

XML 기반의 문서통합 시스템 설계

(Designed of Contents Integration System based on XML)

유 재 규, 박 재 훈, 전 양 승, 정 영 식, 한 성 국
원광대학교 전기전자및정보공학부 컴퓨터공학과
(jkyoo82; pjh98; globaljeon; ysjeong; skhan)@wonkwang.ac.kr

요 약

국내외적으로 정보통신망의 구축이 활발히 진행됨에 따라 이를 이용하여 개인이나 기업과 기관에서 업무의 효율성을 높이기 위하여 전자적으로 문서를 교환하여야 할 필요성이 증대되고 있다. 본 논문에서는 XML기반으로 문서를 통합하는 시스템을 설계하였다. 문서를 개인별로 산출물을 작성하는 문서의 경우 개별적으로 작성한 후 작성된 파일을 종합하여 하나의 문서로 통합하는 작업을 거치게 된다. 이러한 수동적인 방법을 컴퓨터 시스템을 이용하여 통합하는 과정을 자동화하고자 한다. 본 논문에서는 내용과 관련된 태그를 직접 만들 수 있고, XML 문서들을 하나의 큰 문서로 병합할 수 있으며, 어떠한 종류의 응용프로그램과도 통합될 수 있는 범용적 데이터베이스라고도 할 수 있는 XML의 특성을 이용하여 좀 더 확장성이 좋아지게 하였다.

1. 서 론

현재 공공기관, 업체, 학교 등에서는 수많은 형태의 문서 작업이 이루어지고 있다. 하지만 문서 작업 시 사람이 절대적으로 개입되어야 하는 실정이다. 또한 IT 기술이 발전되어지고 있음에도 불구하고 이러한 일련의 작업들은 비효율적인 형태로 진행되고 있다.

이러한 문제점을 해결하기 위하여 CMS (Contents Management System) 등에서 이러한 문제를 해결하고자 하는 시도가 있었지만, 많은 문제점에 봉착하였다. 그 문제점은 문서(콘텐츠)의 통합, 재활용 등이 있다.

따라서 본 논문에서는 XML/Schema를 이용하여 문서의 구조적인 정보 모델링을

하고, 문서의 각 요소들을 기술하기 위하여 XML 문법을 준수하여 표현될 수 있도록 문서 구조 정의를 한다. 이러한 문서의 본 시스템 표준 문서 템플릿을 정의하고, 공동으로 문서를 작업하기 위해서 문서 템플릿을 각각 사용자에게 배포한다.

그리고 각 사용자가 작성하고자 하는 부분을 작성하면, 사용자는 업로드(upload) 인터페이스를 통해서 업로드를 하게 되고, 업로드된 파일은 정해진 템플릿에 복사되어 저장되어 진다. 이때, 사용자가 업로드 하는 문서의 부분을 체크하고, 버전 관리를 해야 한다.

마지막으로 각 작성자가 최종 파일을 올리고서 관리자에게 업로드 완료 메시지를 보내게 되면 관리자는 최종적인 문서의 형태를 점검하는 것으로 문서의 통합 작업이 끝나게 된다.

2. 관련 연구

2.1. XML

초기에 웹은 단순히 정보를 제공하기 위하여, 예를 들면 문자나 그림을 표현하기 위한 도구정도만 사용되어져왔다. 인터넷상의 대부분의 정보는 HTML문서로 구성되어있으며, HTML은 단지 문서의 재현을 위한 정보를 나타내는 하나의 정의된 DTD(Document Type Definition)를 사용하기 때문에 각 문서의 엘리먼트를 의미 있는 정보를 표현하는 기능이 부족하다.

이에 W3C(World Wide Web Consortium)에서는 차세대 웹 문서의 표준으로 XML (Extensible Markup Language)을 1998년에 지정하였다 [1]. XML은 확장성 마크업 언어로서 이름 그대로 HTML같은 고정형식이 아니라 확장이 가능한 언어이며 문서의 내용과 관련된 태그를 직접 정의할 수 있으며 그 태그를 다른 사람들이 사용할 수 있다. XML은 본질적으로 다른 언어를 기술하기 위한 언어, 즉 메타언어이다. 또한 구조적 데이터를 표현할 수 있으며 사용자가 정의한 DTD를 만족하는 트리 구조를 가지고 있어서 XML은 구조적 문서를 표현하는데 유용하다.

또한 XML 기반언어로 RDF/S, RSS, MathML, XHTML, SVG 등이 있으며, 이들 언어들은 단일하게 규정된 방식으로 정의되었기 때문에, 사전정보 없이도 이들 언어로 작성된 문서에 대해 수정이나 적합성 검사를 하는 프로그램의 제작도 가능하다.

2.2. XML 스키마

XML 스키마는 XML 문서의 내용을 제한하고 기술하는 XML언어이다[2,3,4]. XML 문서의 항목과 포함할 수 없는 항목에 대한 규칙을 정의한다. 예를 들어 날짜 필드에 단

어를 입력할 수 없도록 스키마를 정의할 수 있다.

XML 스키마는 DTD의 단점을 보완하기 위해 만들어 졌다. DTD는 내용 모델을 정의하는데 사용되었으며, 다음과 같이 뚜렷한 한계를 가지고 있다[3,4].

- DTD는 XML과 다른 문법을 사용한다.
- DTD는 이름공간(Namespace)을 지원하지 못한다.
- DTD는 데이터형이 제한적이다.
- DTD는 복잡하고 느슨한 확장 메카니즘을 가지고 있다.

XML 스키마는 DTD의 이런 단점을 보완하기 위하여 다음과 같은 특징을 가지고 있다.

- 45가지의 다양한 데이터타입을 지원한다.
- 사용자 정의 데이터타입을 선언할 수가 있다.
- 이름공간을 지원한다.
- 사용자 정의 데이터타입이나 상속을 재정의 한다.
- 속성을 그룹핑한다.
- 모듈화와 재사용이라는 객체 지향적 개념을 제공한다.

2.3. XSLT

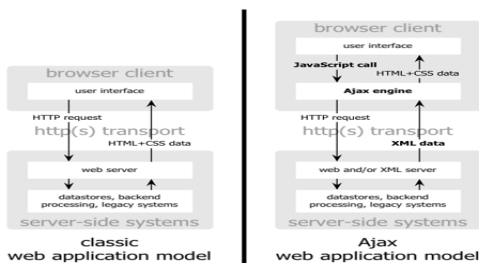
XSLT(XSL Transformations)는 한 XML 문서 구조를 다른 구조를 갖는 XML 문서로 변환하는 방법을 기술하기 위한 표준적인 방법으로서, W3C의 권고사항 중 하나이다[5]. XSLT는 XSL의 확장판이라고도 생각할 수 있다. XSL은 XML 문서, 예를 들어 XML 문서 내에 기술되어 있는 데이터가 어떤 방식으로 웹페이지 내에 표현되어야 하는 가를 보여주는 문서를 형식화하기 위한 언어이다. XSLT는 XML 문서가 다른 데이터 구조로 어떻게 재구성되어야 하는지를 보여준다.

2.4. AJAX

AJAX^{1),2)}는 대화식 웹 애플리케이션의 제작을 위해 다음과 같은 기술 조합을 이용하는 웹 개발 기법을 통칭한다[5,6].

- 문서 표현을 위한 XHTML(또는 HTML)과 CSS표준
- 동적인 화면 출력 및 정보와의 상호작용을 위한 DOM, JavaScript
- 웹 서버와 비동기적으로 데이터를 교환하고 조작하기 위한 XML, XSLT, XMLHttpRequest (XML/XSLT 대신 미리 정의된 HTML이나 일반 텍스트, JSON-RPC를 이용할 수도 있음)

AJAX 기술을 활용한 웹 응용한 ActiveX 기반의 응용과 달리 XML 처리를 할 수 있는 DOM 엔진과 JavaScript 엔진을 가진 대부분의 브라우저나 플랫폼에서 이용가능 가능하다. 또한 비동기적인 데이터 교환을 가능하게, 요청에 대한 서버의 응답을 기다리지 않고 다음 작업이 가능하므로 대기시간이 줄어들고, 이에 따라 서버의 부담을 줄이고 사용자의 체감속도를 높일 수 있다. 또한 이벤트 처리를 할 수 있으므로 효과적인 사용자 인터페이스 구현이 가능하다.



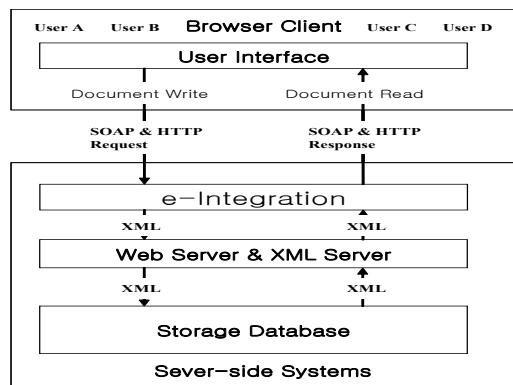
(그림 1) 동기 통신, 비동기 통신

1) AJAX라는 명칭의 기원은 제시 제임스 가렛(Jesse James Garrett)이 2005년 2월 18일 쓴 'A New Approach to Web Application'이라는 에세이에서 'Ajax(Asynchronous JavaScript + XML)'라는 낱말로 이 기술을 소개한 이후로 알려지게 됨
 2) AJAX에 대한 발음은 '에이잭스'나 '에작크스', '아약스', '아작스'등 여러 가지가 있지만, 최근에는 '에이잭스'로 통일되어 가고 있는 추세임

3. 문서통합 시스템 설계

3.1. 문서통합 시스템의 개요

본 논문에서 설계한 XML 기반의 문서통합 시스템은 XML, XML 스키마, XLST, AJAX를 이용하여 시스템을 설계하였다. 그 전체적인 수행 구조는 다음과 같다.



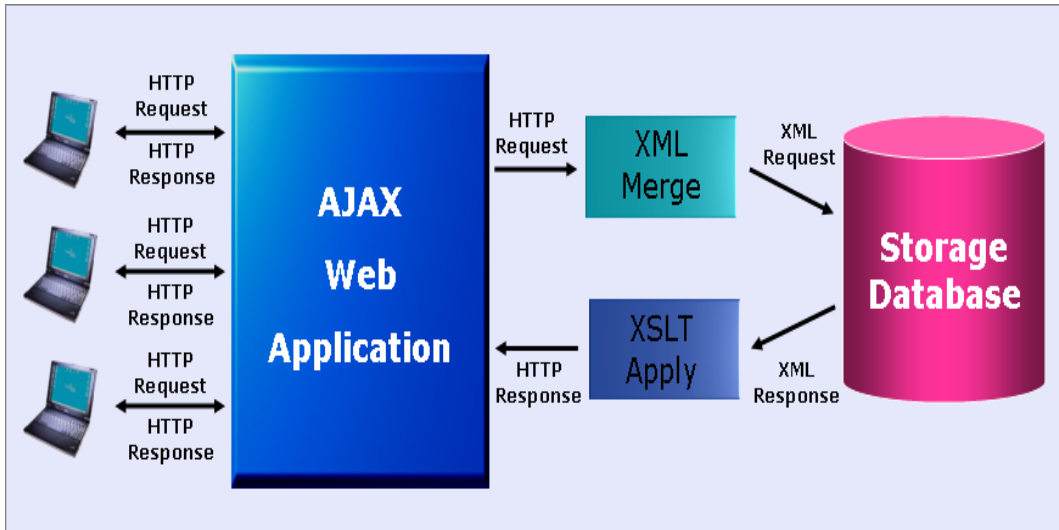
(그림 2) 문서통합 시스템의 구조도

3.2. 통합문서 등록

통합문서 시스템의 서비스를 이용하기 위해서는 통합문서 시스템 문서를 XML과 XSLT를 작성하여 통합문서 서버에 등록하여야 한다. 또한 하나의 XML 문서는 다수의 XSLT(문서의 템플릿)를 등록할 수 있다. 등록하는 과정은 SOAP과 HTTP 프로토콜을 이용하여 서비스에 등록하게 된다. 아래의 (그림 3)은 과정을 간략하게 그림으로 표현하였다.



(그림 3) 통합문서 서버에 통합문서 등록과정



(그림 4) 통합문서 시스템 구성도

3.3. 통합문서 이용

각 사용자는 일종의 AJAX 웹 어플리케이션에 접속하여 문서를 작성, 수정을 한다. 등록된 통합문서 시스템에 접속하여 작성, 수정할 수 있으며, 작성이 완료된 문서는 출력, 모든 문서 파일로 변환하여 준다. 즉 HWP, DOC, PPT, XLS 상에 존재하는 다양한 형태의 파일로 출력이 가능할 수 있다. 최근 이슈가 되고있는 AJAX 기법을 이용하여 사용자와 상호 대화적이기 문서의 추가, 변경이용이하다.

아래의 (그림4)는 통합문서를 이용하는 과정을 간략하게 그림으로 표현하였다.

형식에 맞게 수정 및 보완작업을 거쳐 하나의 문서로 통합해야 하는 비효율적인 과정을 거치게 된다.

각각의 사용자가 자유롭게 만든 문서는 본 통합시스템을 통해 일관된 형식으로 자동 변환 및 보관되어 부가적으로 발생하는 문서의 통합과 관련된 비용 및 시간을 절감할 수 있다. 그리고 각 문서의 일관된 형식을 사용함으로써 키워드 도출이 쉬워지므로 효율적인 문서 검색이 가능해진다.

본 통합시스템은 문서 작성을 위하여 사용자가 개별적으로 대기하지 않고 아무 때나 작성할 수 있으며, 기업 인트라넷에서 업무 보고 시스템을 획기적으로 변화할 수 있다.

4. 결론 및 전망

기존의 문서 작업은 여러 문서를 다수의 사용자가 협업 형태로 하나의 산출물을 도출한다. 이럴 경우 각 사용자가 작성한 문서는 일관성이 없고 새로운 문서를 만들 때마다 마찬가지로이다. 이런 일관성이 없는 문서는 최종 통합 책임자가 별도로 일정한 문서

Acknowledgement

“이 논문은 2006년 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임(원광대학교, 헬스케어기술개발사업단).”

참 고 문 헌

- [1] Extensible Markup Language(XML), <http://www.w3.org/XML/>, 2006.
- [2] 한성국, XML 워크샵, 2004.
- [3] Eric van der Vlist, "Using W3C XML Schema", <http://www.xml.com>, 2000.
- [4] Norman Walsh, "Understanding XML Schemas", <http://www.xml.com>, 2000.
- [5] Jesse James Garrett, "Ajax: A New Approach to Web Applications", <http://www.adaptivepath.com/publications/essays/archives/000385.php>.
- [6] David Heller, Ajax for Designers, 2005



유 재 규
2002~현재 원광대학교 전기전자
및 정보공학부

관심분야: 시맨틱 웹서비스, 온톨로지 공학



박 재 훈
2004 원광대학교 컴퓨터공학과
(공학사)
2005~현재 원광대학교 컴퓨터공학과
(석사과정)

관심분야: 웹서비스, 온톨로지 공학



전 양 승
2001 원광대학교 컴퓨터공학과
(공학사)
2006 원광대학교 컴퓨터공학과
(석사과정)
2006~현재 원광대학교
컴퓨터공학과(박사과정)

관심분야: 시맨틱 웹서비스, 온톨로지 공학, 지능형
e-Business



정 영 식
1993 고려대학교 전산학(박사)
1993~현재 원광대학교 컴퓨터공학부
교수
1997 미시간 주립대학교 전산학과
객원교수
2004 웨인 주립대학교 컴퓨터공학과

객원교수
관심분야: 그리드컴퓨팅, LBS, 분산병렬처리



한 성 국
1979 인하대학교 전자공학과
(공학박사)
1984~현재 원광대학교 컴퓨터공학부
교수
1989 University of Pennsylvania
방문교수

2003~2004 University of Innsbruck와 DERI 연구교수
2004~현재 대한전자공학회 컴퓨터소사이어티 감사
2005~현재 한국정보과학회 호남·제주지부장
관심분야: 시맨틱 웹서비스, 온톨로지 공학, 웹서비스,
의료정보, e-Learning