 수 있어야 한다. 또한 사용자에게 제공되는 데이터는 사용자에게 보다 쉽게 받아들여질 수 있도록 사용자 중심의 정보 구성을 제공해야 한다. 구성데이터에 대한 효율적인 검색과 접근을 제공하기 위해서는 온톨로지의 이용이 필수적이라 할 수 있다. 현재 도면관리시스템에서는 도면명, 제품명, 설계자, 설계날짜, 관련 부서 등의 온톨로지를 이용해 검색에 이용하고 있으며, PDM(Product Data Management) 시스템과 같은 어플리케이션에서도 부품번호, 버전번호, 설계자이름, 승인날짜, 어셈블리구조, 구성데이터 등의 데이터를 가지고 인덱스를 구성해 이용하고 있다. 하지만 이러한 온톨로지에 대한 표현 형식이 모두 달라 시스템의 확장이나 기존의 애플리케이션에서 이미 구성되어 있는 온톨로지에 대한 접근이 용이하지 않다. 또한 레포지터리에 저장된 제품데이터 간의 관계를 기술한 온톨로지가 활용되지 않고 있다. 이러한 온톨로지는 제품의 구성은 물론 설계 의도를 포함하고 있어 지능적인 제품데이터의 활용에 필수적이다.

RDF(Resource Description Framework)는 온톨로지를 웹에서 사용하고 관리하기 위한

표준을 제공한다. RDF는 온톨로지 간의 상호 호환성 제공을 목적으로 W3C에서 제정된 표준으로, 온톨로지의 정의와 저장 및 교환을 위한 표준적인 메커니즘을 제공한다. 특히 온톨로지의 저장과 교환을 위한 포맷으로 XML(Extensible Markup Language) 신택스를 이용함으로써 웹을 통해 쉽게 접근할 수 있고, 서로 다른 시스템간의 정보 교환에 표준적인 데이터 포맷을 제공할 수 있다.

본 연구의 목적은 사용자 중심의 정보 즉, 사용자가 원하는 구조로 구성된 정보를 제공하고 그 서비스를 제공하는 서버에서 구성을 하기 위한 비용을 줄이는데 있다. 이를 위해 정보 서비스의 한 분야인 건축 도면의 설계에 있어서 건축 도면의 설계에 필요한 각 요소들로 온톨로지를 구성하고 온톨로지 정의의 표준인 RDF를 이용하여 추가적인 비용을 줄이면서 제품 온톨로지의 유동적인 구성 방안을 찾는다. 온톨로지 정보를 이용한 그룹 내 자료의 공유를 통해 연구 개발 효율의 향상을 꾀하며, 특히 재구성 가능한 정보 구조를 도입함으로써 원하는 정보를 손쉽게 검색하여 연구 개발 및 전자거래에서 제품정보의 교환을 활성화할 수 있다.

본 논문의 구성은 다음과 같다. 2장에서는 관련 연구에 대해 설명하고 3장에서는 RDF를 이용한 온톨로지에 대해서 설명한다. 4장에서는 온톨로지의 생성 및 재구성 알고리즘에 대해서 설명하고 그 예시를 보여준다. 그리고 5장에서는 시스템 구현 설명 및 예를 보여주고 6장에서 결론을 맺는다.

2. 관련 연구

정보기술에서의 온톨로지는 전자상거래와 같이 지식의 어떤 특정 영역 내에 있는 실체 및 상호작용의 작업 모델을 의미한다. 다시 말해 특정한 도메인 내의 지식을 개념화하고 이를 명세화한 것으로 그 도메인 내에 사용되는 표준 어휘들의 모음이라 할 수 있다. 현재 온톨로지의 연구는 자연어처리와 관련해서 연구 개발 되고 있는데 기존의 다양한 언어 자원으로부터 자연어 처리용 온톨로지를 반자동으로 구축하는 방법을 중심으로 연구되어 한일/일한 기계 번역 시스템에서 어휘 의미 중의성 해소의 효과가 입증된 바 있는 가도카와 시소러스의 개념 분류 체계를 온톨로지의 기반으로 삼아, 여기에서 세종 전자사전과 기계번역 사전, 그리고 대규모 말 뭉치로부터 반자동으로 추출한 다양한 의미 관계들을 추가하여 실용적인 온톨로지를 구축하고 있다[6]. 또한 이에 대규모 온톨로지 구축을 위한 지식 습득의 기본 자료를 국어사전과 백과사전으로 정하고, 국어사전과 백과사전이 가지고 있는 다양한 지식 정보를 온톨로지로 구축하는 방법을 통해서 울산대 NLP 연구소에서는 98년에 10만 여 명사를 대상으로 한국어 명사 의미계층 구조를 자동으로 구축한 바 있으며, 2002년부터 시작하여 한국어 의미망 KSN(Korean Semantic Network)이 구축 중에 있으며, 현재 국어사전과 백과 사전을 이용한 온톨로지를 구축 중에 있다.[1]

RDF는 온톨로지간의 상호 운용성 제공을 목적으로 W3C(World Wide Web Consortium)에서 제정된 표준으로, 서로 다른 시스템간의

정보 교환에 표준적인 데이터 포맷을 제공한다. RDF는 시멘틱 웹과 관련되어 연구가 진행되고 있는데 XML로 표현된 웹 콘텐츠와 이 콘텐츠에 부여된 RDF 메타정보를 관리하기 위한 XML/RDF 콘텐츠 관리 시스템이 개발 중에 있고 정보 통합을 목적으로 RDF를 이용한 웹 온톨로지의 표준화 연구가 진행되고 있다[4]. 또한 전자 상거래에서 각종 서비스 및 보안 응용 프로그램상의 문제를 해결하기 위해 CommerceNet에서 제안한 전자상거래 프레임 워크인 eCo에서 서로 다른 프로토콜 및 시스템간의 상호호환성을 확보하기 위해 비즈니스 웹을 위한 데이터 처리 모델 및 프레임워크 구축과 온톨로지 브로커 시스템에 대한 연구, 그리고 협업 비즈니스를 위한 비즈니스 프로세스 처리 모델에 대한 연구가 진행되었고, 특히 전자 카탈로그와 상품 분류 체계 및 코드를 중심으로 연구가 진행되고 있다[2].

국제 동향으로는 Samizdat와 Sherpa Calender RDF Gateway등이 있는데, Samizdat[15]는 RDF를 기반으로 하여 협업 체계 설정을 위한 엔진을 만드는 프로젝트로, 도서 분류, 등급 편집 및 공동 출판을 지원하며 2003년 12월 라이선스 GPL2 버전 0.5.0을 발표한 바 있다. Sheperpa Calender 프로젝트 [16]은 Sherpa Suite 프로젝트의 한 부분으로 RDF 기반의 연표와 HTML 발표를 Windows에서 누구나 쉽게 논의하고 발표할 수 있도록 하는 연구이고, 2003년 9월에 상업용 베타 트라이얼 버전의 어플리케이션을 발표한바 있다. RDF GateWay[14]는 RDF 데이터베이스, 웹 서버와 어플리케이션을 포함하는

RDF와 같은 데이터의 관리를 MS windows XP/NT/2000 플랫폼에서 관리할 수 있도록 하는 사네틱 웹에 관한 연구 프로젝트이다. 2003년 10월 12일 규칙, 쿼리, N3/N-Triples, 동적 모델, 그래프 병합, 사용자 보안과 HTTP 이벤트 기능을 향상시킨 버전 1.0.3 60 일 한정판 상용 프로그램을 발표하였다.

3. RDF를 이용한 온톨로지

3.1 온톨로지의 스키마 정의

하나의 제품을 기술하기 위한 정보는 여러 개의 하위 목록으로 구분 될 수 있으며, 각각의 하위 목록은 그 자신의 정보를 가지며 그 구성된 정보로부터 메타데이터를 추출하여 온톨로지를 구성할 수 있다. 메타데이터는 정보 데이터를 기술하는 데이터로, 제품 정보의 검색과 관리를 위한 유용하게 사용되는데, 본 논문에서는 건축 도면 설계 문서에서 추출된 정보와, 문서를 등록할 때 부가적으로 추가된 정보들을 바탕으로 생성이 된다. 이 정보들은 경희대학교 졸업논문[8]에서 제안한 건축 도면의 관리 표준을 따르고 있다. 건축 도면 설계 문서에서 추출되는 정보는 문서내의 폰트를 기준으로 추출이 된다. 이렇게 추출된 정보는 온톨로지를 구성하여 RDF로 표현이 된다.

온톨로지를 구성하기 위해서 추출되는 데이터 항목은 스키마를 참조하여 Project Info, Company Info, Designer Info, CD Info, Document Info 정보를 추출하게 된다. 만약

스키마에 해당하는 정보가 존재하지 않는다면, RDF구성에서 해당 스키마는 제외된다. 각 클래스의 정의에 대한 설명은 다음과 같다.

a) Project Info Class : 설계 도면으로 시행되는 공사에 대한 정보를 담고 있다. 포함하는 요소는 다음과 같다.

- JobName : 프로젝트명
- ClassName : 설계분야명
- JobStep : 설계단계명
- Tools : 가공도구
- Designer : 도면 제작자
- ContractData : 계약일
- DeliveryData : 납품일
- Company : 발주처
- Document : 프로젝트의 도면정보
- CD : 도면이 담겨 있는 CD 정보

b) Company Info Class : 공사 계획을 시행하는 발주처에 대한 정보를 담고 있다. 포함하는 요소는 다음과 같다.

- Part : 부서명
- TEL : 전화번호
- Manager : 담당자
- Project : 시행한 프로젝트
- CD Info Class : 제출하기 위한 문서들을 저장하고 있는 CD 정보. 포함하는 요소는 다음과 같다.

- ID : CD 식별자
- Quantity : CD 매수
- Number : CD 번호
- Document : CD에 저장되어 있는 문서 정보

d) Designer Info Class : 도면 설계자에 대

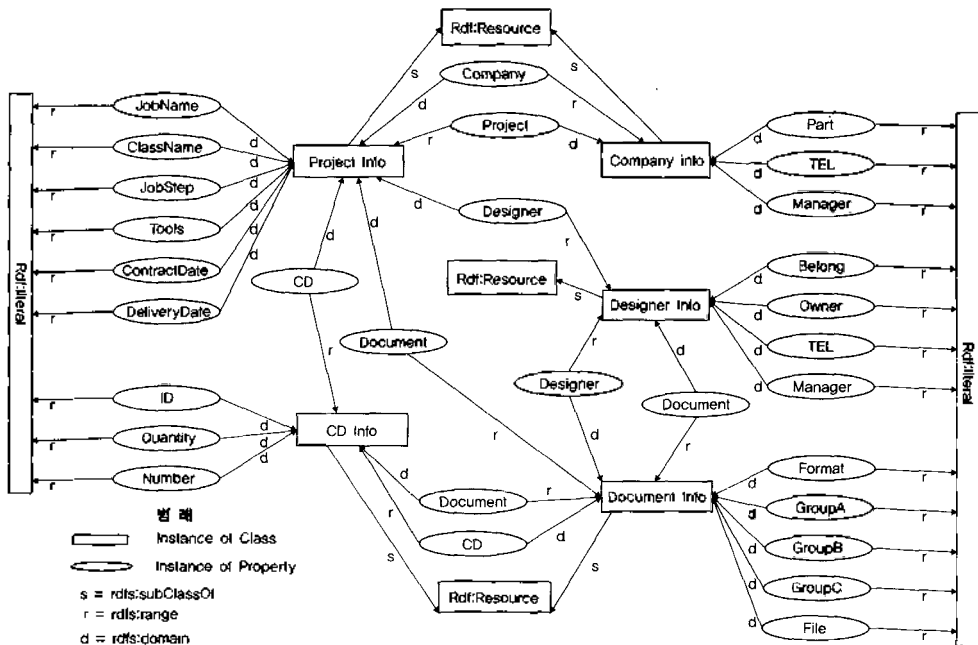
한 정보를 기술하고 있다. 포함하는 요소는 다음과 같다.

- Belong : 소속 회사
 - Owner : 대표자
 - TEL : 전화번호
 - Manager : 담당자
 - Document : 제작한 문서 정보
- e) Document Info Class : 설계 문서의 정보를 포함하고 있다. 포함하는 요소는 다음과 같다.
- Format : 문서의 포맷 형식
 - GroupA : 대분류
 - GroupB : 중분류
 - GroupC : 소분류
 - File : 파일정보
 - CD : 문서를 저장하고 있는 CD 정보

• Designer : 문서의 제작자 정보
 (그림 1)은 각 스키마의 클래스와 관계를 표현하고 있다.

3.2 온톨로지의 생성

스키마를 참조하여 온톨로지를 생성하기 위해서 설계도면 파일을 분석과정이 필요하다. 파일로부터 추출되는 데이터는 파일 정보, 발주처 정보, 계약일, 납품일, 그리고 기타 수치 정보들이 있는데, 이 정보 중에서 파일 정보는 직접 추출 할 수 있는 정보인 반면 형상정보는 파일 본문에 포함되어 있는 내용이기 때문에 이를 추출하기 위해서는 Dr. DWG 라이브러리를 사용하게 된다. Dr. DWG 라이브러리는 DWG 포맷의 분석과 개발을 지원



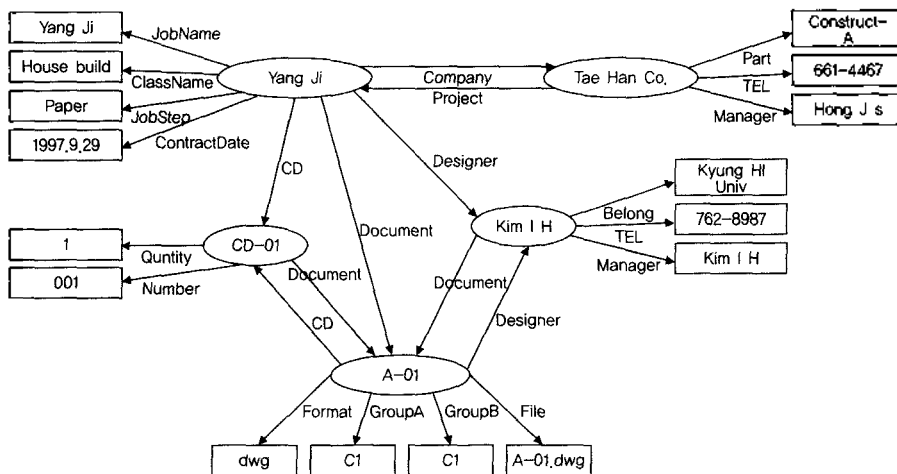
〈그림 1〉 건축 도면의 온톨로지 구조

```

<?xml version="1.0" ?>
- <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:cons="http://islab.inha.ac.kr/~cyberelf/construct-schema#">
- <rdf:Description rdf:about="Yang Ji">
  <cons:JobName>Ynag Ji</cons:JobName>
  <cons:ClassName>House build</cons:ClassName>
  <cons:JobStep>Paper</cons:JobStep>
  <cons:ContractData>1997.9.29</cons:ContractData>
  <cons:company rdf:resource="Tae Han Co." />
  <cons:CD rdf:resource="CD-01" />
  <cons:designer rdf:resource="Kim I H" />
  <cons:document rdf:resource="A-01" />
</rdf:Description>
- <rdf:Description rdf:about="CD-01">
  <cons:Quantity>1</cons:Quantity>
  <cons:Number>001</cons:Number>
  <cons:Document rdf:resource="A-01" />
</rdf:Description>
- <rdf:Description rdf:about="Tae Han Co.">
  <cons:Part>Construct-A</cons:Part>
  <cons:TEL>661-4467</cons:TEL>
  <cons:Manager>Hong I S</cons:Manager>
</rdf:Description>
- <rdf:Description rdf:about="A-01">
  ...
- <rdf:Description rdf:about="Kim I H">
  <cons:belong />
  <cons:manager>Kim I H</cons:manager>
  <cons:TEL>762-8987</cons:TEL>
</rdf:Description>
</rdf:RDF>

```

<그림 2> 건축 도면의 온톨로지를 표현하는 RDF 파일



<그림 3> 온톨로지의 예

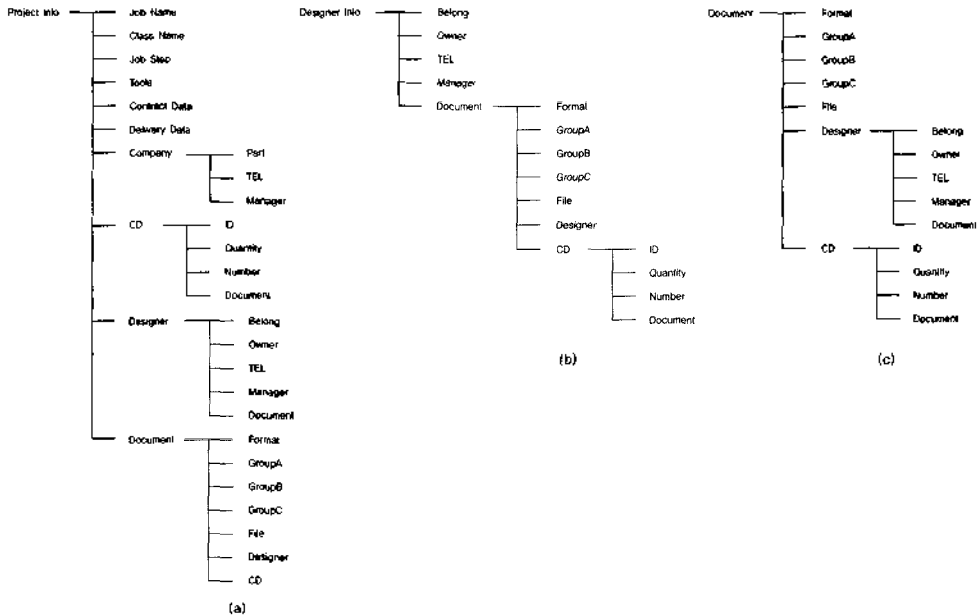
하기 위한 라이브러리로 이 라이브러리를 이용하여 DWG포맷의 파일에 포함된 각 문자열을 추출하는 것이 가능하다. 또한 이 온톨로지는 프로젝트 정보 및 CD 정보와 설계자 정보를 포함하는데 이 정보들은 도면 문서의 등록할 때 등록자가 입력을 해야 하는 부분으로 구성된다. 추출된 문자열 정보 및 사용자 입력 정보를 중심으로 온톨로지가 구성되고 이 정보들은 스키마 정의를 참조하여 RDF 포맷을 생성하게 된다. <그림 2>는 이렇게 생성된 RDF 파일을 나타내고 있다. 그리고 <그림 3>은 RDF로 구성된 온톨로지가 실제로 어떻게 구현이 되어 있는지의 예를 보여주고 있다.

4. 사이트맵의 재구성

4.1 재구성 가능한 정보의 표현 방법

본 연구의 목적은 저장된 온톨로지를 표현할 때 사용자의 목적에 맞게 서로 다른 표현 방법으로 구성하는데 있다. 사용자의 목적에 따라 온톨로지도 사용자에게 맞추어 다르게 표현 되어야 하는데, 예를 들어서 사용자에게 건축 도면 정보를 제공하고자 하는 사이트맵을 구성할 때 사용자의 관심이 건축 공사 프로젝트 일 경우, 제공할 수 있는 서비스는 프로젝트 정보를 포함해서 이 정보와 관련된 발주처 정보, 설계자 정보, CD 정보 및 설계 문서 정보와 그 각각의 정보와 관련된 또 다른 온톨로지 정보들도 연계하여 표현하는 카탈로그를 제공할 수 있다. 한편 사용자의 관심

이 설계자 정보에 있을 때는 도면을 설계한 인명 정보를 중심으로 인명 정보 및 인명 정보와 관련된 문서 정보, CD 정보와 그 각각의 문서 정보 및 CD 정보와 관련된 다른 온톨로지의 정보와 그 관계를 표현 할 수 있는 카탈로그 역시 제공 가능하다. 또한 이미 구성되어 있는 카탈로그의 세부사항 중 하나를 선택하여 그 사항을 중심으로 카탈로그를 재구성 할 수도 있다. 이렇게 함으로써 사용자는 제품 정보에 대해 여러 각도로 접근 가능하기 때문에 정보의 습득 및 그 정보를 이용한 연구 개발의 효율을 높이고 원하는 정보에 빠른 접근이 가능하여 제품데이터의 활용도를 향상 시킬 수 있다. 각각의 사이트맵은 동일한 RDF 파일로부터 추출되어 제공되는데, 이는 각각의 구성을 별도의 파일로 저장하는 방법을 피하여 데이터의 유지 및 관리가 쉽다는 장점이 있고, 각 사이트맵의 구성은 온톨로지의 스키마에 의존하기 때문에 스키마의 적절한 설정에 의해서 다양한 형태로 사이트맵을 재구성 할 수 있다. 사용자에게 제공되는 정보는 정보 재구성 모듈에 의해서 각 목적에 맞는 표현 방법을 구성하여 트리 구조로 재구성하여 표현하게 된다. <그림 4>의 (a), (b), (c) 항목은 서로 다른 관점으로 표현되는 온톨로지 사이트맵의 구성 예를 보여주고 있다. (a)는 Project 정보와 함께 Project 정보를 중심으로 그와 관련된 정보와 그 관계를 제공하며 (b)는 인명 정보를, (c)은 문서 정보를 중심으로 한 정보와 관계를 제공하는 사이트맵의 구성을 보여준다.



〈그림 4〉 사이트맵의 재구성

4.2 온톨로지 재구성 알고리즘

다양한 방법으로 표현하기 위한 온톨로지의 구성은 스키마를 참조한다. 표현하고자 하는 정보는 사용자가 원하는 온톨로지와 함께 관련 있는 다른 온톨로지를 포함하기 때문에, 사용자가 원하는 온톨로지를 루트 노드로 하고 다른 온톨로지를 하위 노드로 재귀적으로 관계를 설정하여 트리를 구성하여 표현하는 방법이 가시적인 효과가 좋다.

온톨로지를 구성하는 위해서는 우선 목적으로 하는 정보와 관련이 있는 엔티티를 검색하여 자료 구조에 저장하고 트리를 구성해야 한다. 트리를 구성하기 위해서는 위치를 결정할 노드와 그 노드의 데이터가 필요하다.

노드를 구성하기 위해서 스키마를 이용한

다. 먼저 root노드를 결정하고 그 하위 노드를 생성하기 위해 root 노드와 관계를 가지는 모든 엔티티를 검색하여 그에 대응되는 엔티티를 찾아 트리의 다음 레벨을 구성한다. 이때 검색된 엔티티는 중복 검색을 방지하기 위해 dataTable 이름을 가지는 테이블에 보관한다. 계속해서, 검색된 엔티티와 관계를 가지는 다른 엔티티가 존재할 경우 다음 검색을 위해서 nextExcTable 이름을 가지는 큐 형식의 자료 구조에 저장을 한다. 큐 형식의 자료구조는 레벨 검색을 통해 가시적 효율이 높은 트리를 구성할 수 있는 장점이 있다. 여기서 생성된 노드를 사용해 RDF 파일로부터 노드의 데이터를 추출하여 해당 노드에 추가한다.

그 후 다음 엔티티를 nextExcTable에서 꺼내어 dataTable에 저장된 엔티티와 중복을 검

사 한다. 이때 중복이 있을 경우 nextExcTable에 있는 다음 엔티티로 진행하고 없을 경우 하위 노드를 구성한다. 이 과정은 nextExcTable에 자료가 남아 있지 않을 때 까지 반복하게 된다.

노드 데이터를 추출하기 위해 RAP-API의 RAPRDF-API[13]을 이용했다. RAPRDF-API는 PHP 환경으로 되어 있기 때문에 웹 서버와 연동하여 프로그램을 하는데 매우 용이하다. RAPRDF-API 라이브러리는 resource에 대해 Object, Subject, predicate값을 검색할 수 있는 find()함수를 제공한다. find()함수는 이 세 요소 중에서 두 인자를 입력으로 했을

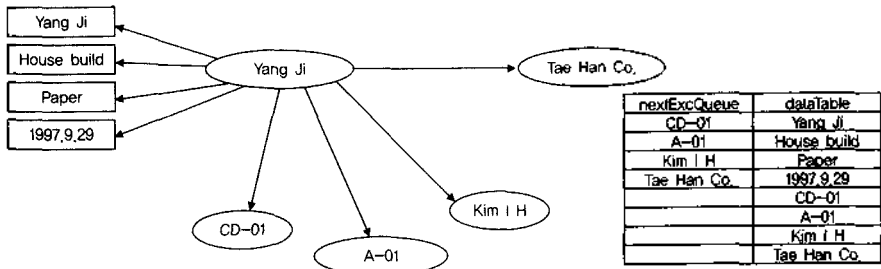
때 나머지 한 요소를 검색해서 반환하는 기능을 제공한다. 해당 노드는 Object에, 네임스페이스는 predicate에 그리고 노드의 정보는 Subject에 각각 해당하게 된다. 따라서 두 데이터 Object와 predicate를 이용하여 RDF로부터 원하는 데이터 값인 Subject값을 추출할 수 있다.

이 과정의 알고리즘은 <그림 5>와 같다. 제품 데이터의 예를 들어, Project Info를 root로 선택할 경우 JobName, Company, ClassName, JobStep, Tools, ContractData, DeliveryData, Designer, Document, 그리고 CD가 엔티티로 추출이 되어 dataTable에

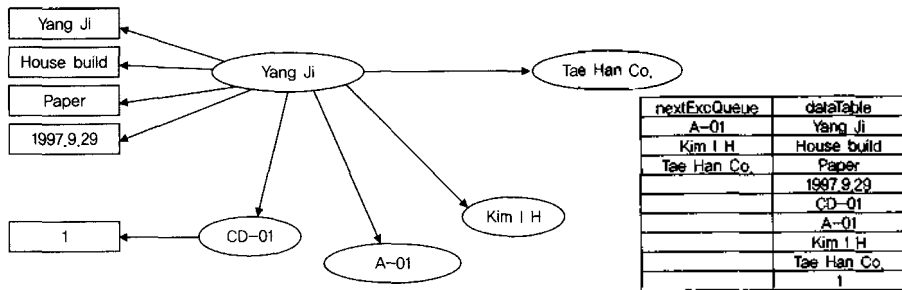
```

Algorithm MakeTree(string)
    Input : 트리 구성 정보 string
    Output : 트리의 루트노드 포인터 treeRootNode
    /* nextExcTable, DataTable : 엔티티 클래스를
    저장하는 자료구조 클래스 Global 변수
    rootNode, nextNode, searchNode, objNode : 엔티티 클래스 변수
    treeNode, treeRootNode : 트리 노드 클래스 변수 */
    rootNode ← string로 구성된 사용자 검색어
    treeRootNode 에 rootNode 값 삽입
    treeNode ← treeRootNode의 포인터
    NextExcTable 초기화
    searchNode ← rootNode
    RDF 문서를 오픈
    do
        for searchNode 와 관계가 있는 엔티티의 개수 do
            objNode ← searchNode와 관계가 있는 엔티티
            if objNode 와 관계가 있는 엔티티가 있다 then
                nextExcTable에 objNode를 Push
                dataTable에 objNode 추가
                object ← objNode 이름
                object 와 네임스페이스로 subject 추출
                subject를 treeNode 클래스로 변환
                treeNode 자식노드에 subject 삽입
            searchNode ← nextExcTable에서 엔티티를 POP
            searchNode와 dataTable 에 있는 엔티티와 중복 검사
            if 중복이 있다. then
                searchNode ← nextExcTable에서 엔티티를 POP
                루트의 처음으로.
            트리에서 searchNode 값을 검색하여 그 노드로 treeNode 포인터 이동
        while (nextExcTable에 자료가 있다.)
    return treeRootNode
    
```

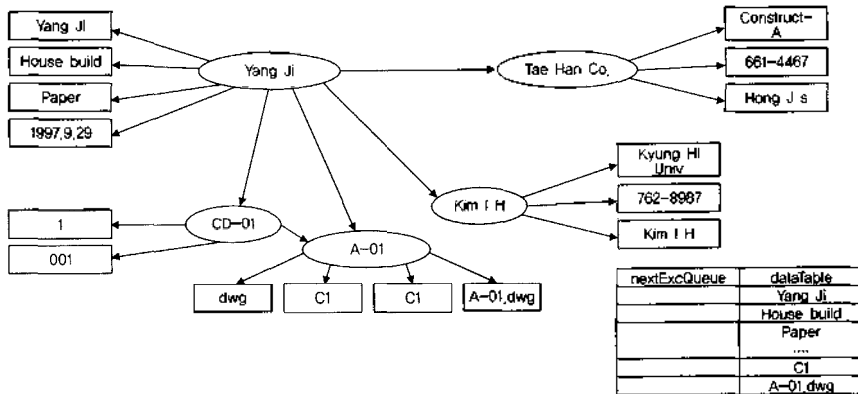
<그림 5> 사이트맵 재구성 알고리즘



〈그림 6〉 트리구성 과정의 예 (1)



〈그림 7〉 트리구성 과정의 예 (2)



〈그림 8〉 트리구성 과정의 예 (3)

저장되고 RAPRDF-API 라이브러리를 통해 RDF로부터 값을 추출하여 트리를 생성

한다. 이때 RDF 데이터에 값이 존재하지 않는 Entity는 무시한다. 이 엔터티 중에서

Designer, Document, CD는 다른 엔티티와 관계를 가지기 때문에 nextExcTable에 저장되어 다음 검색을 위해 준비가 된다(〈그림 6〉). 다음으로 nextExcTable로부터 CD가 팝 연산을 통해 큐로부터 꺼내어 지는데, 이 엔티티를 부모 노드로 하여 다시 검색을 한다. 결과로 CD엔티티와 관계를 가지는 A-01 엔티티가 검색이 되지만, 이 엔티티는 root인 Project Info에서 이미 검색되어진 엔티티이기 때문에 다시 큐에 저장하지 않는다. 이것은 큐 자료구조의 엘리먼트들을 중복검사함으로써 검사가 가능하다(〈그림 7〉).

이런 방법으로 nextExcTable에 저장되는 모든 엔티티를 검색하여 트리를 구성하면 사용자가 원하는 데이터를 root로 가지는 형태의 트리가 생성된다(〈그림 8〉). 재구성된 트리의 노드중 하나를 선택하여 그 노드를 중심으로 새로 트리를 재구성 할 수 있는데, 이를 위해서 선택된 노드가 엔티티인지를 검사하여 엔티티일 경우 그 노드를 Subject로 하여 새로운 트리를 구성한다. 반면 엔티티가 아닐 경우 노드를 Object로 하여 Subject에 해당하는 엔티티를 추출하여 그 엔티티를 중심으로 다시 트리를 재구성 할 수 있다.

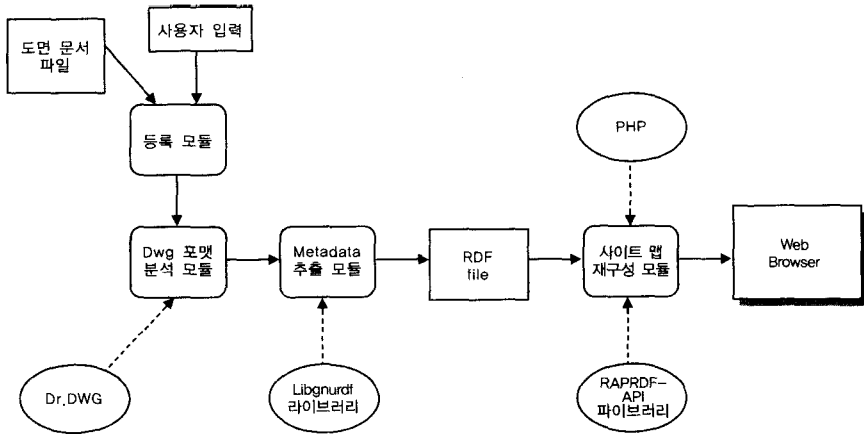
5. 시스템 구현

이 시스템을 구현하기 위해 사용한 도구는 다음과 같다.

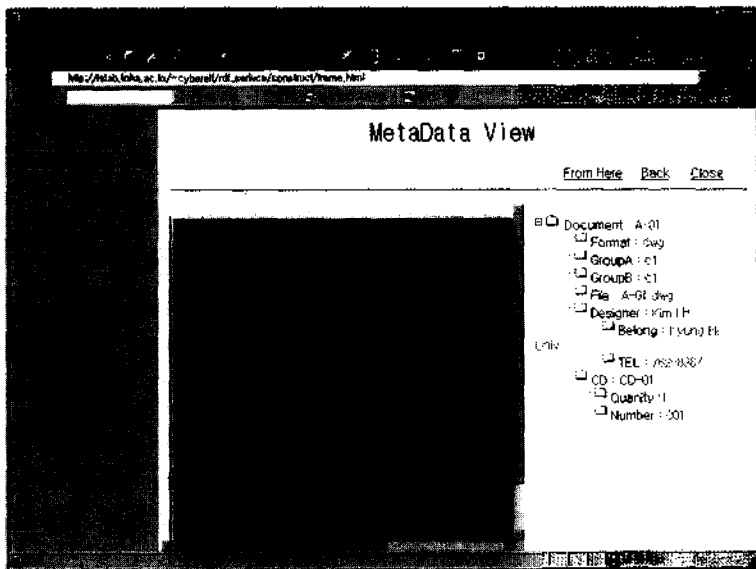
- a) IsaViz : W3C에서 개발된 RDF 생성 및 뷰어로 RDF를 검증하고 테스트하는데 사용

- b) Dr. DWG : ++ 환경으로 되어 있으며 도면 파일인 DWG로부터 각종 정보를 얻어내기 위한 라이브러리
- c) libgnurdf : C 언어 환경으로 되어 있으며 RDF를 생성하거나 RDF에 데이터를 추가하는데 사용
- d) RAPRDF-API : php 환경으로 되어 있으며 RDF에서 값을 추출하는데 사용
- e) 기타 : Linux, php 4.2.3, VisualStudio

시스템은 등록 모듈, 분석 모듈, Metadata 추출 모듈, 정보 재구성 모듈로 구성되어 있다. 등록모듈은 사용자로부터 도면 데이터와 부가 데이터를 입력받는 역할을 하고 STEP-NC 분석 모듈은 ST-developer 라이브러리를 이용해서 등록된 도면 데이터로부터 온톨로지를 생성하기 위한 정보데이터를 추출한다. Metadata 추출 모듈은 DWG 분석모듈이 추출한 데이터와 사용자가 입력한 데이터로부터 온톨로지를 생성한 다음 libgnurdf 라이브러리를 이용해 RDF 파일을 생성하거나 갱신한다. 사이트 재구성 모듈은 PHP라이브러리와 RAPRDF-API 라이브러리를 사용해서 RDF로부터 값을 추출한 다음 표현방법을 구성하고 웹 브라우저에 출력 가능한 데이터를 생성한다. 전체적인 시스템 구성은 〈그림 9〉와 같다. 〈그림 10〉은 실제 예를 보여주는데 사용자가 Project를 중심으로 데이터 검색을 했을 때 서버는 그림의 오른쪽 프레임과 같은 결과를 출력해 준다. 사용자가 메뉴 항목에서 원하는 다른 항목을 선택할 경우, 서버는 사용자가 원하는 구성으로 데이터를 구성한 뒤 사용자에게 제공하게 된다. 사용자가 브라우



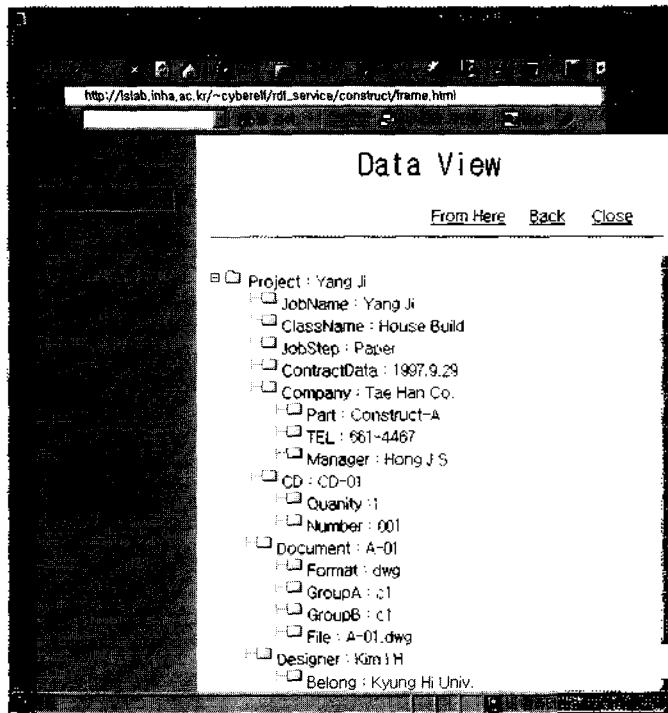
〈그림 9〉 시스템 구성도



〈그림 10〉 도면을 중심으로 한 사이트맵

저에 표시된 내용 중 자세한 것을 보기 원할 경우 해당 데이터를 선택함으로써 해당 사이트로 이동해 자세한 정보를 검색 할 수 있다. 〈그림 11〉은 사용자가 Document에 대한 정보를 선택 했을 때 그 그림을 보여주는 예이다. 만약 사용자가 〈그림 10〉과 같은 정보를

중심으로 RDF 정보를 검색하고자 할 때, 오른쪽 상단 프레임의 From Here 항목을 선택함으로써 다시 재구성된 트리를 보여 줄 수 있다. 만약 그 정보가 하위 트리를 구성할 수 없는 경우 그 트리의 domain 관계에 있는 항목을 중심으로 재구성할 수 있는 From Here



〈그림 11〉 프로젝트를 중심으로 재구성된 사이트맵

메뉴를 사용할 수 있다. 항목을 사용하여 재구성된 뷰는 〈그림 11〉과 같다.

6. 결 론

인터넷의 급격한 확산으로 이용할 수 있는 자원의 양은 비약적으로 증가하고 있다. 또한 이런 자원을 관리하기 위한 온톨로지 역시 비약적으로 증가하고 있는데 이런 온톨로지 정보들을 쉽고 효율적으로 이용하기 위해서 이 온톨로지 정보들을 공통된 포맷으로 구성할 필요가 있고 또한 이렇게 구성된 온톨로지를 검색이 쉽고 효과적으로 접근하기 위한 방법이 요구되고 있다.

본 논문에서는 제품 설계에서 많이 사용하고 있는 DWG¹⁾ 포맷을 이용하여 추출한 메타데이터로 온톨로지를 구성하여 RDF로 저장하는 방법과, RDF로 저장된 온톨로지를 사용자에게 전달하는 방법으로 유동성 있는 사이트맵의 예를 들어서 설명 하였다. 다시 말해서, 온톨로지를 사용하여 사용자에게 제공할 정보를 체계적으로 분류하고 구체화 하였다. 이것은 사용자에게 보다 명확하고 정확한 정보를 제공할 수 있도록 한다. 그리고 RDF와 온톨로지 스키마를 사용해서 사용자에게 제공할 정보를 유동적으로 구성하여 사용자에게 정보를 다양한 각도로 접근할 수 있

1) AutoCAD 어플리케이션의 확장자

도록 하였으며, 서버에서 각 구성을 만드는 데 대한 비용을 절감시킬 수 있도록 하였다.

온톨로지를 이용한 제품의 관리는 사용자가 제품에 대한 다양한 접근을 통해 제품의 유지 보수를 효율적으로 할 수 있기 때문에 보다 높은 이익을 창출할 수 있으며, 제품 데이터의 체계를 사용자가 원하는 관점으로 재구성하여 표현함으로써 제품의 개발이나 관리 면에서 제품의 활용도를 높일 수 있다. 또한 웹 브라우저를 통한 표현은 사용자가 어디서나 제품 데이터에 접속하기 쉬운 인터페이스를 제공한다. 또한, 온톨로지의 저장 포맷을 RDF를 사용함으로써 다른 데이터로의 확장을 용이하게 하고 다른 조직과 그룹 사이에 정보의 호환성을 높일 수가 있다. 재구성할 수 있는 데이터의 구성은 온톨로지에 대한 다양한 접근 방법을 제공함으로써, 데이터의 검색과 접근방법에 효율을 높여 줄 뿐만 아니라 데이터의 유지 및 데이터의 입력에 소모되는 비용을 줄여 줄 수 있게 된다. 본 논문에서는 온톨로지를 추출하는데 DWG 포맷의 도면 데이터만을 사용하였지만, 추출하는 방법을 응용, 확장시켜 다른 분야의 정보를 관리하는 경우에도 활용할 수 있다.

참 고 문 헌

1. 강신재, "자연어 처리를 위한 실용적인 온톨로지 구축", KIPONTO 2003 발표자료, 2003
2. 서울대 이비지니스 기술연구센터, <http://cygnus.snu.ac.kr>
3. 옥철령, "자연어 처리를 위한 온톨로지 - 국어사전과 백과 사전을 이용한 온톨로지 구축", KIPONTO 2003 발표자료, 2003
4. 이강찬, "웹 온톨로지(Web Ontology)의 표준화", 한국정보통신기술협회, 2004
5. 이원희, "시멘틱 웹 기반의 검색 시스템 구조", 워크샵 발표 논문, 부산대학교
6. 정도현, "시멘틱 웹을 위한 온톨로지 언어와 구현사례 연구", 정보관리 연구, vol 34 no.4 pp 87-109, 2003
7. 진연권, "메타데이터를 이용한 본산 STEP 데이터의 웹 인터페이스", 한국 CAD/CAM 학회논문집 제 5권 제 3호 pp231-241, 2000년 9월
8. 최충식, 인트라넷 환경의 설계도서 통합관리시스템 구축에 관한 연구, 경희대학교 석사 졸업 논문, 2001년
9. Dabe Beckett, "RDF/XML Syntax Specification" W3C Proposed Recommendation 15 December 2003
10. Deborah L. McGuinness and Frank van Harmelen eds, OWL Web Ontology Language Overview, W3C Proposed Recommendation 15 Dec 2003
11. Elliotte Rusty Harold, "XML Bible", 정보문화사, 2000년 발행
12. Mark Birbeck, "Professional XML 2nd Edition", 정보문화사, 2002년 발행
13. n@php.<http://navyism.com/main/naboard.php?bd=lecture>
14. RAP - RDF API for PHP v0.7, <http://www.wiwiss.fu-berlin.de/suhl/bizer/rdfapi/>
15. RDF Gateway.<http://www.intellidimension.com/default.jsp?topic=/pages/site/products/rdfgateway.jsp>
16. Samizdat open publishing engine, <http://www.nongnu.org/samizdat/>
17. Sherpa Calendar.<http://www.eventsherpa.com>

저 자 소 개



이희재 (E-mail : cyberelf@islab.inha.ac.kr)
인하대학교 컴퓨터공학과(학사)
현재 인하대학교 컴퓨터정보공학과 석사과정
관심 분야 RDF, DB, Web Programming



유상봉 (E-mail : syoo@inha.ac.kr)
서울대학교 제어제측공학(학사)
아리조나주립대 컴퓨터공학(석사)
피듀대 컴퓨터공학 박사
AT&T Bell 연구소 연구원
현재 인하대학교 컴퓨터공학부 교수
관심 분야 Knowledge & Data Base, EC, 시스템 통합



김인한 (E-mail : ihkim@khu.ac.kr)
1998. 서울대학교 건축학과(학사)
1991. 미국 Carnegie-Mellon대학(석사)
1994. 영국 Strathclyde대 박사
1996~현재 경희대학교 토목건축대학 부교수
관심 분야 CALS Design Database, Digital Design Media