

분산 환경에서 온톨로지 기반의 메타데이터를 이용한 상품 검색 시스템 설계

황치곤^o 김창희 정계동 최영근
 광운대학교 정보통신대학원 정보통신학과
 {duck1052^o, chkim, gdchung, ygchoi}@kw.ac.kr

The design of product search system using the metadata based on ontology at distributed environment

Chi-Gon Hwang^o Chang-Hee Kim Kye-Dong Jung Young-Keun Choi

Dept. of Information communication, Kwangwoon Univ. graduate school of information communication

요약

기존의 이질적인 전자상거래 플랫폼과 시스템을 사용하므로 인해 발생하는 시스템간의 상호 호환 및 운영의 어려움이 있다. 그러나 분산 환경에서 이질적인 특성을 해결하기 위해서 XML을 기반으로 하는 쇼핑몰들의 통합된 정보 검색을 할 수 있는 사이트가 등장하여 각종 쇼핑몰 사이트를 연결하여 통합하는 과정이 진행 중이다. 따라서 상품 검색시 고객에게 메타데이터를 이용하여 선택에 필요한 정보를 제공함으로써 상품을 효율적으로 검색할 수 있다. 본 논문에서는 XML기반으로 분산 환경에서 시스템들의 이질성을 온톨로지 기반의 메타데이터를 이용하여 상품을 검색할 수 있는 시스템을 제안하고, 이 시스템을 보다 효율적으로 검색하기 위하여 일반화 온톨로지, 부품관계지향 온톨로지의 분류방법을 제안하여 분산환경에서 온톨로지 기반의 메타데이터를 이용한 상품 검색 시스템을 설계하였다.

1 장 서론

최근 인터넷 기술의 발전으로 인해 정보량이 급증함으로써 이들 정보자원을 효과적으로 검색하기 위한 방법으로 메타데이터를 이용하여 필요한 정보자원에 정확하게 접근하기 위한 새로운 검색 모형으로 다양한 분야에서 각종 메타데이터가 제안되고 있다. 메타데이터는 정보자원을 효과적으로 검색할 뿐만 아니라 데이터를 재가공 하여 다양한 각종 정보자원에 대한 정보 관리 및 기록 관리를 할 수 있다.[1] 그리고, ISO/IEC 표준에 기반을 둔 메타데이터 레지스트리(Metadata Registry)를 설치하고 새로운 데이터요소나 한정어를 등록 배포함으로써 체계적인 메타데이터를 유통관리 및 통합목록 형태로 구축하여 사용자의 확장검색에서 이용할 수 있다. 실제로 메타데이터란 물리적으로 독립된 형태로 존재하는 정보자원이며 네트워크를 통해 접근할 수 있는 다양한 접근점과 주소를 포함하고 있어 인터넷을 통해 직접 자원에 접근할 수 있다.[2][3] 이러한 데이터들은 분산 환경에서 독립적으로 데이터를 관리함으로써 접근의 제한성, 이질적인 시스템 환경, 구현언어의 다양성 등으로 인하여 데이터 공유 및 통합에 있어서 어려움은 물론 많은 시간과 비용을 투자해야 한다. 따라서 사용자의 요구에 맞게 분산된 데이터를 이용하여 통합적인 결과를 제공하고, 데이터를 공유함으로써 모두에게 편리함과 동시에 효율성을 증가시킬 수 있는 연구가 필요하다.[4][5]

XML은 일반적으로 기업 환경뿐 아니라 웹상에서 데이터 교환을 위한 선도적인 언어로 부각되어 왔다. 이러한 시스템은 서로 간에 포괄적인 접근을 가능하게 해주는 기술로서, 시스템 환경에 제한받지 않고 데이터 교환이 가능하다. 이러한 XML과 메타데이터 기술은 각 시스템으로부터 서로 독립적으로 데이터 교환이 가능하다.[6] 이러한 기술을 기반으로 한 지능적인 상품 검색을 위해 고객이 원하는 때에 원하는 상품 정보를 제공한다는 것은 고객에게 유리 하려고 하는 상품에 대한 정확한 정보를 알아야 한다. 고객의 질문은 대부분 애매모호하므로 질문에 대한 상품의 정확한 정보와 고객이 표현 하지 않는 의미 정보 까지 파악 하여 정보를 검색하는 상품 검색 방법에 관하여 제안 하고자 한다. 따라서 본 논문에서는

- XML기반으로 분산 환경 시스템들의 이질성을 온톨로지 기반 메타데이터를 이용, 상품을 검색할 수 있는 시스템을 제안한다.
- 컴퓨터 상품에 따른 연관관계를 분류한 부품관계지향 온톨로지와 메타데이터 속성을 기반으로 일반화 온톨로지를 제안한다.

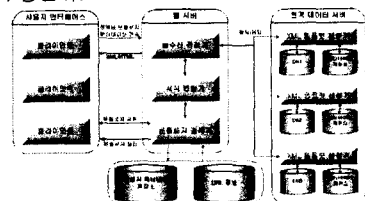
- 온톨로지개념의 연관관계를 이용해 컴퓨터 상품을 효율적이고 쉬운 검색을 위한 방법을 웹에서 사용 가능 하도록 제안한다.

본 논문의 구성은 2장에서 온톨로지 기반의 메타데이터를 이용한 상품검색 시스템을 설명한다. 3장에서 적용 예를 들고, 4장에서 결론 및 향후연구 계획에 대하여 기술한다.

2 장 온톨로지 기반의 메타데이터를 이용한 상품 검색 시스템

2.1 상품 검색 시스템

분산 환경에서 온톨로지 기반의 메타데이터를 이용한 상품 검색 시스템의 구성도는 [그림1]과 같다. 이 시스템은 사용자 인터페이스, 온톨로지 메타데이터 저장소, 웹 서버, 원격 데이터 서버로 구성된다.



[그림1] 시스템 구조도

2.2 사용자 인터페이스

사용자 인터페이스는 클라이언트를 통해 사용자가 질의할 상품에 대한 메타데이터 정보를 온톨로지 검색기에 요청하고, 이 요청된 사항은 온톨로지 메타데이터 저장소를 검색해 클라이언트에 제공해줌으로써 사용자는 그 정보를 이용해 상품을 선택하고, 검색할 수 있게 된다.

2.3 온톨로지 메타데이터 저장소

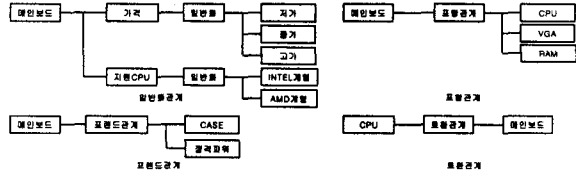
본 논문에서는 상품중 컴퓨터 부품 관계를 기반으로 한 일반화 온톨로지와 부품관계지향 온톨로지 구성된다.

2.3.1 일반화 온톨로지

상품의 본질적인 고유속성은 각 부품을 포함하고 있는 속성을 기반으로 하는 본질적인 고유속성을 세분화하여 개념을 그룹 계층화하여 참고할 수 있도록 하는 것으로 정의한다.

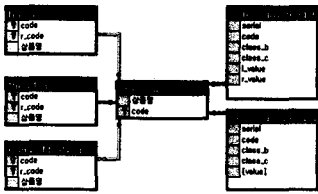
2.3.2 부품관계지향 온톨로지

- (1)포함관계 : 상품의 본질적인 고유속성이 부품이 포함하는 종속관계에 있는 부품군들로 정의한다.
- (2)프렌드관계 : 상품의 본질적인 고유속성이 부품과 동등한 연관 규칙을 가지고 있는 부품군들로 정의한다.
- (3)호환관계 : 상품의 본질적인 고유속성이 부품이 포함하는 부품군으로 포함관계의 반대의 의미로 정의한다. 그러나 그 부품과 호환이 되는 부품군만 설정된다.



[그림2] 온톨로지 분류사례

2.3.3 온톨로지 기반의 메타데이터 구조



[그림3] 온톨로지 기반의 메타데이터 테이블간의 관계

(1)일반화 온톨로지

class_A 테이블은 검색대상 모든 상품을 저장한 테이블이다. setup_grouping_range 테이블은 온톨로지에 대한 속성 정보를 저장한 테이블로 범위 값에 대한 데이터를 저장한다. setup_grouping_value 테이블은 온톨로지에 대한 속성 정보를 저장한 테이블로 나열 값에 대한 데이터를 저장한다.

(2)부품관계 지향 온톨로지

has_a 테이블은 class_A테이블과 포함관계에 있는 상품정보를 저장한 테이블. friend_a 테이블은 class_A테이블과 프렌드관계에 있는 상품정보를 저장한 테이블. compatibility_a 테이블은 class_A테이블과 호환관계에 있는 상품정보를 저장한 테이블.

컴퓨터 상품 검색에서의 입력된 부품에 대한 가격, 제조사, 타입 정도만을 제시하여 고객이 원하는 상품을 간단히 선택할 수 있도록 하고, 그 외에 고객이 구입하고자 하는 상품에 대한 부가적인 정보를 제공하여 고객이 상품을 선택하는데 정확성을 기할 수 있도록 하고 있다.

2.4 웹 서버(Web Server)

2.4.1 송수신 관리자

클라이언트에서 메타데이터 정보를 전송받아, 각 원격 데이터 서버에 전송하기 위해 원격 데이터 서버의 URL 정보를 추가하여 XML 템플릿 생성기로 전송한다.

2.4.2 서식변환기

원격 데이터 서버에서 XML형식으로 전송된 결과를 XSLT를 이용해 HTML형식으로 변환하여 출력하거나, 필요에 따라 XML형식으로도 출력할 수 있다.

[그림4]에서 ①영역은 메인보드에 대한 일반화 결과이고, ② ③은 메인보드와 포함관계에 있는 상품의 결과이고, ④는 프렌드 관계에 있는 상품의 결과이며, XML형식으로 출력한 것이다.

```

1 <?xml version="1.0" encoding="utf-8" ?>
2 <product info="1" >
3   <MAINBOARD 제품명="mainboard" 모델명="intel d865 part" 가격="100000" 제조사="intel" 지향
4   <CPU 제품명="cpu" 모델명="intel pentium4 northwood 2.4" 가격="103000" 제조사="intel" CPU번호
5   <CPU 제품명="cpu" 모델명="intel celeron northwood 2.6" 가격="103000" 제조사="intel" CPU번호
6   <CPU 제품명="cpu" 모델명="intel pentium4 northwood 2.4" 가격="103000" 제조사="intel" CPU번호
7   <CPU 제품명="cpu" 모델명="intel pentium4 northwood 2.4" 가격="103000" 제조사="intel" CPU번호
8   <CPU 제품명="cpu" 모델명="intel pentium4 northwood 2.4" 가격="103000" 제조사="intel" CPU번호
9   <CPU 제품명="cpu" 모델명="intel pentium4 northwood 2.4" 가격="103000" 제조사="intel" CPU번호
10  </product>

```

[그림4] XML결과

2.4.3 온톨로지 검색기

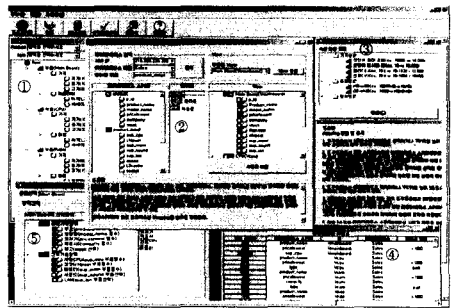
사용자 인터페이스 화면구성을 위한 일반화 항목들을 온톨로지 메타데이터 저장소에서 검색해서 제공한다.

2.5 원격 데이터 서버(RDS : Remote Data Server)

2.5.1 메타데이터 이질성

시스템 관리자가 사용자의 질의가 발생할 때마다 매핑시키는 것은 효율적이지 못하므로 가맹점의 메타데이터 생성은 시스템 구축 초기와 데이터베이스 스키마가 변경될 때 매핑시켜 [그림 1]의 메타데이터 저장소에 저장하여 메타데이터 이질성을 해결했다. [그림5]의 반자동 매핑 생성기는

- ①온톨로지를 기반으로 한 표준 스키마의 메타데이터를 나타내는 표준스키마 원도우
- ②가맹점 데이터베이스의 스키마를 나타내는 쇼핑용 데이터베이스 스키마 원도우
- ③표준 스키마와 데이터베이스 스키마 사이의 이질적인 의미 정보와 문장 정보를 해결하기 위한 속성 이질성 변환 원도우
- ④표준 스키마와 쇼핑용 데이터베이스 스키마 사이의 매핑된 결과를 나타내고 저장하는 매핑 결과 원도우
- ⑤가입 쇼핑물에서 사용자에게 제공할 항목결정을 지원하는 사용자 제공 항목 선택 원도우
- ⑥도움말 원도우로 구성된다.



[그림5] 반자동 매핑 생성기 실행 화면

뷰를 별도 생성하는 이유는 첫째, 사용자의 접근이 분산되게 하고, 둘째, 뷰는 가상테이블로 데이터베이스에서 변경되는 사항이 곧바로 갱신되므로 효율적인 쇼핑물 통제가 가능해지고, 셋째, 검색속도를 향상시킬 수 있으므로 데이터베이스 전체를 검색하는 방법보다는 뷰를 이용하는 것이 유리하다.

2.5.2 XML 템플릿 생성기

```

1 select * from mainboard
2 where mainboard.price < 100000
3 and left(mainboard.chipset, 3) = 'intel'
4
5
6
7
8
9
10

```

[그림6] 메타데이터 저장소에 의해 생성된 XML 템플릿

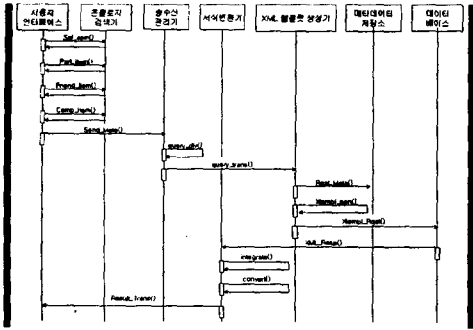
본 논문에서 각 구성요소간의 데이터 이동을 XML 템플릿으로 하고, 이는 XML 템플릿 생성기에서 자동으로 [그림6]과 같이 생성된다. 이는 [그림8]의 사용자 인터페이스 검색내용으로

생성된 템플릿이다. 각 부분을 살펴보면, ①은 XML 선언부이다. ②는 <product>요소에 xml-sql 네임스페이스 선언부이다. ③은 질의의 시작과 끝을 나타낸다. ④는 메인보드 뷰인 vmainboard를 검색하는 질의이다. ⑤의 'for xml auto'문장은 MSSQL의 검색결과가 XML문서로 출력하도록 하는 모드이다.

이 SELECT문의 항목들은 사용자 인터페이스에서 결정되는 것이 아니라 앞에서 나온 반자동 매핑 생성기에서 가맹점 시스템 관리자들에 의해 사용자 제공 항목 선택 윈도우에서 결정되어진 항목들로 이 항목들이 포함되는 시기는 XML 템플릿을 생성할 때이다.

2.6 상품 검색을 위한 시퀀스 다이어그램

시퀀스 다이어그램은 [그림7]과 같은 흐름을 따른다.

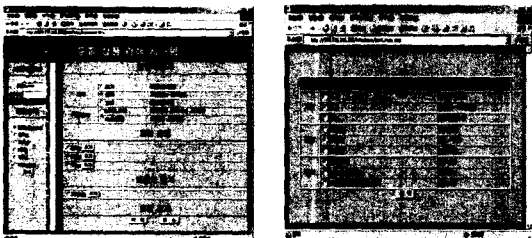


[그림7] 상품검색을 위한 시퀀스 다이어그램

- Sel_Litem(), Part_item(), Friend_item(), Comp_item(): 사용자 인터페이스에서 선택한 부품에 대한 표준 메타데이터를 온톨로지 검색기에 요청하여 사용자 인터페이스 구성.
- Send_meta(): 사용자에 의해 선택된 메타데이터 정보를 송수신 관리기로 전송.
- query_div(): 송수신 관리기로 전송된 메타데이터 정보에 각 가맹점 정보를 추가하여 분산 환경의 각 시스템의 특성에 적합하도록 질의를 분할 생성.
- query_trans(): 분산 환경의 각 가맹점의 시스템에 적합하게 분할된 질의를 XML 템플릿 생성기로 전송하여 템플릿을 생성.
- Rqst_meta(): 가맹점에 매핑되어 저장된 메타데이터 저장소에서 템플릿을 생성하기 위한 항목과 사용자 인터페이스에 출력해야 하는 항목 요구.
- Xtempl_gen(): Rqst_meta()를 통해 전송된 메타데이터로 XML 템플릿 생성.
- XML_Resp(): 검색결과 XML문서를 서식 변환기로 전송.
- Integrate(): 분산 환경의 각 시스템에서 전송된 XML결과 통합.
- convert(): XML 문서를 XSLT로 적용시켜 HTML로 변환.
- Result_Trans(): 통합된 결과 사용자 인터페이스로 전송.

3 적용예

본 논문에서 제안한 온톨로지 기반의 메타데이터를 이용하여 사용자가 원하는 정보가 상품에 대한 연관관계와 도메인의 연관관계를 가지고 있는 각종 정보를 제공할 수 있다.

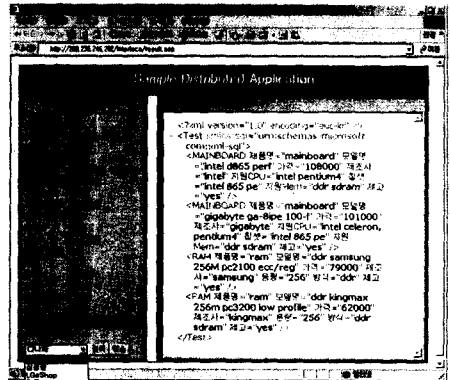


적용예 1 적용예 2
[그림8] 사용자 인터페이스

시스템 적용 환경으로 운영체제는 Windows 2000 server, 데이터베이스 시스템은 MSSQL 2000 server, 언어는 ASP, Visual BASIC, XML을 사용했다.

적용예 1: 클라이언트에서 사용자 ID, 비밀번호, 상품선택, 일반화관계를 선택하고, 포함관계-프렌드관계-호환관계를 추가하여 검색 한다.

적용예 2: 포함관계 상품인 CPU에 대한 검색조건 선택한다.
결과: 조회된 결과



[그림9] 결과

4장 결론 및 향후 연구 계획

본 논문에서는 상품간의 연관 관계에 따른 분류로 부품관계 지향 온톨로지와 메타데이터의 속성을 기반으로 하는 일반화 온톨로지를 제안했고, XML기반으로 분산된 이기종의 시스템들간의 메타데이터를 이용하여 부품 정보보다 효율적이고 쉽게 검색하기 위한 방법을 웹에서 사용 가능하도록 제안했다.

향후 연구는 컴퓨터 상품에 따른 부품관계 뿐만 아니라 여러 상품군으로 확장하여 일반화할 필요성이 있다. 또한, 다중 에이전트를 기반으로 다양한 상품군에 따른 사용자 인터페이스를 자동생성하기 위해 확장된 입출력 인터페이스 에이전트 설계와 학습에 의해 온톨로지를 인공지능으로 갱신하기 위한 수평 에이전트를 더욱 보강하여야 하고, 상품 검색 개인 프로파일 저장으로 관심품목에 대한 추천 에이전트 추가가 요구된다.

참고문헌

- [1] W3C, Resource Description Framework(RDF), <http://www.w3.org/RDF>, 1999.8
- [2] ISO/IEC 11179-1 ISO/IEC 11179-1, "Information technology Part 1 : Framework for the specification and standardization of data elements", 1999.12.
- [3] ISO/IEC 11179-3, "Information technology Part 3: Registry, metamodel and basic attributes Second edition", 2003.3.
- [4] G. Gardarin, A. Mensch, T. Tuyet Dang-Ngoc, L. Smit, "Integrating Heterogeneous Data Sources with XML and XQuery.", Proceedings of 2002 13th IEE(DEXA'02) Workshop, Page 839-846, 2002
- [5] 백두권, 최요한, 박성공, 이정욱, 정동원 "MDR과 Ontology를 결합한 3계층 정보통합 시스템" 정보처리학회 논문지 D 제10-D권 제2호 2003.4
- [6] W3C(World Wide Web Consotium),"W3C Recommendation:Extensible Markup Language (XML)1.0 (Second Edition)", <http://www.w3.org/TR/REC-xml>, 2000.10.