

전자형태 학위논문을 위한 표준 메타데이터 설계

Design on the Standard Metadata for the E-Dissertation

남 영 준(Young-joon Nam)*

목 차

| | |
|---------------------|-----------------------|
| 1. 서론 | 3. 1 전자형태 학위논문의 관리현황 |
| 2. 학위논문의 특성 및 관리 | 3. 2 전자형태 학위논문용 메타데이터 |
| 2. 1 학위논문의 특성 | 3. 3 국내 학위논문용 메타데이터 |
| 2. 2 학위논문의 수집 | 비교 |
| 2. 3 학위논문수집의 한계 | 3. 4 학위논문 메타데이터 표준설계 |
| 3. 전자형태 학위논문의 메타데이터 | 4. 결론 |

초 록

본 연구에서는 국내 전자형태 학위논문의 메타데이터 표준요소를 제안하였다. 제안의 근거는 국내 주요 대학의 메타데이터 표준요소와 NDLTD, 학위논문·원문 공동이용협의회에서 채택하고 있는 표준기술요소를 비교한 데이터를 활용하였다. 이와 같은 표준은 이용자에게 이종간 전자도서관에 수록된 학위논문 자료에 대한 통합정보검색 편의성을 제공할 것이다.

제안한 요소는 더블린 코어의 9개 요소와 학위논문이 갖는 특정요소를 고려한 4개 요소 등 14개 메타요소를 선정하였다. 특히 새로이 확장한 메타 요소는 'thesis.degree.name'을 비롯하여 'thesis.degree.level', 'thesis.degree.discipline', 'thesis.degree.grantor' 등이다.

ABSTRACT

This study aims to develop a standard set for metadata to describe E-dissertation. This standard is designed for interoperability and for convenience of integrated information retrieval. This will allow one to improve existing federated searches, create united databases, and provide greater consistency for researchers searching for thesis and dissertation at different institutions. The fourteen meta elements—nine from Dublin Core and four from local need—are proposed along with newly expanded meta elements, 'thesis.degree.name', 'thesis.degree.level', 'thesis.degree.discipline' and 'thesis.degree.grantor.'

키워드: 학위논문, 메타데이터, 더블린코어
E-dissertation, metadata, Dublin Core

* 중앙대학교 문헌정보학과 교수(namyj@cau.ac.kr)
논문접수일자 2005년 8월 15일
게재확정일자 2005년 9월 15일

1. 서론

전통적으로 국가 차원에서 자국에서 생산되는 학위논문에 대해 완전한 서지데이터를 구축하고, 이와 함께 물리적 형태의 논문도 망라적으로 수집하는 정책을 집행하고 있다. 왜냐하면 각 대학에서 생산하는 학위논문은 또 다른 논문을 생산하는데 중요한 데이터로써 학술적 활용도가 높기 때문이다. 이에 따라 학위논문에 대한 통합관리와 디지털화에 대한 요구도 상대적으로 높아지고 있다. 우리나라의 경우에도, 국립중앙도서관을 비롯하여 국회도서관, 한국교육학술정보원, 서울대학교 중앙도서관 등이 각기 주관이 되어 국내 대학에서 생산하는 석·박사 학위논문 등을 망라적으로 수집하고 웹을 통한 전면적인 서비스를 실시하고 있다. 미국의 경우, UMI는 미국에서 생산하는 모든 학위논문을 체계적으로 수집하여 다양한 형태의 출력물로 이용자들의 요구사항을 충족시키고 있다. 이와 함께 NDLTD(Networked Digital Library of Theses and Dissertations)는 미국 내 주요 대학에서 생산하고 있는 석·박사학위논문을 총체적으로 수집하여, 수집한 논문을 온라인을 통해 자유롭게 열람할 수 있는 통로를 제공하고 있다. 이와 같이 많은 국가들은 자국에서 생산되는 석·박사학위논문을 국가 지식자원의 일부로 간주하여 이를 체계적으로 수집하고, 그 지적 소유물을 최적의 형태로 활용할 수 있도록 시스템을 구축하고 있다.

이와 같은 통합수집과 제공을 위해서는 이에 대한 표준이 우선되어야 하며, 그 가운데 웹기반 검색시스템에 적합한 표준 메타데이터의 개발이 시급하다. 특히 석·박사학위논문의 발행

시기와 형태는 거의 일정하기 때문에 서지기술을 위한 표준 메타데이터의 통합적 효과는 다른 메타데이터에 비해 상대적으로 크다고 할 수 있다. 따라서 국내외 지식자원 관리의 효율성과 온라인상의 이용자 연구·열람 편의성을 지향하기 위해 석·박사학위논문 통합관리의 필요성이 증가되고 있다.

통합관리를 위해서는 각 대학의 자원공유에 대한 개방적인 정책과 통합관리를 위한 표준안이 우선적으로 필요하다. 따라서 지식의 공유를 목적으로 각 대학이 생산하고 있는 학위논문들을 통합할 수 있는 대학 간의 협력과 합의점이 도출되어야 한다. 이를 위해서는 학위논문을 최적의 형태로 표준화할 수 있는 메타데이터의 개발을 비롯하여 각 대학의 현황과 의견을 포괄적으로 수용할 수 있는 방안을 마련해야 한다.

표준의 관점은 학위논문의 디지털 포맷과 국가차원의 학위논문 서지데이터베이스를 구축할 수 있는 메타데이터로 구분할 수 있다. 전자는 기술적 측면으로써 기술의 발전에 따라 변화될 수 있는 것이지만, 후자의 경우는 학위논문이 갖는 고유한 특성(형태적, 대학 정책적)이 변하지 않기 때문에 항구적인 표준이라 할 수 있다. 이에 따라 본 연구에서는 국내에서 생산되는 학위논문의 특성을 수용하고, 국가 차원의 통합서비스를 위해 필요한 표준메타데이터를 제안한다. 이를 위해 우선 학위논문의 통합관리 방안에 대한 이론적 연구를 문헌적 조사를 통해 수행하고자 한다. 이를 위해 국내의 학위논문을 통합적으로 관리하는 기관의 표준과 함께, 국내 대학에서 현재 웹을 통해 서비스하는 온라인열람목록의 메타요소와 실태 현황을 조사하여 분석한다.

2. 학위논문의 특성 및 관리

학위논문은 판매와 배포를 의도하지 않는 대표적인 회색문헌이다. 학위논문은 해당 분야에서 제출자들이 일정한 지적 수준에 도달한 것을 입증하기 위한 연구결과이다. 이러한 관례는 대부분의 국가와 대학에서 동일하게 적용되기 때문에 논문의 제출절차와 포맷은 유사하다. 본 장에서는 학위논문이 갖는 특성과 관리현황을 분석한다.

2.1 학위논문의 특성

우리나라는 고등교육법시행령(대통령령 제 17796호)에 의거하여 학위논문을 제출함으로써 일정 학위를 수여받는다. 동법에서는 심사위원의 엄격한 심사와 공간을 명시하고 있어 학위논문이 갖는 학술적 가치를 부여하고 있다. 즉, 학위논문은 독창성을 갖고, 기술에 있어 정확성을 요구하는 점, 3인내지 5인의 심사위원의 엄격한 심사가 이루어지는 점 때문에 학위논문은 학술 및 연구적 가치를 갖는 1차 자료라 할 수 있다. 이와 같이 일련의 심사규정을 가짐으로써 논문의 저작에 지도교수의 참여가 반드시 이루어져야 하며, 저작자와 함께 중요한 의미를 갖는다. 이러한 특성 때문에 조인숙은 학위논문을 연구보고서의 유형으로 간주하고 있다(조인숙 1993). 또한 출판의 대상이 제한된 범위내에서 이루어지기 때문에 회색문헌과 같은 특성을 갖는다.

학위논문이 갖는 특성을 다음과 같이 서지기술 부분과 출판 부분으로 구분하여 설명할 수 있다.

1) 서지기술적 특성

- 저작은 반드시 독자적으로 이루어지기 때문에 항상 단독저자이며, 저자명 기입형식이 일정하다.
- 학위논문의 주제영역은 학과와 전공이 표시됨으로써 다른 자료에 비해 분명한 주제 분류가 가능하다.
- 학교나 학과, 전공에 따른 출판의 주체가 상이하여도 서지기술요소는 거의 동일하다. 또한 이러한 서지기술요소는 과거에 출판되었거나 미래에 출판될 논문도 동일한 구조를 갖는다. 즉 서지기술요소는 출판주체나 시기에 큰 영향을 받지 않는다.
- 자료조직적인 측면에서 학위논문은 목록기술형식, 작성카드의 종류 등에서 간략화가 가능하다(조호일 1985).
- 출판형식이 획일화 되어 있어 서지정보의 위치가 동일하다.
- 저자의 99% 이상이 내국인이기 때문에 기술상의 어려움이 없고 대부분 첫 번째 저술이기 때문에 전거 통제를 위한 별도의 확인 작업이 필요 없다.
- 서명이나 저자 사항에 한자 기입이 많다.
- 서명이 전문 용어로 구성되어 있어 일반 용어보다는 오타의 가능성이 높다.
- KS CODE에서 지원되지 않는 분자식, 학명, 기호, 부호, 특수문자 등이 많다.
- 영문 대등서명이 많다.
- 서명 내 설명을 위한 괄호 사용이 많다.
- 출판년도와 학위년도가 구분되지 않고 혼용되어 있다.
- 물리적 형태사항의 구성이 거의 동일하다.
- 페이지 기입의 유형이 동일하다(서문 페이지

지가 적고 본문 페이지의 기입이 명확하다).

2) 출판적 특성

- 논문의 발행시기가 거의 동일하다.
- 공식적인 유통경로를 확보하지 않으며, 발행에 따른 정보를 확인하기가 일반 단행본에 비해 상대적으로 어렵다.
- 학위논문의 작성은 대부분 PC에 의해 출판이 이루어짐으로써 최소한의 경비로 디지털화가 상대적으로 용이하다(김남석, 조도희 1995).

이와 같은 특징 때문에 학위논문은 단행본을 비롯하여 다른 자료와 분리하여 별치자료로 취급하는 것이 불가피하며, 학위논문자료의 조직을 위한 이론과 기술을 일반 단행본과는 달리 자료의 입수, 정리, 열람 등에 보다 효율적으로 대처키 위한 방안이 마련되어 적어도 한국내의 대학도서관들간에도 통일성 있는 방법이 마련되어야 한다(김남석, 조도희 1995).

3) 형태적 특성

국내 대학에서 생산되는 학위논문은 학교나 전공에 관계없이 다음과 같은 형태적 특성을 갖는다.

- 학위논문은 표지에 지도교수명과 제목, 저자명, 대학 및 전공명, 학위수준 등과 같은 구조요소를 기술하고 있다.
- 내지를 포함한 본문에는 해당 논문에 대한 대등서명(영어나 한글), 초록(한글 및 영문), 참고문헌이 수록되어 있다.

2.2 학위논문의 수집

각 대학별로 학위논문을 관리하는 기관은 상이하나 최종적인 관리주체는 소속 대학도서관이 일반적인 현상이다. 납본법에 의거하여 각 대학에서 생산하는 학위논문은 출간과 함께 국가기관에 납본을 수행하고 있으나, 현재 국내 어느 기관도 완전한 납본을 이루지 못한 것으로 조사되었다. 이는 여러 가지 이유가 있으나 납본의무자체가 대학의 입장에서 부가적인 업무 부담으로 인식하고 있으며, 또한 납본을 의뢰하는 기관이 상대적으로 많기 때문이다. 이러한 현상에도 불구하고 국가지식의 활용과 관리라는 측면에서 학위논문의 망라적 수집은 국가수준에서 이루어져야 한다. 다음은 학위논문의 형태적 구분에 의해 수집된 방법을 구분한 것이다. 국내 주요기관에서 이루어지는 학위논문의 물리적 형태와 전자 형태의 수집

2.2.1 물리적 형태 학위논문 수집

우리나라에서 주요 기관에서의 학위논문 수집은 법에 의한 납본의 형태로 이루어지고 있다. 납본과 관련된 주요 관련법은 다음 <표 1>과 같다.

우리나라에서는 이와 같은 법적 규정이 있음에도 불구하고 완전한 납본이 이루어지지 않는 것은 완전한 법적 규정과 그에 따른 제재 조치가 명확하게 제시되지 않고 있기 때문이다.

예를 들면, 미래지향적인 도서관법개정시안의 동법 18조에 따르면 학위논문을 납본할 의무가 명확하게 규정되어 있지 않기 때문에 각 대학들은 이 규정의 준수 여부를 의무감이나 강제적 규정으로 인지하지 않고 있다. 또한 법 자

〈표 1〉 우리나라 납본 관련법 일부

| 법령명 | 조문 | 납본처 | 납본대상 | 제출시한 | 법적 규정 |
|------------|------|-----------|-------|-------------|-------|
| 도서관및독서진흥법 | 제17조 | 국립중앙도서관 | 모든 도서 | 발행후 30일 이내 | 강제규정 |
| | 제30조 | 공공도서관 | | | |
| 도서관법개정시안 | 제18조 | 국립중앙도서관 | 모든 도서 | 발행 후 30일 이내 | 강제규정 |
| 국회도서관법 | 제7조 | 국회도서관 | 모든 도서 | 발행후 30일 이내 | 강제규정 |
| 한국교육학술정보원법 | 제18조 | 한국교육학술정보원 | 모든 도서 | 없음 | 입의규정 |

체적으로 발행 또는 제작일을 기점으로 30일 이내에 국립중앙도서관 납본을 의무화하고 있으나, 이에 대한 강제규정이거나 처벌규정이 없기 때문에 이의 수용 여부도 상대적으로 자유롭게 대처할 수 있다. 또한 각 대학의 논문 출간일은 6월이나 12월로 기록되어 있으며, 정식 졸업은 8월이나 익년 2월이기 때문에 30일 이내에 해당 자료를 국립중앙도서관에 납본하는 것은 현실적으로 실행이 불가능하다. 따라서 각 대학들의 실제 납본은 대학의 사정에 따라 자유롭게 이루어지고 있는 실정이다.

2. 2. 2 전자 형태의 학위논문 수집

대부분의 대학도서관에서 수행하는 납본은 도서관의 업무가중의 새로운 형태이다. 현재 많은 정보관리기관에서 각 대학도서관을 대상으로 중복적으로 학위논문에 대한 납본을 종용하고 있다. 한편 전자도서관의 출현에 따라 요구받는 납본의 대상이 인쇄형 자료에서 전자 형태의 학위논문으로 변화되고 있다. 전자 형태의 납본율과 활용율을 높이기 위한 방법으로 기관별로 수집방법을 달리하고 있다.

- 국회도서관은 학위논문 수집의 망라성과 효율성을 높이기 위해 전자정보교류협력을 4년제 대학 194개 학교와 맺고, 자료 및 데이터베이스 교류를 수행하고 있다.

특히 협정 대학과 국회도서관이 각자 구축한 학위논문 관련 원문 데이터의 열람을 원할 경우에, 국회도서관과 협정 대학은 원문 및 서지데이터를 상호 제공하도록 함으로써 전자 형태의 학위논문을 국회도서관에서 수집하고 있다. 특히 희망하는 대학도서관의 경우, 물리적 형태의 자료를 해당 대학으로부터 이관받아 국회도서관에서 전자형태의 학위논문을 구축하는 방법을 통해 전자형태 학위논문을 수집하고 있다.

- 한국교육학술정보원(KERIS)은 대학에서 생산되는 학위논문을 망라적으로 수집하기 위해 '학술논문DB 구축' 사업을 2000년부터 시작하여, 2001년부터 2003년까지는 '국가학술연구DB 구축' 사업을 수행하였다. 원문DB 구축 대상으로는 1차 사업에서 학술논문을, 2차 사업에서는 학술논문 외에 학위논문까지 포함하였으며, 3차 사업에 자동 생성 및 유통체계 시스템을 개발하게 되었다. 이 사업은 2003년 5월부터 11월까지 진행되었으며, 생성 및 유통체계 시스템 개발은 TG인포넷이 맡았다.

이 사업의 목적은 대학 및 교육 유관기관 등 지식 생성기관에서 생산하는 각종 학술자원을 온라인으로 수집, 관리하는 시스템을 개발하는

것이다.

이러한 취지에 따라 한국교육학술정보원은 각 대학 등에서 생산하고 있는 학위논문을 자동적으로 등록할 수 있는 D-Collection 시스템을 개발하여, 자동 학위논문 수집 대상기관으로 다음과 같은 9개의 기관을 선정하여 사업을 수행하고 있다.

시범 서비스 대상으로 KERIS를 비롯하여 교육인적자원부, 한국교육개발원, 한국직업능력개발원, 한국교육과정평가원의 5개 교육 유관 기관과 부산대, 이화여대, 충남대, 성균관대의 4대 대학으로 총 9개 기관이다. 구축 대상 자료는 기관의 특성에 따라 구분하여 수집한다 (KERIS a, 2003).

- 대학 : 학위논문, 학술지논문(대학 발간), 기타 학술자료 등
- 교육 유관기관 : 연구보고서와 정기간행물 등

2. 3 학위논문수집의 한계

국내에서 발행되는 학위논문의 완전한 수집을 완성하기에는 많은 어려움이 수반된다. 이를 어렵게 하는 요소로는 크게 법적인 한계와 대학도서관의 판단으로 구분될 수 있다.

2. 3. 1 법적 제약

전자형태의 학위논문 수집에 어려움이 존재하는 것은 저작권법과 납본관련법의 애매성에 기인한다. 저작권법적으로 학위논문의 권리자는 일차적으로 해당 논문을 제출한 자로 정의할 수 있으나, 실제 이를 지도한 지도교수의 역할에 대한 새로운 해석이 가능하다는 점과 대부

분 국내대학에서는 전자형태의 학위논문 제출을 학교내규로 정하여, 저자의 일부 저작권(열람, 배포권 등)을 학교 당국에 양도하도록 하는 임의 규정서를 내도록 하고 있다. 이는 졸업을 담보한 일련의 강제성을 갖기 때문에 졸업자로서의 약자의 입장을 강압하는 것으로 해석될 수 있다. 즉 언제든지 저작자의 권한을 침해한 것으로 해석될 수 있기 때문에 저작자의 망라적인 동의 확보가 현실적으로 불가능하기 때문에 학위논문 수집의 한계가 될 수 있다.

한편 인쇄형태의 학위논문을 디지털화하여 전자형태의 학위논문이 될 경우에 발행일이나 출간일이 각 대학별로 일주일에서 일년이라는 차이(국회도서관 2004)를 보이고 있기 때문에 30일 이내라는 법적 규정이 유명무실할 수 있다. 또한 전자 형태의 학위논문이 인쇄 형태의 학위논문을 근거로 파생된 독립적인 저작물이기 때문에 별도의 납본대상물로 간주하고 이 형태도 납본이 되는지에 대한 것도 법적으로 명시되어야 한다. 즉 인쇄형태의 자료를 납본함으로써 납본이 완료된 것인지 아니면 별도의 자료로 간주하고 인쇄물의 형태와 동일하게 새로운 납본과정을 거쳐야 하는 것인지를 법적으로 명시할 필요가 있다.

한편 디지털화 자료도 물리적 자료와 동일하게 30일 이내에 납본을 해야 한다면, 국회도서관이나 한국교육학술정보원에서 인쇄본 자료를 수집하여 동 기관이 자신의 자산을 이용하여 학위논문을 전자형태로 변환할 경우에, 디지털화한 자료를 국립중앙도서관에 납본해야 할 의무가 발생한다. 또한 국회도서관도 국회도서관법 제7조 제1-2항에서 강제규정으로 30일 이내에 납본을 강제하고 있다. 따라서 전자

형태의 학위논문을 새로운 납본대상으로 할 경우에 국립중앙도서관이나 혹은 한국교육학술정보원에서 구축한 전자형태의 학위논문파일은 서로 교환 납본해야 하는 법적 해석이 가능하다. 즉, 납본에 따른 주체와 활용도가 명시되지 않은 법적 조건 때문에 오히려 국가 수준이 어려워지고 있다.

2. 3. 2 대학도서관의 판단

학위논문의 관리에 따른 일차 주체는 대학도서관이다. 현재 대학도서관은 법에서 규정하는 납본 의무에 따라 국가기관에 납본을 수행하고 있다. 최근에 이르러 납본은 전자파일의 형태로 이루어지는 추세이며, 해당 자료를 각 도서관에서 운영하는 홈페이지를 통해 자체적으로 원문서비스를 제공하고 있다. 이는 각 대학도서관이 해당 자료에 대한 권한을 행사하려는 의지로 판단된다. 예를 들면, 전자형태 자료에 대한 국가 기관의 처리방안에 대해 설문을 한 결과(한국정보관리학회 2004), 절대 다수의 대학도서관(조사대상의 72%)이 '자관의 도서관에서 디지털 원문자료를 제공하고, 외부 기관(예를 들면, 국회도서관 등)에는 단순히 링크 정보만을 희망한다'로 응답하였다. 또한 자관은 서비스를 제공하지 않고 외부 기관을 통해 자관의 학위논문을 열람하는 것을 희망하는 도서관은 조사대상 가운데에서 한군데도 희망하지 않았다. 이와 같은 조사에서 확인할 수 있는 것은 학위논문의 열람서비스는 각 도서관의 고유 업무로 인식하고 있으며, 자관의 자산이라는 인식이 일반적이란 것을 알 수 있었다. 즉 외부 기관에 대한 전자형태 파일의 납본과 서비스라는 것에 대해 매우 부정적인 것으로 나타났다.

따라서 통합에 대한 필요성을 절실하게 인식하지 않고 있기 때문에, 각 대학도서관의 독특한 정책과 특성에 따라 전자형태의 학위논문 기술 방법을 각 대학의 목록정책에 따라 임의적으로 적용되고 있다. 그렇기 때문에 학위논문에 대한 표준적인 지침이 마련되지 않을 경우에 각 도서관에서 임의적인 편목규정을 적용하고, 궁극적으로는 향후 학위논문의 국가종합목록 작성이 상대적으로 어려워질 수 있다

3. 전자형태 학위논문의 메타데이터

국내에서 발행되는 대부분의 학위논문은 워드프로세서를 통해 저자들이 직접 작성하고 있다. 또한 저자로 하여금 학교당국에 제출하는 학위논문의 형태도 책자형 논문과 함께 전자형태의 논문을 제출하도록 하고 있다. 이는 대학도서관이 책자형 논문과 전자형태의 논문을 의도한 관리방안을 개발해야 함을 의미한다.

3. 1 전자형태 학위논문의 관리현황

국내 대부분의 대학도서관은 자관에서 생산하는 학위논문을 물리적인 형태와 함께 전자형태로 변환하여 원문을 제공하고 있다.

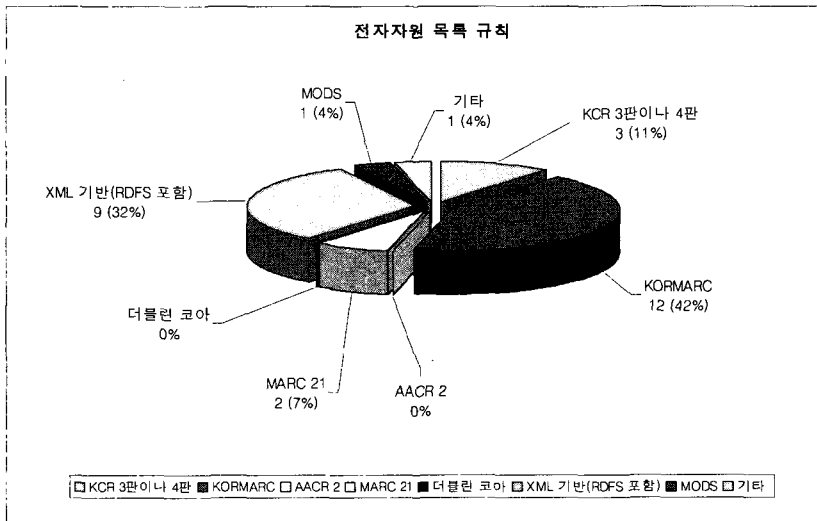
한편 우리나라는 한국목록규칙 4판을 통해 전자자료의 편목을 위한 기술방법을 규정하고 있으나, KORMARC에서는 전자자료의 편목을 위한 기술방법이 아직 표준으로써 개발되지 않은 상태이다. 또한 학위논문의 서지기술방법은 단행본의 기술규칙을 준용하고 있기 때문에 실제적으로 전자형태 학위논문의 통정을 위한 편목규

칙은 각 대학도서관의 정책에 따라 임의적인 해석과 통정을 하고 있다. 즉, 대학 도서관들은 전자 형태 학위논문을 편목하는데 한국목록규칙 4판의 전자자원 기술규칙 가운데 일부를 차용하거나, 단행본의 이형(異形)으로 간주하여 KORMARC 단행본 기술규칙을 적용하고 있다.

〈그림 1〉은 국내 주요 대학도서관 가운데 26개 대학(4년제 대학으로써 석·박사학위논문을 함께 생산하는 대학)을 대상으로 전자자원의 편목규칙을 실제 설문조사한 것이다. 이에 따르면 각 대학 도서관은 전자형태의 학위논문을 포함한 전자자원을 통정하기 위한 목록규칙으로써 KCR을 비롯해 AACR 등 물리적 형태의 기준을 채택하고 있었다. 또한 전자형태의 학위논문을 전자자원으로 간주하여 처리하는 대학도서관은 대부분 XML 기반을 채택하여 소장자원을 관리하고 있었으며, MODS와 같은 기술방식을 채택하여 사용하고 있는 것으로 조사되었다. 이는 현재 개발되고 있는 전자도서관 관련 웹프로

그램이 대부분 XML을 개발 플랫폼으로 사용하고 있기 때문이다. XML은 HTML가 비교하여 여러 장점을 갖고 있으며, 특히 문서와 DTD에 대한 표준 인터페이스를 지원하기 때문에 메타데이터관리에 매우 효율적이다(한국전산원 1999). 또한 많은 도서관이 MARC21과 KORMARC을 일련의 목록기술규칙으로 확대 해석하고 있음도 알 수 있었다.

한편 전자형태의 학위논문관련 서지기술사항이 실제로 웹상에서 적용되는 것을 확인하기 위해 학위논문원문을 웹상에서 제공하는 주요 대학교의 OPAC 항목을 분석하였다(표 2). 이 조사에 따르면 모든 조사 대상의 대학 도서관은 전자형태의 학위논문을 인쇄형태 학위논문과 동일하게 처리하고 있어, OPAC을 통해서는 전자형태와 인쇄형태를 구분하지 않고 있었다. 단 일부대학은 상세서지에서 인쇄본과 대비되는 개념으로써 원문을 열람할 수 있음을 별도의 요소를 표시하고 있었다.



〈그림 1〉 국내 대학도서관의 전자자원 목록규칙(학위논문포함)

한편 전자형태 학위논문에 대한 접근요소는 각 대학의 도서관 정책에 따라 일반단행본 기술 규칙을 기본으로 학위논문의 유형 특성을 반영하여 보다 세부적으로 기술하고 있었다. 특히 '학위논문 주기'에는 학위 구분과 학위 수여년월일을 기술하고, 지도교수명을 일반 주기에 기술하고 있다. 이와 같이 현행 대학도서관 정보 서비스 수준에서는 학위논문만이 갖는 고유한 정보적 특성을 단순하게 주기로 처리하고 있음을 보여주고 있다.

공통적인 부분으로는 '서명/저자' 사항을 비롯하여 5개 요소만이 일치하고 있었으며, 지도

교수명 등과 같은 학위논문만이 갖는 요소는 차이가 발생하였다. 학위논문을 구성한 언어와 목차 등과 같은 요소는 일부 대학에서만 고유하게 제공하고 있어 전자형태의 학위논문관련 서지 기술방법은 대학에 매우 종속적인 현상을 보이고 있었다.

3. 2 전자형태 학위논문용 메타데이터

전자형태 학위논문의 서지기술을 위해 적용하는 메타데이터는 크게 MARC와 XML 기반으로 구분할 수 있다. 국내에서 이 두 요소를 체

〈표 2〉 주요 대학의 학위논문의 OPAC 항목

| OPAC상의 화면요소(상세서지) | A대학 | B대학 | C대학 | D대학 | E대학 | F대학 | 비고 |
|---------------------------------|-----|-----|-----|-----|-----------|-----|------------------|
| 서명/저자 | ○ | ○ | ○ | ○ | ○ | ○ | |
| 대등서명 | | ○ | | | | ○ | |
| 개인저자 | ○ | ○ | ○ | ○ | ○ | ○ | |
| 저자(한자) | | ○ | ○ | | | | |
| 발행사항(발행처,년) | ○ | ○ | ○ | ○ | ○ | ○ | |
| 형태사항 | ○ | ○ | ○ | ○ | ○ | ○ | |
| 학위논문주기 (학위구분, 학과/전공, 학위수여년월) | ○ | ○ | ○ | ○ | ○ | ○ | 학위구분은 석사와 박사로 구분 |
| 수록매체원문(PDF) | ○ | | | ○ | | | PDF자료를 다르게 표현 |
| 전자정보 | | | | | | ○ | |
| 자료유형 | | ○ | ○ | | ○ | | 학위논문 |
| 일반주기 | | | | | ○ | | |
| 서지주기 | | | | ○ | | ○ | |
| 해제주기 | | | | | | ○ | |
| 제출번호 | | ○ | | | | | |
| 지도교수명 (일반주기에 기술함) | ○ | ○ | ○ | | | | |
| 동일자료(복본) | | ○ | | | | | |
| 초록 | ○ | ○ | ○ | | | | |
| 주제분류번호 | | | ○ | | | ○ | |
| 소장정보 | ○ | ○ | ○ | ○ | ○ | ○ | |
| 비통제주제어 (핵심단어 한/영) | | ○ | ○ | | ○ (한글) | | |
| 언어 | | | | | | ○ | |
| 목차 | ○ | | | | | | |

택하는 대표적인 기관의 관리 실태를 조사하면 다음과 같다.

3. 2. 1 전국대학 학위논문원문공동이용협의회 '전국대학 학위논문 원문검색'은 서울대학교가 주도하여 전국 148여 대학으로 구성된 협의체이다. 이 협의체에 가입한 대학들은 협약에 의해 자관의 원문 및 서지 DB를 공동서버에 업로드하고, 협약을 맺은 대학의 이용자들이 학위 논문을 검색하고 자유로이 원문을 볼 수 있도록 구현하였다. 업로드 대상 자원은 각 대학에서 생산한 학위논문과 목록정보이다.

1) 협의체의 목록 정책

협의체에 가입한 대학들은 소속 대학교의 학위 논문을 하나의 서버에서 통합검색을 할 수 있도록 공동 목록 규칙을 채택하고 있다. 협의체 간에 학위논문 서지정보는 MARC형태로 업로드 하되 반드시 856 tag에 원문 링크 정보를 가지는 MARC 데이터를 업로드하도록 규정을 정하였다. 특히 자관에서 이용하는 학위논문 원문 서지 DB가 856 tag의 정보를 가지지 않는 경우에도 별도의 프로그램을 통해서 856 tag를 생성한 후에 공동서버에 해당 서지정보를 업로드하도록 하고 있다. 다음의 예에서 알 수 있는 바와 같이 856 tag를 이용한 통합관리를 위해서는 많은 부분이 수작업을 상대적으로 많이 필요하고 있다는 단점을 확인할 수 있다(최현경, 이창수 2004).

- 로컬 정보 삭제 (등록번호 등)
- 청구기호 (090 tag)은 \$a에 대학명 표기
ex) 서울대 원문

- 자관 제어번호는 035tag로 대학부호를 포함해서 이동
ex) 035 \$a(211032)000000193873
- 원문정보 주소 856(URL) 필수
- KORMARC, MARC21 자료 구분

다음은 단위대학에서 공동 서버로 올려진 서지정보의 한 예이다.

예) '전국대학 학위논문 원문검색'의 예.
005199990306195923▲
008 960903s1996 ulka AK 000a kor▲
035 ▼a(211032)0000193873▲
0410 ▼akor▼beng▲
090 ▼a서울대 원문
1001 ▼a정평가, ▼d1969-▲
24510 ▼a韓國 經濟發展에서 技術集約的 産業의 役割 /▼d鄭平基.▲
260 ▼a서울:▼b서울大學校 大學院, ▼c1996.▲
300 ▼aiii, 40장 :▼b삽도 ▼c26 cm.▲
5020▼a학위논문(석사)--b서울大學校大學院:▼c電氣工學科,▼c1996.▼e오영석▲
504 ▼a참고문헌이 포함되어 있음.▲
856 ▼uhttp://library.snu.ac.kr/pdfview.jsp?bid=193873 \$y이 자료의 원문과 서지정보는 서울대학교에서 제공합니다.

2) 검색요소로써 메타 태그

'전국대학 학위논문원문공동이용협의회'의 첫 번째 사업은 각 대학에서 생산하는 학위논문에 대한 서지정보를 입수하여 모든 도서관이 완전한 학위논문 서지데이터베이스를 보유하는 것이다. 이와 같은 완전한 서지데이터베이스에 기

반하여 학위논문 원문을 웹을 통해 열람할 수 있는 하나의 학위논문 데이터베이스의 구축이 궁극적인 사업의 목표이다.

‘전국대학 학위논문원문공동이용협의회’의 홈페이지에서 공통 검색요소는 미래지향적인 메타데이터 구축을 위한 태그로 채택될 수 있다. 따라서 해당 검색요소를 분석하면 검색항목으로써 논문명을 비롯하여 저자, 전공학과, 수여기관, 주제키워드, 분류번호, 청구기호, 지도교수, 주요어를 사용하고 있다. 이외에도 소장논문을 학위수여년도, 언어, 주제분류, 학위구분(석사, 박사)을 검색의 요소로 사용하고 있었다.

특히 다음과 같은 태그(스타일)를 부착하여 서지정보를 업로드 하도록 권고하고 있어, 검색요소로서 다음과 같은 메타요소를 사용하고 있다.

- <option value="TI" selected>논문명
- <option value="AU">저자
- <option value="MA">전공학과
- <option value="ORG">수여기관

- <option value="AKWD">전체키워드
- <option value="CLA">분류번호
- <option value="CAL">청구기호
- <option value="PF">지도교수명
- <option value="KWD">주요어

3) 협의체의 메타데이터 요소

‘전국 대학교 학위논문원문공동이용협의회’는 기본적으로 제어번호를 비롯하여 KORMARC의 11개의 요소로 구성되어 있다. 이를 MARC 기술요소와 비교하면 다음 <표 3>과 같다.

이 요소들은 크게 자관에서 활용하는 부분(245, 300, 502 등)과 협의체의 공동자원 관리 측면에서 소장처 정보와 원문 데이터제공(090, 035)필드, 링크 필드(856)로 요약될 수 있다.

OPAC의 요소로서 다양한 형태로 표현이 되지만 서지정보의 공유라는 측면에서 핵심 항목을 추출하면, 245 필드와 300, 502 필드와 함께 협의체 형태의 공동자원 관리의 측면에서 소장처 정보와 원문 데이터제공(090, 035)필드, 링크 필드(856)로 요약될 수 있다.

<표 3> 전국대학교 학위논문원문공동이용협의회 필수 요소

| 필수 요소 | 대응 태그 | 비고(속성값/요소값) |
|-----------|---------------|--------------|
| 제어번호 | | 자관고유코드 |
| 서명 | 245 | |
| 저자명 | 100, 245, 745 | 대등서명(745) |
| 전공학과 | 502 | |
| 학위구분 | 502 | |
| 지도교수 | 502 | |
| 수여기관 | 502, 260 | |
| 학위수여년 | 502, 260 | |
| 청구기호 | 090, 035 | 대학교유번호, 원문표시 |
| 형태사항(수록면) | 300 | |
| 웹자원위치 | 856 | |

3. 2. 2 고려대학교

고려대학교 도서관은 XML Schema 방식을 기반으로 하는 “MODS”를 전자정보원의 기술 방식으로 채택하였다. 또한 메타데이터의 기술 정보 이외에 관리정보, 파일정보, 저작권 및 기타 정보를 인코딩하기 위해 동 기관에서 주도하고 있는 “METS” 표준안을 수용하고 있다.

1) MODS(Metadata Object Description Schema)의 개요

MODS는 미의회의 LCND(The Library of Congress' Network Development)와 MARC 표준화국(MARC Standards Office)에서 전문가를 초빙하여 서지데이터 엘리먼트셋을 위한 하나의 스키마를 개발하였다. 이는 각 도서관이 갖고 있는 다양성과 도서관의 활용에 적합하도록 고안되었다(Guenther 2004). XML 스키마의 일종으로써, MODS(Metadata Object Description Schema)는 MARC21의 서지요소를 추출할 수 있는 특징과 원본자료의 기술레코드를 생성할 수 있는 특징을 갖고 있다. MODS는 MARC 필드의 서브셋을 수용하고 숫자보다 용어기준의 태그를 사용한다. 또한 MODS는 W3콘소시움에서 사용하는 XML 스키마 언어로도 표현할 수 있다. 이에 대한 표준은 LCND와 MARC 표준화국에서 관리한다(LC 2005). MODS는 'titleInfo'를 비롯하여 19개의 요소로 구성되며, 주 엘리먼트, 자식 엘리먼트, 엘리먼트의 속성들을 정의한다. 자식 엘리먼트와 자신의 콘텐츠가 산재된 혼합 콘텐츠(mixed content)를 피하기 위해 엘리먼트의 콘텐츠는 최하위 레벨 엘리먼트에 포함된다. 예를 들어 만일 <titleInfo> 엘리먼트가 하위 엘리먼트로

<title>, <partNumber>, <partName>를 갖고 있다면, <titleInfo>는 하위 엘리먼트 들의 래퍼(Wrappe) 태그로만 사용될 뿐 어떠한 문자데이터도 포함하지 않는다. 반면 속성은 모든 레벨의 엘리먼트와 관련될 수 있고, 관련된 엘리먼트와 함께 정의된다(KERIS 2003b). 이와 같은 언어 표현의 단순성 때문에 MODS는 MARC에 비해 학위논문과 같은 학술관련 전자자원 기술에 대해 다음과 같은 장점을 갖고 있다.

- 엘리먼트 집합이 DC에 비해 풍부하다.
- 엘리먼트 집합이 도서관 데이터에 대해 ONIX보다 더 적합하다.
- 스키마가 MARCXML 스키마 보다 최종 사용자에게 더 친숙하다.
- 엘리먼트 집합이 MARC 포맷보다 간단하다.

단점으로는 MODS 스키마는 MARC 21과 완전한 상호 변환이 이루어지지 않는 점이다. 즉, MODS로 변환된 본래 MARC 21 레코드는 태그과정에서 특수성에 대한 손실이나 데이터의 손실 없이 완전히 MARC 21로 다시 변환될 수 없다. 따라서 특정한 MARC 필드는 MODS에서 좀더 일반적인 필드로 변환·적용되어지기 때문에 변환이전의 필드로 복원할 수 있는 낫값을 갖지 않는 경우도 있다(KERIS 2003b). 기존의 MARC 형태로 소장자료를 구축하고 있는 대부분의 국내 도서관에서는 변환에 대한 제한점 극복이 MODS의 범용성을 확보하는 관건이 될 수 있다.

2) 고려대학교의 학위논문 메타데이터

고려대학교 도서관은 학위논문 입력 과정에서 MODS의 상위요소(주엘리먼트) 중 14개를

채택하여 하나의 입력워크시트 폼을 개발하였다. 이 입력워크시트 폼은 인쇄형 자료와 전자형태의 자료를 모두 의도하여 개발한 것으로써, 전자형태의 학위논문의 특성을 수용할 수 있는 요소로 physicalDescription을 사용하고 있다. 다음 <표 4>는 고려대학교 학위논문의 입력 워크시트를 기준으로 서지 메타데이터 항목을 조사한 것이다.

3. 2. 3 한국전산원

한국전산원은 메타데이터로써 광범위한 주제 분야의 정보자원을 간단하게 기술할 수 있는 15개 요소로 구성된 단순 더블링크어를 권장하였다. 이 가운데 Title, Creator, Subject, Publisher,

Date, Type, Identifier의 7개 요소를 필수 요소로 제안하였다. 즉, 모든 분야별 혹은 기관별 확장 메타데이터 스키마를 개발할 때, 이러한 7개 필수 메타데이터 요소가 반드시 포함되도록 하여 통합 메타데이터 검색에 이용할 수 있도록 권장하고 있다. 7개의 필수 요소를 제외한 나머지 요소들은 향후 국가지식정보통합검색시스템에서 제공할 통합 메타데이터 검색에서는 이용되지 않지만, 각 기관 또는 분야별 메타데이터를 구축할 때 정보자원에 대한 보다 풍부한 의미를 기술할 수 있도록 각 기관별 수집 및 구축 자료의 특성을 고려하여 임의적으로 확장할 것을 권장하고 있다. 다음 <표 5>는 한국전산원에서 제안하고 있는 필수 메타데이터 요소이다(한국전산원 2003).

<표 4> 고려대학교 학위논문 입력 워크시트 폼

| 상위요소 | 하위요소 | 속성 | 비고(속성값/요소값) |
|---------------------|---|---|------------------------------------|
| Titleinfo | title subTitle | type -translated | |
| name | namePart affiliation | | |
| typeOfResource | | | "text" |
| genre | | | "학위논문" |
| originInfo | place publisher dateIssued dateCreated | | "고려대학교 대학원" 자동생성 |
| language | languageTerm | | "code" |
| physicalDescription | Form extent digitalOrigin | | "electronic" "born digital" |
| note | | | |
| subject | | | 주제 키워드 |
| classification | | authority | 고대 학위논문분류표 |
| relatedItem | titleInfo<title> originInfo<dateIssued> | type -preceeding -succeeding -host | |
| identifier | | type | 수록DB의 type 속성 |
| location | physicalLocation uri | | 고려대학교 도서관 |
| recordInfo | recordContentSource recordCreationDate recordIdentifier | | 시스템 자동생성 |

〈표 5〉 지식정보자원관리 필수 메타데이터 요소

| 필수 요소명 | 정의 | 사용 가능한 인코딩 스킴 | 요소에 대한 설명 |
|------------|------|---|--|
| Title | 표제 | | - 정보자원에 부여된 제목 - 정보자원을 공식적으로 지칭하는 이름 |
| Creator | 제작자 | | - 정보자원의 내용에 주된 책임을 갖고 있는 개체(사람 또는 단체)의 이름 |
| Subject | 주제 | - KDC - DDC - LCC - UDC - LCSH - MeSH - 자체 주제분류 | - 정보자원의 내용을 간결하게 기술한 정보자원의 주제 - 정보자원의 주제를 표현하는 키워드, 구절, 또는 분류코드로 나타내어지는 데, 통제 어휘나 공식적 분류 스킴에 의한 값을 사용하는 것이 바람직함 |
| Publisher | 발행처 | | - 정보자원을 현재의 형태로 이용 가능하게 만든 데 대한 책임을 갖고 있는 개체(개인 또는 단체)의 이름 |
| Date | 날짜 | - ISO 8601 - W3C-DTF - DCMI Period | - 정보자원이 존재하는 동안 어떤 변화가 발생한 경우의 해당 날짜 - 대체로 정보자원의 생성일이나 정보자원이 현재의 형태로 이용 가능하게 된 날짜와 관련됨 |
| Type | 자료유형 | - DCT1 - 자체 유형분류 | - 정보자원의 범주 또는 유형 - 내용의 일반적인 범주, 기능, 유형, 또는 집성 수준을 표현하는 말로서, 통제 어휘의 사용이 권장됨 |
| Identifier | 식별자 | - URI - DOI - ISBN - ISSN | - 정보자원을 식별하기 위한 고유한 문자열 또는 지시자 |

한국전산원에서 제안하는 필수 메타데이터 요소는 더블린 코어를 기반으로 하고 있으며, 향후 학위논문을 위한 메타데이터에는 이와 같은 7개의 필수 요소가 포함되어야 함을 의미한다.

3. 2. 4 NDLTD (Networked Digital Library of Theses and Dissertations)

미국의 전자형태 학위논문 데이터베이스(ETDs) 구축은 정부의 지원 하에 1990년대 중반부터 버지니아공대(Virginia Tech)를 중심으로 진행되어 왔다. ProQuest가 상업용 목적으로 운영되는 ETDs 중심의 수집 및 관리시스템이라면, NDLTD(Networked Digital Library of Theses and Dissertations)는 국가를 초월하

여 ETDs의 수집과 제공을 선도하는 비영리 시스템이다(NDLTD 2004). 전세계 연구자로 하여금 대학에서 생산되는 학위논문에 무료로 접근할 수 있도록 자발적으로 학위논문을 공개하는 것이다. 이는 학생이 자신의 논문을 직접 업로드하는 방식을 채택하기 때문에 통합관리를 위한 서지기술 및 원문접근을 위한 표준 메타데이터관리의 필요성이 절대적이다.

NDLTD는 석·박사학위논문이 갖는 특성을 고려하여 한국전산원과 동일하게 더블린 코어를 기반으로 하는 자체적인 메타데이터를 개발하였다. 전반부에 16개의 요소는 더블린 코어 요소를 원용하였으며, 후반 4개요소는 학위논문의 특성을 고려하여 임의적으로 확대한 기술 요소이다. 다음 〈표 6〉은 이와 같은 NDLTD의

〈표 6〉 NDLTD의 메타데이터 요소

| 요소 | 설명 | 비고 |
|--------------------------|---|--------------|
| dc.title | 정보원에 주어진 명칭, 논문의 제목 | 필수요소 반복가능 |
| dc.title.alternative | 논문의 다른 제목 | 선택요소 반복가능 |
| dc.creator | 정보원의 내용을 만드는 주요 실무자 | 필수요소 반복가능 |
| dc.subject | 정보원 내용에 대한 주제, 논문의 본문에 나타나는 키워드나 주제 | 필수요소 반복가능 |
| dc.publisher | 정보원의 내용을 전자형태로 변환한 주체, thesis.degree.grantor와 반드시 일치할 필요없음. | 선택요소 반복가능 |
| dc.description | 정보원 내용에 대한 설명, 논문의 경우 초록전문. dc.description은 dc.description.abstract와 바꿔 사용 가능 | 선택요소 반복가능 |
| dc.description.note | 논문 관련 추가 정보 | 선택요소 반복가능 |
| dc.description.release | 버전 설명 | 선택요소 반복가능 |
| dc.contributor | 정보원의 내용 기여자, 보통 공저자, 자문위원, 이사회 회원의 경우에 사용됨. 공저자는 dc.creator 필드에 넣는 것이 더 적합 | 선택요소 반복가능 |
| dc.contributor.role | 작업한 사람의 역할 예시) 자문위원, 이사회 회원, 회장, 부회장, 심사관 | 선택요소 |
| dc.date | 정보원 생성주기와 관련된 날짜, 논문의 제목페이지에 나타남. ISO 8601에 따라 표기, 더블린 코어에서 ISO 8601을 따르도록 권고함. | 필수요소 |
| dc.type | 정보원의 내용에 대한 장르나 특성. 이 필드는 다른 장르와 구별하거나 정보원에 포함된 내용 형태를 확인하기 위해 사용됨. 형태는 "Electronic Thesis or Dissertation"의 경우가 반복적으로 사용되며, http://dublincore.org/documents/dcmi-type-vocabulary/ 에서 형태를 나타내는 표준어휘 검색기능, 학위 및 교육수준 표기는 thesis.degree 필드에서 다름 | 필수요소 반복가능 |
| dc.format | 정보원의 물리적 또는 디지털 형식, 전자 논문이 저장 또는 전달될 전자 형식의 목록을 포함해야 함. 가능하면 MIME 표준 사용(등록된 MIME 타입을 살펴보면 ftp://ftp.sis.edu/in-notes/iana/assignments/media-types/media-types) 형식정보를 사용할 수 없으면 "미정(unknown)"으로 표기, 논문을 전자 형태로 사용할 수 없으며 생략. | 선택요소 반복가능 |
| dc.identifier | 정보원에 대한 명백한 참고, 다운로드나 검색될 URI 제공에 사용됨 | 필수요소 반복가능 |
| dc.language | 정보원의 지적 내용에 대한 언어, 기록된 주요 언어, 다른 언어는 언어 한정자를 사용. Global Qualifiers 참조. 언어명칭은 ISO 6392를 사용해 기록, 언어에 대한 특별한 설명이 없으면 영어임. | 선택요소 반복가능 |
| dc.coverage | 정보원의 범위, 시간연대나 지역에 사용. "수록(coverage)"의 다른 형태를 위해서는 dc.subject 사용 | 선택요소 반복가능 |
| dc.rights | 정보원에 대한 권리정보, 배포, 재판 등에 대한 상황 설명, 시간에 따른 상황변화, 저작권자로 연결 정보 | 선택요소 반복가능 |
| thesis.degree.name | 학위 명칭 | 선택요소 반복가능 |
| thesis.degree.level | 학위 수준 예시) 석사, 박사, 박사후 등 | 선택요소 반복가능 |
| thesis.degree.discipline | 연구분야, 보통 프로그램이나 학과의 명칭이 됨. | 선택요소 반복가능 |
| thesis.degree.grantor | 학위수여 기관, Authority 섹션 참조. | 선택요소 반복가능 |

메타데이터 요소를 조사한 것이다.

새롭게 추가된 4가지 목록요소는 학위명칭에 대한 것과 학위 종류, 전공 및 연구분야, 학위 수여 기관으로 구분한다. 이 4가지 요소는 단행본을 비롯한 다른 형태의 자료에는 존재하지 않는 요소로써 학위논문을 검색하는데 주요한 기준과 접근점이 될 수 있다.

3. 3 국내 학위논문용 메타데이터 비교

국내의 학위논문과 관련된 메타데이터는 크게 MARC에 기반한 것과 MODS와 같은 XML에 기반한 것, 더블린코어에 기반한 것으로 구분할 수 있다. 현재 대부분의 대학 도서관은 KORMARC에 기반하여 전자학위논문 편목업무를 수행하고 있다. 이는 오랫동안 수행되어온 관습이기 때문에 학위논문을 위한 것 외에 일반 자료의 경우도 KORMARC를 수용할 수 있어야 한다. 한편 전자형태 학위논문은 목록정보와 함께 원문정보를 제공하는 것을 목적으로 하기 때문에 이를 지원하는 메타데이터가 필요하다. 따라서 학위논문 표준 메타데이터는 KORMARC의 기술요소를 수용하며, 전자형태 자료의 원문정보와 학위논문이 갖는 형태적 특성을 모두 수

용해야 한다. 즉, 전자형태 학위논문을 위한 메타데이터는 KORMARC를 수용하며 학위논문의 기술요소를 표현할 수 있어야 한다. 다음 <표 7>은 학위논문과 관련된 주요 메타데이터 요소를 상호 비교한 것이다. 위 표에 따르면 조사대상자료 가운데 모든 요소가 일치하는 것은 title을 비롯하여 creator, subject, publisher, date, type, identifier, right, language 등이 있다. 또한 확장요소로는 description의 서브셀인 description.abstract이 있다. 따라서 이와 같은 공통요소는 학위논문을 포함한 모든 자료의 기본서지기술요소라 할 수 있다. 한편 학위논문의 형태적 특성을 수용하기 위해서는 위의 공통요소 이외에 각 표준요소들은 별도의 확장요소를 사용하고 있었다. 즉 더블인 코어의 기본 요소는 일반 자료가 갖는 구조적 특성을 모두 수용하지만 학위논문이 갖는 구조적 특성(학위명, 학위수여기관 등)정보를 제한적으로 표현할 수 있다. MODS와 RDF/XMLs, KORMARC도 학위논문이 갖는 구조적 특성을 제한적으로 수용하고 있었다. 학위논문의 형태적 요소를 대부분 수용하는 것은 NDLTD 요소와 MARC 21, KORMARC 등이었다.

<표 7> 학위논문용 메타데이터 요소 비교표

| 요소 | DC | NDLTD | RDF/XMLs | MODs | Marc-21 | KORMARC |
|-----------------------------|----|-------|----------|-------------|---------|---|
| dc.title | o | o | o | o | 245a | 245 \$a \$x 246 |
| dc.title.alternative | o | o | | | 246 | 245 \$b \$x, 240 \$a, 440 \$a, 740 \$a |
| dc.title.translated | o | | | | 242 | 500 \$j \$i |
| dc.creator (contributor) | o | o | o | o (name) | 100a | 100, 245 \$d, 700, 710 |

〈표 7〉 학위논문용 메타데이터 요소 비교표(계속)

| 요소 | DC | NDLTD | RDF/ XMLs | MODs | Marc-21 | KORMARC |
|------------------------------|----|-------|--------------|-------------------------|---|---|
| dc.subject | ○ | ○ | ○ | ○ | 653a | 050, 056, 080, 082, 085, 600, 610, 611, 630, 650, 651, 653 |
| dc.description. abstract | ○ | ○ | ○ | ○ | 520a | 520 |
| dc.description.note | ○ | ○ | | ○ | any5xx field not otherwise used, e.g.,504 | |
| dc.description. release | ○ | ○ | | | | |
| dc.publisher | ○ | ○ | ○ | ○ (origininfo) | 260a+b | 260 \$ b |
| dc.contributor | ○ | ○ | | ○ (name) | 720a | 245 \$ e, 700, 710 |
| dc.contributor.role | ○ | ○ | | | 720e | |
| dc.date | ○ | ○ | ○ | ○ (origininfo) | 008positions 7-10 | 008/07-14, 260 \$ c |
| dc.type | ○ | ○ | ○ | ○ (type of resource) | leader 6&7:655 | Leader 06, 008/24-25, 256, 502 |
| dc.format | ○ | ○ | ○ | | 856q | 856y |
| dc.identifier | ○ | ○ | ○ | ○ | 856u | 856u |
| dc.language | ○ | ○ | ○ | ○ | 008positions 35-37:[548] | 008/35-37, 041 \$ a |
| dc.coverage | ○ | ○ | ○ | | 651 or 690 | 043, 045, 651 \$ y, \$ z |
| dc.rights | ○ | ○ | ○ | ○ (access condition) | 540 | 020 \$ c |
| dc.source | ○ | | ○ | | | 500, 507 |
| dc.relation | ○ | | ○ | ○ (related item) | | 247, 501, 505, 76x-78x |
| thesis.degree.level | | ○ | | | 502a | 502 |
| thesis.degree. discipline | | ○ | | | 710b | 502 |
| thesis.degree. grantor | | ○ | | | 502a: 710ab | 502,260 |
| thesis.degree.name | | ○ | | | 502a | 502 |
| genre 장르 | | | | ○ | | |
| table of contents 목차 | | | | ○ | 505 | |
| target audience 이용대상 | | | | ○ (DC의 audience) | | |
| classification 분류 | | | | ○ | | |
| location소장위치 | | | | ○ | | |
| extension 로컬정보 | | | | ○ | | |
| recordinfo 레코드정보 | | | | ○ | | |

3. 4 학위논문 메타데이터 표준설계

학위논문 메타데이터의 표준은 전자 형태와 책자 형태에 동시에 적용될 수 있어야 한다. 왜냐하면 많은 대학들이 이미 책자형 학위논문 데이터베이스를 MARC으로 구축하였기 때문이다. 또한 국가적인 통합 데이터베이스를 구축하기 위해서는 범용성을 고려해야 한다. 따라서 학위논문 메타데이터는 KORMARC을 수용하며, 전자 형태의 자료를 고려하고, 학위논문이 갖는 특성을 수용해야 한다. 한편, 더블린코어는 MARC를 비롯한 기존의 메타데이터가 네트워크 상의 정보자원을 표현하는데 구조적인 경직성으로 인하여 많은 비용과 시간이 소요되어 이를 대처할 수 있는 단순한 구조의 형식으로 제안된 것이다(Banski 2002, 전산원 2003b).

이상의 조사에서 학위논문을 편목하기 위해 메타데이터 표준은 다른 메타데이터에 비해 상대적으로 구조와 기술방법이 간략하고, 자관의 목록철학을 수용하기에도 유연한 확장성을 갖는 더블린 코어가 학위논문을 위한 메타데이터에 적합함을 알 수 있었다. 따라서 더블린 코어를 이용하여 기존의 목록규칙과 메타데이터에 대해 최소한 수정을 하고, 국내 대학도서관에서 채택하고 있는 학위논문 관련 서지요소를 수용하는 것이 바람직하다.

이에 따라 본 연구에서는 더블린 코어를 기반으로 하는 학위논문메타데이터의 표준을 설계하였다. 현재 국내 대학도서관에서 서명을 비롯하여 개인저자명, 발행사항, 학위논문주기, 소장정보 등 여섯 가지 요소를 공통적으로 활용하고 있었다. 또한 전자 형태의 학위논문을 위한 메타데이터요소와 관련된 국내외의 주요 규칙을

비교한 결과, title을 비롯하여 creator, subject, publisher, date, type, identifier, right, language, description, abstract 등이 자료구조화를 위한 기본적인 공통요소임을 확인할 수 있었다. 이상의 10가지 요소는 더블린 코어를 비롯하여 NDLTD, RDFs/XML, MODs, MARC21, KORMARC에서도 표현이 가능하다. 현재 국내 대학교의 웹에서 제공하는 OPAC 시스템의 공통요소와 목록규칙의 공통요소를 맵핑한 결과 후자의 10개 요소에 모두 수용될 수 있었다.

한편 OPAC의 학위논문주기에 세분 전개된 하위 요소인 '학위구분'과 '학과/전공', '학위수여년월일', '학위수여기관'은 NDLTD에서 전개하여 사용하고 있는 4개의 요소와 완전하게 일치하고 있다. 한편 DC의 요소인 dc.relation은 본 표준에서 제외하였다. 왜냐하면, 학위논문 특성상 관련자료에의 참조를 표현하는 요소는 상대적으로 사용도가 떨어지기 때문이다. dc.source와 dc.coverage 등과 같은 요소도 학위논문을 구조화하기 위한 요소로써 사용도가 떨어지기 때문에 제외되었다.

학위명과 학위수준, 전공학과, 학위수여기관 정보는 학위논문만이 갖는 매우 특징적인 요소로써 모든 학위논문의 표제면에 나타난 주요 정보원이다. 이는 '전국대학 학위논문원문공동이용협의회'의 필수메타요소에 포함되어 있으며, 주요 대학교의 OPAC 요소에서 선정되었다. 이 요소들은 향후 명실상부한 국가 통합 학위논문 데이터베이스를 구축할 경우에 반드시 포함되어야 할 것으로 판단되었다. 따라서 이에 대응하는 요소로써 thesis.degree.name를 비롯하여 thesis.degree.level, thesis.degree.discipline, thesis.degree.grantor 등 4개의 NDLTD 요소

명을 차용하여 선정하였다. 이와 같이 확장셀을 응용한 것은 궁극적으로 국내 학위논문자료와 함께 북미지역의 학위논문자료가 통합된 국제 통합 데이터베이스로의 발전을 고려한 것이다.

최종적으로 이러한 분석과 논리에 따라 본 연구에서는 아래 <표 8>과 같이 국내 대학교 학위논문의 전자자원 형태를 표현할 수 있는 메타데이터의 표준을 도출할 수 있었다. 이 표준은 더블린 코어를 기반으로 하고 있기 때문에 각 기관별 자료 교환의 변환에 따른 어려움은 없을 것으로 판단된다. 특히 NDLDT에서 제안한 확

장 요소는 MARC21과 연동할 수 있기 때문에 향후 국제적인 완전한 표준으로 활용될 수 있을 것이다.

4. 결론

학위논문은 고도의 연구성과로써 엄정한 심사결과에 의해 발행되는 국가핵심지식기운체 하나이다. 이러한 중요성 때문에 국가차원에서 학위논문의 관리는 지속적으로 이루어졌다. 왜

<표 8> 전자 형태의 학위논문 메타데이터 표준

| 요 소 | 설 명 | 비 고 |
|--------------------------|---|--------------|
| dc.title | 논문의 제목 | 필수요소 반복가능 |
| dc.creator | 논문의 저자 | 필수요소 반복가능 |
| dc.subject | 논문의 내용을 표기하는 주제어(키워드) | 필수요소 반복가능 |
| dc.publisher | 전자형태논문 변환기관 | 필수요소 반복가능 |
| dc.description.abstract | 초록 dc.description 으로 요소명 변경가능 | 선택요소 반복가능 |
| dc.date | 논문수여 연월(일) | 필수요소 |
| dc.type | 형태정보를 제공하며 전자논문이나 학위논문을 표기한다. 장르: 홈페이지 | 필수요소 반복가능 |
| dc.identifier | 정보원에 대한 명백한 참고, 다운로드나 검색될 URI 제공하는데 사용됨 | 필수요소 반복가능 |
| dc.language | 정보원의 지적 내용에 대한 언어. 기록된 주요 언어. 언어에 대한 특별한 설명이 없으면 영어임. | 선택요소 반복가능 |
| dc.rights | 정보원에 대한 권리정보, 시간에 따른 상황변화, 저작권자로 연결 정보 | 선택요소 반복가능 |
| thesis.degree.name | 학위 명칭 | 선택요소 반복가능 |
| thesis.degree.level | 학위 수준(석사, 박사, 박사후 등) | 선택요소 반복가능 |
| thesis.degree.discipline | 학과의 명칭이 됨. | 선택요소 반복가능 |
| thesis.degree.grantor | 학위수여 기관. | 선택요소 반복가능 |

나하면 학위논문은 자관에서 생산하는 수량이 한정되어 있고, 배포를 목적으로 하지 않는 회색문헌의 특성에 따라 이의 발행정보(서지정보)와 원문정보 입수에 많은 일반 자료에 비해 상대적으로 어려움이 많았기 때문이다.

이에 따라 대학도서관들은 새로운 형태의 웹 서비스로써 자관에서 관리하는 학위논문의 원문서비스를 웹을 통해 제공하였다. 이는 필연적으로 전자형태의 학위논문에 대한 새로운 통정 기술을 요구하는 계기가 되었으며, 통합목록을 넘어 통합원문데이터베이스를 지향하는 원동력이 되었다. 본 연구는 이러한 요구를 파악하고, 학위논문이 갖는 특성과 국내외 학위논문 메타데이터 표준을 분석하였다. 분석한 결과를 요약하면 다음과 같다.

1) 학위논문은 다른 자료에 비해 서지기술방법과 출판형태가 매우 일정한 일련의 회색문헌이다.

2) 국가차원에서 전자형태의 학위논문 수집은 국회도서관과 한국교육학술정보원에서 이루어지나, 저작권법과 납본관련법에 의해 망라적 수집이 현실적으로 어렵다. 따라서 각 대학의 단위도서관에서 이루어지는 원문데이터베이스에 대한 링크정보를 제공하는 것이 현실적인 대안이다.

3) 통합적인 학위논문관리를 위해서는 전자형태의 자료를 구조화할 수 있는 메타데이터의 개발이 필요하다.

4) 국내 대학에서 제공하는 OPAC을 분석한 결과, 학위논문 열람을 위한 메타데이터 요소로는 서명/저자사항을 비롯하여 발행사항, 형태사항, 학위논문주기, 소장정보 등이 필수요소로 조사되었다.

5) 국내외 대학 기관에서 학위논문을 위한 메타데이터 요소 가운데 공통적인 것은 더블린 코어요소를 중심으로 title을 비롯하여 creator, subject, publisher, date, type, identifier, right, language 등으로 조사되었다. 또한 확장요소로는 description의 서브셀인 description.abstract이 있었다.

이상의 연구결과를 토대로 본 연구에서는 다음과 같은 기준에 따라 전자형태 학위논문 메타데이터 요소를 개발하였다.

1) KORMARC으로 구축된 요소를 수용하되, KORMARC보다 유연하고 용이한 메타데이터 요소로써 더블린 코어를 기본 셀으로 하였다.

2) 학위논문의 구조특성을 고려한 OPAC의 공통적 요소와 국내외 주요 메타데이터 요소를 비교하여 공통적 요소를 기반으로, 국제적 협력을 고려하여 다음과 같은 14개의 메타데이터 요소를 선정하였다.

- dc.title
- dc.creator
- dc.subject
- dc.publisher
- dc.description
- dc.date
- dc.type
- dc.identifier
- dc.language
- dc.rights
- thesis.degree.name
- thesis.degree.level
- thesis.degree.discipline

- thesis.degree.grantor

본 연구는 학위논문을 위한 메타데이터의 설계로써 실제 각 대학에서 KORMARC으로 구축된 자료에 대한 변환과 함께 이용자들을 위한

새로운 접근점으로써 요소들이 활용되는지에 대한 실험은 이루어지지 않았다. 따라서 향후 제안한 표준에 근거하여 구축된 국가규모의 전자형태 학위논문 데이터베이스를 대상으로 하는 실제적 실험과 평가가 이루어져야 할 것이다.

참 고 문 헌

- 김남석, 조도희. 1995. 한국 대학도서관의 학위논문 관리 및 이용에 관한 연구. 『도서관학논집』, 23: 109-145.
- 김정현. 2000. 전자저널의 목록작성에 관한 연구. 『한국도서관·정보학회지』, 31(1): 187-208.
- 서울대학교. 2002. 『서울대학교 전자도서관 디지털 콘텐츠 구축을 위한 표준화 연구』. 서울대학교.
- 서울대학교. 2002. 『서울대학교 전자도서관 디지털 콘텐츠 구축을 위한 표준화 연구』. 서울대학교.
- 윤희윤. 1987. 大學圖書館의學位論文 蒐集 方案. 『圖書館學論集』, 14: 217-240.
- 조인숙. 1993. 정보원으로서의 학위논문 이용에 관한 연구. 『한국문헌정보학회지』, 24: 205-219.
- 조호일. 1985. 학위논문의 목록기술형식 및 검색 방법고찰. 『도서관논집』, 12: 133-160.
- 최현경, 이창수. 2004. 온라인 전자자료의 협동 목록 시스템에 관한 연구. 『한국도서관·정보학회지』, 35(1): 237-256.
- 한국전산원. 1999. SGML, XML, EDI 통합 및 연계방안. 동원.
- 한국전산원. 2003. 지식정보자원관리 메타데이터 구축 지침. 동원.
- 한국정보관리학회. 2004. 전자파일자료의 데이터베이스 구축 및 공유를 위한 표준화 연구. 연구보고서, 국회도서관.
- Tillett, Barbara. 2003. "AACR2 and metadata: Library opportunities in the global semantic web." *Cataloging & Classification Quarterly*, 36(3-4): 101-119.
- Banski, Erika. 2002. "Implementation of Dublin Core at the University of Alberta Libraries." *OCLC Systems & Services*, 18(3): 130-138.
- KERIS a. 2003. 국내 학위논문 관리의 표준모델 개발에 관한 연구. 한국교육학술정보원.
- KERIS b. 2003. 학술정보 전자자원메타데이터 개발 및 구축방안에 관한 연구. 한국교육학술정보원.
- El Zein, Laila. 2002. "Provision of online access to digital materials in the Temple university libraries: Cataloging issues." *Collection Management*, 26(3): 39-46.
- LC. 2005. MODS Schema & Documentation. Library Congress. [cited 2005. 08.15]

- <<http://www.loc.gov/standards/mods/>>.
- Marko, Lynn, and Christina Powell. 2001. "Descriptive metadata strategy for TEI headers: a University of Michigan library case study." *OCLC Systems & Services*, 17(3): 117-121.
- NDLTD. 2004. About NDLTD. [cited 2004, 11.20]
<<http://www.ndltd.org/about.en.html>>.
- Yee, Raymond, and Rick Beaubien. 2004. "A preliminary crosswalk from METS to IMS content packaging." *Library Hi Tech*, 22(1): 69-81.
- Guenther, Rebecca S. 2004. "Using the Metadata Object Description Schema (MODS) for resource description: guidelines and applications." *Library Hi Tech*, 22(1): 89-98.
- Zhang, Yin, Kyiho Lee, and Bum-Jong You. 2001. "Usage patterns of an electronic theses and dissertations system." *Online Information Review*, 25(6): 370-377.