

교범 XML DTD 설계 및 구현

박세철, 이상훈
국방대학교 전산정보학과
e-mail:psckma@daum.net

The design and implementation of Manual XML DTD

Se-Chul Park, Sang-Hoon Lee
Dept. of Computer Science, Korea National Defence University

요약

XML을 이용하여 다른 조직이나 사용자간에 원활한 데이터 교환과 사용을 위해서는 공통적으로 사용할 수 있는 태그나 용어가 표준화되어야 한다. 현재 군 교범(야전교범 및 기술교범)을 XML을 이용하여 개발하려 노력하고 있으나, 사전연구가 미흡하고 표준이 정해지지 않은 상태에서 각 개별 기관별로 문서구조를 정의하고 태그를 사용함으로써 상호 호환성의 결여 및 차후 변환을 위한 낭비적 요소가 우려된다. 따라서 군 내부에서 XML을 적용한 문서 유형별 표준화가 시급히 요구되고 있다. 본 연구에서는 교범에 대한 XML 문서형정의를 설계하고 국방 표준 교범 XML DTD를 제안한다. 표준으로서의 문서형정의는 전자도서관에서 전문을 구축하는데 이용할 수 있을 뿐만 아니라 대화형 전자식 매뉴얼 구축에 기여할 수 있다.

1. 서론

XML(Extensible Markup Language)은 인터넷 관련 업체들에 의해 사실상의 표준(de facto standard)으로써 객체지향(Object-Oriented) 개념을 근간으로 하고 있으며, 플랫폼과 어플리케이션에 독립적인 데이터 형식이다. XML을 이용하면 문서형정의(Document Type Definition, 이하 DTD)등을 이용한 구조적 데이터 설계, 사용자 정의 태그확장을 이용한 다양한 응용프로그램 지원, 문자열이 아닌 문장구조 또는 특정 영역을 지정한 검색 등 다양한 활용이 가능하다. 이러한 장점으로 전자도서관에서 전문 구축 포맷으로써 다양한 시도가 이루어지고 있다. 이러한 XML을 기반으로 전자문서를 생성하기 위해서는 모두가 공유할 수 있는 표준으로서의 DTD가 우선적으로 정의되어야 한다. XML에서 DTD의 필요성은 문서의 표준화(공유)라는 차원에서 반드시 필요하다. 어느 문서에는 동일한 날짜 요소를

를 'date'라고 하고 다른 문서에서는 '날짜'라고 한다면 서로간에 교환하거나 공유해서 같은 의미로 사용할 수 없다. 따라서 문서구조 표현을 위한 DTD에 관한 연구는 전문 정보의 디지털화에 매우 중요하다.

본 논문에서는 교범(야전교범 및 기술교범)의 XML DTD를 설계하고 국방 표준안을 제안하고자 한다. 이를 위해 제2장에서는 XML DTD 관련 자료를 수집 분석하며 제3장에서는 현행 교범을 교리적인 측면에서 분석하고 적합한 교범 XML DTD를 설계한 후 유효성을 검증한다. 또한 서지시스템을 구현하여 XML DTD를 적용한 원문서서비스가 가능함을 평가하며 제4장에서는 결론 및 앞으로의 추가 연구방향을 결론으로 제시한다.

2. 관련연구

2.1 한국 전자책 문서 표준

```
<!ELEMENT book (metainfo, books)>
<!ELEMENT book (cover?, front?, book+, back?)>
<!ELEMENT book (cover?, front?, (body
%local.book;), back?)>
```

<표 2-1> EBKS ebook 구조

EBKS(Electronic Book of Korea Standard)의 문서 구조 및 메타데이터는 <표 2-1>에서처럼 메타데이터를 표현하는 metainfo와 책들의 집합인 books 요소로 구성되며, books 요소는 다시 cover, front, book, back 요소로 구성되며, 다시 book 요소는 cover, front, body, back 요소로 정의된다.[1]

2.2 ISO 12083 Book XML DTD

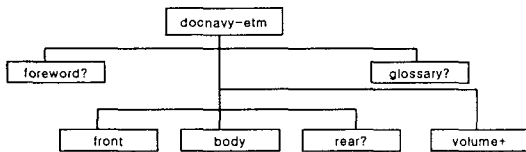
```
<!ELEMENT ebook (front, body, appmat?, back?)>
<!ELEMENT body (part+|chapter)
```

<표 2-2> ISO 12083 ebook 구조

ISO 12083:1998 Book XML DTD[2]에서는 책의 구조를 <표 2-2>에서처럼 전반부, 본문, 부록, 후반부로 정의하고 있으며, 본문은 부와 장으로 정의하였다.

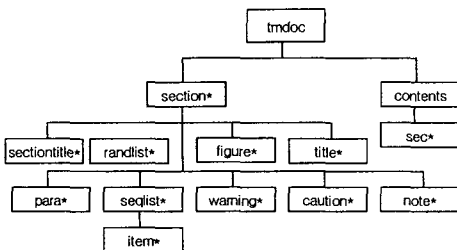
2.3 미 해군 전자식 기술교범 XML DTD(Navy Electronic Technical Manual XML DTD)

<그림 2-1>에서처럼 docnavy-etm은 크게 머리말 (foreword), 교범의 앞부분(front), 본문(body), 후반부(rear)으로 구분하였고, 용어사전(glossary) 및 권 (volume)으로 정의하였다. [3]



<그림 2-1> 미 해군 전자식 기술교범 구조

2.4 해군 전자기술교범



<그림 2-2> 해군 기술교범 구조
해군 전자기술교범의 DTD[4]를 분석하면 <그림

2-2>에서처럼 교범명은 tmdoc요소로 정의하였고, 교범의 하위 요소로 목차인 contents와 절인 section 요소로 구분하였다. 해군교범 XML DTD는 <그림 2-2>의 구조를 적용하여 목차에 링크된 문서를 각각 XML 문서 인스턴스를 생성하는 것으로 이러한 각각의 XML 파일이 모여 하나의 전체 기술교범문서를 구성한다.

3. 교범(Manual) 문서 구조 분석 및 DTD 설계
본 장에서는 기존의 전자책 및 기술 교범 DTD에 대한 연구를 바탕으로 다음과 같이 군 내부에서 통용될 XML DTD를 설계하였다.

3.1 교범의 구조 및 형태

야전교범 및 기술교범의 구조 및 형태는 <표 3-3>과 같다.

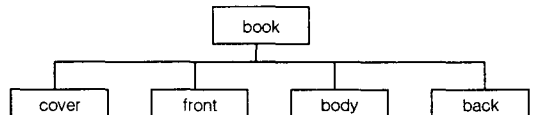
구조	구 성	문서 형태
표 지	· 분류기호 및 제목 · 로고 및 각군명칭	· TEXT · GRAPHIC
목 차	· 장,절 및 항 편성 · 참조페이지	· TEXT
본 문	· 일반사항 · 설치 및 운용지침 · 고장배제 / 정비절차 · 도면과 부품목록	· TEXT/TABLE · PHOTO · GRAPHIC/도면
부 록	· 참고문헌/정비할당표 · 각종 목록표 색인	· TEXT/TABLE · PHOTO · GRAPHIC/도면

<표 3-1> 교범의 구조 및 형태

3.2 교범(Manual)의 DTD설계

3.2.1 교범의 요소 구조

교범의 구조는 <표 3-1> 교범의 구조 및 형태에서 분석한 것을 토대로 <그림 3-1>과 같이 크게 표지면(cover), 전반부(front), 본문(body), 후반부(back)로 정의하였다.



<그림 3-1> 교범(Manual)의 요소 구조

3.2.2 표지면

표지면은 크게 표지전면과 표지이면으로 구성된다. <표 3-2>에서와 같이 표지의 전면은 좌상단에 문헌명과 문헌번호를, 우상단에는 발행

일자를 표기한다. 좌상단 문헌종류와 문헌번호 밑에 문헌의 성격을 구분할 수 있는 색(주황색, 황색, 하늘색)을 삽입한다. 또한 교범 중앙에는 교범명을 하단에는 발행부대로 구성되어 있다. 표지이면에는 일반교범시 교범활용 및 관리시 주의사항으로서 통제구분, 관리규정 및 판권관계와 국외 반출시 준수사항 등을 기재하며 비밀교범시 주의사항 대신 군사보안업무 시행규칙에 준하여 예고문을 기재하며 비밀표기는 통상 흑색을 이용한다.

```
<!ELEMENT cover (coverfront, coverback?)>
<!ELEMENT coverfront (doctypenum |
doctypename | pubdate | bookname | figgr |
pubunit)*>
<!ELEMENT coverback (title | warning |
caution | note)* >
```

<표 3-2> 표지면 구조

3.2.3 전반부

야전교범의 전반부는 <표 3-3>에서처럼 서문, 머리말, 초록, 목차 등의 요소로 구성되어 있다. 요소생성규칙 (+)는 최소한 한 번 이상 나타남을 의미한다. 전반부(front)요소 중 서문, 머리말로써 저자에 의해서 쓰여진 것을 나타내는 preface, 서문, 머리말로써 제3자가 쓴 것을 나타내는 foreword, 서론을 나타내는 intro, 초록을 나타내는 abstract, 그림목차를 나타내는 illuslist, 테이블 목차를 나타내는 tablelist 요소 등이 요소생성규칙 (+)로 그룹화되어 최소한 한 번 이상 발생하며, 목차는 교범의 구조에서 반드시 1회 발생함을 나타낸다. foreword, preface, intro, abstract의 요소에 정의된 paracon 개체는 para, figgr, fnoteref, footnote 요소로 구성되었으며 이는 문장과 그림, 각주 등을 표현할 수 있도록 설계하였다.

```
<!ELEMENT front ((title? | ( foreword |intro|
preface | abstract| illuslist | tablelist )+), toc)>
<!ELEMENT foreword (%paracon;)*>
<!ELEMENT toc (title? , partCont)*>
<!ELEMENT partCont (contTitle*, chapterCont)*>
```

<표 3-3> 전반부 DTD

목차는 교범의 내용을 개괄적으로 볼 수 있는 부분이다. 목차를 나타내는 요소는 toc로 나타내었고 하위 요소는 목차의 제목으로 title을 목차의 1부, 2부

등 부(part)를 나타내기 위해 partcont요소로 나타내었다. partcont의 하위요소로는 제목과 본문의 장을 기술하는 chaptercont요소로, chaptercont의 하위요소로는 제목과 절을 의미하는 sectioncont요소로, sectioncont의 하위요소로는 제목과 항을 의미하는 subsection1cont로 나타내었다. 마찬가지로 subsection6cont까지 동일한 요소로 반복하여 세항, 세세항을 표현하였다. 또한 목차에서 'ID'로 지정된 곳을 link하기 위해 conttitle요소에 rid속성을 정의하였다.

3.2.4 본문

야전교범의 본문은 <표 3-4>에서처럼 크게 부, 장, 절, 항, 세항, 세세항 등으로 구성되어 있으며 각 장은 한 번 이상 발생되고, 각 장의 제1절은 개요로 통일하며 각 장은 최소한 2개 이상의 중속절을 포함한다. 다시 장은 제목, 내용, 절로 구성된다. 절은 다시 절의 제목, 내용, 항으로 구성되는 계속 반복되는 구조로 되어있다. 항 내용을 세분할 필요가 없을 경우에는 세항(細項)을 설치하지 않는다. 본문에서 부, 장, 절, 항 등에서 공통으로 나타나는 요소는 title*, subtitle?, author*, #PCDATA, fnoteref, xref, xlink, citation, emphasis, subscript, supscrt, extref, dialink, seqlist, randlist, deflist, footnote, figgr, foldout, external, table, chart 등으로 교범에서 나타나는 모든 형태를 표현할 수 있도록 설계하였다.

```
<!ELEMENT body (part+ | chapter+)>
<!ELEMENT part (title* , subtitle? ,
(author*, (%paracon;)*)*, chapter+)*>
<!ELEMENT chapter (title* . subtitle? ,
(author*, (%paracon;)*)*, section+)*>
```

<표 3-4> 본문 구조

3.2.5 후반부

야전교범의 후반부는 <표 3-5>에서처럼 부록과, 배포표 및 교범 개선건의 및 오류보고 등의 양식, 그리고 교범의 표지후면에 연구담당관, 감수관 및 출판정보 등의 요소로 구성된다.

```
<!ELEMENT back (appendix+, afterword)>
<!ELEMENT appendix(title*, subtitle*,
subsection1)*>
<!ELEMENT afterword (%sectionframe:|idinfo)*>
<!ELEMENT idinfo ( (title | subtitle | edition
| pub | print | pubdate |
printdate)+, writer*, director*)>
```

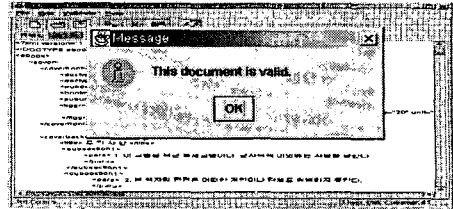
<표 3-5> 후반부 구조

따라서 후반부는 크게 부록을 나타내는 appendix와

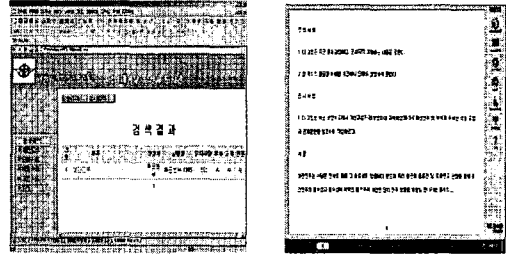
부록 외에 양식, 출판정보등을 나타내는 afterword 요소로 정의하였다.

3.2.6 공통요소

공통요소로는 문서내의 강조 부분을 나타내는 emphasis요소, 위첨자 및 아래첨자는 각각 supscprt와 subscprt요소, 외부 참조를 나타내는 xref요소, 도표 및 주석에 대한 참조는 각각 dialink와 ftnote 요소로 정의하였다. 순차 리스트 및 랜덤 리스트를 나타내는 요소로는 각각 seqlist와 randlist로 정의하였고 그림 및 책 사이에 접어놓은 지도 등을 나타내기 위해 figgr 및 foldout 요소로 정의하였고 테이블은 제목으로title, 테이블 그룹으로 tgroup로 정의하였으며 각주, 테이블, 수식 및.차트 등을 나타내기 위해 각각 footnote, table, formula, chart 요소로 정의 하였다. 파싱되지 않은 외부 데이터인 동영상은 각각 audio 및 video요소로 정의 하였다.



<그림 3-2> 유효성 검증(validation) 결과



<그림 3-3> 전자도서관 원문 검색결과 화면

```

<!ELEMENT emphasis %text;>
<!ELEMENT supscprt (#PCDATA)>
<!ELEMENT subscprt (#PCDATA)>
<!ELEMENT xref %text;>
<!ELEMENT dialink (#PCDATA)>
<!ELEMENT ftnref EMPTY>
<!ELEMENT ftnote (para+)>
<!ELEMENT randlist (title?, item+)>
<!ELEMENT seqlist (title?, item+)>
<!ELEMENT item (#PCDATA | ftnref | xref
|xlink | emphasis |subscprt | supscprt | extref |
dialink|para|warning | caution | note)*>
<!ELEMENT figgr (title? | sbtitle? | (figcont |
subfig?|subfoldout)*)*>
<!ELEMENT table (title?, (tgroup)+)>
<!ELEMENT chart (title?, (tgroup)+)>
<!ELEMENT fnoteref (#PCDATA)>
<!ELEMENT footnote (#PCDATA | para)*>
<!ELEMENT formula (#PCDATA | title)*>
<!ELEMENT external (title?)*>
    
```

<표 3-6> 공통 요소

3.3 교범 XML DTD의 유효성 검증(validating)

야전교범 DTD의 유효성을 검증하기 위해 야간전투교범 및 기술교범에 대해 설계된 DTD구조에 따라 XML 문서를 작성한 후 MSXML이라는 XML파서를 사용하여 유효성을 검증하였다. 검사결과 야전교범 DTD는 유효한 문서(Valid Document)임이 입증되었다.<그림 3-2, 그림 3-3>

4. 결론 및 향후 연구과제

본 논문에서는 교범 XML DTD를 각 군(육군, 해군, 공군)의 교범을 분석한 자료를 바탕으로 설계하였으며, 설계한 교범 XML DTD는 전자도서관 전문 구축을 위한 문서 포맷으로 뿐만 아니라 대화형전자기술교범(IETM)에도 적용할 수 있도록 설계하였다. XML 문서 포맷과 관련하여 현재 각 업체마다 다양한 문서형 구조(DTD)를 사용하고 있으며, 그 결과 여러 종류의 뷰어와 변환 도구의 개발이 요구된다. 따라서 국방 표준으로서의 XML DTD 선정은 필수적이며, 전자도서관의 디지털화를 촉진하기 위해 관련 응용프로그램의 개발과 기존 디지털문서(아래한글,워드 등) 으로부터 XML 포맷으로의 변환 툴의 개발 등이 이루어져야 한다.

참고문헌

- [1] EBK 표준화위원회 워킹그룹, <http://orange.yonsei.ac.kr/ebook>, 2001.
- [2] ISO 12083 Book XML DTD, <http://www.xmlxperts.com/xmlbookdtd.htm>, 1998.
- [3] Navy Electronic Technical Manual XML DTD, <http://navysgml.dt.navy.mil/dtdfosi/netmxmltdt.html>, 2001.
- [4] 해군전자기술교범, 국방인트라넷 [http:// 34.1.20 . 11 /](http://34.1.20.11/),2001.