

# 한국전자책 문서표준(EBKS)의 개발 및 활용방안

○  
 손원성\*, 고승규\*, 이경호\*\*, 김성력\*\*\*, 임순범\*\*\*\*, 최윤철\*

\*연세대학교 컴퓨터과학과, \*\*미국 국립표준기술원(NIST) 객원 연구원  
 \*\*\*숙명여자대학교 문헌정보학과, \*\*\*\* 숙명여자대학교 멀티미디어학과

## The Development and Application of EBKS (Electronic Book of Korea Standard)

Won-Sung Sohn\*, Seung-Kyu Ko\*, Kyong-Ho Lee\*\*, Sung-Hyuk Kim\*\*\*,  
 Soon-Bum Lim\*\*\*\*, Yoon-Chul Choy\*

\*Dept. of Computer Science, Yonsei University, \*\*IT Lab. National Institute of Standards and Technology,

\*\*\*Dept. of Library and Information Science, Sookmyung Women's University,

\*\*\*\*Dept. of Multimedia Science, Sookmyung Women's University

### 요 약

전자책에 대한 사회적 관심과 다양한 장점에도 불구하고 현재 상이한 전자책 문서 포맷을 사용한 결과 전자책 콘텐츠에 대한 교환 및 공유를 어렵게 하여 결과적으로 전자책의 시장 활성화를 가로막는 가장 큰 장애가 되고 있다. 미국과 일본에서는 이미 관련 표준안을 제정한 바 있으며, 한국에서도 2001년 한국전자책 컨소시엄(EBK)의 표준화 분과위원회를 통하여 전자책 산업 표준인 한국전자책문서표준(EBKS)을 제정하였다. 한국에서의 전자책 표준 EBKS는 HTML 및 PDF보다 다양한 장점을 제공하는 XML을 기반으로, 전자책 콘텐츠에 대한 정확한 교환, 한국 출판환경의 특징 반영, 생성 용이성, 관련 표준과의 호환성 제공 등을 기본 방향으로 한다. 한편 EBKS는 DTD 측면에서 고정된 문서구조, 확장 매커니즘 제공, 세로 쓰기, 다단 편집, 그리고 한글 및 고어 특징 반영등과 같은 특징을 포함하며, 스타일의 경우 교환을 위해서는 XSL-FO를 이용하여 콘텐츠와 스타일을 분리하고 있으며, 출력을 위해서는 XSL-FO를 권고한다. 그리고 EBKS 메타데이터는 사용의 용이성, 교환 및 호환성등을 제공하기 위하여 한정어(qualifier)를 사용하지 않는 dc-metadata와 확장을 위한 x-metadata로 구성된다. 한편 본 논문에서는 이러한 EBKS의 활용으로 XSL-FO를 이용한 전자책 콘텐츠에 대한 정교한 출력 결과를 제시한다.

### 1. 서론

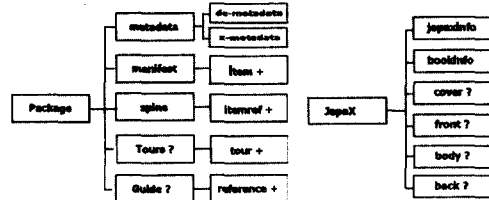
전자책이란 책의 콘텐츠를 디지털 형태의 정보로 가공 및 저장한 출판물을 의미하며, eBook, e-텍스트, 온라인북, 파일북 등 다양한 이름으로 불리고 있다[1,2,3]. 이러한 전자책은 기존의 인쇄된 책과 비교하여 유통 과정의 단순화 및 재고 부담 절감, 멀티미디어 기능 제공, 다양한 콘텐츠 출력 기능, 그리고 저장 관리의 효율성등의 장점들을 제공할 수 있다[1,2,3].

한편 전자책 문서 포맷과 관련하여 현재 XML(eXtensible Markup Language), HTML, PDF(Portable Document Format), 기타 자체 개발한 포맷 등이 사용된 결과, 전자책 콘텐츠에 대한 교환 및 공유가 어렵고 이는 반드시 해결되어야 한다. 따라서 미국[8]과 일본에서는 전자책 문서 포맷의 표준화에 대한 필요성을 인식하고 각각 OEB 출판 구조 명세(Open eBook Publication Structure Specification)[9]와 JepaX[7]를 제정한다.

OEB PS는 OEB 패키지(OEB Package)와 OEB 문서(OEB Document)로 구성되어 있으며[그림 1] 콘텐츠의 배포 및 출력을 목적으로 한다. 실제 전자책 내용을 표현하는 OEB 문서는 OEB 사양에 부합되는 기본(basic) 문서와 임의의 XML문서인 확장(extended) 문서로 나눌 수 있다. 기본 문서는 OEB 사양에서 정의된 구문만을 이용하여 작성된 문서로 기본적으로 XHTML 1.0과 유사한 구조를 갖는다. 확장 문서는 임의의 문서 구조를 수용할 수 있는 XML 문서를 의미한다. 그러나 이와 같은 OEB 문서에서는 구체적인 문서 구조를 정의하지 않음으로 인해 전자책 교환이나 가공 시 모호한 경우가 발생할 수 있는 단점이 있다.

JepaX는 콘텐츠 생성이나 교환 포맷으로 사용하는 것이 목적이며

OEB와는 달리 배포를 고려하지 않고 있다. JepaX의 구조는 다음 [그림 1]과 같이 jepaxinfo, bookinfo, cover, front, body, back으로 구성되어 있다. 그러나 JepaX도 OEB PS와 마찬가지로 임의의 사용자 정의 엘리먼트가 무분별하게 발생할 수 있기 때문에 명확한 전자책 콘텐츠 교환을 보장하기 어렵다는 단점을 포함하고 있다.



[그림 1] OEB Package와 JepaX의 전체 구조

한국에서도 지난 2000년 한국전자책 컨소시엄(EBK)[5]의 표준화 분과위원회[6]를 통하여 2001년 4월 한국전자책문서표준(EBKS)을 제정하였다. EBKS는 전자책 콘텐츠에 대한 정확한 교환을 목적으로 제정되었으며 이는 한국 전자책 업계에서 가장 시급히 요구되는 사항을 반영한 결과이다. 한편 EBKS는 상호운영성, 수용성, 확장성, 응용성, 공개성, 간결성 등을 보장하기 위하여, 구조 확장 및 콘텐츠와 스타일 분리가 어려운 HTML 및 부가정보 및 구조 정보 표현의 부재, 재사용성이 떨어지는 PDF의 단점을 극복할 수 있는 동시에 위의 이슈들을 모두 만족하는 XML을 한국전자책문서표준(EBKS)의 기본 포맷으로 결정하였다[7].

또한 EBKS는 미국 OEB PS와 일본 JepaX와는 달리 고정된 문서 구

조를 포함하고 있으며, 한국에서 발간되는 대부분의 책의 구조를 지원하고 있다. 그리고 한국 문헌의 특징인 세로쓰기 및 다단편집, 그리고 고어 등을 고려하였으며, 기타 문서 구조에 대한 확장 방법을 제공할 수 있다.

EBKS 메타데이터는 더블링크어의 15개 기본 엘리먼트를 기반으로 하며 특히 한정어(Qualifier)를 사용하지 않는다. 이는 메타데이터 작성을 가급적 쉽게 하기위한 의도이며, 전자책 업체에서의 다양한 응용을 위하여 확장 방법인 x-metadata를 지원한다.

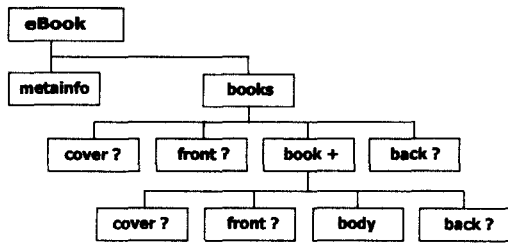
EBKS에서의 한글 인코딩은 OEB, JepaX와의 호환을 위하여 Unicode를 기반으로 하며 인코딩 방식은 UTF-8 또는 UTF-16을 지원하도록 한다. 또한 폰트가 제공되지 않는 코드를 표시하기 위하여 한국어 어문규정 가운데 문장부호법에 있는 빠짐표(□)를 사용하고, 언어 표기는 DC의 "Language"와 RFC1766을, 동시에 언어 이름은 ISO639 표준을 따른다. 한편 폰트에서 지원되지 않는 한자(고어)나 EBKS 엔터티에 정의되지 않은 특수 문자 표기를 위해서 사용자 정의가 가능한 x-character셋을 제공한다.

다음장에서는 지금까지 살펴본 EBKS의 상세 내용에 대하여 살펴본다.

2. EBKS DTD의 상세구조

2.1 eBook 문서 구조

EBKS의 루트 엘리먼트인 ebook 엘리먼트는 [그림 2]에서처럼 메타데이터를 표현하는 metainfo와 책들의 집합인 books 엘리먼트로 구성되어 있다.



[그림 2] EBKS의 전체 구조

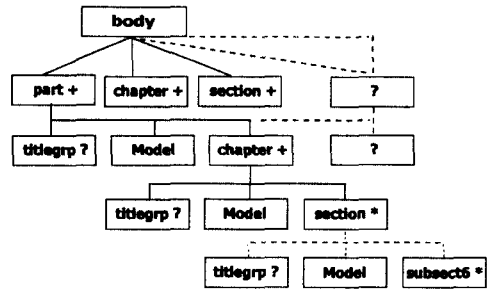
2.1.1 Books, Cover, Front, Back

EBKS로 작성된 전자책 문서는 한 권 이상의 책(book)으로 구성될 수 있으며, 이는 모음집(books) 혹은 진집 형태의 서비스를 지원하기 위함이다. 일반적으로 cover 요소는 종이책의 앞과 뒤 표지를 구성하는 부분이다. 그러나 전자책에서의 커버는 그 형태가 다양하게 변경되어 구성될 수 있으며, 특히 이미지로 구성된 figure 혹은 출판사에 대한 상세정보들은 필수적으로 포함되어야 한다. Front는 책의 본문이 시작되기 이전의 머리말, 감사의 글, 서문 등의 요소로 구성되는 부분이다. 한편 책의 컨텐츠 순서를 표현하는 목차(toc)도 front에 나타나며, 특히 EBKS에서의 toc 엘리먼트는 교환시에는 공백(EMPTY) 엘리먼트로 처리된다.

2.1.2 Body, Part, Chapter, Section

body는 전자책의 실질적인 컨텐츠를 기술하는 부분으로 서비스하러

는 책의 종류에 따라 part, chapter, section 엘리먼트 중 하나를 선택하여 컨텐츠의 Top level 엘리먼트를 구성할 수 있다[그림 3].



[그림 3] EBKS의 body 전체 구조

Part[그림 3] 요소는 컨텐츠의 body부분에서 매우 중요하게 사용될 수 있으며, 그 이유는 대부분의 본문 구조는 Part 요소를 이용하여 상위 레벨의 구조를 정의할 수 있기 때문이다. Chapter[그림 3]는 일반적으로 어떤 한 주제가 일단락이 되지만 전후와의 연관성을 갖고 있는 책의 한 부분으로서, 타이틀 성격의 구조 혹은 내부적으로 section 요소를 포함한 경우 등으로 나타난다.

EBKS의 section[그림 3] 요소는 하부로 총 6개 레벨의 subsection으로 구성되며, 이는 대부분의 경우 6레벨의 subsection 구조로 처리가 가능하기 때문이다.

2.1.3 EBKS의 확장 메커니즘

전자책의 구조를 정의할 때에 가장 다양한 구조가 나타나는 부분은 body이하 부분이라 할 수 있다[그림 3]의 ? 경우. 따라서 EBKS DTD에서는 [그림 4]와 같이 body 부분뿐만 아니라 모든 엘리먼트 부분에서 구조를 확장할 수 있는 메커니즘을 제공하고 있다.

```

<ENTITY % localBook "" >
<ELEMENT book (cover?, front?, (body %local.book;), back?) >
    
```

[그림 4] EBKS에서의 엘리먼트 확장 예

2.2 eBook Content 엘리먼트

본 절에서는 EBKS DTD의 Content 정의와 관련된 다양한 요소들 중 공통속성, Paragraph, 참조모델, 리스트, 테이블 등에 대하여 설명한다.

2.2.1 Common Attribute, Paragraph, xcharacter

EBKS의 모든 엘리먼트는 공통속성(Common Attribute)을 갖는다. 그 기능은 가장 먼저 엘리먼트 참조를 위한 id, 엘리먼트 구분을 위한 메타 정보 부여 기능의 role, 사용언어를 뜻하는 lang, 가로, 세로쓰기를 정의한 hdir, vdir 등이 있다. Paragraph model은 본문이나 기타 엘리먼트에서의 실제 컨텐츠 데이터를 처리하며, 이를 위하여 #PCDATA 및 여러 inline 요소들 및 float 요소들을 포함하고 있다. EBKS에서는 표현이 불가능한 문자들을 xcharacter 엘리먼트로 표현할 수 있다. 이는 한국 출판물에서는 기존 폰트체계로 표현이 불가능한 고어(한문)나 사용자가 특수 문자를 직접 정의하여 사용하는 경우를 위하여 xcharacter란

엘리먼트를 제공하고, 컨텐츠 교환 시 이에 대한 명확한 결과를 보장하기 위함이다.

**2.2.2 Float, Reference, List, Table 모델**

EBKS 문서의 모든 부분에서 발생할 수 있는 요소들은 float 모델로 정의된다. 이는 제한된 구조로 정의하기 어려운 주석, 이미지, 수식, 참조 모델 등이 포함된다. 또한 EBKS 문서 내부 및 외부의 참조를 위한 참조모델은 내부 참조인 xref, 외부 자원 참조를 위한 uri, 등으로 구성된다. 리스트는 논리적으로, 또는 스타일 측면에서 매우 중요한 요소이다. 따라서 EBKS의 리스트 모델은 다양한 리스트 형태를 지원하는 다양한 type 속성을 통하여 정의되며, 특히 기존 웹 문서와의 호환성을 고려하여 CSS의 속성과 호환성을 유지하도록 고안되었다.

EBKS의 테이블 모델은 CALS 테이블을 지원한다. CALS 테이블 모델은 Column Specification과 Span Specification을 이용하여 각 Entry에 특정 의미를 부여할 수 있기 때문에 셀에 대한 의미정보 부여, 재사용, 검색 등을 위한 장점을 제공할 수 있다.

**2.3 메타데이터**

EBKS 메타데이터는 국내의 메타데이터 표준 및 전자책 메타데이터들을 비교하고 한국 전자책 여건 및 국내에서 사용되고 있는 메타데이터 특성을 고려하여 적합한 메타데이터 표준을 작성하였다. EBKS 메타데이터 표준은 간결성, 사용의 용이성, 일관성, 확장성, 기타 메타데이터와의 상호운용성을 원칙으로 한다. 따라서 위 원칙을 최대한 수용하는 국제 표준인 더블린 코어 메타데이터 표준을 적용하여, 더블린 코어의 15개 엘리먼트 및 x-metadata 엘리먼트를 반복해 사용하여 더블린 코어에서 표현되지 않는 전자책 상거래 및 검색 등을 위한 정보를 추가할 수 있도록 하였다.

**3. XSL-FO를 이용한 EBKS의 출력**

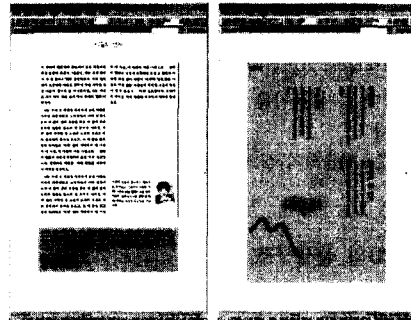
EBKS에서는 스타일 정보를 교환의 대상으로 간주한다. 즉 스타일 정보는 전자책 컨텐츠 출력을 위한 논리 구조 및 스타일정보를 포함할 수 있기 때문에 문서 교환시 EBKS 문서와 스타일 정보를 모두 포함하여 컨텐츠 출력 결과에 대한 최소한의 레이아웃을 보장하기 위함이다. 이를 위하여 EBKS는 XSL-FO를 이용하여 컨텐츠와 스타일을 분리하는 동시에 전자책 교환시 별도 혹은 두가지 모두를 교환하도록 한다. 한편 전자책 스타일과 관련해서는 기존 CSS1 및 CSS2, 그리고 DSSSL 등을 고려할 수 있으나, XSL-FO가 한국 출판환경의 특성인 세로쓰기, 다단표기 및 기타 다양한 출력기능을 모두 제공할 수 있기 때문에 EBKS에서는 XSL-FO를 교환 및 출력 포맷으로 사용할 것을 권고하고 있다.

다음의 예제 [그림 6][그림 7]은 잡지 및 영어 학습서를 EBKS를 이용하여 작성한 XML 문서를 XSL-FO를 이용하여 다단, 세로쓰기, 테이블, 리스트 및 다양한 스타일을 적용한 결과를 보여주고 있으며, 그 결과 XSL-FO는 기존 CSS 보다 미려한 출력 결과를 보장할 수 있음을 알 수 있다.

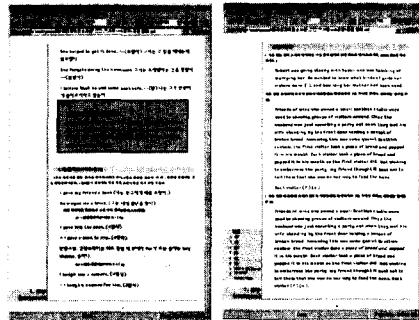
**4. 결론**

EBKS는 전자책 컨텐츠에 대한 정확한 교환을 목적으로 제정되었으

며 이는 한국 전자책 업계에서 가장 시급히 요구되는 사항을 반영한 결과이다. 이를 위하여 EBKS는 미국 OEB PS와 일본 JapaX와는 달리 고정된 문서구조를 포함하고 있으며, 한국에서 발간되는 대부분의 책의 구조를 지원하고 있다. 또한 한국 문헌의 특징을 최대한 고려하였으며, 다양한 문서 구조를 반영하기 위한 구조 확장 방법을 제공한다. EBKS는 DTD 측면에서 고정된 문서구조, 확장 매커니즘 제공, 세로쓰기, 다단 편집, 그리고 한글 및 고어 특성 반영등과 같은 특성을 포함하며, 스타일의 경우 교환을 위해서는 XSL-FO를 이용하여 컨텐츠와 스타일을 분리하고 있으며, 특히 출력에 위해서는 다양한 기능의 XSL-FO를 권고하고 있다. 본 논문에서는 이러한 장점을 실제 예제를 통하여 확인하였다.



[그림 5] 잡지 출력의 예



[그림 6]영어 학습지 출력 예

**참고문헌**

- [1] 문화관광부, “한국전자책 산업 발전방안 연구,” 2000.
- [2] 문화관광부, “디지털시대의 전자책(e-Book) 발전방향,” 2000.
- [3] 손원성, 고승규, 최윤철, 외 “전자책 관련 기술과 발전 동향,” 한국멀티미디어학회지 제5권 제1호, 한국멀티미디어학회, 2001.3
- [4] 손원성, 고승규, 최윤철, 외 “XML에 기반한 한국 전자책 문서 표준,” 정보처리학회지 제8권 제3호, 한국정보처리학회, 2001.5
- [5] EBK(e-Book of Korea) Consortium, “A Study of Korean Standardization of eBook documents,” Technical Report, 2001.
- [6] EBK(e-Book Korea) Consortium, <http://www.ebk.or.kr>, 2001.
- [7] Japanese Electronic Publishing Association (JEPA), <http://www.jepa.or.jp>.
- [8] Open eBook Forum (OEBF), <http://www.openbook.org>, 2001.
- [9] Open eBook Publication Structure 1.0, Open eBook Forum (OEBF), <http://www.openebook.org>, 1999.