

문서 구조 정의 언어를 이용한 XML 문서 변환 시스템의 설계 및 구현

조영환
데이콤 e-Biz기술팀
yhcho@dacom.net

Design and Implementation of XML Translator Using XMeta Language

Young-Hwan Cho
e-Biz Engineering Team, DACOM

요 약

최근 들어 인터넷 기반의 B2B 전자상거래가 확산되고 데이터 교환을 위한 전자문서 형식으로 XML의 효용성에 대한 인식이 높아지면서, 전통적인 방식의 B2B 전자거래 모델인 EDI의 한계를 극복한 XML/EDI, ebXML 등의 전자상거래 모델이 등장하게 되었다[1]. 또한 이러한 XML 기반 전자상거래 모델에서는 XML 전자문서를 활용한 데이터 교환을 기반으로 하고 있다. 하지만 아직까지 각 기업에서 관리되고 있는 전자상거래 데이터의 대부분이 비 XML 데이터 형식으로 이루어져 있어 XML 전자문서를 활용하는데 현실적으로 어려움이 있다. 따라서 이러한 문제를 해결하기 위해서 본 논문에서는 기존의 전자상거래 데이터를 XML 기반 전자상거래 시스템에서 활용하기 위해서 비 XML 데이터의 구조를 기술하는 XMeta 언어를 정의하여, 이를 기반으로 다양한 데이터 포맷을 원활히 지원할 수 있는 XML 문서변환기를 설계 및 구현하였다.

1. 서론

정보통신 기술의 발전과 인터넷 사용의 확산은 디지털 경제 시대를 가속화하고 있으며, 기업의 정보 거래 모델의 방식과 관행을 근본적으로 변화시키고 있다. 또한 W3C에 의해서 제안된 차세대 웹 문서 표준인 XML(eXtensible Markup Language)[2]은, 문서의 내용과 구조 정보를 기술하기 위하여 설계된 자기 서술적(self describe) 언어로서, 복잡한 구조를 갖는 데이터를 표현할 수 있고 표준 인터페이스와 공개 소프트웨어를 이용하여 쉽게 처리할 수 있기 때문에 기업의 정보 거래를 위한 데이터 교환 표준으로 활용되고 있다. 이러한 환경 아래 등장한 XML/EDI[3], ebXML(Electronic Business XML)[4] 같은 전자거래 모델은 인터넷과 XML을 기본 요소로 하고 있다. 그러나 아직까지 기업 Legacy 시스템은 비 XML 문서를 정보의 기본 형식으로 사용하기 때문에 XML 문서를 이용하여 XML 기반 전자거래

모델과 상호 운영하는 것이 불가능하다.

이러한 문제점을 해결하기 위하여 Legacy 시스템에서 사용중인 비 XML 문서를 XML 문서로 변환하여 사용할 필요성이 제기되고 있으며, 이러한 이유로 플랫폼 파일을 XML 문서로 변환하는 방법[5], 데이터베이스 시스템의 데이터를 XML 문서로 변환하는 방법[6][7]등의 연구가 이루어지고 있다. 그러나, 현재까지 연구된 방법을 보면 문서 변환 응용 프로그램에서 읽게되는 원시(Source) 문서의 구조와 변환 후 생성될 목적(Destination) XML 문서의 구조가 응용 프로그램에 내포되어 있어서 다양한 형태의 원시 문서 구조나 목적 XML 문서 구조를 지원하기 위해서는 프로그램의 수정이 없이는 불가능한 유연성이 부족한 구조이다.

따라서, 본 논문에서는 프로그램의 수정 없이 다양한 형태의 문서 구조를 효율적으로 지원하기 위해 원시 및 목적 문서의 구조를 정의할 수 있는 XML

언어를 설계하여, 이것을 메타데이터로써 이용하는 XML 문서 변환 방법을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 XML Schema[8] 기반의 문서 구조 정의 언어의 설계에 대해서 설명하고 3장에서는 XML 문서 변환기의 설계 및 구현에 대해서 기술한다. 4장에서는 결론 및 향후 과제를 제시한다.

2. XMeta: 문서 구조 정의 언어

2.1 XML Schema

XML이 최초 발표되었을 때는 XML 문서의 구조를 정의하기 위한 구문 규칙으로 DTD(Document Type Definition)를 사용하였다. 그러나 XML 문서가 데이터를 표현하는 표준으로 활용되기 시작하면서 SGML(Standard Generalized Markup Language)을 위해서 만들어진 DTD의 구문만으로는 표현의 한계가 나타나게 되었다. 따라서 DTD의 단점을 보완하기 위해서 W3C에서는 XML 문서의 구조를 정의하기 위한 구문 규칙으로 XML Schema를 발표하였다. XML Schema는 XML 구문 규칙을 따르며, 45개의 다양한 데이터형을 지원하고 사용자 정의 데이터형을 선언할 수 있으며, 명시적인 반복횟수의 지정이 가능하다. 또한, DTD 대신 XML Schema를 사용함으로써 XML 문서 내의 데이터 유효성 검증을 XML 구문처리를 이용해서 문서 구문 해석 단계에서 할 수 있으므로 XML 문서 처리 응용 프로그램을 개발하기에 많은 이점이 있다. 이러한 장점 때문에 본 논문에서도 XML Schema를 이용하여 문서 구조 정의 언어인 XMeta를 설계하였다.

2.2 XMeta 언어의 설계

본 논문에서는 다양한 형태의 비 XML 문서를 XML 문서로 효율적으로 변환하기 위해서 비 XML 문서의 메타데이터를 이용하여 XML 문서로 변환하

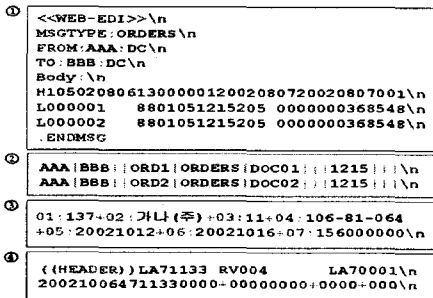


그림 1. 플랫폼 파일의 구조

는 방법을 제안한다. 이를 위해서 비 XML 문서 구조 정의 언어인 XMeta를 설계하였으며, XMeta 스키마는 플랫폼 파일이나 데이터베이스에 존재하는 데이터의 구조를 XMeta 언어에 맞게 정의한 XML 문서로서, 비 XML 문서의 메타데이터 역할을 한다.

XMeta 언어를 설계하기 위해서 분석한 플랫폼 파일 문서들의 일반적인 구조는 [그림 1]과 같으며 이들은 다음과 같은 특징을 갖고 있다.

- ① 데이터 집합은 한 개 이상의 그룹을 갖는다.
- ② 그룹은 한 개 이상의 레코드를 갖고 반복적으로 나타날 수 있다.
- ③ 레코드는 데이터 필드를 갖는다.
- ④ 그룹, 레코드, 필드는 시작과 끝을 표시하는 태그를 가질 수 있다.(From, .ENDMSG 등)
- ⑤ 그룹, 레코드, 필드는 각각을 구분하는 구분자를 가질 수 있다.(+, |, ;, \n 등)
- ⑥ 필드는 문자 필드와 숫자 필드가 있으며 가질 수 있는 최대, 최소 길이가 정해져 있다.
- ⑦ 필수 사용 데이터 필드가 있으며, 공백은 0 또는 Space로 채울 수 있다.

위와 같은 특징을 갖고 있는 비 XML 문서의 구조를 정의하기 위해서 Group, Record, Field, DBInfo 엘리먼트를 가지는 XMeta 언어를 설계하였다.

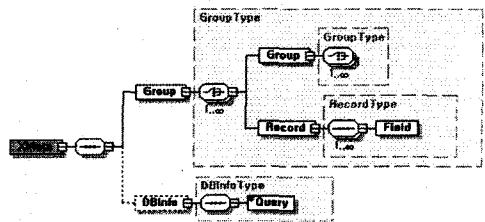


그림 2. XMeta 구조

각 엘리먼트는 [그림 3]과 같은 속성을 가지고 비 XML 문서의 구조를 정의한다.

Attributes	Group	Record	Field	DBInfo
Name	Name	Name	Name	DBVendor
SourceType	ChildOrder	ChildOrder	DBFieldName	DBServerName
Encoding	GroupSep	RecordSep	DataType	DBName
	Header	Header	FieldSep	DBDriver
	Trailer	Trailer	Filter	DBServerIP
	MaxOccurs	MaxOccurs	Justified	DBServerPort
	MinOccurs	MinOccurs	Mandatory	UserName
	MapToXml	MapToXml	MaxLength	Password
			MinLength	Encoding
			MapToXml	Decoding

그림 3. 엘리먼트별 속성

XMeta 엘리먼트는 XMeta 스키마의 루트 엘리먼트이며 변환 후 생성될 XML 문서의 루트 엘리먼트의 이름, 원시 문서의 종류와 Encoding 방법을 정의하는 속성을 가진다. Group 엘리먼트는 자식 엘리먼트로 Group 엘리먼트와 Record 엘리먼트를 가질 수 있으며, Group의 이름, 자식 엘리먼트의 사용 형태, 사용 횟수, Group을 구분하는 구분자 등을 정의하는 속성을 가진다. Record 엘리먼트는 한 레코드를 표현하기 위한 엘리먼트로서 레코드의 이름, 시작 태그와 끝 태그 등을 정의할 수 있으며 Field 엘리먼트는 데이터 필드를 표현하기 위한 엘리먼트로서 데이터의 속성을 정의한다. 마지막으로 데이터베이스의 정보를 정의하는 DBInfo 엘리먼트와 Query 엘리먼트가 있다.

3. XML 문서변환 시스템 설계 및 구현

3.1 시스템 구성도

XML 문서 변환 시스템의 구성도는 [그림 4]와 같다. 시스템의 구성은 XMeta 스키마 생성을 위한 XMeta Schema Builder와 XML 문서 변환 엔진인 XML Translator로 이루어져 있다. XMeta Schema Builder는 문서 변환을 위한 준비 단계에 사용되는 것으로 플랫폼 파일이나 데이터베이스의 구조를 가져와서 그에 맞는 XMeta Schema를 생성하는 Windows 어플리케이션이다.

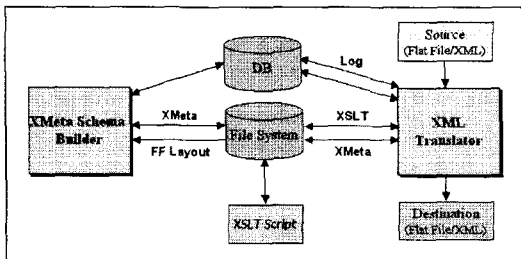


그림 4. XML 문서 변환 시스템 구성도

XML Translator는 문서 변환 준비 단계에서 만들어져 File System에 저장되어 있는 XMeta Schema와 XSLT를 이용하여 원시 문서를 변환하고자 하는 목적 문서로 변환시키는 XML 문서 변환 엔진이다.

3.2 XML Translator

XML Translator는 [그림 5]와 같은 구성으로 이루어져 있다. XTR UA는 사용자 Request를 받아서 처리하고 그 결과를 돌려주는 역할을 하는 클래스

이며, FF2XML, XML2FF, XML2XML, XML2PDF 클래스는 문서 변환 클래스이다.

플랫 파일을 목적 XML 문서로 변환하는 방법은 먼저, FF2XML 클래스가 XMeta 스키마 파일을 이용하여 XMeta 스키마에 정의된 XML 문서를 생성한 후, XML2XML 클래스가 FF2XML의 결과로 생성된 XML 문서와 XSLT를 이용하여 목적 XML 문서로 변환한다. 반대로 XML 문서를 플랫 파일로 변환하는 방법은 두가지 방법을 제공한다. 하나는 XML 문서를 XSLT를 이용하여 변환하는 방법이고, 다른 하나는 플랫 파일을 XML 문서로 변환하기 위해서 정의했던 XMeta 스키마를 이용하여 XML 문서를 플랫 파일로 변환하는 방법이다. XML2XML과 XML2PDF의 변환을 모두 XSLT 또는 XSL-FO를 이용한다.

Log Manager는 변환 결과를 데이터베이스 기록하고 Store Manager는 변환에 사용된 원시 문서와 목적 문서를 File System에 저장한다. XML 문서 변환 시스템의 관리 기능은 Servlet/JSP/Applet를 이용하여 웹서버를 통해서 제공한다.

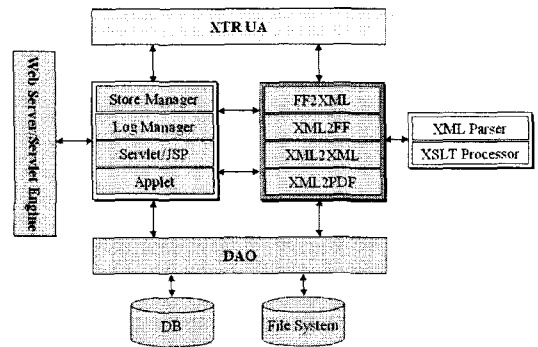


그림 5. XML Translator의 구성

3.3 시스템 구현

본 논문에서 개발한 문서 변환 시스템은 XMeta Schema Builder의 경우 Java Swing을 이용하여 Windows용 어플리케이션으로 개발하였으며, XML Translator 서버는 Unix 운영체제 환경 하에서 Java, Servlet, JSP 등을 이용하여 구현하였다. 웹서버와 서블릿 엔진은 Apache와 Tomcat을 사용하였으며 XML 문서 프로세싱을 위해서 Apache Xerces, Xalan, FOP를 사용하였다.

XMeta Schema Builder의 사용자 인터페이스는 [그림 6]과 같다. 그림 왼쪽의 엘리먼트 트리는 생

성중인 XMeta Schema의 구조를 보여주는 것이고, 오른쪽 상단은 각 엘리먼트의 속성 값을 정의하는 테이블이다. 그리고 오른쪽 하단에서는 각종 이벤트를 보여줌과 플랫폼 파일 레이아웃 및 데이터베이스 테이블 스키마를 보여줌으로써 엘리먼트의 속성을 정의하는데 도움을 줄 수 있도록 구현하였다.

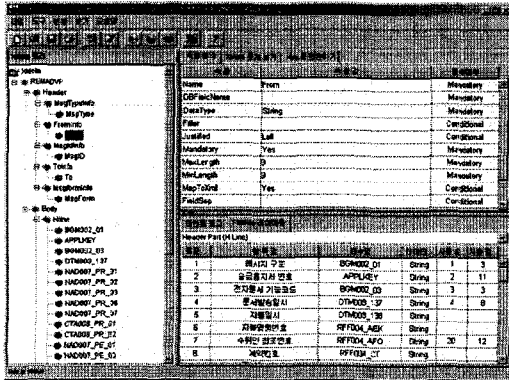


그림 6. XMeta Schema Builder 사용자 인터페이스

XML 문서 변환 실험을 위해서 사용된 문서의 예는 [그림 7]과 같다.

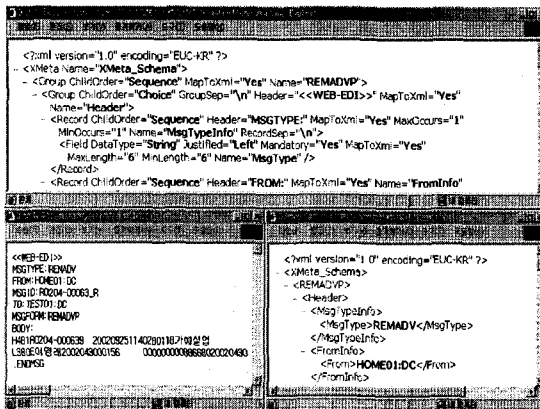


그림 7 플랫폼 파일, XMeta 스키마, 변환된 XML 문서

왼쪽 하단의 문서가 변환하고자 하는 플랫폼 파일 문서이며, 상단의 XML 문서가 플랫폼 파일 문서의 구조를 정의한 XMeta 스키마이다. 이 두 개의 문서를 이용해서 변환되어 생성된 문서가 바로 오른쪽 하단의 XML 문서이다.

4. 결론 및 향후 과제

기업의 정보 거래를 위한 환경이 인터넷의 보급과 XML의 등장으로 개선되면서 이를 기반으로 한

새로운 전자 거래 모델인 XML/EDI, ebXML 등이 등장하였으나 아직까지 이러한 XML 기반 전자 거래 모델을 이용하여 XML 전자 문서를 교환하는 것은 아직까지 대부분 기업들의 Legacy 시스템에서 비 XML 문서를 사용하고 있는 이유 때문에 현실적으로 어려움이 있다. 그래서 이러한 문제를 해결하기 위해서는 Legacy 시스템의 데이터를 XML 문서로 변환하는 것이 필요하다.

따라서 본 논문에서는 다양한 플랫폼의 비 XML 문서를 변환하기 위하여 Java 기반의 XML 문서 변환 시스템을 설계 및 구현하였다. 그리고 다양한 Legacy 시스템의 데이터 구조를 효율적으로 지원하기 위해서 XMeta 언어를 설계하여 이를 메타데이터로 이용하는 방법을 사용함으로써 다양한 문서 구조를 지원하기 위해서 문서 변환 모듈을 새로 개발하거나 변경하는 문제를 해결할 수 있었다.

향후 보다 다양한 형태의 비 XML 문서 구조의 문서를 실험함으로써 XMeta 언어의 적합성을 검증할 필요가 있다.

참고문헌

- [1] 조영환, 인현철, 임경미, "P2P기반의 전자상거래 모델을 위한 XML/EDI 시스템" 한국정보처리학회 춘계학술발표대회 논문집 9권 2호, 2002
- [2] W3C, Extensible Markup Language(XML) 1.0, REC-XML-19980210, <http://www.w3c.org>, 1998
- [3] XML/EDI Group, XML/EDI(XML/Electronic Data Interchange), <http://www.xmledi-group.org>
- [4] ebXML.org, electronic business XML(ebXML), <http://www.ebxml.org>
- [5] Dong-Soo Lim, Ji-Hoon Kang, Woo-Yong Han, "An XML Adapter for Non XML-based e-Commerce Systems to Interoperate with XML-based e-Commerce Systems", Proc. SCI 2000, July, 2000
- [6] 장우혁, 김홍식, "XML 기반의 효율적인 데이터 저장관리를 위한 DB2XML 변환 Wrapper의 설계", 한국정보과학회 춘계학술발표대회 논문집 28권 1호, pp 106~108, 2001
- [7] 최광미, 박수영, 정채영, "XML기반 EDI 문서교환을 위한 미들웨어 설계 및 구현", 한국정보처리학회논문지 B 제9-B권 6호, 2002
- [8] W3C, "XML Schema Part 0: Primer", <http://www.w3.org/TR/xmlschema-0/>, 2001