

RDF를 이용한 제품데이터 온톨로지 표현과 활용

이희재*, 유상봉**

Representation and Application of Ontology for Product Data Using RDF

Lee, H. J.* and Yoo, S. B.**

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 design and tooling information included in STEP-NC files is focused as an example. By analyzing the relationship among the product 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 product data repositories. In this example, the users can select the view that he or she is interested in (e.g. product, tool, person). With such various views of an product data repository, the users can access the specific data more effectively.

Key words : RDF, Ontology, Metadata, Product data, Reconfigurable Sitemap

1. 서 론

사용자가 원하는 설계데이터에 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 없이 설계자뿐만 아니라 일반 사용자도 어디에서나 쉽게 원하는 데이터를 검색하고 접근할 수 있어야 한다. 데이터에 대한 효율적인 검색과 접근을 제공하기 위해서는 온톨로지의 이용이 필수적이라고 할 수 있다¹⁾. 현재 도면 관리 시스템에서는 도면명, 제품명, 설계자, 설계날짜, 관련 부서 등의 온톨로지를 이용해 검색의 효율을 높이고 있으며, PDM 시스템과 같은 애플리케이션에서도 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 않을 뿐만 아니라 레포지터리에 저장된 제

품데이터 간의 관계를 기술한 온톨로지 또한 활용되지 않고 있다. 이러한 온톨로지는 제품의 구성은 물론 설계 의도를 포함하고 있어 지능적인 제품데이터의 활용에 필수적이다²⁾.

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 스택스를 이용함으로써 웹을 통해 쉽게 접근할 수 있고, 서로 다른 시스템간의 정보 교환에 표준적인 데이터 포맷을 제공할 수 있다. 또한 RDF는 공개적으로 개발되어 왔기 때문에 누구나 별도의 라이선스를 취득하지 않고 사용할 수 있으며, 별도의 비용을 지불하지 않아도 된다. RDF가 다른 포맷에 비해 가지는 장점은 Table 1에 나타나 있다.

본 연구에 사용되는 데이터는 제품 가공 정보를 기술하는 STEP-NC(STandard for the Exchange of Product Model Data-Numeric Control) 데이터이다³⁾. 연구의 목적은 STEP-NC의 정보로부터 온톨로지 정보를 구축하고 온톨로지 정의 표준인 RDF를 이용하여 추가적인 비용을 줄이면서 제품 온톨로지의 유동적인 구성 방안을 찾는 데 있다. 위에서 언급한 RDF의

*인하대학교 대학원 컴퓨터정보공학과

**중신화원, 인하대학교 컴퓨터공학부

- 논문투고일: 2004. 05. 25

- 심사완료일: 2004. 10. 06

Table 1. 온톨로지 저장 및 관리 방법의 비교

	RDF	데이터베이스	MS content Management Server
저장 방법	XML 문법	자체 포맷	자체 포맷
데이터 공유	공통된 포맷으로 공유가 용이	다른 시스템 사이의 공유에 별도의 비용 요구	다른 시스템 사이의 공유에 별도의 비용 요구
데이터 보존	시스템 의존도가 낮음	시스템 의존도가 높음	시스템 의존도가 높음
정보의 구성	스키마에 따른 동적인 구성	동적인 구성 지원하지 않음	어플리케이션 레벨의 동적 구성
확장성	모듈에 맞는 스키마 제작으로 확장용이	데이터 항목에 대한 테이블 변경으로 확장이 어려움	확장을 위한 모듈 제공
가격	XML 기반의 공개 개발 도구	고가의 어플리케이션	고가의 어플리케이션

상점은 RDF가 연구의 목적을 위한 적절한 사료라고 판단된다. 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 전자 상거래에서 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 전자 상거래와 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. 온톨로지의 연구는 자연어 처리와 관련해서 진행되고 있는데 다양한 자연어 처리용 온톨로지를 반자동으로 구축하여 전자 사전, 기계 번역 사전 등을 이용한 한 일 기계 번역 시스템을 중심으로 활발하게 이루어지고 있다^[1]. 또한, 제조 설계분야에는 PSL(Process Specification Language) 온톨로지 프로젝트 및 NIST(National Institute of Standards and Technology)의 Ontology-Based Manufacturing Standard 등을 들 수 있다^[2]. PSL 온톨로지 프로젝트는 생산 과정 정보를 교환하는데 필요한 용어와 개념을 명확하게 제공하는 것을 목적으로 하는 프로세스를 위한 프로젝트이다. STEP(STandard for the Exchange of Product Model Data) 프로젝트가 상호 호환성을 위해 동일한 도메인 내의 정보에 대해 어휘적인 면에 초점을 맞춘 프로젝

트임에 비해, PSL 프로젝트는 서로 다른 소프트웨어의 환경으로 이루어진 복잡한 생산 과정을 위한 프로젝트이다. NIST 각종 기술 사항 및 표준화 작업을 위해 만들어진 기관으로 Ontology-Based Manufacturing Standard에 관한 프로젝트를 진행하고 있다. 이는 기술 생산과정을 비휴일 머신을 통하여 정보의 공유를 하기 위한 프로젝트로이다.

RDF는 온톨로지간의 상호 운용성 제공을 목적으로 W3C(World Wide Web Consortium)에서 제정된 표준으로, 서로 다른 시스템간의 정보 교환에 표준적인 데이터 포맷을 제공한다^[3]. RDF는 시맨틱 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 전자 상거래 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^[5].

국제 동향으로는 Samizdat와 Sherpa Calender RDF Gateway 등이 있다. Samizdat는 RDF를 기반으로 하여 협업체제 설정을 위한 엔진을 만드는 프로젝트로, 도서 분류, 등급 편집 및 공동 출판을 지원하며 2003년 12월 라이선스 GPL2 버전 0.5.0을 발표한 바 있다^[6]. Sheperpa Calender 프로젝트는 Sherpa Suite 프로젝트의 한 부분으로 RDF 기반의 연표와 HTML 발표를 Windows에서 누구나 쉽게 논의하고 발표할 수 있도록 하는 연구이고, 2003년 9월에 상업용 배타

트라이얼 버전의 어플리케이션을 발표한 바 있다^[11]. RDF GateWay는 RDF 데이터베이스, 웹 서버와 어플리케이션을 포함하는 RDF와 같은 데이터의 관리를 MS windows XP/NT/2000 플랫폼에서 관리할 수 있도록 하는 시멘틱 웹에 관한 연구 프로젝트이다^[10,12]. 2003년 10월 12일 규칙, 질의, N3/N-Triples, 동적 모델, 그래프 병합, 사용자 보안과 HTTP 이벤트 기능을 향상시킨 버전 1.0.3 60일 한정판 상용 프로그램을 발표하였다.

3. RDF를 이용한 온톨로지

3.1 온톨로지의 스키마 정의

하나의 제품을 기술하기 위한 정보는 여러 개의 하위 목록으로 구분될 수 있으며, 각각의 하위 목록은 정보를 가지고 메타데이터를 추출하여 온톨로지를 구성할 수 있다. 메타데이터는 정보 데이터를 기술하는 데이터로, 제품 정보의 검색과 관리를 위해서 유용하게 사용된다.

본 논문에서 메타데이터는 STEP-NC파일에서 추출한 정보와 함께 STEP-NC를 등록하면서 부가적으로 입력해야 하는 정보들을 바탕으로 추출된다. 추출되는 요소 중에는 특히 STEP-NC의 헤더정보와 STEP-NC에 기술된 Working_step 엔티티를 중심으로 추출이 된다.

STEP의 각 엔티티는 라인 번호와 함께 값을 가지는데 값은 분자열이거나 다른 엔티티의 참조이다. 엔티티의 참조는 참조 대상의 엔티티의 라인 번호를 기술함으로써 이루어진다. 예로써, STEP-NC의 엔티티 중 Machining_Workingstep 엔티티는 가공해야 할 형상과 그 가공을 위한 가공 도구 정보를 담고 있는 엔티티를 참조하고 있다. STEP의 데이터는 ST-developer를 이용하여 추출할 수 있다. Fig. 1의 예제에서 Machining_Workingstep 엔티티는 세 번째 인자로 #14라인 번호를, 네 번째 인자로 #24라인번호의 엔티티를 참조하고 있는데, 세 번째 인자는 가공 형상을, 네 번째 인자는 가공 도구정보를 포함한다. Fig. 1은 Round_Hole 가공을 Center_drilling을 통해서 이루어짐을 나타내고 있다. 이렇게 추출된 온톨로지는 RDF(Resource Description Framework)로 표

현이 된다.

온톨로지를 구성하기 위한 단계로서 그 첫 번째는 스키마를 참조하여 Model, Feature, Person, Tool 들을 추출하는 과정이다. 만약 스키마에 위에서 언급된 정보들이 존재하지 않는다면 RDF를 구성할 때 각 스키마들은 추출에서 제외된다. 여기서 추출되기 위한 정보들은 클래스로 분류가 되는데, 각 클래스의 정의에 대한 설명은 다음과 같다.

a) Model Class : 제품 모델에 대한 정보를 기술한다. 제품 모델은 제품을 구성하는 하나의 재체이다. 포함하는 요소는 다음과 같다.

- ▲ material : 제품의 재료정보
- ▲ schema : 제품에 사용된 스키마
- ▲ Image_Name : 제품의 가시적 정보를 포함하는 파일
- ▲ Regist_Info : 모델의 등록정보
- ▲ Designer : 모델 제작자
- ▲ Feature_Info : 모델에 포함된 형상 정보
- ▲ File_Info : 모델이 기술된 파일

b) Person Class : 제품과 관련된 인명에 대한 정보를 담고 있다. 포함하는 요소는 다음과 같다.

- ▲ PersonName : 사람의 이름
- ▲ PersonMail : 사람의 메일 주소
- ▲ PersonOrg : 사람의 소속 기관
- ▲ Persion_ID : 사람의 식별자
- ▲ PersionFile : 사람이 등록한 파일

c) File Class : 제품이 기술된 파일에 대한 정보를 담고 있다. 포함하는 요소는 다음과 같다.

- ▲ File_Name : 파일명
- ▲ File_ID : 파일 식별자
- ▲ File_desc : 파일 간략 설명
- ▲ File_Class : 파일의 종류
- ▲ org : 파일이 만들어진 기관
- ▲ Author : 파일의 제작자

d) Feature Class : 제품에 포함된 형상에 대한 정보를 담고 있다. 포함하는 요소는 다음과 같다.

- ▲ Feature_Works : 형상의 가공방법
- ▲ Feature_ID : 형상의 식별자
- ▲ Feature_bottom : 형상의 바닥상태
- ▲ Feature_Depth : 형상의 가공 깊이
- ▲ FeatureFile : 형상이 포함된 파일
- ▲ ModelInfo : 형상을 포함하는 모델

e) WorkingStep Class : 형상의 가공 정보를 담고 있다. 포함하는 요소는 다음과 같다.

```

-----
#10=MACHINING_WORKINGSTEP Working Step: Round Hole_3.#14.#24.#12;
#14=ROUND_HOLE3 Round hole: #51.#24.#38.#22.#18.#16;
#24=CENTER_DRILLING#36.#34.#13.#1.#51.#32.#30.#28.0.0.0.#26
-----

```

Fig. 1. STEP-NC의 machining_workingstep 엔티티.

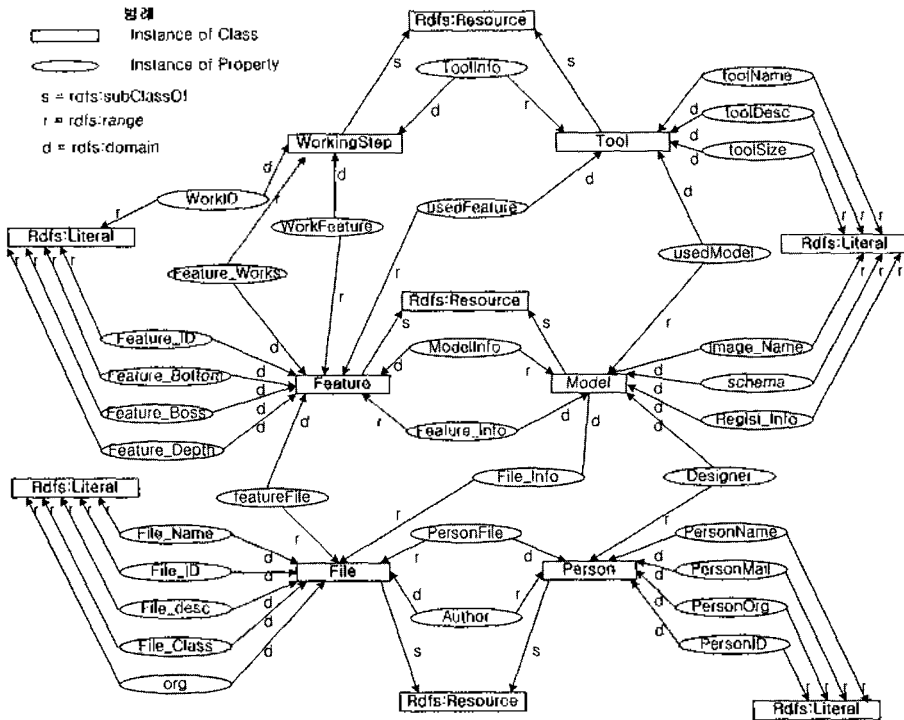


Fig. 2. 온톨로지의 스키마.

- ▲ ToolInfo : 가공 도구에 대한 정보
 - ▲ WorkID : 가공 식별자
 - ▲ WorkFeature : 가공이 포함된 형상
- f) Tool Class 가공 도구에 대한 정보를 담고 있다.
- ▲ toolName : 도구 이름
 - ▲ toolDesc : 도구 설명
 - ▲ toolSize : 도구 크기
 - ▲ usedModel : 도구가 사용된 모델
 - ▲ useFeature : 도구가 사용된 형상

Fig. 2는 각 스키마의 클래스와 관계를 표현하고 있다.

3.2 온톨로지의 생성

스키마를 참조하여 온톨로지를 생성하기 위해서 STEP-NC 파일을 분석하는 과정이 필요하다. STEP-NC 파일로부터 추출되는 데이터는 '소속기관', '제작자', '스키마', 그리고 '형상정보' 정보들이 있는데, 이 정보 중 '소속기관', '제작자', '스키마' 정보는 헤더 파일에 포함되어 있는 정보인 반면 '형상정보' 정보는 파일 본문에 포함되어 있는 내용이기 때문에 이를 추출하기 위해서 STEP-developer 라이브러리를 사용하

게 된다.

STEP은 이기종간의 상호 호환성을 위해 여러 파트로 나누어져 있고 각각의 파트마다 스키마가 존재한다. ST-developer 라이브러리는 STEP 포맷의 분석과 개발을 지원하기 위한 라이브러리로, 기본적으로 AP201, AP202, AP203등의 파트에 대한 스키마를 제공한다. STEP-NC를 분석하기 위해서는 STEP-NC 스키마가 필요한데, STEP-NC 스키마는 ST-developer를 제작한 STEP tools사에서 제공하고 있고, 기타 NC를 다루는 다른 곳에서도 구할 수 있으며 현재도 갱신 중에 있다. ST-developer는 MS Visual Studio의 MFC와 연동하기 때문에 Express 포맷의 스키마를 컴파일 하여 스키마에 포함된 각각의 클래스를 C++ 언어의 클래스로 변환을 해 주는 기능을 제공한다. 변환된 클래스는 클래스에 해당 엔티티가 가지는 정보를 반환하는 멤버 함수를 가지고 있어 값을 추출할 수 있다. 반환되는 정보의 타입은 스키마에 포함된 다른 클래스 타입이거나 C 언어가 제공하는 기본타입이다. 이 클래스를 생성할 때 ST-developer가 제공하는 함수를 통해서 트리 구조의 자료구조를 생성하고 STEP-NC 데이터를 메모리에 적재한다.

제공되는 기본 함수 및 클래스는 다음과 같다.

a) 클래스

- ▲ RoseDesign 클래스 : STEP-NC 디자인에 대한 클래스이다. 현재 편집하고 있는 STEP-NC 문서를 나타낸다. 문서는 멤버함수인 useDesign()을 통해서 사용되는데, 함수의 인자는 문서의 파일명이다. saveDesign() 함수는 RoseDesign을 파일로 저장하며 destroyDesign() 함수는 디자인을 메모리에서 제거한다.
- ▲ RoseDeisgn 클래스 : 이 클래스는 노드 하나를 가리키는 포인터 클래스이다. 멤버함수인 traverse() 함수는 노드를 탐색할 때 이전에 탐색되었던 노드는 중복해서 탐색을 하는 것을 막는 함수이다. domain() 함수는 엔티티의 이름을 인자로 전달 받는데 RoseDesign에서 전달받은 인자를 찾아 주는 역할을 하는 함수이다. 엔티티가 복수인 경우에는 멤버 함수인 next() 함수를 통해서 다음 엔티티로 포인터를 이동시킬 수 있다. domain() 함수와 next() 함수는 RoseObject 타입의 클래스를 반환한다.
- ▲ RoseObject 클래스 : RoseDesignCursor와 같이 RoseDesign의 노드를 가리키는 포인터 클래스이다. RoseDisignCursor가 노드 탐색을 위해 사용되는데 반해 RoseObject 클래스는 가리키고 있는 노드를 가지고 작업을 하기 위한 클래스이다.
- ▲ Entitys : ST-developer에 의해 컴파일된 스키마는 각 엔티티에 대한 클래스를 생성한다. 각 클래스는 각 엔티티 요소가 참조하거나 가지고 있는 값이나 다른 엔티티 클래스를 반환하는 멤버 함수를 가지고 있다.

b) 기본함수

- ST-developer는 STEP을 위한 많은 기본함수를 제공한다. 하지만 STEP 데이터로부터 필요한 데이터를 추출하는 데는 스키마로부터 컴파일된 클래스의 멤버함수를 이용하기 때문에, 실제로 쓰게 되는 기본함수들은 다음과 같다.
- ▲ ROSE_DOMAIN() : 스키마로부터 생성된 엔티티 클래스를 인자로 받으며 그 엔티티의 도메인을 반환하여 다른 함수에서 그 도메인을 사용하거나 검색하게 하는 역할을 한다.
 - ▲ ROSE_CAST() : 엔티티 클래스와 RoseObject 클래스를 인자로 받으며 RoseObject 클래스를 엔티티 클래스로 캐스팅 변환을 해 준다. 변환하는 엔티티는 같은 슈퍼 클래스를 가지거나, 슈퍼클래스와 서브클래스의 관계만 가능하다.

STEP-NC 파일로부터 생성된 메타데이터는 스키마를 참조하여 RDF 모델을 생성한다. Fig. 3은 생성된 RDF 파일을, Fig. 4는 RDF로 구성된 온톨로지의 인스턴스 예를 보여주고 있다. Fig. 4에서 타원은 RDF 스키마에서 정의한 클래스들의 인스턴스에 해당한다. 즉 Model, Feature, Tool, WorkingStep, Person, Filedml 인스턴스이며 각 노드에 대한 id 값은 STEP-NC 파일의 엔티티 이름으로 표현하였다. 그리고 아크 위에 나타나는 요소들이 아크가 시작되는 리소스가 가지는 속성들이며 아크가 끝나는 요소들이 그 속성이 가지는 값이 된다. 속성이 취하는 값이 문자열일 경우는 사각형으로 표시된다.

STEP-NC 파일에는 나타나지 않아 인스턴스 그림에서 표현되지 않는 메타데이터 요소들은 실제 XML 인코딩에서 빈 태그를 사용하였다. 이 정보들은 RDF 파일 포맷으로 저장된다. RDF 파일로 저장을 자동적으로 하기 위해서 RDF 생성모듈이 필요하다. RDF 생성 모듈은 RDF 라이브러리를 사용하여 RDF를 생성하는 프로그램으로 메타데이터 자료를 입력받아 RDF 파일을 출력한다. RDF 라이브러리는 libgnurdf를 사용했는데 libgnurdf는 GNUpdate 에서 개발한 라이브러리로 XML 라이브러리인 libxml2.0을 기반으로 하고 있으며 C 프로그램 언어를 지원한다. XML 라이브러리를 기반으로 하여 XML을 관리하는데 필요한 함수를 사용할 수 있다.

```
<?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:step="http://islab.inha.ac.kr/step-schemas#" >
  <rdf:Description rdf:about="Objects">
    <step:ob_name>Model_R_H</step:ob_name>
    <step:ob_name>O_S</step:ob_name>
  </rdf:Description>
  <rdf:Description rdf:about="Model_R_H">
    <step:file_info rdf:resource="20031213oh.stp"/>
    <step:image_name>sample1.jpg</step:image_name>
    <step:material>steel</step:material>
    <step:designer>Young Tae Hyung</step:designer>
    <step:schema/>
    <step:regist_info/>
    <step:feature_info rdf:resource="R_H_Hole"/>
  </rdf:Description>
  <rdf:Description rdf:about="Milling1">
    <step:tool rdf:resource="MILL_20MM"/>
  </rdf:Description>
  <rdf:Description rdf:about="HOLE1 22MMDrilling1">
    <step:tool rdf:resource="SPIRAL DRILL"/>
  </rdf:Description>
  <rdf:Description rdf:about="O_S_Slot">
    <step:boss>NULL</step:boss>
    <step:bottom>Planar Bottom</step:bottom>
    <step:feature_id>SIDE Slot1</step:feature_id>
    <step:feature_works rdf:resource="SIDE Slot1_Milling1"/>
    <step:depth>50 MM</step:depth>
    <step:featureFile rdf:resource="20030203oh.stp"/>
  </rdf:Description>
</rdf:RDF>
```

Fig. 3. 생성된 RDF파일의 예.

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 한다.

예를 들어서 사용자에게 STEP-NC 제품정보를 제공하고자 하는 사이트맵을 구성할 때를 들 수 있다. 사용자의 관심이 모델정보 검색일 경우, 제공할 수 있는 서비스는 모델 정보를 포함해서 모델 1 정보와 관련된 형상정보, 사용자 정보 및 등록된 STEP-NC 파일 정보와 그 각각의 정보와 관련된 또 다른 온톨로지 정보들도 연계하여 표현하는 카탈로그를 제공할 수 있다. 한편 사용자의 관심이 도구 정보에 있을 때는 가공에 사용되는 도구 정보를 중심으로 도구 정보 및 도구 정보와 관련된 형상 정보 모델 정보와 그 각각의 모델 정보 및 형상 정보와 관련된 다른 온톨로지의 정보와 그 관계를 표현할 수 있는 카탈로그를 제공도 가능하다. 다시 말해서 사용자의 의도대로 정보의 레벨을 결정할 수 있으며, 정보를 정렬하는 것이 가능하다. 이것은 일반적으로 정보를 구성하여 사용자에게 제공하는 Visualization 개념과는 다르게 사용자가 능동적으로 정보를 구성하는 것으로 사용자 중심의 서비스를 제공할 수 있다는 장점이 있다. 사용자에게 유리한 정보 구성은 사용자가 여러 각도에서 정보를 검색하는 것을 가능하도록 하기 때문에 정보의 습득 및 그 정보를 이용한 연구 개발의 효율을 높이고 원하는 정보에 빠른 접근이 가능하여 제품데이터의 활용도를 기존에 비해 향상시킬 수 있는 장점이 있다.

각각의 사이트맵은 하나의 RDF 파일로부터 각각의 작업을 거쳐서 다양한 구성으로 변환되어 제공된다. 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 재구성할 수 있다는 장점이 있다.

Fig. 5는 서로 다른 관점으로 표현되는 온톨로지 사이트맵의 구성 예를 보여주고 있다. Fig. 5의 a는 Model을 중심으로 그와 관련된 정보와 그 관계를 제공하며 b는 도구 정보를, c는 인병정보를 중심으로 한 정보와 관계를 제공하는 사이트맵의 구성을 보여준다.

4.2 온톨로지 재구성 알고리즘

다양한 방법으로 표현하기 위한 온톨로지의 구성은 스키마를 참조한다. 사용자에게 전달되는 정보는 사용자가 원하는 정보와 함께 그 정보와 관련 있는 다른 정보를 포함하기 때문에, 사용자가 원하는 정보를 루트 노드로 하고 다른 정보를 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 이용한다. 먼저 root 노드를 결정하고 그 하위 노드를 생성하기 위해 root 노드와 관계를 가지는 모든 엔티티를 검색하여

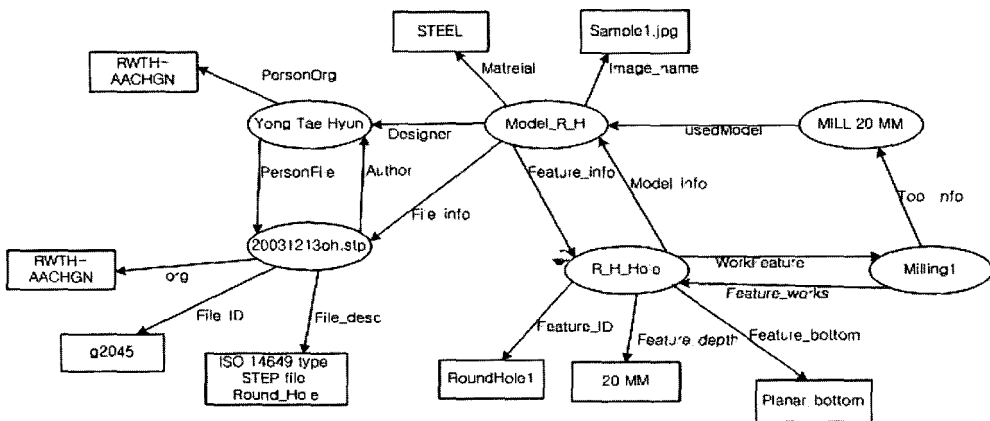
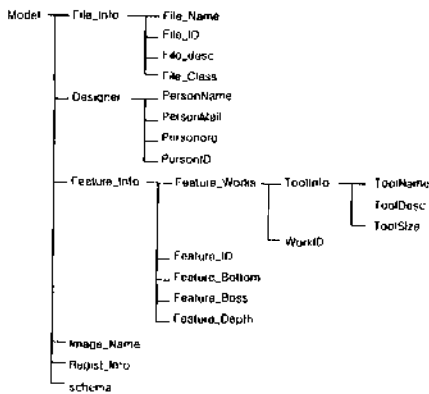
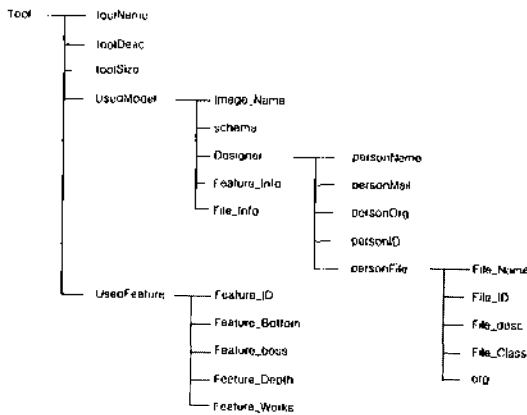


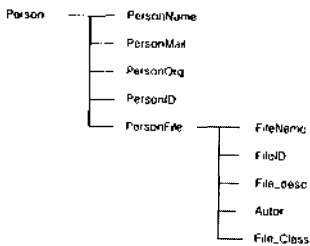
Fig. 4. 온톨로지 인스턴스의 예.



(a) 모델 중심의 구성



(b) 가상 도구 중심의 구성



(c) 인명 중심의 구성

Fig. 5. 목적별 사이트맵의 구성.

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 자료구조에 저장한다. 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 검사 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.

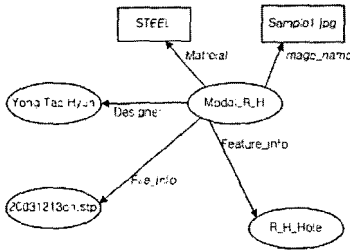
노드 데이터를 추출하기 위해 RAP-API의 RAP-RDF API for PHP를 이용했다. RAP-RDF API for PHP는 PHP 환경으로 되어 있기 때문에 웹 서버와 연동하여 프로그램을 하는데 매우 용이하다. RAP-RDF API for PHP 라이브러리는 resource값에 대해 Object, Subject, predicate값을 검색할 수 있는 find() 함수를 제공한다. find() 함수는 이 세 요소 중에서 두 인자를 입력으로 했을 때 나머지 한 요소를 검색해서 반환하는 기능을 제공한다. 해당 노드는 Object에, 네임스페이스는 predicate에 그리고 노드의 정보는 Subject에 각각 해당하게 된다. 따라서 두 데이터 Object와 predicate를 이용하여 RDF로부터 원하는 데이터 값인 Subject 값을 추출할 수 있다. 이 과정의 알고리즘은 Fig. 6과 같다.

Algorithm MakeTree(string)

```

Input : 트리 구성 정보 string
Output : 트리의 루트노드 포인터 treeRootNode
/* nextExcTable, DataTable : 엔티티 클래스들
   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 의 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
루프의 처음으로.
트리에서 searchNode 값을 검색하여 그 노드로 treeNode 포인터 이동
while (nextExcTable에 자료가 있다.)
return treeRootNode
    
```

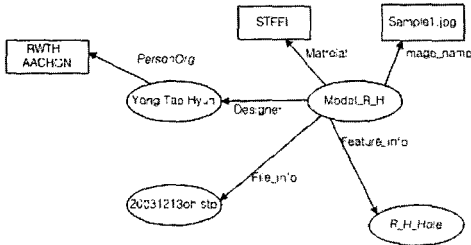
Fig. 6. 재구성 알고리즘.



(a) Model 엔티티 검색

엔티티 저장 자료구조

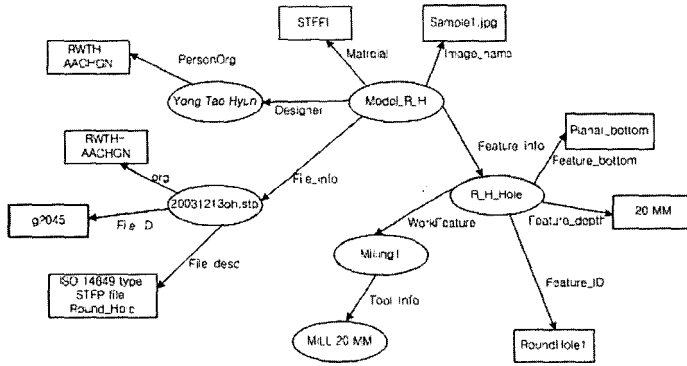
nextExcTable	dataTable
Designer	Image_Name
File_Info	Material
Feature_Info	Designer
	File_Info
	Feature_Info



(b) Designer 엔티티 검색

엔티티 저장 자료구조

nextExcTable	dataTable
File_Info	Image_Name
Feature_Info	Material
	Designer
	File_Info
	Feature_Info
	PersonOrg



(c) 검색 완료

엔티티 저장 자료구조

nextExcTable	dataTable
	Image_Name
	Material
	...
	Feature_Bottom
	Feature_Work
	Tool_Info

Fig. 7. 트리 구성 과정 예.

제품 데이터를 사용하여 작업을 하는 경우를 예를 들어 설명하면 다음과 같다. Model을 root로 선택할 경우 Image_Name, Material, Designer, File_Info, 그리고 Feature_Info가 엔티티로 추출이 되어 dataTable에 저장되고 RAP-RDF API for PHP 라이브러리를 통해 RDF로부터 값을 추출하여 트리를 생성한다. 이 엔티티중에서 Designer, File_Info, Feature_Info는 다른 엔티티와 관계를 가지기 때문에 nextExcTable에 저장이 된다(Fig. 7a). 다음으로 nextExcTable로부터 Designer가 팝 연산을 통해 큐로부터 꺼내어 지는데, 이 엔티티를 부모 노드로 하여 다시 검색을 한다. 결과로 Designer와 관계를 가지는 PersonOrg가 검색 되어 dataTable에 저장이 되고 RDF로부터 데이터를 추출하여 트리에 추가된다. PersonOrg는 다른 엔티티와

관계를 가지지 않으므로 nextExcTable에는 저장이 되지 않는다(Fig. 7b). 이런 방법으로 nextExcTable에 저장되는 모든 엔티티를 검색하여 트리를 구성하면 사용자가 원하는 데이터를 root로 가지는 형태의 트리가 생성된다(Fig. 7c). 이 트리는 정보의 표현 알고리즘에 의해 HTML로 변환되고 추가의 이미지를 삽입하여 사용자에게 제공되게 된다.

5. 시스템 구현

이 시스템을 구현하기 위해 사용한 도구는 다음과 같다.

- a) ST-Developer : STEP tools사에서 개발된 라이브러리로 입력된 STEP-NC 파일을 분석하고 데

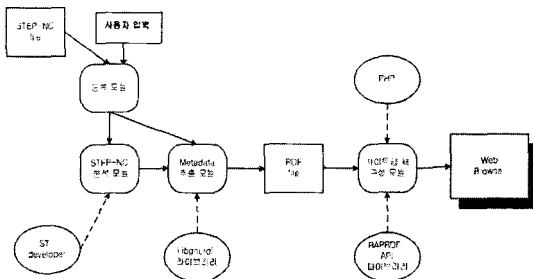


Fig. 8. 시스템 구성도.

이터를 추출하는데 사용

- b) IsaViz : W3C에서 개발된 RDF 생성 및 뷰어로 RDF를 검증하고 테스트하는데 사용
- c) libgnurdf : c 환경으로 되어 있으며 RDF를 생성하거나 RDF에 데이터를 추가하는데 사용
- d) RAP-RDF API for PHP : php 환경으로 되어 있으며 RDF에서 값을 추출하는데 사용
- e) 기타 : Linux, php 4.2.3, VisualStudio

본 논문의 시스템은 등록 모듈, 분석 모듈, Metadata 추출 모듈, 정보 재구성 모듈로 구성 되어 있다. 등록 모듈은 사용자로부터 STEP-NC 데이터와 부가 데이터를 입력받는 역할을 하고 STEP-NC 분석 모듈은 ST-developer 라이브러리를 이용해서 등록된 STEP-NC 데이터로부터 온톨로지를 생성하기 위한 정보 데이터를 추출한다. Metadata 추출 모듈은 STEP-NC 분석모듈이 추출한 데이터와 사용자가 입력한 데이터로부터 온톨로지를 생성한 다음 libgnurdf 라이브러리를 이용해 RDF 파일을 생성하거나 갱신한다. 사이트 재구성 모듈은 PHP 라이브러리와 RAP-RDF API for PHP 라이브러리를 사용해서 RDF로부터 값을 추출한 다음 표현방법을 구성하고 웹 브라우저에 출력 가능한 데이터를 생성한다. 전체적인 시스템 구성은 Fig. 8과 같다. Fig. 9는 실제 예를 보여주는데 사용자가 Model을 중심으로 데이터 검색을 했을 때 서버는 그림의 오른쪽 프레임과 같은 결과를 출력해 준다. 오른쪽 프레임에는 RDF의 정보가 담고 있는 각 메타데이터들의 관계를 표현하고 있다. 각각의 메타데이터는 온톨로지로 구성된 정보들이며 실제 데이터를 이루는 데이터의 부분 정보이다. 사용자가 메뉴 항목에서 원하는 다른 항목을 선택할 경우, 서버는 사용자가 원하는 구성으로 데이터를 구성한 뒤 사용자에게 제공하게 된다. 사용자가 브라우저에 표시된 내용 중 자세한 것을 보기 원할 경우 해당 데이터를 클릭 함으로

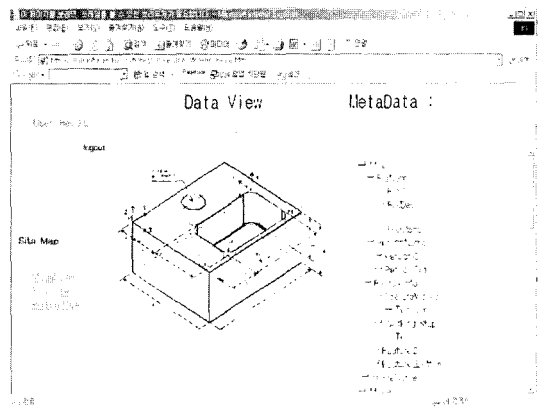


Fig. 9. 모델 중심의 정보 뷰.

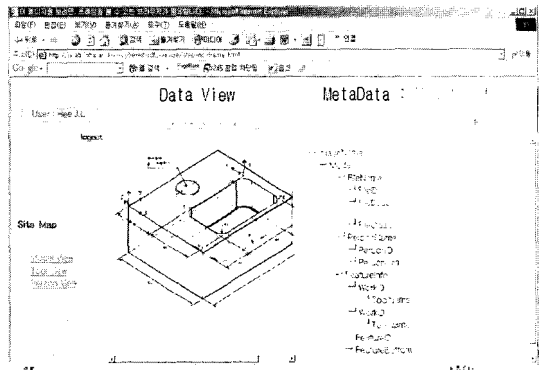


Fig. 10. 재구성된 사이트맵.

써 해당 사이트로 이동해 자세한 정보를 검색 할 수 있다. Fig. 10은 사용자가 Image_Name에 대한 정보를 클릭했을 때 그 그림을 보여주는 예이다. Fig. 10에서 메타데이터들은 사용자의 요구에 맞게 재 정렬 되었다. 사용자가 원하는 정보인 Image_Name 정보의 우선순위가 높기 때문에 Image_Name 정보를 가장 상위 정보로 하여 구성이 되어 있고, 이 정보를 중심으로 데이터들이 검색되어 있다

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 요구

되고 있다.

본 논문에서는 제품 가공 데이터 표준인 STEP-NC 데이터를 이용하여 STEP-NC 데이터에서 추출한 메타데이터로 온톨로지를 구성하여 RDF로 저장하는 방법과, RDF로 저장된 온톨로지를 사용자에게 전달하는 방법으로 유용성 있는 사이트맵의 예를 들어서 설명하였다. 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.

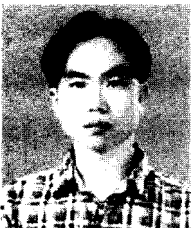
본 논문에서는 온톨로지를 추출하는데 STEP-NC 데이터만을 사용하였지만, 추출하는 방법을 응용, 확장시켜 다른 분야의 정보를 관리하는 경우에도 활용할 수 있다. 본 논문에서 구현한 시스템은 메타데이터를 추출하여 RDF를 생성하고 RDF를 이용하여 재구성 가능한 사이트맵을 제공한다. 이 방법은 일반 DB나 MS Contents Management 제품에 비해 가격이나 호환성면에 장점을 가지고 있다. 이는 제품의 연구, 개발비용을 줄이고 생산 단가를 낮추는 이익을 제공한다.

감사의 글

본 연구는 2003년도 인하대학교 교내 연구비의 지원으로 수행되었음.

참고문헌

1. 강신재, "자연어 처리를 위한 실용적인 온톨로지 구축", KIPONTO 2003 발표자료, 2003.
2. 김동수, "전자상거래 프레임워크 표준화 동향", KT Standardization Trends 제18호, 2003.
3. 서석환, "STEP-NC 기술 소개", 제3차 STEP-NC 기술 산학협동 세미나 자료집, 2001년 9월 6일.
4. 이강찬, "웹 온톨로지(Web Ontology)의 표준화", 한국정보통신기술협회, 2004.
5. 진연관, "메타데이터를 이용한 분산 STEP 데이터의 웹 인터페이스", 한국CAD/CAM 학회논문집, 제5권, 제3호, pp. 231-241, 2000년 9월.
6. 최무라, "제품 정보의 검색에 온톨로지 활용 방법", 한국 CAD/CAM 학회논문집, 제6권, 제4호, 2001년 4월.
7. Craig Schlenoff, Rob Ivester, and Amy Knutilla, "A robust Process ontology for manufacturing systems integration", Proceedings of 2nd International Conference on Engineering Design and Automation, August 1998.
8. Dmitry Borodaenko, "Model for Collaborative Decision Making Based on RDF Reification", <http://www.nongnu.org/samizdat/>, 2004.
9. Graham Klyne, "Resource Description Framework (RDF): Concepts and Abstract Syntax", W3C Recommendation 10 February 2004.
10. Jim Rapoza, "Intellidimension App Shows Power of RDF", eWeek article, April 14, 2003.
11. Libby Miller, "RDFiCal: iCalendar in RDF", ILRT Semantic Web Technical papers, 2003.
12. Shelley Powers, "Practical RDF", O'Reilly, 2003.



이 희 재

2002년 인하대학교 전기전자컴퓨터 공학과 졸업(공학사)
2004년 인하대학교 대학원 컴퓨터정보공학과 졸업(공학석사)
관심분야: XML, RDF, Web DB, IT



유 상 봉

1982년 서울대학교 제어계측공학과 학사
1986년 Arizona 주립대학교 전기 및 컴퓨터공학과 석사
1990년 Purdue 대학교 전기 및 컴퓨터공학과 박사
1989년 AT&T Bell 연구소 연구원
1990년 삼성전자 컴퓨터부문 선임연구원
1992년~현재 인하대학교 컴퓨터공학부 교수
관심분야: 지식공학, STEP, GIS, Web Data처리