

7. Scheme을 이용한 XML 기반의 데이터 관리 웹 에어전트 시스템 설계 및 구현

전자통신공학과 유 선 영
지도교수 임 재 흥

인터넷이 널리 보급됨에 따라 지식과 정보 교류의 기반이 웹(web)으로 옮겨지고, 일반인에게 익숙한 웹에서 정보를 추출하거나 가공하고, 정보를 공유하거나 교환하려는 요구가 증가하는 추세이다. 현재 웹 문서를 작성하는 하이퍼텍스트 생성 언어(HTML : HyperText Markup Language)는 초기 출판을 목적으로 탄생된 생성 언어로 웹 브라우저(browser) 화면에서 보여준 문서의 모양을 정의하는 표현 위주의 언어이다.

HTML은 정보 축적과 추출하는 방법이 비효율적이고, 추출한 정보의 가공이 어렵다. 이러한 HTML의 단점을 보완할 수 있는 새로운 형태의 언어가 월드 와이드 웹 컨소시엄(W3C ; World Wide Web Consortium)에서 1996년 웹 문서의 표준으로 표준 범용 생성 언어(SGML ; Standard Generalized Markup Language)를 기반으로 하는 확장성 생성 언어(XML ; eXtensible Markup Language)를 제안되었다.

XML은 문서 내용에 대한 구조와 의미를 기술하고 문서 하나를 여러 형태로 보여줄 수 있으며, 콘텐츠(contents)를 기반으로 데이터를 필터링(filtering)하거나 응용 목적에 맞게 재구성할 수 있다. XML이 단순한 콘텐츠에서 데이터베이스로까지 그 적용 분야가 확장되면서 XML로 표현된 정보를 효율적으로 저장하고 관리하는 방법에 대한 관심이 증대되고 있다. 오늘날 인터넷이 확산되면서 웹을 이용하여 정보를 검색하거나 인터넷을 통한 구매 활동이 활발해지면서 기업은 웹을 이용하여 데이터베이스를 관리한다.

이에 따라 웹에서 데이터베이스를 관리하거나 데이터베이스에 저장되어 있는 내용을 웹 문서로 표현하고, 서로 다른 데이터베이스에 저장되어 있는 내용을 웹에서 교환할 수 있는 시스템에 대한 연구되고 있다. 기존의 데이터베이스 관리 시스템(DBMS ; DataBase Management System)에서 저장된 데이터를 웹에서 교환하기 위해서는 표현 위주의 HTML 보다는 문서 구조에 대한 정보를 가지고 있는 XML이 더 적합하다. 그러나 XML은 구조상 기존 데이터베이스와 많은 차이점이 있기 때문에 XML 문서와 데이터베이스 사이의 매핑(mapping)이 중요하다.

따라서 XML을 매개체로 하여 데이터베이스에 저장된 데이터를 XML로 변환시키고, 변환된 XML 데이터를 웹에서 효율적으로 정보를 관리하는 방법과 XML 데이터를 데이터베이스에 저장하기 위한 매핑 방법, XML 데이터를 데이터베이스에 저장하여 데이터베이스에서 데이터를 관리하기 위해 사용하는 기술을 적용하여 XML을 관리하는 방법에 대한 연구와 웹에

서 데이터를 쉽게 검색하고 가공할 수 있는 시스템이 필요하다.

본 논문은 데이터베이스에 저장된 데이터를 XML 문서로 변환하고 XML 문서를 데이터베이스에 저장하는 규칙을 제안하여 웹에서 데이터베이스와 XML 사이의 데이터를 교환하고 관리하는 웹 에이전트 시스템을 구현하였다. 구현한 웹 에이전트 시스템은 데이터베이스에 저장된 데이터를 XML로 변환하거나 XML을 데이터베이스에 저장하거나 웹을 이용하여 초보자도 쉽게 데이터베이스에 접근하여 데이터베이스의 종류에 상관없이 데이터베이스를 관리할 수 있고, 저장된 데이터를 스킴(Scheme) 언어의 리스트로 표현하여 사용자에게 익숙한 웹 브라우저로 정보 추출 또는 검색을 할 수 있도록 구현하였다.

현재 XML과 연동하여 데이터베이스를 관리할 수 있는 상용 제품으로 펜타 시스템의 타미노(Tamino)와 오라클 제품이 있다. XML 문서를 XML 포맷으로 저장하고 관리하는 타미노는 클라이언트와 서버간의 연결을 위하여 별도의 소프트웨어와 라이선스를 구매해야 하고, 데이터베이스 데이터를 교환할 수 있는 오라클 제품은 오라클에서 제작한 XSU(XML SQL Utility)를 가진 오라클 데이터베이스와 다른 데이터베이스간의 정보 교환만 가능하다. 그러나 구현한 웹 에이전트 시스템의 차이는 별도의 소프트웨어가 없이 인터넷에 연결되어 있는 웹 브라우저를 사용한다는 점이 타미노와 다르고 오라클 데이터베이스에 저장된 데이터만을 다른 데이터베이스로 전달할 수는 있는 오라클 제품과는 달리 웹 서버가 가지고 있는 JDBC의 드라이버에 의해 연결 가능한 데이터베이스가 결정되므로 모든 데이터베이스에 대한 JDBC가 개발된다면 웹 서버는 모든 데이터베이스간의 정보 교환이 가능하게 된다. 또한, 스킴을 이용하여 데이터베이스를 관리하고 사용자가 원하는 다양한 검색 조건을 데이터베이스 질의문이 아닌 웹 브라우저 화면에서 조건을 입력하거나 선택하여 다양한 데이터 처리를 쉽게 할 수 있다. 또한 서로 다른 데이터베이스간이 데이터 교환을 하기 위해 XML을 매개체로 하도록 만든 본 논문의 웹 에이전트 시스템은 데이터베이스에 저장된 데이터 전체를 복제하거나 필요한 데이터만 선택적으로 복제할 수 있어 현재 상용화되고 있는 제품과 차별화된다.