

# EJB 컴포넌트 기반으로 한 XML Repository 시스템 설계

정상혁<sup>0</sup>, 이정수, 주경수

순천향대학교 전산학과

grotest@hanmail.net, jungsoc6@hitel.net, gssoojoo@asan.sch.ac.kr

## Designing XML Repository System based on EJB Components

Sang-Hyuk Jung<sup>0</sup>, Jung-Soo Lee, Kyung-Soo Joo

Dept. of Computer Science, College of Engineering Soonchunhyang Universit

### 요 약

웹 정보의 교환과 공유를 위해 XML 문서의 필요성이 증가함에 따라 XML 문서의 저장과 검색에 대한 많은 연구들이 현재 진행되고 있다. 본 논문에서는 XML 문서를 효과적으로 저장, 검색하는 XML Repository 시스템을 설계한다. XML Repository 시스템은 XML 문서를 관계형 데이터베이스에 저장하고, 관계형 데이터베이스에서 검색하여 XML 문서로 생성할 수 있다. 또한 XML 응용 시스템과 관계형 데이터베이스 시스템 사이에 원활한 연계를 위해서, XML DTD를 관계형 데이터베이스 스키마로의 변환이 가능하다. 본 논문에서는 EJB 컴포넌트를 이용하여 XML 문서를 저장, 검색, 변환 할 수 있는 XML Repository 시스템을 설계하고자 한다.

### 1. 서 론

XML은 사용하는 플랫폼에 독립적이며, 문서 정보의 전송과 교환이 쉽고, 문서의 풍부한 의미를 그대로 나타낼 수 있는 장점을 지닌 웹 표준 언어이다. XML은 시스템 독립적인 포맷으로 데이터를 전송하기 위한 공통 전송 기술로 사용될 수 있다. 웹에서의 자유로운 데이터 교환을 위해, XML은 문서 자체에 문서의 구조를 기술하고 있으며, 사용자가 문서의 구조를 원하는 대로 정의할 수 있게 한다. 이러한 구조적 유동성은 어떤 형식의 데이터도 XML 형태로 기술될 수 있도록 해줌으로써 웹에서 운용되는 모든 데이터가 통합저장 및 처리될 수 있는 기반을 제공한다[1][4].

본 논문의 2절에서는 관련연구를 기술하고, 제3절에서는 XML Repository 시스템에 대한 설계를 다루며, 제4절에서는 마지막으로 결론을 기술한다.

### 2. XML Repository 시스템

#### 2.1 개요

XML 문서의 생성은 XML 문서 편집기나 각종 응용 프로그램에 의해서 생성될 수 있다. 이렇게 생성된 XML 문서는, 그 양이 적을 경우 간단한 파일시스템으로 저장 및 삭제등의 운용이 가능하나, 그 양이 많아지게 될 경우, 파일 시스템만을 가지고 필요한 파일을 수정/삭제 및 검색해내는 것이 어렵게 된다.

기존의 저장관리 시스템 중에는 효율적으로 다량의 텍스트 문서의 유지 관리를 위해 개발된 시스템들이 있다. 하지만 이러한 시스템들은 문서의 구조적인 정보를 중요하게 생각하지 않고 단순히 문서내의 내용만을 유지 관리하기 때문에 XML과 같은 구조화된 문서의 저장 및 관리 시스템으로는 적합하지 않다. 따라서 이러한 이유 등으로 인해 XML을 위한 전용의 DBMS가 필요하게 된다[3].

#### 2.2 기능

XML 문서 저장관리기가 가져야 할 기능은 크게 6가지로 나누어 볼 수 있다[3].

첫째로 DTD와 XML 문서 Loading 기능이다. 이는 문서 저장관리기가 문서의 생성을 담당하고 있지 않기 때문에 다른 응용으로 인해 생성된 XML 문서 및 DTD를 저장 관리기로 가져오는 역할을 담당한다.

둘째로 Parsing과 Validation 기능이다. 이는 여러 엘리먼트로 구성된 XML 문서를 파싱하고 엘리먼트 단위로 문서를 처리해 주기 위해서 필요한 기능이며, 추가적으로 Valid 한 문서를 검증하기 위해서 필요하다. 파싱에는 문서의 파싱뿐만 아니라 DTD의 파싱 기능도 있어야 한다.

셋째로 저장 기능이다. 이는 저장 관리기의 기본적인 기능으로 추후 검색 요청이나 문서의 변경 요청이 있을 경우, 이를 처리해 주기 위해 저장하는 기능이다. 문서뿐만 아니라 해당 문서의 DTD를 저장할 수 있는 기능도 지원되어야 한다.

넷째로 검색기능이 지원되어야 한다. 이러한 기능에는 Element, XML 문서의 내용(Content)부분 검색 기능, Parent, Child, Sibling, Ancestor, Descendant 등의 구조정보 검색 기능, 속성의 이름 및 값에 대한 검색 기능 등이 제공되어야 한다.

다섯째로 변경 기능이 지원되어야 한다. 저장된 DTD의 삭제 기능, XML 문서의 삭제 기능, 엘리먼트와 속성의 추가, 삭제, 수정 기능 등이 필요하다.

여섯번째로는 순회기능이 필요하다. 검색 기능과의 차이점은 검색기능이 정적인 자료의 추출인데 반해, 순회 기능은 동적으로 자료에 접근하는 기능이다. 순회 기능으로 현재 접근중인 엘리먼트 또는 속성에 대한 기준노드 유지기능과, 기준노드의 이름, 값, 타입등의 기준노드 특성 추출 기능, 그리고 기준노드를 DTD에서 문서로, 문서에서 엘리먼트로 또는 이 반대로 바꾸어 주는 노드 레벨에 대한 순회 기능이 필요하다. 이러한 노드 레벨에 대한 순회뿐만 아니라, 엘리먼트의 경우 같은 레벨

간에 존재하는 부모, 자식, 형제 노드들간의 순회 기능도 필요하다.

### 2.3 유형

XML 문서는 DTD가 있는가 없는가에 따라 문서를 나누어 볼 수 있으나 이는 문서의 성격을 결정짓지는 못한다. 문서의 성격이란 XML 문서가 과연 어떠한 목적에 사용되어 질 것인가를 나타내는 것으로 크게 데이터-중심 XML 문서와 문서-중심 XML 문서로 나누어 볼 수 있다. 이러한 문서의 유형에 따라 응용들이 달라지고 저장관리기도 이에 영향을 받게 된다.

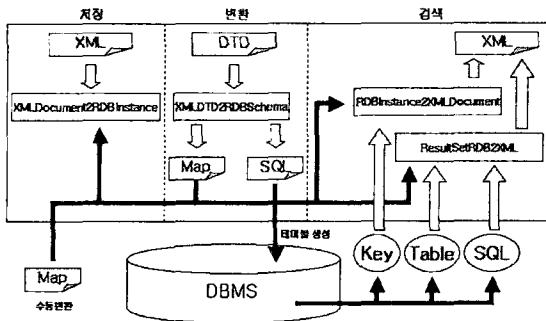
데이터-중심 XML 문서는 매우 정형적인 구조를 가지며, 내용 부분의 구조가 서로 혼재되어 있는 양이 적거나 없다. 형제들간의 순서가 중요하지 않고, 전자 상거래용 문서, 비행 스케줄, 레스토랑의 메뉴등의 문서가 이런 문서에 해당한다. 또한 기계적인 입장에서 생성 및 관리되는 것이 특징이다[2][3].

이에 반해 문서-중심 XML문서는 매우 혼재된 형태의 내용을 갖는 비정형적 구조로써, 대부분의 형제 엘리먼트간의 순서가 중요한 의미를 가진다, 책, 편지, 광고등의 문서들이 이런 유형에 해당하며, 대부분 사람의 손에 의해 작성되고 읽혀지며 사용되게 된다[2][3].

## 3. XML Repository 시스템 설계

### 3.1 시스템 구조

다음 그림과 같이 XMLDBMS에서 제공하는 기능들을 EJB 컴포넌트화하여 XML Repository 시스템을 구현한다.

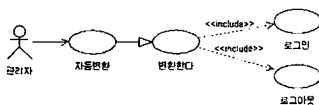


(그림 1) XML Repository 시스템

본 논문에서는 4개의 세션빈을 사용하여 XML Repository 시스템을 설계하였다. XML Repository 시스템은 크게 변환, 저장, 검색의 3부분으로 나누어진다.

### 3.2 변환단계

XML 응용시스템과 관계형 데이터베이스 시스템 사이에 원활한 연계를 위해서는, XML DTD를 관계형 스키마로의 변환할 필요가 있다.

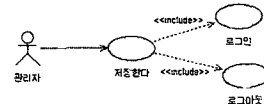


(그림 2) 변환 유즈케이스

변환단계는 자동 변환과 수동 변환으로 이루어진다. 자동변환은 XMLDTD2RDBSchema EJB 컴포넌트에 의해 변환되며, 수동변환은 사용자가 직접 Map문서를 작성하여 XML Repository 시스템에 적용하는 것을 말한다. Map문서란 XML 문서와 관계형 데이터베이스의 테이블 사이에 존재하는 관계를 기술한 XML 문서이다. SQL 문서는 CREATE문으로 자동 생성하여, 그 SQL 문서 토대로 사용자가 직접 데이터베이스 시스템에 테이블을 생성한다.

### 3.3 저장단계

유효한 XML 문서와 Map 문서를 입력을 받아 Map 문서 형태에 맞게 XML 문서의 내용(Content)을 관계형 데이터베이스 테이블에 저장한다.



(그림 3) 저장 유즈케이스

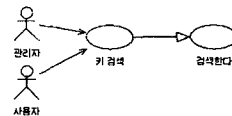
저장단계에서는 주의할 점은 XML 문서의 데이터 중복이다. 즉, 데이터베이스에 있는 데이터와 입력받은 새로운 XML 문서가 동일한 내용일 경우 중복이 발생하므로 이를 해결하여야 한다. XML Repository 시스템에서 사용하는 EJB 컴포넌트는 XMLDocument2RDBInstance 이며, 이 컴포넌트는 XML 문서와 Map 문서를 입력받는다.

### 3.4 검색단계

검색은 3가지 방법으로 검색할 수 있다. 첫 번째 방법은 키(key)를 이용하여 검색하는 방법, 두 번째 방법은 하나의 테이블에 있는 데이터를 모두 검색하여 XML 문서를 생성하는 방법, 마지막으로 직접 사용자가 SELECT문을 이용하여 XML 문서를 생성할 수 있다.

#### 3.4.1 키(Key)를 이용하여 검색

관계형 데이터베이스 테이블이 하나이상인 해당 테이블에서 키(Key)값을 이용하여 전체 테이블들을 검색한다. 검색한 데이터들은 입력받은 Map 문서에 의해 XML 문서로 생성이 된다.



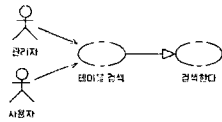
(그림 4) 키(Key) 검색 유즈케이스

XML Repository 시스템에서 사용하는 EJB 컴포넌트는 RDBInstance2XMLDocument 이다. 이 컴포넌트는 검색하고자 하는 테이블, 키(Key)값, Map 문서와 검색한 후 XML를 파일 형태로 생성하기 위해 파일 이름을 입력하여야 한다.

#### 3.4.2 테이블 검색

테이블 검색은 하나의 테이블에 있는 데이터들만 모두 검색

하여 XML 문서를 생성할 경우에 사용한다. 예를 들어 테이블에 10개의 레코드가 존재할 경우, XML 문서 안에 동일한 엘리먼트가 10개 존재한다.

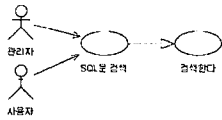


(그림 5) 테이블 검색 유즈케이스

XML Repository 시스템에서 사용하는 EJB 컴포넌트는 ResultSetRDB2XML 이다. 이 컴포넌트는 검색하고자 하는 테이블, Map 문서와 검색한 후 XML을 파일 형태로 생성하기 위해 파일 이름을 입력하여야 한다.

### 3.4.3 SQL 문 이용하여 검색

XML Repository 시스템에서 직접 SELECT 문을 이용하여 검색할 수 있다. 예를 들면 조인하거나 조건문을 사용하여 검색한 후 특정 Map 문서 형태로 XML 문서를 생성할 수 있다.



(그림 6) SQL 문 검색 유즈케이스

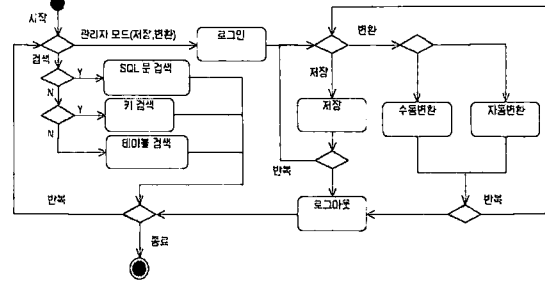
XML Repository 시스템에서 사용하는 EJB 컴포넌트는 ResultSetRDB2XML 이다. 이 컴포넌트는 직접 SELECT문을 작성하고, Map 문서와 검색한 후 XML을 파일 형태로 생성하기 위해 파일 이름을 입력하여야 한다.

### 3.5 자바 빈

XML Repository 시스템은 수행을 위해 필요한 자바빈들이 존재하는데, 그 중에서 가장 중요한 자바빈은 파일시스템을 담당하는 Conversion 클래스이다. 이 클래스는 File 형태의 문서를 String형으로 변환해 주는 기능과, 반대로 String형 문서를 파일로 생성하는 기능을 제공한다. 이밖에 저장 단계에서 데이터 중복 방지를 담당하는 클래스, 데이터베이스 시스템에 접근할 수 있도록 도와주는 클래스 등이 존재한다.

### 3.6 처리과정

XML Repository 시스템에 대한 처리는 그림 7과 같이 비즈니스 모델링으로 전체적인 흐름을 보여준다.



(그림 7) 처리과정

본 논문에서 제시한 유즈케이스 토대로 전체적인 처리과정을 보여준다. 여기서 저장과 변환은 관리자만 접근할 수 있도록 로그인 유즈케이스를 포함한다. 검색은 사용자와 관리자 모두 접근이 가능하므로 로그인 유즈케이스는 필요치 않다.

## 4. 결론

XML은 현재 웹에서 구조화된 정보나 반-구조화된 정보를 교환하기 위한 표준 마크업 언어로 채택되어 가고 있다. 한편 웹에서 정보교환을 위한 XML 문서의 소스 데이터는 Legacy 데이터베이스와 연결되어야 한다[4]. 본 논문에서는 XML 문서를 효과적으로 저장, 검색하기 위한 XML Repository 시스템을 설계하였다. XML Repository 시스템은 XML 문서를 관계형 데이터베이스에 저장하고, 다시 검색하여 XML 문서로 생성할 수 있다. 본 논문에서 설계된 XML Repository 시스템은 기업 간 웹 정보의 교환 및 공유를 위한 하나의 해결방안이 될 수 있다.

### 참고문헌

- [1] Joo Kyung-Soo, "A Design of Middleware Components for the Connection between XML and RDB", 2001 IEEE International Symposium on Industrial Electronics Proceedings, Pusan, Korea, June 2001
- [2] 박성진 "XML 데이터베이스" 인터넷정보학회지 한국인터넷 정보학회 2001년 9월 제2권 3호 38-46
- [3] 연제원, "XML문서의 효율적 검색 및 변경을 위한 저장관리의 설계 및 구현", 충남대학교 대학원 석사학위논문, 2000년 2월
- [4] 이상태; 주경수, "객체모델을 이용한, XML DTD의 ORDB 스키마로의 변환", 한국데이터베이스학회, 정보기술과 데이터베이스 저널, pp105-116, 2001
- [5] 이상화; 왕지현; 윤보현; 박성진, "XML Repository 시스템의 저장구조에 관한 연구", 한국인터넷 정보학회 추계학술발표대회 논문지, 제2권 2호, 190-197
- [6] 김경주; 조남규, UML Components 컴포넌트 기반 소프트웨어 명세를 위한 실용적인 프로세스, 인터비전, 2001, pp75-77