

인터넷 자원에 대한 서지레코드 기술에 관한 연구

A Study of the Description of Bibliographic Records on the Internet

남 태 우, 중앙대학교 문헌정보학과

Nam Tae-Woo, Dept. of Library & Information Science, Chung-Ang University

이 명 규, 대불대학교 도서관

Lee Myoung-Gyu, Library, Daebul University

인터넷의 자원에 접근하기 위해서는 그 정보를 조직화할 필요성이 있으며, 조직화된 자료들은 디지털화되어 서비스될 수 있도록 기술이 발전하고 있다. 따라서, 현존하는 도서관의 편목 기술과 절차를 응용하여 인터넷 자원에 적합한 OCLC 매뉴얼과 네트워크상의 자원에 대한 목록을 메타데이터로 기술한 DUBLIN core 기술법을 비교분석하여, MARC 포맷으로의 변환을 살펴 본 바, 더블린 코아 형태로 제공되는 메타데이터 요소는 MARC포맷의 각 필드로 변환 할 수 있어야 하며, 더블린 코아와 같은 간략형 서지레코드의 요소를 더욱 개발해야 한다. 또한, 여러 유형의 메타데이터 통합에 대한 방법으로 MARC에 대한 수정 및 새로운 필드의 마련이 필요하며, 표준기술규칙의 마련으로 정확하고 대량의 데이터를 구축하며, 특히 이러한 기술법을 익힌 전문가들의 등장이 필요할 것이다.

1. 서론

인터넷을 통해서 이용할 수 있는 가치있는 정보가 많고, 또 이 정보를 이용하는 이용자들이 급속히 증가하고 있다. 따라서 이러한 인터넷의 자원에 접근하기 위해서는 그 정보를 조직화할 필요성이 있으며, 조직화된 자료들은 디지털화되어 서비스될 수 있도록 기술이 발전하고 있다.

따라서, 여기에서는 현존하는 도서관의 편목 기술과 절차를 응용하여 인터넷의 많은 자료들을 디지털화하여 서비스 할 수 있도록 서지 레코드를 생성하는데 효율적인 방법을 제시하고 있는 AACR2R 제9장에 기술된 컴퓨터 파일의 기술규칙을 응용하여 인터넷 자원에 대한 서지기술의 지침서 역할을 하는 OCLC 매뉴얼과 네트워크상의 자원에 대한 목록을 메타데이터로 기술한 DUBLIN core 기술법을 비교하며, MARC 포맷으로의 변환을 살펴보고자 한다.

2. 인터넷 자원의 기술 역할

인터넷 자원은 컴퓨터 파일에 대한 원격접근으로 이용하기 때문에, 인쇄 형태의 딸림자료가 없다. 또한 별개의 내부 파일로 이용된다고 할 지라도, 그 정보를 담고 있는 레이블이 부착된 컨테이너가 없다.

여기에서는 이러한 자원에 대한 기술과 다양한 메타데이터 기술법과 자원의 기술에서 정보제공자(surrogate) 및 메타데이터의 역할에 대해 살펴보고자 한다.

디지털 자료에 대한 미래의 서지적 접근 아래 깔려있는 기본적인 가정은, 네트워크가 되어 있는지 되어 있지 않든, 정보제공자의 요구는 미래의 정보 네트워크에서 감소하기 보다는 오히려 증가할 것이라는 점이다. 정보제공자는 실제적 자료를 기술하고 그것에 접근하는 방법을 탐색자에게 알리는 레코드를 편목/색인하고 있다. 정보제공자는

중요한 문헌의 특성과 관계성에 대한 확실성(구별성)을 풍부하게 상세히 기록하거나 문헌의 존재와 위치를 나타내는 주요한 기능을 간단하게 기록한다. 자동 데이터컬렉터(로봇)는 계속적으로 자유롭게 이용할 수 있는 정보를 수집하고 색일할 것이다. 그러나 어떤 이유에서든지, 잠재적인 자료 전개를 위한 서지적 접근 구조는 정보제공자에게 의존할 것이다.

메타데이터는 문서와 객체에 관한 문헌이다. 자료를 기술하고 그 자료가 위치한 곳을 나타낸다. 그리고 그 자료를 성공적으로 이용하기 위해서 필요한 것을 기록한다. 이러한 데이터 요소는 해당 도큐먼트나 객체에서 필드와 태그에 새겨 둘 수 있거나, 정보제공자 레코드에 기록해 둘 수 있다. 전반적으로 메타데이터는 자유로운 형식일 수 있거나, 문자 그대로 메타데이터를 구성하고 축약하는 방법을 정의한 수백가지 방안이 있는 규칙을 규정할 수 있다.

도서관, 학술 협회, 정부기관 그리고 상업적인 단체에서 생산한 다수의 메타데이터 방안이 있다. AACR2, MARC포맷, LC분류법 및 주제분석도구, 국립의학도서관, Forest Press(DDC)와 같은 방안은 범위가 크고 넓게 이용된다. 그리고 TEI-conformant texts에서 TEI header를 포함한 the Text Encoding and Initiative(TEI) Guidelines for Electronic Text Encoding and Interchange; SGML문서 유형 정의의 the Encoded Archival Description(EAD); 지도와 지리공간 자료의 고유 특성을 수용하기 위해 U.S. Federal Geographic Data Committee가 발전시킨 the Content Standards for Digital Geospatial Metadata(CSDGM) 등과 같이 여타의 방안은 전문적인 기관에서 발전시켰다.

각각의 이러한 방안은 특정한 기관의 이해, 정보 자료 필요, 도큐먼트와 같은 객체를 기술하기 위한 고유한 필요조건에서 구성되었고, 그 분야와 밀접하게 관련된 전문가들이 발전시켰다.

3. 더블린 코아

OCLC/NCSA 메타데이터 워크샵에서 데이터에

대한 간단한 자료 기술법으로 알려진 더블린 코아는 인터넷과 같은 디지털 환경에서 저자(creator)나 정보제공자가 인터넷 자원에 대해 기술할 수 있도록 최소한의 데이터 요소만을 포함한 메타데이터 기술 방안이다. 제정 당시에는 의미있는 최소의 수(13개)로 유지된 더블린 코아 메타데이터는 1997년 기술(Description)과 Rights management의 두 요소를 추가하여 15개의 요소로 구성되었으며, 쉽게 작성하고 넓게 적용하기 위해 크게 3개의 그룹-1) 접근점(서명, 주제, 식별자, 저자, other agent), 2) 식별을 용이하게 하는 정보(출판사, 일자, 객체 유형, 형식, 언어, 범위), 3) 기술된 자료와 다른 자료를 연결짓는 객체(관계, source)-을 기초로 해서 나눌 수 있다.

더블린 코아의 요소에 대한 정의를 각각 살펴보면, 주제(subject)는 '저작에서 언급되는 논제'로 정의하며, 표제(title)는 '객체의 이름', 저자(author)는 '저작의 지적내용에 일차적으로 공헌한 사람', other-agent는 '편집자, 삽화가, 사진작가 등 본문내용에 있어 부차적이지만, 중요한 역할을 하는 역할자를 기술하기 위한 곳', 발행처(publisher)는 '객체를 만드는 역할을 한 기관명', 일자(date)는 '객체가 현재의 형태를 갖추게 된 일자', 식별자(identifier)는 '객체를 확실히 구분하기 위해 사용되는 문자열 또는 숫자', 객체타입(object-type)는 '객체의 범주나 유형', 형식(form)은 '객체에 관한 데이터 표현 방식', 언어(language)는 '내용물의 언어', source는 '기술된 객체가 추출된 객체', 범위(coverage)는 '객체의 시공간적 특징을 기술하며 검색하기 위한 시공간 범위를 제시하는 요소' 등으로 정의할 수 있다.

더블린 코아 데이터 요소는 외부 참조(편목규칙 또는 전거파일)의 이용을 배제하면서 단지 본질적인 속성을 기술하며, 부가적인 전문적 정보를 포함하여 확장할 수 있다. 더블린 코아 데이터 요소는 낱말 배열이 독립적이어서 요소의 정의와 내용기술시에 특정 코딩방식에 구애를 받지 않는다. 더블린 코아 데이터 요소는 선택적(생략 가능)이고 반복적이다. 더블린 코아 데이터 요소는 일반적으로 정의를 이해한 후에 의미를 전달하기 위해 자격 부여자를 통하여 수정할 수 있다.

만일, 이러한 데이터 요소 세트가 공식적으로 인정된다면, 다시말해 하나의 기준이 된다면, 적어도 4가지 다른 용도로 쓰일 수 있을 것이다.

①저자와 출판사들은 자신들의 데이터에 메타데이터를 작성하도록 권장할 것이다. ②저작도구나 네트워크출판자들로 하여금 이러한 세트를 위한 소프트웨어를 개발토록 하여, IP가 더욱 쉽게 메타데이터를 제공할 수 있게 될 것이다. ③IP에 의해 생산된 메타데이터는 보다 상세한 목록이나 기술을 위한 기초가 될 것이다. ④만일 특정한 이익단체가 더욱 세부적인 정보를 요구할 지라도, 일반적인 코어세트가 여러 분야에서 통용될 수 있으리라는 것을 확신할 수 있다.

더블린 코아의 전문가들로부터 도움을 받은 개발자들은 표준의 메타데이터를 구체화하여 이런 유형의 정보를 만드는데 익숙하지 않은 정보 창작자와 생산자를 위해서 데이터세트를 만드는데 도움을 주는 형식을 만들 것으로 예상된다.

4. OCLC 매뉴얼

OCLC 매뉴얼은 미국 교육부 지원의 '인터넷 자원의 목록 구축' 프로젝트를 돕기위해 OCLC에서 개발한 "인터넷 자원 편목(Cataloging Internet Resources)"에 대한 안내서를 의미한다. 인터넷에서 접할 수 있는 자료의 서지 기술 생성에 적합하도록 편목규칙과 MARC 내용 표시를 조사하였다. 그 조사 결과로 인터넷 자료의 서지 기술 지침서를 1995년에 발행하고 1997년에 제2판을 발행하였다.

이 매뉴얼은 표준화된 형식으로 정확한 기술 및 접근점과 더불어 상세한 서지레코드를 생성하는 "AACR2"을 기초로 하여 지침서를 작성하였으며 이것은 복잡한 규칙의 문장이 들어 있지 않고 "AACR2R"의 수정안 및 의회도서관 규칙 해설(LCRIs)도 이용하였다. 특히, 근래에 자주 접하는 컴퓨터 파일, 전자 잡지 등을 편목하는데 있어서 적용시켜 왔던 '영미편목규칙 제2판 개정판(AACR2R)'의 제9장 '컴퓨터 파일'을 참조하였다.

또한, 컴퓨터 환경에서 서지 레코드 정보를 입력하기 위한 표준 포맷이 MARC포맷인데 미국에

서 이용되는 USMARC를 적용하였으며, 변형된 MARC는 OCLC MARC로 이용하였다.

OCLC 매뉴얼은 AACR2를 이용하여 서지기술을 규정한 코드와 태그를 고려해야 하기 때문에, AACR2의 순서대로 기술하면서 각 사항 및 항목에 MARC 표시기호와 식별기호를 함께 기술하고 있다.

우선, AACR2R과 OCLC 매뉴얼과의 큰 차이점을 살펴보면, AACR2R의 제1사항에서 기술하는 '표제 및 저자표시사항'에서 '종합표제가 없는 자료'에 대한 기술과 제5사항인 '형태기술사항'에 대한 기술이 OCLC매뉴얼에서는 필요하지 않다는 것이다.

5. 더블린코아와 OCLC매뉴얼 비교

더블린 코아와 OCLC 매뉴얼의 기술요소를 USMARC로 매핑하여 비교하면 <표 1>과 같다.

이와 같이, <표 1>에 나타난 것처럼, 더블린 코아 기술법은 OCLC 매뉴얼 및 USMARC보다는 간략한 서지 기술법이라는 것을 알 수 있다. 따라서, 더블린 코아의 요소에서 누락된 판차사항 등을 보완하여 세밀한 기술법이 되기 위해서는 메타데이터의 기술요소가 개발되어야 할 것이다.

6. 결 론

이상과 같이 서로 다른 유형의 기술법, 즉 더블린 코아와 OCLC 매뉴얼을 비교 분석해 본 결과 다음과 같이 요약할 수 있다.

첫째, 더블린 코아 형태로 제공되는 메타데이터 요소는 MARC포맷의 각 필드로 변환 할 수 있어야 한다. 오히려 세밀한 정보를 제공하기 위해서는 더블린 코아와 같은 간략형 서지레코드 보다는 이미 보편화된 MARC포맷으로 더욱 자세한 정보를 실을 수 있도록 필드를 개발해야 한다.

둘째, 더블린 코아는 접근 정보로 강조되고 있으나 원정보를 기술하는 데에는 MARC포맷보다는 부족하다. 그리고 저자가 제공하는 메타데이터 보다는 전문적으로 규정한 서지 레코드의 기술에 따라 작성된 메타데이터가 더 나올 것이다.

<표 1> 더블린 코어와 OCLC 매뉴얼의 USMARC

DUBLIN core	OCLC manual	USMARC
Subject	subject headings classification	653 \$aindex term-uncontrolled 650 subject added entry-topical term 050 LCCN, 082 DDCN
Title	9.1B title proper 9.1D parallel titles 9.1E other title information	245 \$atitle 245 \$bremainder of title 245 \$bremainder of title
Author	9.1F statement of responsibility	100 \$amain entry-personal name 110 \$amain entry-corporate name 245 \$cstatement of responsibility 700 \$aadded entry-personal name 710 \$aadded entry-corporate name
Description	9.7B17 summary	520 \$asummary
Publisher	9.4 publication, distribution	260 \$bname of publisher
Other Agent	9.7B6 statements of responsibility	536 funding information note
Date	9.4 publication, distribution	260 \$cdate of publication
Object Type	9.3B file characteristics 9.7B8 file characteristics 9.1C general material designation	256 \$acomputer file characteristics 516 type of computer or data note Leader/06 type of record, 245 \$hmedium
Form	9.7B16	530 additional physical form available note
Identifier	9.7B19 numbers	020 ISBN, 022 ISSN, 856 \$uURL
Relation	12.7B7 relationships with other serials 1.7B15 reference to published descriptions	580 linking entry complexity note 772 parent record entry, 773 host item entry 780 preceding entry 524 preferred citation of described materials note
Source	1.7B22 combined notes relating to the original	534 original version note 786 data source entry
Language	9.7B2 language/script	041 language code, 546 language note
Coverage	9.7B1a nature and scope	spatial: 034 \$acoded cartographic mathematical data 255 \$ccartographic mathematical data temporal: 045 \$atime period of content 513 \$btype of report and period covered note
Rights	9.7B20 copy being described, restrictions on use	506 restrictions on access note 590 local notes

마지막으로, 여러 유형의 메타데이터 통합에 대한 방법으로 MARC에 대한 수정 및 새로운 필드의 마련이 필요하며, 표준기술규칙의 마련으로 정확하고 대량의 데이터를 구축하며, 특히 이러한 기술법을 익힌 전문가들의 등장이 필요할 것이다.

7. 참고문헌

- AACR, 2nd ed., 1988 rev. Amendments 1993.
- Caplan, Priscilla and Guenther, Rebecca. Metadata for Internet Resources: Dublin Core Metadata Elements set and its mapping to USMARC. *CCQ22*(3/4), 1996. pp. 43-58.

- *Cataloging Internet Resources: A Manual and Practical Guide*. Nancy B. Olson, ed. Dublin, Ohio: OCLC Online Computer Library Center, c1995. ISBN 1-55653-189-3.
- ————. 2nd ed. c1997. ISBN 1-55653-236-9.
- Michelson, A. *Why a CNI white paper Framing the problem*. Unpublished speech given at the 1995 ASIS Annual Meeting, October 9, 1995, Chicago, IL.
- Weibel, S. Metadata : The foundations of resource description. *D-lib Magazine*. 1995.
- Younger, Jennifer A. Resources Description in the Digital Age. *Library Trends*, Vol. 45, No. 3 Winter 1997. pp. 462-481.