

웹 온톨로지를 이용한 공급자와 판매자의 데이터 통합 모델

김영근*, 김수경*, 안기홍*

*한밭대학교 정보통신전문대학원 컴퓨터공학과

e-mail:sosha@nate.com, kimsk@nate.com, khahn@hanbat.ac.kr

The Data integrated model of the Supplier and Retailer using the Web-Ontology

Yeong-Geun Kim*, Su-gyeong Kim*, Kee-Hong Ahn*

*Dept of Computer Engineering, Hanbat National University

요 약

정보통신의 발전은 경제 구조에 많은 변화를 가져왔으며, 향후에도 많은 변화를 예고하고 있다. 가장 대표적인 것은 경제 흐름의 근간이라 할 수 있는 유통의 변화일 것이다. 이것은 통신망 위에 구축되는 가상의 공간에서 이루어지는 유통을 지칭하는 것으로, 빠른 속도로 발전하고 있다. 그러나 무수히 많은 상품들을 각 사이트별 상품 데이터베이스를 구성하는 것은 매우 불합리한 단순 반복 작업이라 하겠다. 이런 문제를 해결하기 위해 시멘틱 웹의 RDF(Resource Description Framework)언어를 이용하여 웹 온톨로지를 구성하여 공급자와 판매자간 상품 데이터를 공유하게 된다면, 전자상거래에 효율성을 가져올 것이라 생각된다. 따라서 본 논문에서는 RDF의 소개와 RDF 웹 온톨로지 편집기를 제작하고, 구성된 RDF를 참조하여 간단한 쇼핑몰을 구성하여 온톨로지를 통한 데이터의 효율적 상호운영성(Interoperability)을 증명하여, 현재 제안되고 있는 상품 카탈로그 분야의 발전에 기여하고자 한다.

Key words : RDF, Semantic web, Ontology, Interoperability

1. 서 론

전자상거래 시장의 성장에 따른 전자상거래 사이트의 구축 기술 또한 많은 발전을 해왔다. 그러나 완성된 전자상거래 사이트에 제품을 등록한다는 것에는 큰 발전을 이루지 못하고 있으며, 단순한 입력 작업을 통해 이루어지고 있다. 전자상거래 사이트의 이러한 데이터 입력 작업에는 다음과 같은 문제가 있다. 첫째, 전자상거래 업체별 상품 데이터베이스 구성으로 인한 비효율성. 둘째, 제품에 대한 전문가가 아닌 비전문가의 왜곡된 정보 제공 가능성 및 이로 인한 소비자, 판매자, 공급자간 분쟁의 소지. 셋째, Off-Line에서 생산자의 권장 소비자 가격 표시와 같은 최소한의 시장 통제력을 행사할 수 없으며 전적으로 판매자에게 의존하는 경향. W3C에서는 시멘틱 웹에 대한 연구가 활발하며, 그 중 웹에서 기계가 이해할 수 있는 정보를 교환하는, 어플리케이션에 상호 운용성이 제공되는 RDF를 제공하고 있다[1]. 이것을 이용하여 위에서 언급하고 있는 데이터베이스의 데이터 중복입력으로 인한 비효율성을 제거 하고, 상호 운영함으로써 문제점을 해결하기 위한 방법을 제

시하고자 한다.

논문 구성은 2장 기술적 고찰, 3장에서는 생산자와 판매자간의 데이터 통합을 위한 MRDIM(Maker and Retailer Data Integrate Model)시스템을 제안하였고, 4장에서는 3장의 제안 시스템 구현하였고, 마지막으로 5장에서는 본 논문에서 결론 및 향후 연구 방향에 대해 기술하였다.

2. 기술적 고찰

2.1 일반적 전자상거래 사이트의 상품 데이터베이스구성 방법
사이트 개발 일정상 데이터베이스 프로그래밍이 완료된 후 상품 명세를 상품 데이터베이스에 입력하는 단순 반복 작업을 하며, 상품 정보의 수정시에도 해당 상품을 등록하고 있는 모든 사이트의 상품 관리자가 동일한 작업을 수행한다. 특별한 경우에는 다음과 같은 경우가 있다. 조달청의 경우는 상품정보 데이터베이스를 공유시스템을 통해 해당사이트와 공유하거나, 이미 만들어진 상용화된 상품 데이터베이스를 사오는 경우도 있으나 양자는 모두 제한적이며, 활용성이 떨어진다.

2.2 전자상거래 표준화 통합포럼(ECIF)의 제안 표준안
 ECIF에서는 2000년 10월 전자 카탈로그 기술위원회 발족을 시작으로 2004년 3월 상호운용성 간담회까지 많은 기술적 논의와 제안을 하고 있다. ECIF의 표준안 중 전자 카탈로그 부분은 다음과 같다. 상품 기술에 대한 표준화, 카탈로그교환 및 전송을 위한 표준, 카탈로그조작을 위한 표준이 있다[3]. 카탈로그 교환 및 전송을 위한 표준은 포맷 표준과 전송 표준이 있으며, 포맷 표준은 상품 정보를 담고 있는 문서 형태에 관한 표준으로, 전자 카탈로그 표준의 궁극적 목적이 상호운용성이라고 볼 때 매우 중요한 것이다. 본 연구는 ECIF제안 중 포맷 표준과 관련이 있다. 2001년 전자 카탈로그 적용지침에 따르면 포맷 표준으로 XML을 사용하고 있다. XML 스펙은 데이터와 표현 부분이 분리되어 네트워크트래픽을 적게 발생시키며, 이식성 등이 뛰어나 이기종간 정보 교환에 우수하기 때문이다. 그러나 당시 제기된 문제는 업종 간 XML 태그에 대한 공통의 합의가 이루어지지 않았다는 것이다. 이런 문제에도 불구하고 IT업체들은 cXML, xCBL, BizTalk같은 XML 프레임 워크를 제시하였다.

2.3 W3C에서 제안하고 있는 시맨틱 웹

W3C에서는 웹에 있는 정보를 기계가 읽을 수는 있지만, 이해할 수 없고, 정보가 방대하여 수작업으로 정보를 이해하고 처리하는 것은 불가능하므로 이를 위해 RDF를 만들었다. RDF의 포괄적 목표는 특정 어플리케이션 영역에 관한 어떤 가정도 하지 않으며, 또 어느 어플리케이션 영역의 의미도 정의하지 않으면서, 자원을 기술하기 위한 방식을 정의한 것이다[5]. 이 방식의 정의는 영역과 중립적이어야 하며, 또 이 방식은 어떤 영역에 관한 정보를 기술하기에도 적절해야 한다. W3C의 목표에서 알 수 있듯이 XML 태그에 대한 합의가 없더라도 차후 의미적 해석을 통해 데이터 태그에 대한 합의를 유도할 수 있으며, 리소스에 대한 공유를 가능하게 해주어 데이터의 상호운영이 가능하도록 하는 명세를 제공하고 있다. 본 논문에서는 W3C에서 제공하고 있는 RDF를 이용하여 전자 카탈로그의 RDF와 어휘를 만들어 실제 제기되고 있는 문제점의 해결을 제시하고자한다.

3. 시스템 제안

상품 데이터베이스 구축에 따른 비효율성이 각 개별 기업 및 국내, 국외의 표준화 단체에 의해 제기되고 있으며 이를 해결하기 위한 다각도의 노력이 이루어지고 있다. 본 연구에서는 표준화된 기술을 이용하여 문제 해결을 위한 시스템을 제안하고자 한다. 상품 카탈로그는 W3C에서 제안하고 있는 상호운영성의 특징을 갖으며, 데이터를 기계가 이해할 수 있도록 규정하고 있는 RDF로 작성하였다. 상품 카탈로그의 상호운용측면은 여러 측면이 있을 수 있으나, 본 논문에서는 공급자 측의 RDF를 판매자 측에서 참조하는 것으로 데이터 중복 입력의 비효율성을 해결하고, 공급자와 판매자간 데이터 연동을 도

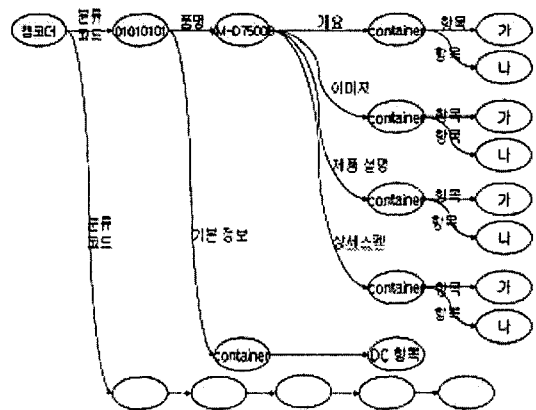
모하여 공급자의 상품 명세 수정시의 불편함을 제거하고자 한다. 공급자 측이 작성한 상품 명세를 직접 판매자가 참조하기 때문에 서론에서 언급한 상품에 대한 전문성, 정보 왜곡문제 또한 해결 되리라고 본다.

3.1 제조업자에 의한 제품 RDF작성

제조업자가 만드는 RDF문서는 연구의 목적을 달성하기에 충분한 예가 될 수 있도록 실제 제품에 대해 모델링을 하였다. 제조업자 1개, 분류는 대분류, 중분류, 소분류로 나누었으며, 제품은 소분류에서 각 두개를 선정하여 4개의 제품에 대해 RDF문서를 작성하였다. RDF문서의 기본 사항은 Dublin Core에서 정의한 요소집합을 사용하였으며, 제품에 대한 일반 사항은 표준으로 정해진 것이 없어 임의로 정하였다. <표 1>은 Index Resource에 포함될 내용을 정리한 것이다. 향후 표준화 작업이 진행되면 표준에 맞는 문서로 수정하여 적용할 수 있다.

<표 1> Index RDF의 내용

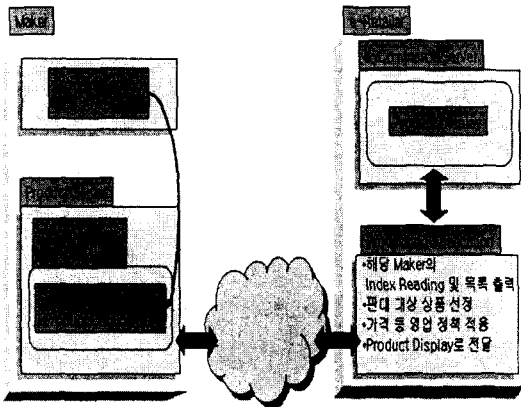
Maker	대분류	중분류	소분류	비고
삼성전자	음향 & 영상	캠코더	VM-D7500B	
			VM-D7500S	
		디지털 카메라	DigiMaxV3	
			DigiMaxV4	



(그림 1) 제품 정보 그래프

(그림1)은 본 논문에서 활용하기 위해 계층 모델을 작성한 것으로 RDF의 트리플 구조로 표현한 것이다. 제품 카탈로그에 표시할 항목을 모델 그래프도 정리하여 각 제품별 디스크립션에서부터 리소스를 정하여 제품이 하나의 리소스를 이룰 수 있도록 설계하였으며, 리소스는 현재 제품 생산자가 웹 사이트에 표시하고 있는 자료를 토대로 어휘를 정의하여 리소스의 내부에 첨가하였다.

3.2 MRDIM(Maker and Retailer Data Integrate Model)



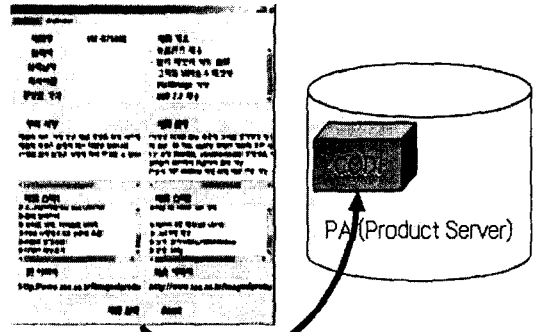
(그림 2) 제조업자와 판매자간의 데이터 통합 모델

(그림 2)은 제조업자와 판매자간의 데이터 통합 모델을 나타낸 것이다. MRDIM모델은 RDF에 익숙하지 않은 제조자가 RDF 데이터를 생성하기 위한 편집기와 카탈로그를 판매자에게 배포하기 위한 PS(Product Server)를 제안 하고 있으며 PS안에는 외부에 공개되는 CODi(Catalog Open Directory)가 있다. 판매자 측에는 PA(Processing Application)이 있다. PA가 하는 일은 첫째, CODi안에 있는 Index를 읽어와 해당 URI에 있는 상품 RDF 목록을 표시한다. 둘째 판매자는 자사 관련 상품을 선별한다. 셋째, 자사의 영업 정책에 따른 가격, 이벤트를 설정한다. 넷째, e-Commerce Server의 제품 진열 항목에 전달한다.

4. 시스템 구현

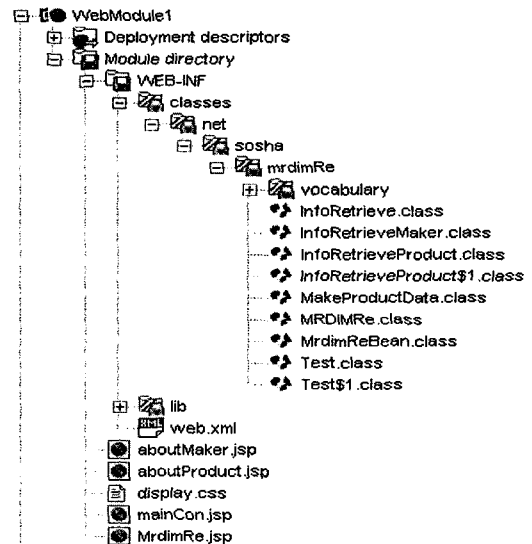
이번 장에서는 시스템의 구현에 대하여 첫째, 제조업자 측의 에디터 제작과 PS의 CODi 구성 및 설치, 둘째, e-Retailer측의 e-Commerce Server PA제작, 셋째, 완성된 MRDIM 시스템에 최종 소비자의 접속 장면을 설명한다. 이와 같은 시스템을 구현하기 위한 기반 기술로는 휴렛팩커드사의 Jena와 썬사의 자바기술을 이용하였다. 최종 소매자의 전자상거래 사이트의 웹 서버로는 아파치재단의 톱캣을 사용하였다.

내부적 모듈은 서블릿을 사용하여 네트워크상의 데이터를 교환하였으며, 내부 수행 알고리즘은 자바빈즈를 사용하여 작성하였다. 제조자 측의 카탈로그 에디터는 3장에서 언급한 <표 1>, (그림 1)에 준하여 입력 양식을 만들어 제조자 측의 상품 카탈로그 작성자가 RDF문서 생성을 인지하지 못하는 상황에서 자연스러운 입력이 가능하도록 제작하였다. 전자 상거래 소매업자 측은 판매하고자 하는 제품의 생산자 사이트의 주소를 입력하는 것만으로 제조업자의 생산 제품을 조회하여 판매하고자 하는 제품을 자신의 전자상거래 사이트에 진열할 수 있도록 하였다.



(그림 3) 에디터와 CODi

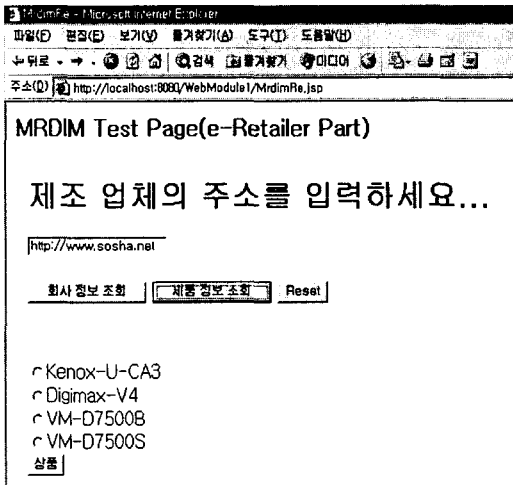
(그림 3)은 에디터에서 RDF지식이 없는 일반부서 직원이 상품 카탈로그를 입력하여 PA의 CODi에 RDF파일을 위치시키는 그림이다. 이처럼 위치된 RDF제품 카탈로그는 소매업자가 접근할 수 있도록 개방적이며, 이곳에는 생산자가 생산하는 제품 정보와 제품을 생산하는 생산자에 관한 기본적인 정보가 들어 있다. 현재는 간단히 회사이름, 사업자 등록번호, 회사 이미지, 대표자, 대표자 이미지의 정보를 가지고 있지만 향후 최종 판매자와 상품 거래 약정, 장려금 수여여부, 각 계약별 거래 특약 정보를 추가하여, 최종판매자의 판매 정책, 가격정책, 제품 프로모션 등에 활용할 수 있으며, 제품 구매 의사 결정에도 활용할 수 있을 것이다.



(그림 4) 웹 모듈의 구성도

(그림 4)는 소매업자 측 모듈구성도로서 상품 전시를 위한 측면은 MrdimRe.jsp, mainCon.jsp, aboutMaker.jsp, aboutProduct.jsp로 구성되어 있다. 어휘는 Maker, Product 두개가 있으며, PA부분의 주요 클래스를 살펴보면,

MRDIMRe에서는 납품 가능 제품 목록을 보여 줄 것인지, 회사소개 부분을 보여 줄 것인지, 상품 카탈로그를 보여 줄 것인지를 결정하여 해당 클래스로 필요한 인자를 넘겨주며, InfoRetrieveMaker는 회사 정보에 관한 데이터를 추출하고, InfoRetrieveProduct는 제품 정보에 대한 데이터를 추출하여 재구성한다. 이 시스템에 의하면 최종 소비자는 제품 제조업자의 정보를 중간 소매업자의 전자상거래 서버를 통해 재구성된 정보를 보게 되는 것이다.

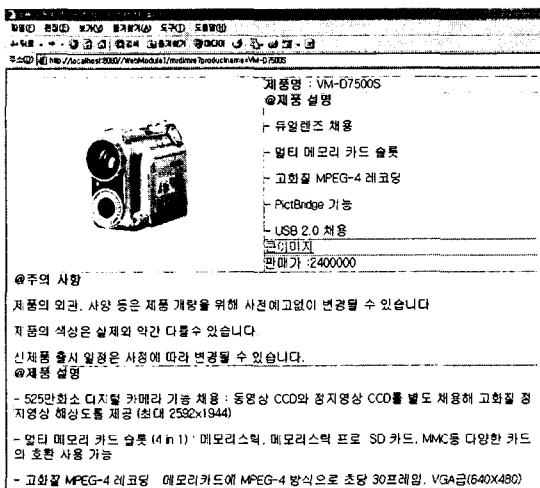


(그림 5) 최종 판매업자(e-Retailer)의 판매 제품 선택

진열 정책에 맞추어 화면을 구성하고, 가격정책에 맞추어 판매가격을 정하게 된다. (그림 6)은 이렇게 선택된 제품을 진열 및 가격 정책에 맞추어 데이터를 재구성하여 표시하고 있는 그림이다. 지금까지 구현된 시스템에서도 알 수 있는 것처럼 데이터는 제조업자에게 있고 데이터를 정보로서 가공하는 모듈은 소매업자에게 있다. 이 시스템은 제조업자가 자신의 공장도 가격을 고정시킬 수 있으며, 제품 스펙을 임의로 수정할 수 있으며, 수정된 정보는 자신의 데이터를 이용하는 소매업자 모두에게 동시에 변경 적용된다. 소매업자는 상품 데이터베이스 구성을 클릭 한번으로 해결할 수 있게 된다.

5. 결론 및 향후 연구 방향

본 논문에서는 첫째, 정보 통신의 발전에 의해 변화하고 있는 유통구조에 대해 언급하였으며, 이중 사이버 공간상에 제품 진열의 문제점에 대해 언급하였다. 둘째, 이를 해결하기 위한 기술적 대안으로 RDF에 대해 고찰하였다. 셋째, 문제 인식과 학문적 고찰을 통하여 해결 방안 제시를 하였다. 넷째, 제시된 해결방안을 구현하여 실증하였다. 판매될 상품에 대한 데이터는 제조업자가 가지고 있어 데이터 통제가 가능하다는 것이고, 최종 소비자는 제조업자가 가지고 있는 데이터를 소매업자가 재생산한 정보로서 보게 된다는 것으로 쇼핑몰 A를 통해 B사의 C 상품을 A쇼핑몰의 판매 조건으로 정보 조회한다는 것이다. 이것은 RDF의 목적과 부합한다고 할 수 있겠다. 향후 제조업자에서 출발한 데이터를, 중간상을 거쳐 최종소비자의 회계장부까지 데이터 통합을 계획하고 있다.



(그림 6) 소매업자의 전자 상거래 사이트에 진열

(그림 5)는 소매업자가 제조업자의 사이트의 CODi에 접근하여 제조업자가 생산하고 있는 제품의 목록을 표시하고 있는 화면으로, 소매업자가 판매하고자 원하는 제품을 선택하면, 제조업자의 데이터와 자신의

참고문헌

[1] Shelly Powers "Practical RDF" 1rd Ed. O'Reilly & Associates, Inc.
 [2] HP Labs SemanticWebResearch . "Jena - A Semantic Web Framework for Java". <http://www.hpl.hp.com/semweb/>, 2003
 [3] "전자 카탈로그 표준화 가이드라인" 한국 전자 거래 협회. 2003
 [4] Sheila A. McIlraith, and David L. Martin, "Bringing Semantic to Web services," Intelligent System, IEEE, vol. 18, no1, pp90-93, 2003
 [5] www.w3.org/RDF