

전자책 라이브러리를 위한 메타데이터 개발에 관한 연구*

A Study on the Development of Metadata for eBook Library

하진희(Jin Hye Ha)**, 김성혁(Sung Hyuk Kim)***,
임순범(Soon Bum Lim)****

초 록

전자책 서비스업체는 전자책에 대한 충분한 데이터를 제공하고 있지 않기 때문에 대부분의 도서관은 전자책의 목록정보를 도서관의 목록에 자동으로 다운로드할 수 없다. 본 논문은 이러한 문제점을 해결하기 위하여 다양한 메타데이터 간의 호환성 및 상호운용성을 확보하기 위하여 전자책 라이브러리를 위한 메타데이터를 개발하였다. 이를 위해 전자책 서비스업체에서 사용하는 메타데이터, KS X 6100 메타데이터, 더블링크어, MARC, TEI Header 등을 비교 분석하여 공통의 메타데이터 요소들을 도출하였다. 도출된 공통의 메타데이터 요소를 핵심 기술 요소로 정의하고, 그 외 전자책 고유 특성을 나타내는 메타데이터 요소를 상세 및 추가 기술 요소로 정의하였다.

ABSTRACT

The information about eBook cannot be added automatically to library catalog because the eBook Service Provider does not provide enough eBook metadata. Therefore, this paper was developed the eBook library's metadata in order to maintain the compatibility and interoperability among various metadata standard. In trying to accomplish this, we have comparatively analyzed the eBook service provider's metadata, KS X 6100 metadata, Dublin Core, MARC, and TEI Header, and extracted common metadata elements from them. We defined these common metadata elements as a core element and added other elements that uniquely describe eBook characteristics as a detailed and additional elements for eBook metadata.

키워드: 전자책, KORMARC, 메타데이터, eBook metadata, metadata

* 본 논문은 2002년도 정보통신부지원 기초기술연구지원사업 연구과제 “전자책 라이브러리를 위한 표준간의 변환 및 인증에 관한 연구”의 일부로 수행되었음 (C1200210000047).

** 숙명여자대학교 대학원 문헌정보학과(h77j03h28@hanmail.net)

*** 숙명여자대학교 정보과학부 교수(ksh@sookmyung.ac.kr)

**** 숙명여자대학교 정보과학부 조교수(sblim@sookmyung.ac.kr)

■ 논문 접수일 : 2003. 6. 30

■ 게재 확정일 : 2003. 9. 1

1 서 론

1.1. 연구의 목적 및 필요성

지식정보사회에서 정보 이용자는 종이 매체뿐만 아니라 디지털 매체 자료 같은 다양한 유형의 정보를 검색하여 이용하기를 원한다. 그러나 기존의 도서관에서 활용하고 있는 MARC 형태로 작성된 목록 데이터는 도서 목록을 위한 국제적인 표준이긴 하지만, 인터넷 기반 디지털 도서관이 급속히 발전함에 따라 등장한 웹 자원 및 멀티미디어 자원에 대한 목록 기술 형식으로는 매우 비효과적이다.

이미 여러 도서관에서 전자책을 사용하기 시작하여 학생, 연구자 및 기타 독자들에게 언제 어디서나 전문(full-text) 전자책 서비스를 받을 수 있도록 하고 있다.

그러나 아직 전자책 서비스업체에서 판차사항에 관한 충분한 데이터를 제공하고 있지 않기 때문에 도서관에서는 목록작성자가 실제로 전자책을 볼 때까지는 목록에 전자책을 위한 레코드를 추가할 수 없다. 전자책 유통에 있어 중요한 요소인 이용자와 도서관을 고려한다면, 전자책 정보를 표현하는 목록 데이터 정보의 중요성을 간과할 수 없다.

현재 전자책 열람 서비스를 제공하고 있는 도서관들은 단지 전자책 타이틀 정도만 제공할 뿐이다. 뿐만 아니라 도서관에서 전자책 목록 데이터 작성을 위해 여러 가지의 전자책 뷰어(viewer)를 다운

받거나 구입해야 하는 등의 번거로움을 감수해야 한다.

현재 여러 전자책 서비스업체가 운영하고 있는 시스템들은 각기 개별적으로 개발되고 운영되고 있다. 따라서 메타데이터 표준의 개발을 위해 소요되는 경비뿐만 아니라 호환성 문제를 해결하기 위한 추가적인 노력이 필요한 실정이다.

그러므로 다수의 메타데이터 표준의 일관성을 유지시키고 각 메타데이터 표준에 존재하는 메타데이터 요소들의 의미를 명확히 하여 메타데이터의 호환성 및 상호운용성 향상을 위한 전자책 라이브러리를 위한 메타데이터의 개발은 매우 시급하다.

본 논문은 전자책 정보의 원활한 이용을 위해 기존 종이책 도서관에서 사용하는 메타데이터와 전자책 서비스업체에서 사용하는 메타데이터의 특성을 반영하여 전자책 라이브러리를 위한 메타데이터를 개발하였다.

1.2. 연구 방법 및 제한점

전자책 라이브러리 메타데이터를 개발에 앞서 먼저 전자책 라이브러리의 구축을 위한 기본 구성요소를 분석하여 제시하였다. 전자책 라이브러리를 위한 메타데이터 개발을 위해서 기존 전통적인 도서관에서 사용하고 있는 메타데이터 표준인 MARC, Dublin Core, TEI header, 전자책 메타데이터 표준인 KS X 6100, OeBPS, JepaX, 그리고 전자책 서비스업체에서 전자책 정보를 제공하기 위해 사

용하는 메타데이터를 분석하였다. 전자책 서비스업체는 EBK 표준화위원회 워킹그룹에서 활동한 업체를 중심으로 국내 대표 전자책 서비스업체 5곳을 선정하였다. 선정한 업체는 (주)바로북닷컴 (www.barobook.com), (주)와이즈북토피아(www.wisebook.com), 하이북스토어(www.hiebookstore.com), 드림북(www.dreambook.co.kr), 에버북닷컴(www.everbook.com)이다.

기존 종이책 도서관에서 전자책 목록 데이터를 효율적으로 작성하기 위해 향후 KORMARC 으로의 변환 작업이 수행되어야 하며 이를 위해 본 연구에서는 그 초기 단계로서 본 연구에서 개발한 전자책 메타데이터와 KORMARC 레코드와의 비교표를 작성하여 제시하였다.

2 이론적 배경

2.1 메타데이터 표준

2.1.1 MARC

MARC(Machine Readable Cataloging)은 기계가독형 목록으로 도서관에서 상호 '정보의 공유'를 위해 서지정보를 전달하는 정보 교환용 표준 형식이다. 이는 각각의 내부 형식에서 표준 형식으로 변환하여 외부에 서지정보를 보내거나 외부에서 표준형식으로 서지정보를 받아서 내부 형식으로 변환하여 사용하기 위한 형식이다. MARC은 현존하는 메타데이터 형식 가운데 가장 널리 사용되고 발달된 형식

이다.

MARC은 1965년 MARC I이 완성된 후 이를 수정·보완하여 1968년 MARC II(LC MARC)가 제정됨으로써 도서관 자동화의 중요한 개념으로 등장하게 되었다(Heery 1996). 미국 의회 도서관에서 개발한 MARC II는 미국의 국가적 표준인 USMARC 뿐만 아니라 국내의 KORMARC 등 세계 각국의 MARC 형식 개발의 기초가 되었다.

최근 미국 의회도서관과 캐나다 국회도서관은 공동으로 USMARC과 CAN/MARC 형식을 통합한 MARC 21을 제안하였으며, 이는 다양한 유형의 자료를 기술하는 통합 포맷으로 사용하고 있다.

MARC으로 표현된 전자책은 기존의 도서관 등의 분야에서 이미 사용하고 있는 형태로 메타데이터를 표현하기 때문에 전자책 메타데이터로 고려해 볼 수 있으나 일반 사용자가 작성하기 어렵다는 단점이 있다(문화관광부 2000).

현재 KORMARC은 KCR4가 최종 확정되면 자료 유형에 따라 구분되어 있는 것을 통합서지용 KORMARC에 모두 수용하기로 결정하고, 현재 통합서지용 KORMARC 형식(안)은 MARC 21을 바탕으로 기존에 발표된 KORMARC 형식들을 모두 통합하는 방식으로 작성 중에 있다(이계선 2002).

2.1.2 Dublin Core

더블린코어는 기술(descriptive) 메타데

이터로서, 네트워크 환경에서 웹 자원을 기술하고 접근하는데 필요한 15개의 기본 엘리먼트로 구성되어 있으며 필요에 따라 엘리먼트의 확장을 통해 특정한 응용분야에 적용이 가능하다.

그러나 다양한 분야에 단일 형식의 메타데이터만을 적용하는 것이 실제로는 불가능한 경우가 있어 더블린코어에서는 한정어를 사용한 확장방안을 제시하고 있다. 이 한정어는 상호운용성을 최소한으로 고려하였으며 응용 분야에 따라 부가적인 한정어 사용이 가능하다(김진아 외 2001).

2.1.3 TEI header

TEI Header는 전자문헌 작성과 교환을 위한 가이드라인으로 1987년 뉴욕 Pough-keepsie에서 개최된 회의에서 제시된 몇 가지 원칙에서 출발하였다. 이에 1994년 4월 TEI P3가 발행되었고, 2002년 6월에 TEI P4: Guidelines for Electronic Text Encoding and Interchange가 발간되었다.

TEI Header는 기존에 전자문서 작성 시 서로 상이한 인코딩(encoding) 방식에 따른 낭비를 제거하고 다양한 유형의 문헌에 적절한 코딩 집합을 기술하기 위한 전자문헌기술 방식이다.

전자문헌의 서지정보와 비서지정보를 수록하고 있는 TEI Header에서 순수한 서지정보만은 MARC 레코드와 매우 유사한 형식으로 구성되어 있다(조윤희 2003).

TEI Header의 다양한 요소는 핵심세

트(core sets), 기본세트(base sets), 부가 세트(additional sets), 그리고 보조세트(auxiliary sets)'의 태그 집합으로 정리할 수 있다.

2.2 전자책 메타데이터 표준

2.2.1 KS X 6100

전자책 메타데이터는 전자책 자체에 대한 정보와 전자책과 관련한 권리정보, 즉 지적 재산권의 행사 및 이동 정보와 상거래 행위와 권리에 대한 정보도 제공할 수 있다.

KS X 6100 메타데이터는 국제 표준인 더블린코어 메타데이터 표준을 적용하여 더블린코어의 15개 엘리먼트를 반복적이면서 선택적으로 사용하도록 메타데이터를 작성하였다. 그리고 OEB 메타데이터 방식과 같이 x-metadata 엘리먼트를 반복해 사용하여 더블린코어에서 표현되지 않는 전자책 상거래 및 검색 등을 위한 정보를 추가할 수 있도록 하였다(손원성 외 2001).

2.2.2 OeBPS

미국 전자책 표준 포럼인 OeBF (Open e-Book Forum)에서 제정한 OeBPS(The Open eBook Publication Structure)는 전자책의 콘텐츠, 구조 그리고 표현을 위한 XML 기반 표준이다.

2002년 8월 27일 발표된 버전 1.2는 버전 1.01의 기본 OeBPS 엘리먼트 확장과,

일부 엘리먼트의 삭제 및 네임스페이스 추가 등 새롭게 제정한 것이 아닌 기본 버전을 수정·보완한 것이다.

기존의 표준 및 메타데이터들이 전자책 정보에 관한 모든 사항을 기술 할 수 없으므로 OeBPS는 이미 제정된 표준들을 서브셋(subset)으로 명세함으로써 가능하면 상세히 기술하려고 한다. 또한 전자책 서비스업체에 필요한 상호운용성을 위해 기존 표준에서는 명세하고 있지 않는 제한요소를 두었다 (OeBPS 1.2 2002).

OeBPS 메타데이터는 더블링크어를 기록하는 dc-metadata 엘리먼트와 부가적인 메타데이터를 기록하는 x-metadata 엘리먼트로 구성되어 있다.

2.2.3 JepaX

일본전자출판협회(JEPA: Japanese Electronic Publishing Association)에서 제정한 전자출판물의 표준 교환용 포맷인 JepaX는 각 전자책 서비스업체에서 전자책을 배포하기 전단계로서 문서를 배포할 시에 변환기를 거쳐 최종 판매형식으로 변환할 수 있도록 정의한 XML 기반의 표준이다(이민석 2002).

JepaX 문서는 루트 엘리먼트 하위에 <bookinfo> 엘리먼트를 사용해 문서의 메타정보를 표현하며 이 엘리먼트는 생략할 수 없다. 이 엘리먼트 하위에는 자체 정의한 9개의 엘리먼트를 이용하여 전자책 메타정보를 제공한다.

3 전자책 라이브러리 개요

3.1 전자책 라이브러리의 성격 및 기능

전자책 라이브러리는 단순히 전자책 정보를 모아 놓은 저장소 이상의 의미를 가져야 한다. 따라서 일반적인 디지털 도서 관과는 차별성을 드러낼 수 있어야 한다. 이러한 차별성을 전자책이라는 자료 형태의 특징에서 구체화하면 다음과 같다.

3.1.1 전자책 정보의 중심

전자책은 흔히 eBook이라고 불리며 기존의 책과는 달리 파일 형태의 출판물을 전용 뷰어(viewer)를 통해 컴퓨터나 전용 단말기로 읽는 디지털 출판물을 말한다.

지식과 정보의 수집과 축적, 저장은 전통적으로 중요한 도서관의 기능으로 취급되었다. 그러나 디지털 환경에서 이러한 전통적인 도서관의 역할은 보다 적극적인 의미를 가지게 된다. 물리적인 실재를 가지는 책에 대한 정보의 의존성이 점점 감소되고 인터넷, 각종 시청각자료, 데이터베이스, 전자저널, 전자책 등 다양한 형태의 정보원들이 점점 중요시되면서 디지털 형태의 자료를 관리하기 위한 새로운 전략이 필요하게 되었다.

이에 전자책 라이브러리는 전자책 정보를 표현하는 다수의 메타데이터 표준을 일관되게 유지시키고 각 메타데이터 표준에 존재하는 메타데이터 요소들의 의미를 명확히 하여 메타데이터의 호환성 및 상

호운용성 향상을 위한 전자책 정보의 중심이 되어야 한다.

3.1.2 정보 이용자의 정보에 대한 자유로운 접근과 이용 보장

디지털 기술을 이용한 정보서비스의 궁극적인 목표는 정보 이용자가 원하는 자료를 원하는 시간에 원하는 장소에서 아무런 제약 없이 이용할 수 있도록 하는 것이다. 지식정보사회에서 정보 이용자는 종이 매체뿐만 아니라 디지털 매체 자료 같은 다양한 유형의 정보를 검색하여 이용하기를 원하고 있다. 전자책 라이브러리는 정보 이용자에 대한 이러한 제약 없는 정보 서비스를 제공할 수 있도록 해야 한다. 정보 이용자에게 정보에 대한 자유로운 접근과 이용을 보장하는 것은 결코 비현실적인 일은 아니다.

도서관의 전통적인 자료는 MARC, 더블링크어 등과 같은 메타데이터가 개발되어 대부분 도서관들이 이를 활용하고 있다. 그러나 이러한 MARC 기반의 목록 데이터는 도서 목록을 위한 국제적인 표준이긴 하지만 인터넷 기반 디지털 도서관이 급속히 발전함에 따라 증가하는 웹 자원 및 멀티미디어 자원에 대한 목록 기술 형식으로는 매우 비효과적이다.

그러나 전자책 라이브러리를 통해 전자책에 대한 정보를 상세하고 일관되게 표현할 수 있으므로 정보 이용자는 전자책 원문에 대한 정교한 검색과 부분 대역 등과 같은 편리한 기능을 제공 받을 수 있다.

3.1.3 전자책 서비스업체의 관리 비용 절감

현재 여러 전자책 서비스업체가 운영하고 있는 시스템들은 각기 개별적으로 개발되고 운영되고 있다. 따라서 필요한 메타데이터 표준을 제정하기 위해 많은 시간과 투자를 해야 한다. 그러므로 메타데이터 표준의 개발을 위해 소요되는 경비뿐만 아니라 호환성 문제를 해결하기 위한 추가적인 노력이 필요한 실정이다.

그러나 전자책 라이브러리를 통해 전자책 서비스업체는 전자책 정보를 하나의 레지스트리(registry)에 등록시킴으로써 공간, 시간 등 다양한 관리 비용을 절감할 수 있다.

3.1.4 도서관 사서의 업무 효율화

국내의 전자책 서비스업체가 전자책 라이브러리에 등록된 전자책 정보를 목록작성자가 다운로드만 받아 정보 이용자에게 제공하도록 할 수 있다. 따라서 전자책 정보를 검색하고자 하는 이용자들의 다양한 문의들을 신속하고 효율적으로 처리할 수 있게 한다. 뿐만 아니라 도서관 상호대차나 원문복사 서비스까지 효과적으로 수행하여 이용자들이 쉽게 이용할 수 있도록 지원한다.

3.2 전자책 라이브러리를 위한 핵심 요소

전자책 라이브러리는 전자책에 대한 정보 및 콘텐츠를 관리하는 일종의 디지털 도서관이다. 다양한 형태, 다양한 정보,

다양한 접근법 등 획일하지 않은 특성을 지닌 디지털 도서관의 가장 중요한 요소는 상호운용성(interoperability)이다. 전자책 라이브러리 역시 가장 핵심적인 요소는 상호운용성이라고 할 수 있으며 이를 위해 다음의 세 가지 요소가 요구된다.

첫째, 메타데이터 간 호환이다. 메타데이터 정보와 소프트웨어 인터페이스는 이질적인 구문론과 의미론을 자동으로 호환하여 표현할 수 있어야 한다. 이는 개방되고 복합적인 환경에서 소프트웨어 상호간에 혹은 정보의 서로 다른 폼(form)으로 공유가 가능하게 한다.

둘째, 레지스트리이다. 메타데이터 레지스트리, 메타서비스 레지스트리, 그리고 메타매핑(데이터 포맷과 스키마의 일반적인 매핑) 레지스트리 등이 구축되어야 한다.

셋째, 변환틀이다. 모든 이용자들이 자원을 공유하고 이용하기 위해서 데이터 변환을 위한 소프트웨어가 반드시 개발되거나 이용이 가능하도록 조치해야 한다. 이들 레지스트리와 변환틀은 상호운용을 위한 시스템의 개발 및 운영을 원활하게 하는 하부구조로서 우선 개발되어야 하는 필수 요소이다.

4 전자책 라이브러리를 위한 메타데이터 개발

4.1 메타데이터 요구조건

다양한 형태, 다양한 정보, 다양한 접근

법 등 다양한 특성을 지닌 디지털 도서관의 가장 중요한 요소는 상호운용성이다. 전자책 라이브러리는 전자책에 대한 정보 및 콘텐츠를 관리하는 일종의 디지털 도서관으로서 상호운용성은 중요한 요소 중의 하나이다.

전자책 메타정보를 기술할 전자책 라이브러리 메타데이터는 이 상호운용성 뿐만 아니라 간결성, 획일성, 확장성, 응용성, 수용성, 공개성의 특징을 가져야 한다. 뿐만 아니라 전자책의 특징을 잘 표현할 수 있어야 하며 동시에 검색이 용이해야 한다.

2003년 4월 9일 전자책도서관통합컨소시엄(KOBLA)이 출범함에 따라 통합솔루션 공급, 컨소시엄 이름으로 전자책 공급, 전자책 기술 최첨단 적용, 그리고 전자책 콘텐츠 공유가 가능하게 된다. 따라서 전자책 라이브러리를 위한 메타데이터는 이에 부합할 수 있도록 호환성 및 상호운용성 그리고 확장성을 고려해야 한다.

본 논문에서 제안한 전자책 라이브러리 메타데이터는 국제 표준인 더블린코어를 고려하고 동시에 국내 여건을 반영해야 한다.

4.2 메타데이터 분석

메타데이터에 대한 단일화된 하나의 국제 표준은 없다. 더블린코어와 같이 비교적 단순한 메타데이터에서 TEI Header와 같이 더욱 복잡하고 풍부한 메타데이터까지 복잡성과 풍부성 수준이 다른 여러 메타데이터 표준들이 폭 넓게 사용되고 있

〈표 1〉 주요 전자책 서비스업체 제공 메타데이터 비교

업 체		바로북닷컴	하이북 스토어	드림북	와이즈북 토피아	에버북 닷컴
요 소						
	서 명	○	○	○	○	○
저 자	저 자	○	○	○	○	○
	번 역 자	-	-	-	○	-
출 판 사 항	출판사명	○	○	○	○	○
	주 소	-	-	-	○	-
	전화번호	-	-	-	○	-
	팩 스	-	-	-	○	-
	제 작 자	-	-	-	○	-
	공 급 자	-	-	-	○	-
	출 판 일	○	-	○	○	○
	제 작 일	-	-	-	○	-
	가 격	○	○	○	○	○
	파일크기	○	-	-	-	○
	파일유형	○	-	○	-	○
	지원리더	-	-	○	-	○
	표지 이미지	○	○	○	○	○
	내용 소개	○	○	○	-	○
	저자 소개	○	○	-	-	○
	목 차	○	○	○	○	○
	리 뷰	○	-	○	-	○
	ISBN	-	-	-	○	○
	수상내역	-	-	○	-	-
	미리보기	-	-	○	○	-
	요구사항	-	○	-	○	-
기 타	디자이너	-	-	-	○	-
	담당자	-	-	-	○	-

다(Heery 1996).

본 논문에서는 국내 전자책 서비스업체 5개를 선정하여 각각 전자책 정보를 기술하기 위해 사용하는 주요 메타데이터 요소들을 분석하였다. 이와 함께 메타데이터 표준으로 널리 사용되고 있는 메타데

이터 중에서 그 범위와 영향을 기반으로 3개의 메타데이터-KORMARC, 더블린코어, TEI Header-를 고려하고, 전자책 메타데이터 표준으로는 국내 표준인 KS X 6100을 고려하였다.

〈표 1〉은 선정된 5개 업체별로 제공하

〈표 2〉 메타데이터 표준 요소 비교

요 소 \ 메타데이터	더블린코어	MARC	TEI Header	KS X 6100
Subject	○	○		○
Title	○	○	○	○
Author(Creator)	○	○	○	○
Publisher	○	○	○	○
Publication Place		○	○	
Other agent	○	○	○	○
Date	○	○	○	○
Object type	○	○		○
Format	○	○	○	○
Identifier	○	○	○	○
Relation	○	○	○	○
Source	○	○	○	○
Language	○	○	○	○
Description	○		○	○
Coverage	○	○		○
Right	○			○
Abstract		○		
Edition		○	○	○
Notes		○	○	
Classification		○	○	
Keyword			○	
Availability			○	
Cost			○	○
Extent(size)			○	○

고 있는 메타데이터 요소를 분석한 결과이다. 그 결과, 전자책 정보를 기술하는데 전자책 업체에서 사용하는 공통의 메타데이터는 총 6개이며 이 요소는 5개 업체 모두에서 사용하고 있다. 그 요소는 '서명', '저자', '출판사명', '가격', '표지이미지', '목차'이다.

국내 전자책 표준인 KS X 6100 메타데이터는 전자책 정보를 등록하기 위해 제작되었으며, 이미 전자책 서비스업체들

사이에서 사용되고 있는 표준이다. KS X 6100 메타데이터는 간결성, 사용의 용이성, 일관성, 확장성 기타 메타데이터와의 상호운용성을 원칙으로 제작되었다.

〈표 1〉의 원칙을 최대한 수용하는 KS X 6100 메타데이터는 국제 표준인 더블린코어 메타데이터를 적용하여 더블린코어의 15개 엘리먼트를 반복적이면서 선택적으로 사용하고, OeBPS 메타데이터 방식과 같이 x-metadata 엘리먼트를 반복

해 사용하여 더블링크어에서 표현되지 않는 전자책 상거래 및 검색 등을 위한 정보를 추가할 수 있도록 하였다.

〈표 2〉는 메타데이터 표준인 더블링크어, KORMARC, TEI Header 그리고 KS X 6100의 메타데이터 요소를 비교한 결과로서 기존 연구¹⁾에서 제시한 결과를 수정하여 반영하였다.

본 논문에서 제안하는 전자책 라이브러리를 위한 메타데이터는 전자책 서비스업체에서 실제로 사용하고 있는 메타데이터를 고려하여 이를 포함하고 있는 KS X 6100 메타데이터를 수용하되, KORMARC, 더블링크어, TEI Header 를 고려하여 일부 엘리먼트를 추가하였다.

4.3 메타데이터 정의

본 논문에서 제안하는 전자책 메타데이터는 앞에서 분석한 과정을 통해 다음과 같은 기준으로 선정하였다.

첫째, 자료의 형태 구별 없이 모든 자료를 포괄할 수 있는 메타데이터 요소를 선정하였다. 이는 시스템간 상이한 메타데이터 요소 중 최소한의 공통된 필수 데이터 요소를 제시하여 시스템간 상호운용성을 확보하기 위함이다.

둘째, 여러 메타데이터 표준에서 전자책 고유 특성을 잘 표현하는 일반성이 높은 메타데이터 요소를 핵심 기술 요소와

상세 기술 요소로 선정하였다. 이는 기존 메타데이터 요소를 최대한 수용하여 시스템간 연동이나 정보공유가 가능하도록 하기 위함이다.

셋째, 전자책 메타데이터는 기존 메타데이터 요소를 최대한 수요함과 동시에 변화하는 환경에 적절한 요소를 수용할 수 있는 여지가 있어야 한다.

이상의 일반적인 기준을 근거로 선정된 전자책 라이브러리를 위한 메타데이터는 크게 세 그룹으로 구성되어 있다. 하나는 '핵심 기술 요소'로서 선정한 5개의 전자책 업체에서 모두 사용하고 있는 메타데이터 요소 6개가 이에 해당한다(〈표 1〉 참조). 다른 하나는 '상세 기술 요소'로서 선정한 3개의 메타데이터 표준과 국내 전자책 메타데이터 표준인 KS X 6100에서 공통으로 사용되고 있는 메타데이터 요소 11개 중 핵심 공통 요소와 겹치는 3개의 메타데이터 요소-Title, Author, Publisher-를 제외한 8개이다(〈표 2〉 참조).

메타데이터간 비교표를 통해 선정한 14개 메타데이터 요소는 전자책 라이브러리를 위한 최소한의 메타데이터이다. 그 외에 전자책의 고유 특성을 나타낼 수 있는 요소나 기타 사용 목적에 따라 확장하여 사용할 수 있는 '추가 기술 요소'로서 모두 12개이다.

1) 조윤희, 2003. 디지털콘텐츠 메타데이터 포맷의 비교 연구 『한국문헌정보학회지』 37(2): 135-152.

〈표 3〉 전자책 라이브러리를 위한 메타데이터 요소

	메타데이터 요소	반복여부	필수/선택
핵심 기술 요소	<ul style="list-style-type: none"> - 서명(title) - 저자(author) - 출판사명(publisher) - 가격(price) - 표지 이미지(image) - 목차(toc) 	반복	필수
상세 기술 요소	<ul style="list-style-type: none"> - 기타 에이전트(other agent) - 연도(date) - 포맷(format) - 식별기호(identifier) - 관계(relation) - 출처(source) - 언어(language) - 기술(description) 	반복	해당시 필수
추가 기술 요소	<ul style="list-style-type: none"> - 출판일(publish date) - 파일유형(type) - 파일크기(size) - 지원리더(reader) - 내용소개(abstract) - 저자소개(author description) - 서평(review) - 미리보기(preview) - 시스템사양(system) - 주제(subject) - 대상독자(audience) - 판사항/버전(version) 	반복	선택

4.3.1 핵심 기술 요소

핵심 기술 요소는 총 6개이며 선정된 5개 전자책 서비스업체에서 공통으로 제공하고 있는 메타데이터 요소이다. 각 메타데이터 요소는 '서명(title)', '저자(author)', '출판사명(publisher)', '가격(price)', '표지 이미지(image)', '목차(toc)'이다. 이 요소는 필수 항목으로 반복 사용이 가능하다.

4.3.2 상세 기술 요소

상세 기술 요소는 총 8개이며 선정된 3개의 메타데이터 표준인 더블린코어, MARC, TEI Header 그리고 국내 전자책 메타데이터 표준인 KS X 6100을 비교하여 공통적으로 사용되고 있는 8개의 메타데이터 요소이다. 공통으로 사용되고 있는 메타데이터는 총 11개였으나, 그 중 '서명', '저자', '출판사명'은 핵심 기술 요

소에 포함되므로 이 세 요소를 제외한 8개만이 해당된다. 각 메타데이터 요소는 '기타 에이전트(other agent)', '연도(date)', '포맷(format)', '식별기호(identifier)', '관계(relation)', '출처(source)', '언어(language)', '기술(description)'이다. 이 요소는 해당 시 필수 항목으로 반복 사용이 가능하다.

4.3.3 추가 기술 요소

추가 기술 요소는 총 12개이며 핵심 기술 요소와 상제 기술 요소에서 출현하지 않은 전자책 고유 특성을 반영할 수 있는 메타데이터 요소이다. 각 메타데이터 요소는 '출판일(publish date)', '파일유형(type)', '파일크기(size)', '지원리더(reader)', '내용소개(abstract)', '저자소개(author description)', '서평(review)', '미리보기(preview)', '시스템사양(system)', '주제(subject)', '대상독자(audience)', '버전(version)'이다. 이 요소는 선택 항목으로 반복 사용이 가능하다.

4.4 KORMARC 과의 비교

본 연구에서 개발한 전자책 메타데이터와 국내 도서관에서 실제로 사용하고 있는 KORMARC 레코드와 비교를 시도하였다.

비교를 위해 다음사항을 고려하였다. 첫째, 개발된 전자책 메타데이터를 중심으로 일대일 또는 일대다로 KORMARC 레코드를 연결하여 비교하였다. 둘째, KORMARC에서 전자책을 지원하는 레코

드를 중심으로 연결하여 비교하였다. 셋째, KORMARC의 리더와 디렉토리는 비교 범위에서 제외하였다. (단, '파일유형(type)'의 경우는 제외)

비교에 사용한 KORMARC은 국립중앙도서관에서 발간한 한국문헌자동화목록형식 단행본용(1993), 비도서자료용(1996)과 연속간행물용(1995)을 참조하였다.

이를 바탕으로 제안한 전자책 메타데이터 요소와 KORMARC 레코드를 비교한 비교표가 다음의 <표 4>와 같다.

5 결 론

본 논문에서는 전자책 콘텐츠 및 메타데이터의 효율적인 관리 및 검색을 위하여 전자책 라이브러리를 위한 메타데이터를 개발하였다. 현재 전자책 서비스업체에서 사용하는 메타데이터를 기반으로 더블링크어, MARC, TEI Header 그리고 KS X 6100 메타데이터 요소를 비교 분석하여 제안한 전자책 라이브러리 메타데이터는 전자책의 특성을 잘 표현하고 있다.

본 연구에서 개발한 메타데이터를 통해 전자책을 이용하고자 하는 이용자뿐만 아니라 도서관 사서 그리고 전자책 서비스업체 모두가 효율적인 검색, 열람 및 관리가 가능해질 것이다.

전자책 라이브러리의 핵심 요소인 상호운용성을 위해 개발된 메타데이터 요소와 도서관에서 사용하는 표준 메타데이터인 KORMARC 레코드를 비교하였다. 이는

〈표 4〉 전자책 메타데이터 요소와 KORMARC 레코드와의 비교

전자책 메타데이터 요소		KORMARC 레코드	
핵심 기술 요소	서명 (title)	기본표목-통일서명 축약표제 통일서명 본서명 부출표목-통일서명 부출표목-본서명과 다르게 부출 되는 서명	130 ▼a 210 ▼a 240 ▼a 245 ▼a 730 ▼a 740 ▼a
	저자 (author)	기본표목-개인명 기본표목-단체명 기본표목-회의명 서명저자사항-첫번째저자 서명저자사항-두번째이하저자 부출표목-개인명 부출표목-단체명 부출표목-회의명	100 ▼a 110 ▼a 111 ▼a 245 ▼d 245 ▼e 700 ▼a 710 ▼a 711 ▼a
	출판사명 (publisher)	발행사항-발행지, 배포지 -발행자, 배포자 -제작지, 인쇄지 -제작자, 인쇄자	260 ▼a▼b▼e▼f
	가격 (price)	국제표준도서번호-가격 기타 표준부호-가격	020 ▼c 024 ▼c
	표지 이미지 (image)	전자적 위치 및 접속	856 ▼u
	목차 (toc)	목차 내용주기	505 ▼a▼b▼c
	기타 에이전트 (other agent)	발행/배포사항	260 ▼b▼f
상세 기술 요소	연도 (date)	발행사항-발행년, 배포년, 인쇄년 발행예정일자 촬영/녹음일시와장소주기 복제주기-복제일자 동시발간저록-발행년	260 ▼c▼g 263 ▼a 518 ▼a 533 ▼d 777 ▼d
	포맷 (format)	형태사항 물리적매체 자료의 구조와 배열 컴퓨터파일, 데이터유형주기 복제주기-복제형태사항 전자적위치및접속-전자형식유형	300 ▼a▼b▼c 340 ▼a▼b▼d▼e▼f 351 ▼a▼b▼c 516 ▼a 533 ▼e 856 ▼q
	식별기호 (identifier)	ISBN ISSN 기타표준부호-ISRC -UPC -ISMN -EAN -SICI 발행자번호	020 ▼a 022 ▼a 024 0▼a 024 1▼a 024 2▼a 024 3▼a 024 4▼a 028 ▼a

〈표 4〉 전자책 메타데이터 요소와 KORMARC 레코드와의 비교(계속)

전자책 메타데이터 요소	KORMARC 레코드		
상 세 기 술 요 소	관계 (relation)	원본주기-원본서명	534 ▼t
		원저저록-서명	765 ▼t
		번역저록-서명	767 ▼t
		이판저록-서명	775 ▼t
		시스템 사항에 관한 주기	538 ▼a
		총서사항/부출표목-서명	440 ▼a
		총서사항/부출되지않거나 다르게 부출되 는 총서표제	490 ▼a
		상위총서저록-서명	760 ▼t
		모체레코드저록-서명	772 ▼t
		선행저록-서명	780 ▼t
		총서부출표목-통일총서명	830 ▼t
		형태사항-딸린자료	300 ▼e
		하위총서저록-서명	762 ▼t
		기본자료 저록-서명	773 ▼t
		후속저록-서명	785 ▼t
		이용가능한다른형태자료주기	530 ▼a
		기타형태저록-서명	776 ▼t
보유판/특별호저록-서명	770 ▼t		
비특정적관계저록-서명	787 ▼t		
인용주기-색인지, 초록지명	510 ▼a		
출처 (source)	원저작자-원표제에 관한 주가 원서명, IS SN, ISBN	507 ▼t ▼x ▼z	
	참조정보원주기	581 ▼a	
언어 (language)	언어부호	041 ▼a ▼b ▼g ▼h	
	언어주기	546 ▼a	
기술 (description)	일반주기-관계	500 ▼b	
	합철본주기	501 ▼a	
	학위논문주기	502 ▼a ▼b ▼c ▼d	
	서지내력주기	503 ▼a	
	서지주기	504 ▼a	
출판일 (publish date)	발행사항-발행년, 배포년, 인쇄년	260 ▼c ▼g	
추 가 기 술 요 소	파일유형 (type)	리더06-07, 국립중앙도서관 서지레코드 제어번호	
	파일크기 (size)	컴퓨터파일특성	256 ▼a
		형태사항	300 ▼a
		전자적위치및접속-파일크기	856 ▼s
	지원리더 (reader)	일반주기-자료의 특성 사항	500 ▼s
	내용소개 (abstract)	요약, 초록, 해제주기	520 ▼a ▼b ▼c
	저자소개 (author des cription)	-	-
	서평 (review)	표제저작자사항-잡정보	245 ▼c
미리보기 (preview)	전자적 위치 및 접속	856 ▼u	

〈표 4〉 전자책 메타데이터 요소와 KORMARC 레코드와의 비교(계속)

전자책 메타데이터 요소		KORMARC 레코드	
추가 기술 요소	시스템사양(system)	시스템 사항에 관한 주기 컴퓨터파일로 접근하는 시스템 세목	538 ▼a 753 ▼a▼c
	주제(subject)	미국국회도서관청구기호	050 ▼a
		국립중앙도서관청구기호	052 ▼a
		한국십진분류기호	056 ▼a
		국제십진분류기호	080 ▼a
		듀이십진분류기호	082 ▼a
		기타 분류기호	085 ▼a▼b
		주제명부출표목-개인명	600 ▼a
		주제명부출표목-단체명	610 ▼a
		주제명부출표목-회의명	611 ▼a
주제명부출표목-통일표제	630 ▼a		
주제명부출표목-일반주제명	650 ▼a		
대상(audience)	이용대상자 주기	521 ▼a	
판사항/버전(version)	판사항-판표시	250 ▼a▼b	

도서관 측면에서 보면 작성된 비교표를 통해 본 연구에서 개발한 전자책 메타데이터의 이해를 도울 수 있게 하였다.

본 논문에서 개발한 전자책 메타데이터를 통해 전자책 서비스업체는 일관된 전자책 정보의 등록이 가능하고, 이용자는 원문 정보까지 확대하여 정교한 검색을 할 수 있다. 또한 도서관은 전자책에 대한 표준 목록 데이터를 다운로드 받아 사용함으로써 간과되거나 구체적으로 처리되지 못하였던 전자책 정보 구축과 고도의 전문성을 가진 분야에 대한 지식 정보 까지도 제공할 수 있어 사서들의 업무 효율화를 가져올 수 있다.

향후 연구에서는 더블링크어 외에도 SOIF, IAFA template, GILS, RFC1807, PICS 등 다양한 네트워크 자원의 기술방법과 비교하는 것을 고려해야 한다. 더 나아가 KORMARC 뿐만 아니라 MARC

21, DC, DC21, ONIX 등 다양한 표준 메타데이터 요소로의 변환에 관한 연구가 필요하다.

참 고 문 헌

- 김진아, 김태수. 2001. 동영상 정보의 메타데이터 구축에 관한 연구, 『情報管理學會誌』, 18(4): 15-35.
- 문화관광부. 2000. 『한국전자책(eBook) 산업 발전방안 연구』. 연구보고서.
- 손원성, 고승규, 이경호, 김재경, 김성혁, 임순범, 최윤철. 2001. 한국 전자책 문서표준(EBKS)의 개발. 『情報管理學會誌』, 18(2): 255-271.
- 이민석. 2002. 『eBook』, TTA Journal 제81호: 73-78.

- 이재선. 2002. KORMARC의 현재와 미래. 『KORMARC. KOLASII. KOLIS-NET』(제40회 전국도서관대회 워크숍 자료집, 2002. 9, 대구전시컨벤션센터): 1-18.
- 조윤희. 2003. 디지털콘텐츠 메타데이터 포맷의 비교 연구. 『한국문헌정보학회지』, 37(2): 135-152.
- Heery. 1996. "A review of metadata: a survey of current resource

description formats"
<[http:// www.ukoln.ac.uk/metadata/desire/overview/review_toc.htm](http://www.ukoln.ac.uk/metadata/desire/overview/review_toc.htm)>.

- Open eBook Publication Structure Specification Version 1.2. 2002. NewYork: OeBF
<[http:// www.openebook.org/oebps/oebps1.2/index.htm](http://www.openebook.org/oebps/oebps1.2/index.htm)>.

K C I