

메타데이터의 의미론적 확장에 관한 연구

Study on the Semantic Extension of the Concept of Metadata

남 태 우(Taewoo Nam)*

이 승 민(Seungmin Lee)**

목 차

- | | |
|---------------------|----------------------|
| 1. 서론 | 3.3 메타데이터의 개념적 혼란 |
| 2. 메타데이터의 기원 및 발전사 | 4. 메타데이터의 특성 분석 |
| 2.1 메타데이터의 기원 | 4.1 일반적인 측면에서의 메타데이터 |
| 2.2 메타데이터 용어의 역사 | 4.2 도서관계에서의 메타데이터 |
| 3. 현재 정보환경에서의 메타데이터 | 4.3 메타데이터의 의미론적 확장 |
| 3.1 메타데이터의 급증 | 5. 결론 |
| 3.2 메타데이터의 의미의 변화 | |

초 록

현재의 정보환경에서, 메타데이터는 정보자원을 기술하고 관리하는데 있어서 강력한 도구로 인식되고 있다. 하지만, 그 적용의 광범위함으로 인해서 그리고 메타데이터가 지닌 복합적인 의미로 인해서 메타데이터가 무엇인지에 대한 명확한 설명이 어렵게 되었다. 이는 메타데이터를 단순한 서지기술이나 서지적 구조와 구분짓는 것을 어렵게 하고 있으며, 이로 인해 메타데이터가 무엇인지에 대한 혼란이 가중되고 있다. 이러한 개념적인 혼란을 없애고 메타데이터가 지닌 의미를 확인하기 위해서, 본 연구에서는 기존의 메타데이터에 대한 정의에 기반하여 도서관계에 적용할 수 있는 메타데이터의 본질적인 개념을 재정립하였다.

ABSTRACT

In the current information environment around the library community, metadata is recognized as a sophisticated and powerful tool that can manage and represent information resources. However, with the discreet use of the concepts of metadata and metadata standard, there is no clear boundary that differentiates metadata standard from simple resource description and traditional bibliographic structure, leading to confusion as to what a metadata and metadata standard is. To consider these issues, this research discussed what metadata and metadata standards are based previous definitions of metadata. Based on those definitions, the fundamental concept of metadata is reestablished to be properly used in the library community.

키워드: 메타데이터, 메타데이터 역사, 도서관 목록

Metadata, History of Metadata, Library Cataloging

* 중앙대학교 문헌정보학과 교수(namtw@cau.ac.kr)

** 중앙대학교 문헌정보학과 시간강사(seungminator@gmail.com)

논문접수일자: 2010년 10월 20일 최초심사일자: 2010년 10월 25일 게재확정일자: 2010년 11월 9일
한국문헌정보학회지, 44(4): 373-393, 2010. [DOI:10.4275/KSLIS.2010.44.4.373]

1. 서론

정보기술의 발달 및 광범위한 보급으로 인해, 디지털 형태의 전자적 정보자원의 수가 계속하여 급격하게 증가하고 있으며, 이들 정보자원을 공유하고 교환하는 방법 또한 개선 및 향상되어 왔다. 이러한 정보환경의 변화로 인해, 디지털 자원이 기존의 인쇄형태 자료와 대등한 위치에 놓이거나 점차적으로 이를 대체하는 경향이 나타나고 있다. 하지만, 디지털 정보자원은 전통적 인쇄자료와는 상당히 다른 특징들을 지니고 있으며, 이로 인해 MACHINE READABLE CATALOGING(MARC) 등과 같은 인쇄자료를 위해 고안된 전통적인 방식의 서지도구들은 디지털 정보자원을 기술, 조직, 관리하는데 있어 여러 가지 한계에 부딪치게 되었다. MARC는 전통적인 인쇄자료를 관리하기 위해 고안되었기 때문에 디지털 자원의 수많은 이질적인 세부사항까지 다 표현할 수는 없다. 게다가, MARC 레코드를 생성하기 위해서는 그 복잡한 구조로 인해 전문적인 지식이 필요하며, 엄청난 양의 시간과 노력이 필요하다. 이러한 이유로, MARC과 같은 전통적인 서지도구는 그 양이 급속하게 증가하는 디지털 자원을 기술하고 관리하기 위한 형식으로는 적절하지 않게 되었다. 이로 인해 디지털 자원을 기술하기 위한 메타데이터에 대한 필요성이 제기되었다.

메타데이터는 웹상의 디지털 정보자원을 기술하고 관리하기 위한 효과적인 서지도구의 하나로 인식되고 있다. 일반적으로 메타데이터는 여러 가지 상황에서 다양한 방식으로 적용되는 복합적인 개념이다. 또한, 여러 가지 상이한 목적을 충족시키기 위해 다양한 메타

데이터 표준들이 개발되어 왔다. 하지만, 더블린 코어(Dublin Core)가 디지털 형식이나 물리적인 형식을 지닌 정보자원 모두를 포괄하는 것으로 메타데이터의 개념을 확장한 이후, 단순한 서지기술이나 서지데이터 구조와 메타데이터 표준 사이의 명확한 의미적 구분이 없어지게 되었으며, 이는 메타데이터 표준이 무엇인지에 대한 혼란을 가져오게 되었다. 이러한 개념적인 혼란은 여러 커뮤니티들이 자신들의 고유한 목적을 충족시키기 위하여 수많은 이질적인 메타데이터 표준들을 개발하게 하는 결과를 초래하게 되었다. 메타데이터 표준의 다양화는 특정 유형의 정보자원에 대한 상세한 기술을 가능하게 한다는 장점을 지니고 있기는 하지만, 이질적인 메타데이터 표준의 무분별한 생성은 다른 메타데이터 표준들과의 호환을 저해하는 등의 여러 가지 문제를 야기시키고 있다. 또한, 각각의 커뮤니티는 커뮤니티가 지닌 고유한 목적과 특성에 따라 기술형, 관리형, 구조형 메타데이터 등 여러 가지 유형의 메타데이터 표준들을 개발하여 사용하고 있다. 하지만, 현재 사용되는 대부분의 메타데이터 표준들은 구조화 된 여러 가지 요소들을 추가함으로써 이러한 다양한 기능들을 하나의 표준 안에서 포괄적으로 수행하고 있다. 이는 결국 메타데이터 표준들이 각각 독립적으로 존재하게 되어 여러 가지 유형의 정보자원을 효과적으로 기술하는 것이 아닌 서지레코드의 중복이나 호환성 없는 데이터의 작성 등과 같은 문제를 야기시키고 있다. 이러한 문제를 해결하기 위해서는 메타데이터가 무엇인지에 대한 명확한 논의가 있어야 한다. 메타데이터가 지닌 본질적인 개념에 대한 정립이 없다면 메타데이터는

데이터를 기술하는데 있어서 혼란을 가중시키게 될 것이며, 신뢰성 없는 서지레코드의 폭발이라는 상황을 초래할 수 있다.

이에 본 연구에서는 메타데이터의 개념을 그 기원으로부터 다시 살펴보고 기존의 메타데이터에 대한 여러 가지 정의를 고찰하여 메타데이터가 지닌 본질적인 의미를 재정립하고자 한다. 이를 기반으로 문헌정보학 분야에 적합하게 사용될 수 있도록 메타데이터의 의미를 확장하고자 한다.

다만, 메타데이터는 여러 가지 상황과 목적에 따라 그 의미가 다양한 방식으로 적용되고, 생성되는 메타데이터 표준 또한 특정 커뮤니티의 목적을 충족시키기 위한 형태로 나타나고 있기 때문에, 본 연구에서는 메타데이터 표준의 유형을 구분하지 않고 메타데이터 표준 생성의 기반이 되는 메타데이터의 본질적인 개념에 대해 논하는 것으로 연구의 범위를 한정하고자 한다.

2. 메타데이터의 기원 및 발전사

디지털 자원의 급증 및 정보 유형의 다양화로 인해, 메타데이터는 문헌정보학 분야에서 가장 널리 이용되는 용어 가운데 하나가 되고 있다. 하지만, 이 용어가 널리 사용되면서 여러 가지 다양한 상황에서 메타데이터에 대한 수많은 정의들이 내려지게 되었으며, 이는 메타데이터가 무엇인지에 대한 잘못된 인식과 혼란을 초래하는 결과를 가져오게 되었다. 이러한 상황에서 메타데이터가 지닌 의미를 명확히 정립하기 위해서는 이 용어가 어떻게 생겨나게 되

었는지 그리고 어떤 과정을 거쳐 현재의 의미로 사용되고 있는지를 살펴보는 것이 선행되어야 한다.

2.1 메타데이터의 기원

메타데이터는 그리스어의 *meta*와 라틴어의 *datum*이 합쳐진 용어이다. 메타데이터라는 용어에 대한 설명이나 정의는 이 두 가지 합성된 용어들로부터 형성된 것이며, 이 용어들 또한 본래의 의미와는 다르게 그 의미가 변형되어 사용되었다(Gill 2000).

어원적으로, (*μετά* - 에서 유래된) *meta*-라는 용어는 고대 그리스어에서 비롯되었다. 이 용어는 여러 가지의 의미를 지니고 있으며, 과거에는 단독으로는 사용되지 않았다. *meta*라는 용어가 하나의 단어로 사용될 경우 “-사이 에, -와 함께, -의 뒤에”라는 의미를 지니게 되며, 접두사로 사용될 경우에는 “-의 위에, -를 넘어서는” 또는 “-보다 포괄적인, -보다 근본적인”이라는 의미로 사용된다(Merriam-Webster Dictionary 2007). 현대 라틴어와 영어에서 *meta*라는 용어는 “초월적이거나, 높은, 또는 보다 본질적인 특성”을 나타내는 것으로 사용되고 있다(Hillmann 2005). 하지만, 이와 같은 어원적이고 사전적인 정의들은 정보자원을 기술하기 위한 도구로서의 메타데이터를 설명하는 것은 아니다. 또한, 이 용어는 여러 학문분야에서 적용되고 있으며, 적용되는 상황에 따라 여러 가지 다른 의미로 사용되고 있다. 형이상학에서 사용될 경우, *meta*라는 용어는 “보다 높은 수준의 특성 또는 보다 근본적인 종류”를 지칭하며, 컴퓨터공학 분야에서는 “보다 근원적인

정의 또는 기술사항"을 의미하는데 사용된다(Public Library Association 2007). 최근에는 "다른 개념으로부터의 추상적인 관념"이라는 의미를 나타내는데 사용되기도 한다(Taylor 2007). 이를 종합해 보면, *meta*라는 용어는 직접적으로 볼 수는 없지만, 어떤 영향력을 발휘하며 근본적인 배경에 존재하는 것의 일부분이라는 의미를 함축적으로 표현하는 것으로 볼 수 있다(O'Sullivan 2007). 이와 같이, 이 용어가 지닌 근본적인 개념은 사전적인 의미에 비해 보다 구체적이며, 용어가 적용되는 특정 상황에 적합하도록 변형되어서 사용되고 있다.

*data*는 라틴어 *datu*에서 유래되었는데, 이는 "어떤 주어진 것" 또는 "사실이라고 알려진 것" 등을 의미하며, 추리나 추론을 위한 기반으로 사용되는 사실적인 정보를 지칭한다(Oxford English Dictionary 2007). 이러한 사전적인 의미와 유사하게, McCrank는 수집되거나 추론되는 가장 작은 정보의 단위에 주어지는 사실이나 사항이라고 데이터를 정의하고 있다(McCrank 2002, 627). 컴퓨터공학 분야에서의 *data*는 단순한 사실이나 사실적인 정보 또는 주어진 기록, 특히 분석을 위해 조직된 정보라는 의미를 지니고 있다(Taylor 2007). 이러한 의미에 기반해서, *data*는 데이터베이스 안에 수록된 텍스트, 이미지 또는 기록 등과 같은 특정 형태의 정보로 그 의미가 확장되어 사용되고 있다(Hanks 2007).

메타데이터라는 용어의 의미는 *meta*와 *data*가 지닌 의미의 조합에 그 기반을 두고 있다. 넓은 의미에서 보면, 메타데이터는 무언가에 대해서 설명을 하고 그에 대한 정보를 제공해 준다는 점에서 일종의 데이터라고 볼 수 있다.

하지만, 메타데이터의 의미는 단순한 사실적 정보로서의 데이터 보다는 조금 더 고차원적인 것이다. 또한, 메타데이터는 단독으로 사용되지 않고 그것이 기술하는 무언가와 항상 관련되어 있다는 점에서 데이터와도 차이가 난다. 이러한 차이점들은 *meta*라는 의미를 추가하면서 생겨난 복합적인 특성에 기인한다. 이러한 특성으로 인해서, 메타데이터는 하나의 간단한 문장으로는 명확하게 설명될 수 없는 복잡한 개념이 되었다. 하지만, 메타데이터를 조합하는 두 가지 근본적인 용어의 의미에 기반해서 보면, 메타데이터는 "다른 데이터(혹은 정보)를 수반하는 데이터(혹은 정보)"라고 간단하게 설명할 수 있다(Althem and Palmer 2002). 이러한 의미는 오랜 기간에 걸쳐 변화, 변용되면서 우리가 현재 사용하는 문헌정보학 관련 용어에 포함되었으며, 상황에 따라 여러 가지 다른 의미로 적용되고 있다.

2.2 메타데이터 용어의 역사

(대문자 M으로 시작하는) 메타데이터라는 용어는 일련의 데이터를 기술하기 위한 목적으로 1969년 Jack E. Myers에 의해서 만들어진 용어이다(Caplan 2003, 1). 그는 자신이 개발한 메타모델(*MetaModel*)을 실행하는데 필요한 소프트웨어 제품의 이름으로 이 용어를 사용하였으며, 이 제품을 개발하고 판매하기 위해 설립한 자신의 회사(The Metadata Company)에 메타데이터라는 이름을 붙였다. Myers가 메타데이터(*Metadata*)라는 용어를 사용하기 전까지는 메타데이터(*metadata*), 메타 데이터(*meta data*) 혹은 메타-데이터(*meta-data*) 등의 용

어는 존재하지 않았다(Greenberg 2005, 19).

메타데이터라는 용어가 처음으로 인쇄되어 사용된 자료는 The Metadata Company에서 만든 제품 소개서이며, 이 소개서에서 처음으로 메타데이터라는 용어가 공식적으로 사용되었다. 또한, 1986년에 이 용어는 The Metadata Company의 상표로 등록되었다(U.S. Trademark Registration No. 1,409,206). 이 상표등록에 따르면, 하나의 단어로 사용되는 메타데이터라는 용어는 이 회사의 현재 제품이나 앞으로 나올 제품을 의미하는 것으로 사용되었으며, 이 용어의 보다 일반적인 의미는 메타데이터(metadata) 혹은 메타-데이터(meta-data)로 표시되어야 한다고 명시되어 있다. 이때까지는 메타데이터라는 용어에 특별한 의미가 부여된 것은 아니었다.

메타데이터라는 용어에 특별한 의미를 부여한 것은 1988년 미국 항공우주국(National Aeronautics and Space Administration: NASA)에서 발행한 『Directory Interchange Format Manual』 초판이라고 알려져 있다(Caplan 2003, 1). 하지만, NASA에서 발행한 매뉴얼 이전에도 메타데이터에 특별한 의미를 부여해서 사용한 예가 여러 연구에서 나타나고 있다.

1973년, Bo Sundgren은 박사학위논문에서 메타데이터라는 용어에 특별한 의미를 부여하여 사용하였다(Sundgren 1973). 그는 통계정보 및 자신의 논리적 정보 데이터 베이스 모델(infological data base model)에 수록된 일련의 데이터를 표현하기 위해서 메타-정보(meta-information)와 메타데이터(metadata)의 개념을 제안하였다. 이 모델에서, 그는 메타-정보를 “정보 위의 정보(information on information)”

라고 정의하고 있으며 “데이터베이스에 수록된 정보의 내용에 대한 양질의 정보 및 이외의 정보”라고 설명하였다. 그는 또한 양질의 정보 및 이외의 정보를 메타데이터라고 설명하면서, 메타데이터를 “메타-정보를 표현한 데이터”라고 정의하였다(Sundgren 1973, 104). 이 정의에 따르면, 메타-정보는 데이터베이스 내의 정보가 수록하고 있는 내용을 표현하는 정보이며 메타데이터는 메타-정보를 표현한 데이터이다. 이러한 메타-정보와 메타데이터라는 개념에 대해 Sundgren은 각기 다른 표현으로 생각했으나, 이 두 가지 개념은 모두 정보가 수록하고 있는 내용에 대한 데이터 또는 정보라고 정의될 수 있다.

2년 후, Sundgren은 메타-정보와 메타데이터의 개념을 보다 상세하게 구체화하였다. 그는 데이터 베이스에 수록된 정보의 내용을 정보(information)와 메타-정보(meta-information)의 두 가지 유형으로 구분하였다. 그는 또한 “정보 위의 정보”를 메타-정보의 하나의 중요한 카테고리 정의하였는데, 이는 “데이터 베이스에 수록된 정보의 내용에 대한 양질의 정보 및 이외의 정보”를 의미하는 것으로 사용되었다(Sundgren 1975, 59). 이러한 측면에서, 그는 “정보 위의 정보”를 데이터의 표현으로 간주하였으며, 메타-정보를 그 데이터의 표현에 대한 정보로 설명하였다. 그는 또한 메타-정보는 다른 정보에 의해서도 표현된다고 설명하면서 이 정보를 메타데이터라고 명명하고 “메타-정보를 표현한 데이터”라고 정의하였다(Sundgren 1975, 59). 이러한 맥락에서, 메타데이터는 정보 위의 정보, 메타-정보, 정보에 대한 데이터의 표현으로 정의될 수 있다.

1977년, Date에 의해서 메타데이터는 “데이터에 대한 데이터”로 정의되었다(Date 1977, 149). 그에 따르면, 데이터 사전(data dictionary)은 특별한 데이터베이스이며 “데이터에 대한 데이터”는 이 데이터 사전 안에 수록되는 것이라고 주장하였다. 그는 메타데이터를 “데이터에 대한 데이터”로 지칭하면서, 단순한 미가공 데이터라기 보다는 데이터베이스 시스템에 수록된 다른 개체를 기술하는 것으로 메타데이터를 설명하였다(Date 1977, 149). 이러한 의미에서, 메타데이터는 데이터베이스 내의 다른 개체를 기술하는 데이터의 한 유형이며 데이터 사전의 일부분으로서의 기능을 한다고 볼 수 있다.

McCarthy가 정의한 메타데이터 역시 “데이터에 대한 데이터”라는 정의에 기반을 두고 있으며, “데이터의 내용에 대해 체계적으로 기술한 정보 및 여러 가지 방식으로 검색되고, 조정되며, 표현될 수 있는 조직체계”라고 메타데이터를 정의하였다(McCarthy 1982, 235). 그는 또한 메타데이터를 데이터베이스에 의해서 검색, 조정, 표현될 수 있는 데이터 기술(description)로서의 기능을 하는 통계적 데이터베이스 관리 시스템의 일부로 설명하였다. McCarthy가 정의한 메타데이터의 가장 큰 특징은 기존의 메타데이터에 대한 정의에 구조라는 개념을 추가함으로써 메타데이터의 정의를 확장한 것이다. 이러한 관점에서 보면, 메타데이터는 단순하고 구조화되지 않은 데이터 기술이 될 수도 있고, 여러 가지 데이터베이스를 통제하는데 사용되는 데이터 사전과 같은 구조화되고 복잡한 것이 될 수도 있다. 이러한 구조적인 측면과 함께, 메타데이터는 대상이 되는 정보자원의 여러 가지 다른 유형이나 데이터 추론에 있어서의 다양한

논리적인 수준에도 적용될 수 있다.

이전의 정의들과는 다르게, Durrell은 메타데이터를 “기술된 개체의 확인, 발견, 평가, 관리를 도울 수 있는 정보를 수록하고 있는 구조화되고 부호화 된 데이터”라고 정의하였다(Durrell 1985, 3). 이는 메타데이터 자체의 의미 보다는 메타데이터가 어떻게 사용되는지에 중점을 둔 기능적인 정의이며, 메타데이터를 데이터 운영상에서 편찬되고 유지되는 데이터의 한 유형이라고 보는 것이다. 하지만, 메타데이터는 데이터베이스 안에 수록된 가공되지 않은 데이터에 대한 논리적인 관점을 제공해 준다는 점에서 가공되지 않은 데이터와는 구별되어야 한다.

1986년, Shelley는 신속하면서도 간단하게 데이터와 정보에 접근하고자 하는 새로운 정보요구가 발생하고 있다고 주장하면서, 이러한 정보요구를 충족시키기 위해서 메타-정보라고 하는 새로운 유형의 정보의 필요성에 대해 주장하였다. 메타-정보는 “정보에 대한 정보”라고 정의된다(Shelley 1986, 116). 이 정의는 이전에 사용되었던 정의와 유사하지만, 변환된 데이터와 정보를 추가함으로써 기존의 정의를 확장하고 있다. 따라서, 이러한 메타-정보의 개념은 데이터 기술 뿐만 아니라 재가공되고 재처리된 데이터 및 통합된 데이터 세트까지 포함할 수 있게 되었다.

반면, Krockel과 Westbrook은 메타데이터를 데이터를 변환하는 하나의 과정으로 보고 있다. 이들 역시 메타데이터를 ‘데이터에 대한 데이터’라고 정의하고 있기는 하지만, 여러 정보자원으로부터 축적된 데이터를 기계가독형의 형태로 변환하는 기능에 중점을 두고 있다(Krockel and Westbrook 1987, 373). 데이터

변환을 실행하기 위해서는 데이터베이스에 수록된 데이터가 지니고 있는 각각의 요소들이 정보단위의 구조, 즉 메타데이터에 의해서 분석되고 기술되어야 한다. 따라서, 이 정의에서는 메타데이터를 데이터라고 볼 수는 없으며, 오히려 메타데이터는(수치, 테이블, 그래픽 등의) 축적된 데이터를 기계가독의 형식으로 변환하는 구조화된 코딩 시스템의 일부로서 다루어지고 있다.

1988년, NASA는 『Directory Interchange Format Manuals』에서 메타데이터를 “일련의 데이터에 대한 기술”이라고 정의하였다(NASA 1990, 1; NASA 1991, 1). 이 정의에 따르면, 메타데이터는 데이터 셋을 기술하는 정보라고 볼 수 있다. 이 정보는 이용자 안내서에 수록된 데이터 및 디렉토리, 목록, 인벤토리 등에 수록된 데이터 셋에 대한 기술사항 등을 포함하게 된다. NASA는 데이터 셋을 기술하는 정보는 어

떤 것이든지 모두 다 메타데이터로 간주하였으며, 정보 단위 사이의 관계를 정의하는데 필요한 모든 정보를 포함하는 것으로 메타데이터의 범위를 확장하였다(NASA 1990, F-10; NASA 1991, F-12). 이 정의는 데이터 기술뿐만 아니라 정보 사이의 관계를 수용하는 것까지 메타데이터의 개념을 확장한 것이다.

지금까지 언급한 메타데이터의 정의의 변화를 시대별 및 기능별로 정리하면 다음 <표 1>과 같다.

<표 1>에 나타난 바와 같이, 메타데이터는 “데이터에 대한 데이터” 혹은 “정보에 대한 정보”라는 포괄적인 정의로부터 시작하여 데이터베이스에 수록된 각각의 데이터에 대한 기술로 그 의미가 구체화되었으며, 그 이후로는 메타데이터라는 용어가 적용되는 상황에 따라 구조화된 데이터, 데이터 변환의 과정 등으로 그 의미가 변환되어 사용되었다. 이러한 정의들은

<표 1> 메타데이터에 대한 정의의 변화

연 도	연구자	정 의	기 능
1969	Myers	소프트웨어 제품	--
1973	Sundgren	정보 위의 정보	데이터베이스에 수록된 정보의 내용에 대한 정보
1975	Sundgren	정보 위의 정보	데이터의 표현에 대한 정보
1977	Date	데이터에 대한 데이터	데이터베이스에 수록된 다른 개체를 기술하는 데이터의 한 유형
1982	McCarthy	데이터의 내용에 대해 체계적으로 기술한 정보	데이터베이스에 의해서 검색, 조정, 표현될 수 있는 구조화 된 데이터 기술
1985	Durrell	데이터베이스 내의 개체의 특성을 기술하는 데이터	데이터 운영상에서 유지되는 구조화되고 부호화 된 데이터
1986	Shelley	정보에 대한 정보	데이터 기술뿐만 아니라 재 가공되고 재처리 된 데이터 및 통합된 데이터 셋
1987	Krockel & Westbrook	데이터베이스에 수록된 데이터가 지니는 기본적인 정보단위	데이터를 기계가독형으로 변환하는 과정
1988	NASA	일련의 데이터에 대한 기술	데이터 셋을 기술하는 모든 정보 및 정보 사이의 관계를 수용하는 데이터

메타데이터를 특정 정보 혹은 정보단위에 대한 기술로서의 기능을 하는 것으로 간주하였으나, 이러한 제한적인 의미의 메타데이터는 이전의 정의에 '데이터 사이의 관계를 기술'하는 기능을 추가함으로써 그 의미가 확장되어 사용된 것이다.

3. 현재 정보환경에서의 메타데이터

3.1 메타데이터의 급증

일반적으로 "데이터에 대한 데이터"라고 정의되는 메타데이터는 웹상의 디지털 자원을 효과적으로 기술하기 위한 도구로 고안되었다. 이용어는 데이터 기술사항을 지칭하기 위해서 컴퓨터 공학자 및 데이터베이스 관련 분야에 의해서 만들어지기는 했으나, 메타데이터의 근본적인 개념은 정보를 기술하고 조직하기 위한 최초의 시도 이후로 도서관계에서도 오래전부터 존재해 왔던 것이다.

도서관 목록은 도서관계에서 사용되는 메타데이터의 가장 대표적인 예이다. 목록의 목적은 이용자들로 하여금 해당 자료에 손쉽게 접근할 수 있도록 하기 위해 각각의 자료가 지닌 세부적인 특성들을 표준화 된 방식으로 기술함으로써 도서관 장서에 포함된 정보자원에 관한 정보를 제공해 주는 것이다. 목록과 유사하게, 메타데이터, 특히 서지적 메타데이터 역시 정보자원의 세부사항을 일관성 있는 방식으로 기술함으로써 정보자원을 확인하고 관리하기 위한 도구로서의 기능을 한다. 정보자원의 기술은 메타데이터 요소를 통해서 완성되는데, 이

들 요소들은 구조화 된 체계를 지닌 메타데이터 표준 내에서 그 의미와 기능이 정의되고 규정된다. 각각의 요소들은 저자, 서명, 주제 등과 같은 특정 정보자원이 지닌 여러 가지 다른 측면들을 표현하지만, 일련의 데이터 요소들은 전체적으로 메타데이터 표준을 구성하게 되며 정보자원을 기술하기 위한 표준화 된 방식을 제공해 준다.

하지만, 웹상에서 이용할 수 있는 정보자원의 유형이 급격히 다양해지면서, 하나의 메타데이터 표준으로 모든 정보자원을 관리하는 것은 거의 불가능하게 되었다. 하나의 메타데이터 표준은 특정 유형의 정보자원을 기술하는데 적합한 데이터 요소들을 제공해 줄 수는 있다. 하지만, 모든 유형의 다양한 정보자원을 모두 다루기 위해서는, 메타데이터 표준은 엄청나게 많은 수의 고도로 상세한 요소들을 수록하고 있어야 하며, 이들 요소들을 체계적으로 조직하기 위한 상당히 복잡한 구조가 필요하게 된다. 이러한 복잡성을 피하기 위해 메타데이터 요소의 수를 제한하게 되면, 수록하고 있는 데이터 요소의 한계로 인해서 여러 커뮤니티의 다양한 요구를 충족시키기 어려워질 것이며, 그렇기 때문에 특정 커뮤니티와 관련된 모든 정보자원을 기술하는데 있어서 여러 가지 한계에 직면하게 될 수 있다. 결국, 이러한 상황은 특정 커뮤니티의 고유한 목적을 충족시키기 위해 특성화 된 수많은 메타데이터 표준을 개발하는 결과를 초래하게 되었다. 이러한 메타데이터 표준의 다양화는 많은 수의 이질적인 정보자원을 조직하고 관리하기 위해서는 필수불가결한 것이라고 볼 수 있으며, 각각의 정보자원 유형을 위해 개발된 특정 메타데이터 표준

을 이용하는 것이 실질적인 정보자원을 다루는데 있어서는 보다 효과적일 수도 있다. 이러한 특정화 된 메타데이터 표준을 사용하게 되면, 하나의 메타데이터 표준으로 모든 정보자원을 기술하는 것에 비해 보다 상세하고 정확하게 정보자원을 기술할 수 있게 된다. 또한, 맞춤형 데이터 요소를 이용한 특정화 된 메타데이터 표준을 통해서 각 커뮤니티의 고유한 요구도 충분히 충족시킬 수 있게 된다.

하지만, 여러 가지 메타데이터 표준을 사용함으로써 얻어지는 장점에도 불구하고, 메타데이터 표준의 다양화는 고유한 특성을 지닌 실질적인 메타데이터 표준의 급증이라는 문제를 일으키고 있다.

3.2 메타데이터의 의미의 변화

데이터베이스 스키마를 기술하는 정보를 지칭하기 위해 메타데이터라는 용어가 사용된 이후, 메타데이터는 데이터베이스 시스템 내의 하위 시스템으로 여겨져 왔으며, 간단하게 “데이터에 대한 데이터” 혹은 “정보에 대한 정보”라고 정의되고 있다.

“데이터에 대한 데이터”로서의 메타데이터는 초창기에는 단순한 데이터 기술을 지칭하는 것이었지만, 이는 통계학 분야 등과 같은 다른 분야에도 적용되어 왔으며, 그 의미는 데이터베이스에 수록된 데이터나 데이터 셋을 기술하는 정보를 지칭하는 것으로 점차 변형되고 확장되었다. NASA의 정의 이후, 메타데이터라는 용어는 여러 분야에서 광범위하게 사용되기 시작하였으며, 각 분야의 목적을 충족시키기 위해서 서로 다른 방식으로 재해석되었다. 또

한, 메타데이터의 개념은 기술된 데이터나 정보, 변환된 데이터, 또는 정보단위 사이의 관계를 표현한 것과 같은 모든 유형의 정보를 포함하는 것으로 확대되었다. 이렇게 일관성 없이 메타데이터라는 용어를 사용함으로써 메타데이터의 정의가 난무하게 되어 메타데이터의 의미에 대한 혼란을 가져오는 결과를 초래하게 되었다. 그 결과, 메타데이터의 개념은 복잡함 설명 없이는 명확하게 이해될 수가 없게 되었으며, 메타데이터라는 용어를 둘러싼 모호함이 가중되게 되었다.

3.3 메타데이터의 개념적 혼란

메타데이터는 웹상의 디지털 정보자원을 기술하고 관리하기 위해 일련의 요소들을 정의하여 메타데이터 스키마 형식으로 구축되었다. 메타데이터 스키마가 특정 커뮤니티에서 광범위하게 사용되면 이는 메타데이터 표준으로 정립되어 해당 분야에서 하나의 표준적인 메타데이터 스키마로 사용되어 왔다.

문헌정보학 분야에서 사용된 최초의 메타데이터 표준은 1995년에 개발된 더블린 코어(Dublin Core Element Set)이다. 메타데이터라는 개념은 MARC 이후로 도서관계에서도 계속해서 사용되어 왔지만, 더블린 코어 이전에는 전자정보를 기술하기 위한 표준은 거의 개발되지 않았다. 더블린 코어가 발표된 이후 문헌정보학 분야에서도 여러 가지 메타데이터 표준들이 사용되어 왔지만, 더블린 코어 이전에 발표된 실질적인 메타데이터 표준으로는 TEI Header가 유일하다(표 1 참조). 1994년에 CDWA의 초판이 발표되기는 했지만, 2000년에

CDWA 2.0 버전이 발표될 때까지는 CDWA는 널리 사용되는 표준이 아니었다.

〈표 2〉에 나타난 바와 같이, 더블린 코어 이전까지의 메타데이터 표준들은 웹상의 디지털 정보자원만을 대상으로 구축되었지만, 현재는 전자정보 뿐만 아니라 비전자적 형식으로 된 정보자원과 유사한 자료를 기술하기 위한 것으로 발전해 왔다(그림 1 참조).

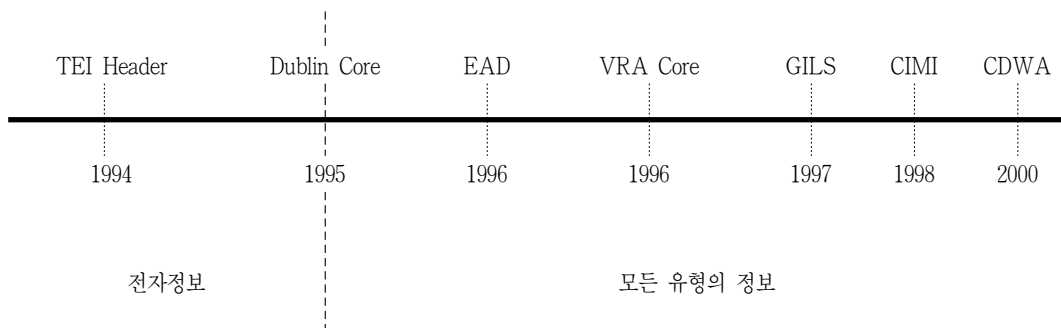
〈그림 1〉에서 나타난 바와 같이, 대상이 되는 정보자원의 유형이 변화함과 동시에 이들 메타데이터의 목적과 기능 또한 다양해지게 되었다. 이로 인해 전자형태의 정보자원을 기술하는 것 이외의 여러 가지 기능을 지닌 많은 메타데이터 표준이 발표되었지만, 이는 메타

데이터가 무엇인지에 대한 개념적인 혼동을 가져오게 되었다. 또한, 웹상의 디지털 정보자원을 위해 구축된 가장 보편적인 메타데이터 표준인 더블린 코어가 1997년 메타데이터 표준의 대상이 되는 정보자원을 그 형식에 상관 없이 모든 유형의 정보자원으로 확장한다고 발표하게 되면서 이러한 개념적인 혼란은 더욱 가중되었다.

이 개념적인 혼란을 해결하지 않은 상태에서 여러 커뮤니티들이 여러 가지 상이한 메타데이터 표준들을 계속해서 개발해 왔다. 이들 메타데이터 표준들의 대부분은 특정 목적을 충족시키기 위한 것이며 다른 표준들과 호환되지 않는 고유한 특성들을 지니고 있다. 이러한 이질

〈표 2〉 메타데이터 표준의 생성

메타데이터 표준	구축연도	발표연도	대상 정보자원
TEI Header	1987	1994	전자정보
더블린 코어	1995	1995	네트워크상의 정보
EAD	1995	1996	고문서
CDWA	1994	1994(2000)	예술작품
VRA Core	1995	1996	시각정보
GILS	1994	1997	정부간행물
CIMI	1995	1998	문화 정보



〈그림 1〉 메타데이터 표준의 대상 정보자원의 유형 변화

성은 각각의 표준이 지닌 구조와 체계를 구성하는 고유한 데이터 요소들로부터 비롯되고 있으며, 특정 커뮤니티의 고유한 목적을 충족시킬 수 있도록 적합화 된 것이다. 이로 인해서, 둘 또는 그 이상의 표준에 수록된 요소들이 특정 정보자원의 동일하거나 유사한 특성을 표현하는데 사용된다 하더라도, 각각의 표준에 수록된 요소들이 지닌 고유한 의미 때문에 하나의 메타데이터 표준으로 작성된 레코드는 다른 표준들과 쉽게 공유될 수가 없다. 또한, 많은 메타데이터 표준들은 특정 커뮤니티를 위해서 혹은 특정 목적을 충족시키기 위한 방식으로 개발되었기 때문에, 특정 커뮤니티가 개발한 메타데이터 표준을 다른 커뮤니티가 적용하여 사용하는 데는 많은 어려움이 따르게 된다. 이러한 상황은 2006년에 이미 22개의 서지적 메타데이터 표준이 사용되고 있다는 UKOLN의 연구결과에서도 잘 나타나고 있다(Caplan 2006, 11; Yu, Lu, and Chen 2003, 155). 이 결과는 단순한 서지적 데이터 구조와 메타데이터 사이의 개념적인 구분이 없이 특정 정보자원을 기술하는 것이면 모두 다 서지레코드로 인식되는 상황을 초래함으로써 메타데이터의 의미와 기능을 왜곡시키고 있으며, 정보자원을 기술하고 관리하기 위해 메타데이터를 적용하는데 있어서의 어려움을 야기시키고 있다.

4. 메타데이터의 특성 분석

1977년 Date의 정의 이후, 메타데이터는 “데이터에 대한 데이터”로서 널리 인정받게 되었지만, 이 정의는 메타데이터가 무엇인지에 대

한 이해를 돕기에는 그리고 메타데이터가 지닌 복합적인 특성을 명확하게 설명하기에는 모호한 점이 많다. 이를 해소하기 위해 여러 분야에서 메타데이터의 개념을 구체화하고 세분해서 정의해 왔다. 이렇게 새롭게 적합화 된 개념들은 크게 메타데이터의 일반적인 측면과 도서관계에서의 메타데이터의 두 가지 그룹으로 나누어서 생각해 볼 수 있다.

4.1 일반적인 측면에서의 메타데이터

여러 커뮤니티에서는 해당 커뮤니티가 처한 특정 상황에 적합하도록 그리고 자신들의 고유한 목적을 충족시킬 수 있도록 메타데이터의 특정 측면에 중점을 두어 메타데이터에 대한 정의를 내리고 있다. 이러한 정의들은 “데이터에 대한 데이터”라는 메타데이터의 공통된 개념에 영향을 받고 있기는 하지만, 이들 각각의 정의들은 메타데이터의 서로 다른 측면에 초점을 맞추고 있다.

4.1.1 웹자원의 기술로서의 메타데이터

웹상에서 이용할 수 있는 디지털 자원의 수가 급격하게 증가하게 되면서 메타데이터는 데이터베이스에 수록된 데이터 셋을 기술하는 정보로서, 그리고 이들 웹 자원을 기술하고 관리하기 위한 도구로서 사용되기 시작했다(Caplan 2003, 3). 하지만, Berners-Lee나 Qin과 같은 일부 연구자들은 메타데이터로 기술하는 데이터의 범위를 웹상의 전자자원으로 한정하여 사용할 것을 주장하였다. Berners-Lee는 웹 자원에 대한 정보를 지칭하는 용어 혹은 일반적으로 기계가 처리할 수 있는 유형의 정보로서 메

타데이터를 한정하여 정의하였다(Berners-Lee 1997). 그는 이 정의의 핵심으로 기계가독형이라는 개념을 언급하였는데, 이는 컴퓨터 상의 소프트웨어는 기계가독형의 정보만을 처리할 수 있기 때문이다. 따라서, 메타데이터는 기계가 이해하여 처리할 수 있는 웹 자원에 대한 기술이라고 정의될 수 있다. 또한 Qin은 메타데이터를 “디지털 개체에 대한 자동색인을 수용할 수 있으며 따라서 네트워크 상의 정보자원을 보다 쉽게 검색할 수 있도록 해주는 것”이라고 주장하면서 웹상의 하이퍼텍스트 문서에 포함된 기계가독형의 기록이나 개체에 대한 기록화라고 설명하고 있다(Qin 2000, 84). 이러한 관점에서 보면, 메타데이터는 네트워크 상의 정보자원 검색을 지원하기 위해 웹 상의 기계가독형 정보를 기술하는데 사용되는 것이라고 볼 수 있다.

이들 정의에서 보면, 메타데이터는 기계가독형 자원, 전자자원, 네트워크 자원, 그리고 하이퍼텍스트 문서에 포함된 자원 등과 같은 웹상의 디지털 자원을 기술하는데 사용되는 것으로 한정되고 있다. 이들 정의들은 컴퓨터공학 관점에서 바라본 제한적인 정의이기는 하지만, 메타데이터라는 개념의 근원에 기반하여 메타데이터의 의미와 기능을 보다 구체적으로 정확하게 설명해 주고 있다.

4.1.2 구조화 된 데이터로서의 메타데이터

이전의 정의들에 따르면, 메타데이터는 웹상의 디지털 정보자원에 대한 데이터로 그 의미가 한정되고 있다. 이러한 제한된 관점에 기반해서, Qin은 메타데이터는 기계가 처리할 수 있는 특정 구조를 지닌 여러 가지 형식으로 작

성된 데이터라고 설명하고 있다(Qin 2006). 이러한 의미에서, 메타데이터는 구조화 된 형식에 포함된 데이터에 대한 데이터 혹은 정보를 의미하게 된다. Qin의 정의와 유사하게, Day는 구조화 된 메타데이터는 기계가독형이 될 수 있으며 여러 가지 다양한 프로세스를 지원할 수 있다고 주장하였다(Day 1999). 보다 구체적으로, Borbinha는 메타데이터를 다른 정보자원에 대한 구조화 된 정보라고 정의하였다(Borbinha 2004). 이러한 정의에 따르면, 메타데이터는 구조화되어야 하며 구조화되지 않은 메타데이터는 기계적 처리를 지원할 수 없기 때문에 아무런 의미를 지니지 않게 되는 것이다.

이와는 달리 기능적인 관점에서의 정의들도 제안되었는데, 메타데이터는 정해진 개체와 관련되어 있는 기능을 지원해 주는 개체에 대한 구조화 된 데이터라는 것이다(Greenberg 2003, 245). 이 정의에 따르면, 메타데이터는 정보자원 개체에 대한 구조화 된 표현을 통해서 정보자원을 조직하고 관리, 검색하기 위한 강력한 도구로서의 기능을 하게 된다. 이러한 정의들은 메타데이터의 역할과 기능을 구체화하고 있다는 점에서 이전의 정의들에 비해 보다 정확하게 메타데이터의 의미를 설명하고 있다. 또한, 이 정의들은 구조화되어 있지 않거나 기계가독형 데이터가 아닌 것들을 메타데이터로 간주하지 않음으로써 메타데이터 자체의 범위를 명확하게 설정해 주고 있다.

4.1.3 데이터 기술로서의 메타데이터

초창기의 메타데이터는 전자형태의 정보자원을 기술하는 정보를 나타내기 위해서 고안되었다. 이러한 개념은 현재 전자적인 네트워크

환경에서 정보를 효과적으로 검색하는데 있어 필수적인 것으로 인식되고 있으며, 그렇기 때문에 정보검색 과정의 기반으로 인정받고 있다(El-Sherbini 2001). 데이터의 기술은 메타데이터라는 개념이 발생한 이후 가장 오랫동안 지속된 메타데이터의 기본적인 기능이다. 데이터 기술로서의 메타데이터는 정보자원의 특성과 특징들을 기술함으로써 원 데이터의 의미를 (사람과 기계 모두를 포함한) 이용자에게 이해 시키는데 중점을 두고 있다. 이러한 의미에서, 메타데이터는 웹상의 정보를 구성하는 다른(데이터) 개체들이 지닌 필수적인 특성들을 기술하기 위한 구조화 된 기술사항을 지칭한다. 이 데이터 기술은 원 데이터가 지닌 세부사항들을 기술하고 설명해 주는 또 다른 유형의 데이터이며, 해당 데이터의 구체적인 표현으로서의 기능을 하게 된다.

메타데이터는 또한 데이터의 추출을 의미하며, 보다 세부적으로는 낮은 수준의 데이터를 기술해 주는 보다 높은 수준의 데이터 추출이라고 정의되기도 한다(El-sherbini 2001; Sen and Jacob 1998). 또한, El-Sherbini와 Klim은 메타데이터를 데이터 저장소, 데이터 어플리케이션, 시스템 등에 존재하는 보다 높은 수준에서의 데이터 추출로 설명하고 있다(El-Sherbini and Klim 2004, 240). 이러한 정의들은 원 데이터의 내부에 존재하는 세부사항에 대한 기술로서 메타데이터의 기능을 설명하고 있다. 즉, 메타데이터는 데이터가 어플리케이션이나 시스템과는 독립적으로 관리되도록 해주는 데이터 추출이라는 의미를 지니고 있다. 이러한 의미에서, 메타데이터는 실질적인 데이터는 아니지만, 실제 데이터와 직접적 혹은 간접적으로

관련된 정보를 제공해 주게 된다.

4.1.4 메카니즘으로서의 메타데이터

컴퓨터공학 또는 공학분야에서의 메타데이터의 정의는 일반적인 정의에 비해 보다 구체적인 측면을 다루고 있다. 이들 분야는 기계가 데이터를 처리할 수 있도록 하기 위한 메카니즘으로서의 메타데이터의 기능에 중점을 두고 있으며, 가공되지 않은 원래의 데이터를 기계가 이해할 수 있는 데이터로 변환하는 과정으로서 메타데이터를 설명하고 있다.

이러한 측면에서, 메타데이터는 웹 자원에 대한 기계가 인식할 수 있는 정보 혹은 하이퍼 텍스트 안에 수록된 다른 사항들에 대한 정보로 간주된다(Qin 2000). 이러한 정의는 원래의 웹 자원이 표현되기 위해서는 기계가 인식할 수 있는 형태로 변환되어야 한다는 의미를 함축하고 있다. El-Sherbini에 따르면, 메타데이터는 전자형식의 문서나 데이터를 기술하는데 사용되는 방법으로 간단하게 정의하고 있는데, 여기서 의미하는 방법은 정보가 지닌 의미를 표현하기 위한 메카니즘을 지칭하는 것이며 이는 정보의 검색, 이해, 이용 등을 촉진하기 위한 기능을 하는 것이다(Terge 1998).

이러한 정의는 정보가 지닌 의미를 기술하기 위한 메카니즘으로서의 메타데이터에 중점을 둔 것이며, 데이터나 정보 그 자체를 반영하는 것은 아니다. 오히려, 이는 데이터의 구조 혹은 메타데이터 생성의 과정과 관련된 것이다. 따라서, 메카니즘으로서의 메타데이터는 데이터 기술이라는 기능을 뒷받침하는 데이터 변환 프로세스를 의미하는 것이라고 볼 수 있다.

4.2 도서관계에서의 메타데이터

메타데이터는 1995년에 더블린 코어가 발표된 이후 문헌정보학 분야에서는 하나의 중요한 용어로 자리잡게 되었다. 메타데이터가 지닌 근본적인 의미인 “데이터에 대한 데이터”라는 개념은 도서관의 장서를 기술하고 조직하기 위한 최초의 노력 이후로 도서관계에서도 계속 존재해 왔던 개념이다. 목록과 메타데이터 사이의 이러한 개념적인 유사성으로 인해, 도서관 목록은 일종의 메타데이터로 여겨지기도 한다. 메타데이터와 목록은 개념적인 구분 없이 수십 년 동안 공존하는 개념으로 함께 사용되어 왔다. 하지만, 메타데이터의 개념은 도서관이라는 환경에 맞도록 수정되고 적합화되고 있다. 이러한 변화된 개념은 일반적인 의미의 메타데이터 뿐만 아니라 목록이라는 개념과도 구분되는 것이다.

4.2.1 목록으로서의 메타데이터

문헌정보학 분야에서는 도서관 목록과 메타데이터가 지닌 개념적인 유사성 때문에 이 두 가지 도구를 동일한 것으로 간주하기도 한다. Hodge에 따르면, 메타데이터와 목록은 정보자원을 기술한다는 동일한 목적을 지닌 하나의 개념이라고 고려되고 있다(Hodge 2005). 인터넷 환경의 개발 및 발전과 함께, 목록이라는 용어는 메타데이터라는 용어에 의해서 대체되어 왔으며, 이 새로운 환경에서 변화된 것은 단지 그 용어이며 개념은 동일하다는 것이다. Baker 또한 도서관 목록과 메타데이터를 동일한 것으로 여겼는데, 이 두 가지는 그 형식은 다르지만 모두 정보자원을 기술하는 기능을 하고 있으며

정보자원에 담겨진 세부적인 데이터를 제공해주는 것이라고 주장하였다(Baker 2003). 따라서, 이러한 관점에서 보면 메타데이터는 전통적으로 사서들이 목록에 입력했던 정보에 대한 인터넷 시대의 용어일 뿐인 것이다.

목록과 메타데이터의 근본적 원리는 동일하기 때문에, 메타데이터와 목록을 동일시하는 것은 어떤 측면에서는 이치에 맞는 것일 수도 있다. 또한, 메타데이터와 목록은 정보자원의 검색을 용이하게 하고 다른 여러 가지 관련된 기능들을 수행하기 위해 일련의 구조화 된 기술 데이터를 제공한다는 동일한 목적을 지니고 있다. 하지만, 이러한 견해와는 달리, 이들 두 가지 도구 사이를 구분하고자 하는 연구도 수행되었다. 이 가운데 가장 두드러진 것으로는 기술되는 정보자원의 유형에 따라 목록과 메타데이터를 구분하는 것이다. 즉, 목록은 주로 물리적인 개체를 기술하기 위한 것인 반면, 메타데이터는 전자형식의 자원만을 대상으로 하는 것이다(El-Sherbini 2004, 240; Heery 1996). 하지만, 전자형식의 자원이 급격하게 증가함에 따라, 도서관 장서에도 이러한 새로운 유형의 자원이 대규모로 포함되게 되었다. 이러한 환경적 변화는 도서관계로 하여금 도서관 운영의 구성요소로서 메타데이터를 채용하도록 하였는데, 이는 전통적인 도서관 목록은 새로운 유형의 자원을 충분히 기술할 수 없으며 메타데이터는 웹상의 전자자원을 적절하게 기술하고 관리할 수 있는 도구로서 활용될 수 있기 때문이다(Day 2001; Tozer 1999, 10). 이것이 바로 메타데이터와 목록이 같은 개념인지 아닌지에 대한 개념적인 혼란을 가져오게 된 주된 이유이다. 이러한 관점에서 보면, 메타데이터와

목록은 서로 다른 개체를 기술하기 위한 도구이며, 정보자원을 기술하는데 있어서 상이한 형식과 새로운 인코딩 언어를 사용하는 구조라고 볼 수 있다.

이러한 차이점에도 불구하고, 메타데이터는 도서관 환경에서 사용하는 전통적인 의미의 목록과 상반되는 개념은 아니다. 또한, 메타데이터와 목록의 차이는 실질적인 것은 아니며, 기술하는 대상이 되는 정보자원의 형식은 다르지만 메타데이터나 목록이라는 접근방식 모두 정보자원을 기술하는데 있어서 동일한 원리를 적용하는 것이다. 도서관계에서는, 도서관과 관련된 정보자원을 기술하는데 사용될 경우 메타데이터를 목록의 변형된 형태로 간주하고 있다. 이와 반대로, 목록 레코드 내의 특정 데이터 필드는 전자자원을 기술하기 위해 고안되었기 때문에, 목록을 일종의 메타데이터의 한 유형으로 볼 수도 있다(Heery 1996). 따라서, 메타데이터와 목록은 도서관과 관련된 여러 가지 유형의 정보자원을 기술하는 상황에서 각각이 지닌 단점을 상호보완하는 도구로 이해될 수 있다.

4.2.2 정보자원의 대용물로서의 메타데이터

문헌정보학 분야에서 사용되는 메타데이터는 정보자원을 기술함으로써 다양한 도서관 활동을 지원해 주기 때문에, 메타데이터의 개념은 모든 유형의 정보자원을 기술해 놓은 모든 유형의 정보를 포함하는 것으로 확장되어 사용되고 있다. 메타데이터를 정보자원의 대용물로 간주한다면 이러한 관점은 일면 타당할 수 있다.

이러한 대용물의 개념은 도서관계에서도 여러 가지 측면에서 채택되어 사용되어 왔다. 도서관 목록의 기본 단위는 정보자원을 표현해

주는 대용물로서, 해당 정보자원을 기술함으로써 간단한 형식으로 정보자원에 대한 데이터를 전달해 주는 기능을 하게 되었다(Burnett, Ng, and Park 1999, 1210). 이러한 의미에서, 도서관계에서 사용되는 대용물이라는 개념은 정보자원에 담겨진 데이터나 정보에 대한 표현을 의미하는 것으로 변환되어 사용되고 있다.

이러한 정의에 기반해서 보면, 대용물은 단순한 데이터 기술 뿐만 아니라 정보자원의 대체물로서의 기능을 하며, '데이터에 대한 데이터'로서의 메타데이터와 유사한 것으로 생각해 볼 수 있다. 따라서, 메타데이터는 정보자원에 대한 구체화 된 표현 혹은 정보자원의 대용물을 생성해 주는 정보로 정의될 수 있다(Milstead and Feldman 1999). 도서관 목록 또한 정보자원의 대용물을 생산하는 것으로 간주될 수 있다. 목록에 의해 작성된 서지레코드는 정보자원 속에 담겨진 데이터와 정보를 제공해 주는 것이며 해당 자원의 대용물로서의 기능을 할 수 있다.

비록 메타데이터와 목록은 모두 정보자원의 대용물을 생산해 내기는 하지만, 메타데이터에 의해 생성된 대용물은 전자형태의 정보자원을 대상으로 하는 반면, 목록에 의해 생성된 대용물은 물리적 자원을 대상으로 한다는 점에서 이 두 가지 도구의 기능을 구별해 볼 수 있다. 하지만, 이러한 차이점에도 불구하고, 현재의 정보환경에서는 원 정보자원이 물리적 형태를 지니고 있다 하더라도 대부분의 대용물들은 전자적인 형태로 생성되고 있다. 따라서, 일부 연구자들은 도서관계에서 사용되는 메타데이터는 도서관과 관련된 모든 유형의 정보자원에 대한 서지적 대용물로 사용된다고 주장하면서

메타데이터의 의미를 확장하고 있다(Coyle 2005, 60; Dillon 2001).

4.3 메타데이터의 의미론적 확장

메타데이터는 적용되는 상황에 따라 여러 가지 다양한 의미를 포괄하게 되는 복합적인 개념이다. 또한, 메타데이터라는 개념은 메타데이터의 어떤 측면이 고려되는지 그리고 어떤 목적을 충족시키기 위해 사용되는지에 따라 그 의미와 기능이 다르게 적용되기도 한다. 문헌정보학 분야에서도 고유한 목적을 충족시키기 위해 메타데이터의 개념을 적합화하여 사용해 왔다. 하지만, 메타데이터가 지닌 다양한 세부적인 측면 때문에, 문헌정보학 분야 내에서도 메타데이터가 무엇인지에 대한 명확하면서도 널리 인정받을 수 있는 정의를 제시하지 못하고 있다. 따라서, 이전에 제시된 정의에 기반해서, 각 정의에 공통적으로 사용된 의미를 확인하고 문헌정보학 분야에 적합한 개념으로 확장함으로써 문헌정보학 분야에 적용할 수 있는 메타데이터의 의미를 보다 명확하게 도출해 내었다. 이 메타데이터의 개념은 어느 한 문장으로 간단하게 설명할 수 있는 것이 아니며, 메타데이터가 지닌 근본적인 개념에 기반하여 메타데이터가 수행하는 기능으로 나누어서 확장해 볼 수 있다.

첫째, 메타데이터는 정보자원에 대한 정보를 기술하고 표현해 주어야 한다. 메타데이터는 일반적으로 정보자원이 지닌 특성을 기술해 주는 데이터에 대한 데이터를 의미한다. 즉, 메타데이터는 실질적인 데이터는 아니지만, 원 데이터와 직접적 혹은 간접적으로 관련된 정보를

제공해 주는 일종의 데이터이다. 이러한 정보를 제공해 주기 위해서, 메타데이터는 특정 정보자원에 대한 표현이나 대응물을 생성해야 한다. 정보자원에 대한 표현이 없다면, 메타데이터는 아무런 의미가 없게 된다.

둘째, 문헌정보학 분야에서 사용되는 메타데이터는 모든 유형의 정보자원을 다룰 수 있어야 한다. 메타데이터는 본래 전자형태의 정보자원을 기술하기 위해 고안되기는 했지만, 그 기능은 전자형태 자원 이외의 형태를 지닌 자원까지도 포함하는 모든 유형의 정보자원을 기술할 수 있는 것으로 그 의미가 확장되어 사용되고 있다. 현재까지도, 메타데이터는 전자형태의 정보자원을 기술하는데 보다 적합한 것으로 인식되고 있다. 하지만, 문헌정보학 분야에서 사용하는 의미로는 메타데이터와 도서관 목록 사이의 구분이 엄격하게 적용되지는 않고 있다. 또한, 메타데이터가 전자형태의 정보자원에만 국한되어서 사용된다면 이는 도서관의 활동을 충분히 지원할 수 없게 된다. 따라서, 문헌정보학 분야에서 사용되는 메타데이터는 유형에 상관없이 모든 정보자원을 다룰 수 있어야 한다.

셋째, 메타데이터는 기계가 이해할 수 있고 처리할 수 있는 구조화 된 정보이어야 한다. 이는 메타데이터는 임의적으로 표현된 데이터 집합이 되어서는 안된다는 것을 의미한다. 메타데이터는 메타데이터 레코드가 체계적으로 저장될 수 있는 구조에 따라 명확하게 구조화 되어야 하며 자동화 된 도서관 활동을 효과적으로 지원하기 위해 기계에 의해 연속적으로 처리될 수 있어야 한다.

이상과 같이 확장된 의미의 메타데이터는 특정 정보자원이 무엇인지, 그 자원이 무슨 내용

을 담고 있는지를 명확하게 설명할 수 있을 때만이 메타데이터로서의 의미를 지니게 된다. 특히, 문헌정보학 분야에서 적용되는 메타데이터는 정보자원을 명확히 기술할 수 있어야 하며, 메타데이터를 적용하는 환경에서 메타데이터로 하여금 어떤 역할을 하도록 기대하고 있는지에 따라 메타데이터 레코드에 특별한 의미를 부여할 수 있어야 한다. 따라서, 문헌정보학 분야에서의 메타데이터가 본질적인 의미에 맞도록 기능하기 위해서는 정보자원에 보다 많은 의미를 부여할 수 있어야 하며, 문헌정보학의 요구를 충분히 충족시키기 위해서 보다 유용하게 사용될 수 있어야 한다.

5. 결론

웹으로 대표되는 정보통신기술의 발전 및 이에 따른 전자형태의 정보자원의 급증으로 인해, 현재의 정보환경은 웹 이전의 시대와는 급격히 다른 형태를 지니게 되었다. 이러한 환경적 변화는 전통적 인쇄자료를 관리하기 위한 도서관 목록의 기능적 한계를 가져오게 되었으며, 디지털 자원의 기술을 위한 보다 정밀하고 강력한 메타데이터에 대한 필요성을 가져오게 되었다. 하지만, 메타데이터에 대한 급증하는 요구로 인해, 메타데이터의 개념에 대한 명확한 정의가 내려지지 않은 상태에서 수많은 메타데이터 표준들이 생겨나게 되었다. 이는 결국 단순한 서지기술과 메타데이터 사이의 구분을 어렵게 만들고 있으며, 메타데이터가 무엇인지에 대한 개념적인 혼란을 초래하게 되었다. 이러한 상황은 특정 목적을 수행하기 위한 메타데

이터 표준들의 무분별한 생성을 야기시킴으로써, 메타데이터 고유의 의미와 기능을 저해하는 결과를 가져오고 있다. 또한 메타데이터는 그 개념이 지닌 복합적인 특성으로 인해서 메타데이터가 무엇인지에 대해 간단하고 명확하게 정의하는 것은 거의 불가능한 일이 되었다. 또한, 메타데이터는 상당히 넓은 범위의 의미 혹은 메타데이터가 사용되는 상황에 따라 다르게 사용되는 함축적인 의미들을 포함할 수도 있다. 이러한 의미적·개념적인 혼란을 정리하기 위해, 본 연구에서는 메타데이터가 무엇인지에 대해 지금까지 내려진 정의에 기반해서 그 의미를 포괄적으로 고찰함으로써 문헌정보학 분야에서 사용되는 메타데이터의 개념을 재정립하고 그 역할과 기능을 문헌정보학의 상황에 적합하도록 확장하였다.

메타데이터는 “데이터에 대한 데이터”로서 정의되며, 다음과 같은 사항을 충족시켜야 한다.

첫째, 메타데이터는 특정 정보자원에 대한 정보를 기술하고 표현해야 한다. 메타데이터는 실질적인 데이터는 아니지만, 원 데이터와 직접적 혹은 간접적으로 관련된 정보를 제공해주어야 하며, 해당 정보자원에 대한 표현이나 대용물을 생성해야 한다.

둘째, 메타데이터는 모든 유형의 정보자원을 다룰 수 있어야 한다. 현재 도서관계에서는 메타데이터와 도서관 목록 사이의 구분이 엄격하게 적용되지는 않고 있으며, 상호보완적인 도구로서 기능을 하고 있다. 따라서, 메타데이터가 전자형태의 정보자원에만 국한되어서 사용된다면, 이는 도서관의 여러 가지 기능을 충분히 지원할 수 없게 된다.

셋째, 메타데이터는 기계가 이해할 수 있는

처리할 수 있는 구조화 된 정보이어야 한다. 메타데이터는 서지레코드가 체계적으로 저장될 수 있는 구조에 따라 명확하게 구조화 되어야

하며, 자동화 된 도서관 기능들을 효과적으로 지원하기 위해 기계에 의해 연속적으로 처리될 수 있어야 한다.

참 고 문 헌

- [1] Altheim, M., & Palmer, S. B. 2002. *Augmented Metadata in XHTML*. Sun Microsystems, Inc. [online]. [cited 2010. 10. 9].
<<http://www.altheim.com/specs/meta/NOTE-xhtml-augmeta.html>>.
- [2] Baker, T. 2003. *DCMI Usage Board review of application profiles*. [online]. [cited 2010. 10. 10].
<<http://dublincore.org/usage/documents/profiles/index.shtml>>.
- [3] Berners-Lee, T. 1997. *Metadata architecture*. [online]. [cited 2009. 12. 3].
<<http://www.w3.org/DesignIssues/Metadata.html>>.
- [4] Borbinha, J. 2004. "Authority control in the world of metadata." *Cataloging & Classification Quarterly*, 38(3/4): 105-116.
- [5] Burnett, K., Ng, K. B., & Park, S. 1999. "A comparison of the two traditions of metadata development." *Journal of the American Society for Information Science*, 50(13): 1209-1217.
- [6] Caplan, P. 1995. "You call it corn, we call it syntax-independent metadata for document-like objects." *The Public-Access Computer Systems Review*, 6(4): 19-23.
- [7] Caplan, P. 2003. *Metadata fundamentals for all librarians*. Chicago, IL: American Library Association.
- [8] Coyle, K. 2005. "Understanding metadata and its purpose." *The Journal of Academic Librarianship*, 31(2): 160-163.
- [9] Date, C. J. 1977. *An introduction to database systems*, 2nd edition. MA: Addison-Wesley Publishing Company.
- [10] Day, M. 1999. *Metadata and electronic information*. [online]. [cited 2010. 10. 11].
<<http://www.ukoln.ac.uk/metadata/presentations/circe/birmingham.html>>.
- [11] Day, M. 2001. "Metadata in a nutshell." *Information Europe*, 6(2).
- [12] Dillon, M. 2000. "Metadata for web resources: How metadata works on the Web." *In Proceedings of Library of Congress Bicentennial Conference on Bibliographic Control for the New Millennium: Confronting the Challenges of Networked Resources and the Web*.

- November 2000. Washington, D.C.
- [13] Durrell, W. R. 1985. *Data Administration: A Practical Guide to Data Administration*, McGraw-Hill.
- [14] El-Sherbini, M. 2001. "Metadata and the future of cataloging." *Library Review*, 50(1): 16-27.
- [15] El-Sherbini, M., & Klim, G. 2004. "Metadata and cataloging practices." *The Electronic Library*, 22(3): 238-248.
- [16] Fillia, M., et al. 2002. "MetaDL: A digital library of metadata for sensitive or complex research data." In *Proceedings of ECDL 2002*, 16-18 September 2002. Rome, Italy: Pontifical Gregorian University.
- [17] Gill, T. 2000. Metadata and the World Wide Web. In *Introduction to Metadata: Pathways to digital information*. Online edition, version 2.1. [online]. [cited 2010. 10. 11]. <http://www.getty.edu/research/conducting_research/standards/intrometadata/metadata.html>.
- [18] Gill, T., Gilliland, A. J., & Woodley, M. S. 2000. *Introduction to metadata: Pathway to digital information*. Online edition, version 2.1. [online]. [cited 2010. 10. 9]. <http://www.getty.edu/research/conducting_research/standards/intrometadata/metadata.html>.
- [19] Greenberg, J. 2003. "Metadata and the World Wide Web." *Encyclopedia of Library and Information Science*, 72: 1876-1888. New York: Marcel Dekker, Inc.
- [20] Greenberg, J. 2005. "Understanding metadata and metadata schemes." *Cataloging & Classification Quarterly*, 40(3/4): 17-36.
- [21] Hakala, J. 1997. "Dublin Core in 1997: A report from Dublin Core metadata workshop 4 & 5." *Nordinfo-Nytt*, 20(3/4): 10-22.
- [22] Hanks, D. 2007. *All about metadata: An instruction for digital photographers*. [online]. [cited 2010. 9. 30]. <<http://www.davidhanks.com/AllAboutMetadata.pdf>>.
- [23] Heery, R. 1996. "Review of metadata formats." *Program*, 30(4): 345-373.
- [24] Hillmann, D. 2005. *Dublin Core Metadata Initiative: Using Dublin Core*. [online]. [cited 2010. 9. 13]. <<http://dublincore.org/documents/usageguide/>>.
- [25] Hodge, G. 2005. "Metadata for electronic information resources: From variety to interoperability." *Information Services & Use*, 25(1): 35-45.
- [26] Hunter, J. 2001. "MetaNet: A metadata term thesaurus to enable semantic interoperability between metadata domains." *Journal of Digital Information*, 1(8).

- [27] Krockel, H., & Westbrook, J. H. 1987. "Computerized materials-information systems." *Philosophical Transactions of the Royal Society of London, Series A, Mathematical and Physical Sciences*, 322(1567): 373-391.
- [28] Marchionini, G. 2007. "Toward multimedia surrogates." *Los Angeles Chapter of the American Society for Information Science and Technology Talk*, 12 January 2007.
- [29] McCarthy, J. L. 1982. "Metadata management for large statistical databases." *In Proceedings of the Eighth International Conference on Very Large Data Bases*, 8-10 September 1982. Mexico City, Mexico. 234-243.
- [30] Merriam-Webster Dictionary. 2007. Online edition. [online]. [cited 2010. 8. 30]. <<http://www.mw.com/dictionary/interoperability>>.
- [31] Milstead, J., & Feldman, S. 1999. "Metadata: Cataloging by any other name." *Online*, January 1999: 24-31.
- [32] National Aeronautics and Space Administration (NASA). 1990. *Directory Interchange Format Manual*, Version 3.0. Greenbelt, Maryland: NASA/GSFC National Space Science Data Center (NSSDC).
- [33] National Aeronautics and Space Administration (NASA). 1991. *Directory Interchange Format Manual*, Version 4.0. Greenbelt, Maryland: NASA/GSFC National Space Science Data Center (NSSDC).
- [34] O'Sullivan, M. 2007. The word "META: meaning and pronunciation." *In HTML code tutorial*. [online]. [cited 2010. 9. 6]. <http://www.htmlcodetutorial.com/document/index_tagstagsupp_14.html>.
- [35] Oxford English Dictionary. 2007. Online Edition. [online]. [cited 2010. 8. 30]. <http://dictionary.oed.com/cgi/entry/00307096?query_type=word&queryword=meta&first=1&max_to_show=10&sort_type=alpha&result_place=4&search_id=rpyw-Lc3m3ds-3152&hilite=00307096>.
- [36] Public Library Association, 2007. *Metadata: Always more than you think*. [online]. [cited 2010. 9. 3]. <<http://www.pla.org/ala/pla/plapubs/technotes/metadata.cfm>>.
- [37] Qin, J. 2000. "Representation and organization of information in the Web Space: From MARC to XML." *Informing Science*, 3(2): 83-88.
- [38] Qin, J. 2006. "Metadata as the underpinning of sustainable and effective access to scientific data." *In Proceedings of 20th International CODATA Conference: Scientific data and knowledge within the information society*, 22-25 October 2006. Beijing, China.
- [39] Sen, A. & Jacob, V. S. 1998. "Industrial-strength data warehousing." *Communications of*

- the ACM*, 41(9): 28-31.
- [40] Shelley, P. E. 1986. "Resolution Committee report." *In Proceedings of the 3rd International Conference on Geoscience Information*, 1-6 June 1986. Adelaide, South Australia, vol. 2: 116-117.
- [41] Sun Microsystem. 2002. *Digital Library Technology Trends*. Santa Clara, CA: Sun Microsystem, Inc.
- [42] Sundgren, B. 1973. *An infological approach to data bases*. Doctoral dissertation. School of Economics, University of Stockholm, Sweden.
- [43] Sundgren, B. 1975. *Theory of data bases*. 1st edition. New York: Petrocelli/Charter.
- [44] Taylor, A. G. 1999. *The organization of information*. Englewood, Colorado: Libraries Unlimited Inc.
- [45] Taylor, C. 2007. "Metadata's many meanings and uses: Briefing paper." *Catalogue & Index*, February 2007.
- [46] Terje, B. 1998. "A semantic modeling approach to metadata." *Internet Research: Electronic Networking Applications and Policy*, 8(5): 377-386.
- [47] Tozer, G. 1999. *Metadata management for information control and business success*. Norwood, MA: Artech House.
- [48] Weibel, S. 1997. "Discovering online resources. The Dublin Core: A simple content description model for electronic resources." *Arts and Humanities Data Services*. [online]. [cited 2010. 9. 10]. <http://ahds.ac.uk/public/metadata/disc_01.html>.
- [49] Wendler, R. 1999. "Branching out: Cataloging skills and functions in the digital age." *Journal of Internet Cataloging*, 2(1): 43-54.
- [50] Yu, S., Lu, K., & Chen, R. 2003. "Metadata management system: Design and implementation." *The Electronic Library*, 21(2): 154-164.

