

청소년 과학정보 메타데이터 요소 및 데이터베이스 구축 연구

A Study on Developing Metadata Elements and Database of the Science Information for Youth

곽 승 진(Seung-Jin Kwak)*

목 차

- | | |
|--------------------------|-------------------------|
| 1. 서론 | 3. 청소년 과학정보 메타데이터 DB 구축 |
| 1. 1 연구의 필요성 및 목적 | 3. 1 웹 자원 정보 수집 |
| 1. 2 연구의 방법 | 3. 2 메타데이터 작성 절차 |
| 2. 청소년 과학정보 메타데이터 요소 설계 | 3. 3 메타데이터 데이터베이스 구축 |
| 2. 1 메타데이터 요소의 도출과정 | 4. 메타데이터 시스템의 타당성 평가 |
| 2. 2 메타데이터 형식 참고 자료 | 4. 1 평가의 개요 및 절차 |
| 2. 3 참고한 메타데이터 세트의 비교 분석 | 4. 2 피실험자의 인구통계학적 특성 |
| 2. 4 메타데이터 요소의 추출과정 | 4. 3 타당성 평가 결과 |
| 2. 5 도출된 메타데이터 세트 | 5. 결론 |

초 록

본 연구는 메타데이터 시스템이 가지는 효율성을 기반으로 청소년을 대상으로 하는 인터넷 상의 과학정보 메타데이터 요소 추출 및 데이터베이스 시스템을 구축하여 이용자 중심의 타당성 평가를 목적으로 한다. 청소년 과학정보 메타데이터 기술요소 추출은 더블린코어 요소를 기준으로 국내외의 주요 메타데이터 프로젝트의 연구 자료를 비교하여 6개의 필수요소와 4개의 부가요소로 구성된 청소년 대상 과학정보 메타데이터 기술요소를 도출하였다. 이를 토대로 청소년을 대상으로 하는 과학정보 메타데이터 데이터베이스를 구축하였으며, 이용자 중심의 평가를 통해 피실험자 대부분이 과학문제 해결에 도움이 된다고 하였다. 이는 본 연구에서 제시하는 청소년 과학정보 메타데이터 서비스 시스템이 청소년들의 과학정보에 대한 접근성을 향상시킬 수 있음을 보여준다고 할 수 있다.

ABSTRACT

This study intends to design a metadata service system of the science information for youth on the web based on the efficiency of the metadata system. Metadata scheme of the science information for youth on the web has been developed and designed metadata collection on the basis of the previously designed classification system. Metadata scheme of the science information for youth, which is consisted of six essential elements and four additional elements, has been brought out, compared to not only Dublin Core Metadata Element Set, but also the main studies related to domestic and foreign metadata projects. Based on the results of it, metadata database of the science information has been designed and it is expected to be applicable to metadata service system of the science information for youth on the web.

키워드: 과학정보, 메타데이터, 메타데이터 요소, 청소년

Science Information, Metadata, Metadata Elements, Youth

* LG상남도서관 기획관리팀장(ksj@lg.or.kr)
논문접수일자 2004년 2월 18일
게재확정일자 2004년 3월 17일

1. 서론

1.1 연구의 필요성 및 목적

과학기술력이 21세기 국가 경쟁력을 좌우하는 중요한 요소 중의 하나로 인식되면서, 최근 정부와 과학관련 기관들이 청소년들의 과학과목 기피와 학력저하 현상의 사회 문제를 개선하고 이공계 인력의 육성과 확보를 위해 다양한 정책과 정보서비스를 개발하고 있다. 특히 청소년기부터 과학에 대한 관심과 흥미를 유도하기 위하여 과학영재 프로그램 등의 과학정책이 촉진되고 있으며, 이에 발맞추어 청소년을 대상으로 특성화된 온라인 과학정보 또한 한국과학문화재단 등에서 의욕적으로 생산하고 있다.

그러나 현재 각 기관에서 생성되고 있는 수많은 청소년 대상의 과학정보를 체계적으로 통합 제공하기 위한 메타데이터 구축 사례 및 연구가 질적으로 부족한 상황이며, 이로 인해 정보는 양적으로 풍부하지만 필요한 정보를 찾기 위해서는 많은 시간이 필요하고, 정확한 정보를 찾는 데 어려운 문제점이 있다.

자동색인 웹 검색엔진을 통한 정보검색의 지원이 정보홍수로 대변되는 정보의 양적 문제를 극복하기 위한 팔목할만한 발전이지만, 특히 주제검색에 있어서 너무 많은 정보가 출력되는 등의 문제를 해결하기에는 여전히 미흡한 실정이다. 이처럼 정보검색이 어려워진 요인의 하나는 현재 널리 사용되고 있는 텍스트 표현 마크업 언어인 HTML이 문서의 다양한 출력에는 매우 유용하지만 구조관계 태그가 거의 없다는 점에서 찾을 수 있다. 이러한

단점을 보완하기 위해 문서구조의 구체적인 표현이 가능한 XML(Extensible Markup Language)에 관한 논의가 활발히 진행되고 있으며, 오늘날 XML 문서들이 점점 증가해 감에 따라 검색 효과에도 상당한 진전이 이루어지고 있다. 그러나 XML은 문서 작성에 드는 시간과 노력이 많이 투입되어야 하는 단점 때문에 아직 보편화되어 있지는 않다. 또한 XML이 웹문서의 표준적 마크업 언어로 자리 잡고, 그 결과 문서의 구체적 속성에 의거한 자동색인 작업이 궤도에 오른다고 해도 인터넷 정보의 기하급수적 팽창에 따른 문제들을 해결하기 위한 노력은 계속되어야 할 것이다. 왜냐하면 너무 많은 자료가 인터넷에 산재해 있고 용어가 통일 되지 않아 이용자의 노력에도 불구하고 정보탐색 결과의 질이 보장될 수 없는 상황이 단순한 마크업 언어의 개선을 통해 해소된다고 보기는 힘들기 때문이다.

따라서 대량의 인터넷 정보를 체계적으로 선별하고 편목하여 신속하고 정확한 검색을 이루고자 하는 메타데이터(Metadata) 시스템은 다른 관점에서 인터넷자원의 공유와 이용을 돕는 효과적 시도라 할 수 있다(고영만, 오삼균 1999). 오늘날 특정 이용자들의 정보요구를 고려해 등장한 야후, 네이버 등의 대형 포털사이트(Portal Sites)들의 기술적 배경이 되는 도구로서의 메타데이터의 필요성과 유용성에 대한 추론이 가능하다.

청소년들이 일반인에 비해 정보기술에 대한 이해와 능력이 미숙한 측면을 고려한다면, 메타데이터 시스템 구축에 대해 더욱 개인화된 고려가 필요할 것이며, 효율적인 과학정보 메타데이터를 구축하기 위해서는 이용자의 관심

분야나 관심 키워드에 관련된 과학기술정보만을 제공할 수 있는 서비스개발에 초점을 맞추어야 한다. 이러한 개인화의 요구를 충족시키기 위해서는 이용자의 수준에 맞춘 적합한 분류체계의 개발과 이용자의 검색효율성을 고려한 메타데이터 요소의 설계가 필수적이라 할 수 있다.

1. 2 연구의 방법

바람직한 청소년 대상 과학정보 메타데이터 시스템을 개발하기 위한 기술요소 설계와 데이터베이스 구축 및 평가를 위해 본 연구는 다음과 같은 방법으로 연구를 진행하였다.

첫째, 청소년의 검색효율성을 고려한 메타데이터 요소 설계를 위해 국내외 관련 메타데이터 요소를 선별하여 조사하고, 더블린코어를 기준으로 상호 매핑하여 공통요소를 추출한다. 공통요소를 기반으로 필수와 선택 요소를 선별 및 추가하여 설계안을 제시한다.

둘째, 웹디렉토리 분류 체계는 따로 설계하지 않고 기 개발된 연구 결과를 반영하고, 이를 토대로 웹사이트 수집 로봇인 SWING을 이용하여 자동 수집과 필요에 따라 수동 수집을 병행하여 메타데이터 컬렉션 DB를 구축한다.

셋째, 구축된 청소년 과학정보 메타데이터 시스템을 실제 이용자인 과학교사와 우리나라 학제를 기준으로 청소년에 해당하는 중고등학교 학생을 대상으로 평가한다.

2. 청소년 과학정보 메타데이터 요소 설계

2. 1 메타데이터 요소의 도출과정

전통적으로 메타데이터는 데이터에 대한 데이터(Data of Data)라는 의미로 정의되지만, 오늘날에는 주로 네트워크상의 디지털 형태의 정보 자원에 대한 소재나 내용을 기술(Description)하고 또 네트워크 환경에서 이를 탐색하고 이용하기 위한 요소들의 집합이라는 의미로 통용된다. 청소년 대상의 과학정보 메타데이터 요소는 이용대상자인 청소년의 검색 효율성과 서비스 제공자의 메타데이터 작성의 편의성 및 관리의 효율성을 고려한 메타데이터 추출을 모색하였다.

청소년 대상 과학정보 메타데이터 요소를 도출하기 위한 과정은 다음과 같다. 첫째, 국내외 관련 메타데이터 요소 세트를 조사한다. 둘째, 청소년 대상 과학정보 메타데이터 기술 요소를 도출하기 위해 사용할 메타데이터 세트를 선별한다. 셋째, 선별된 각각의 메타데이터 세트를 구성하는 요소와 속성들을 나열하여 더블린 코어를 기준으로 상호 매핑하여 공통 요소를 추출한다. 넷째, 도출된 공통 요소 가운데 필수적이라 판단되는 요소의 선별 및 불필요한 요소를 배제한다. 다섯째, 공통 요소 이외의 요소 가운데 필요하다고 판단되는 요소를 선별하고 추가한다.

2. 2 메타데이터 형식 참고 자료

청소년 대상 과학정보 메타데이터 구축을

위한 메타데이터 요소를 도출하기 위하여 국내의 관련된 메타데이터를 조사하고 선별하는 과정을 통해 기초가 된 메타데이터 세트는 <표 1>과 같다. 아래 선별된 메타데이터 세트와 같이 전 세계적으로 메타데이터 개발의 기반이 되는 더블린 코어와 교육자료 메타데이터 프로젝트에서 성공적인 사례로 평가받는 IEEE LOM, GEM과 같은 해외 메타데이터 프로젝트 그리고 우리나라의 대표적인 학술정보 메타데이터를 구축한 한국교육학술정보원(KERIS)의 사례를 참고하였다.

1) Dublin Core

웹 자원을 기술하기 위한 대표적인 메타데이터로 데이터의 호환성을 유지하고, 네트워크 자원의 기술에 필요한 일련의 데이터 요소를 규정하며, 이들 자원의 신속한 검색을 목적으로 1995년 OCLC와 NCSA(National Centre for Supercomputer Applications)가 더블린(Dublin)에서 개최된 워크숍에서 합의된 메타데이터를 더블린코어라고 한다. 더블린코어는 자원의 소재위치와 이 자원을 획득하는데 필요한 데이터만으로 요소를 구성하고, 안전이나 검증, 구독 등과 관련된 요소는 제외하였다. 더블린코어에서 네트워크 자원의 식별에 필요한 데이터요소만으로 제한한 기본 의도는

데이터요소의 단순화를 통해서 유용성과 이용의 용이성을 확보하고, 필요시 다양한 목적에 따라 확장하기 위한 것이다. 본 연구에서 더블린코어는 인문사회연구회 산하기관의 DB 목록 요소 및 타 메타데이터 세트의 매핑을 위한 기준 도구로서 사용하였다.

2) IEEE Learning Object Metadata (LOM)

급속히 증가하는 교육 분야의 디지털 자원의 효율적인 관리와 검색의 용이성 증대를 위해 IEEE에 의해 개발된 메타데이터 세트이다. LOM은 다양한 형태의 교육 자료들에 필요한 최소한의 속성만을 정의함으로써 폭 넓은 적용성과 확장성을 갖도록 하는 데에 초점을 두었다. 특히 교육 분야의 디지털 자원이 급속도로 증가하고 있는 상황에서 그에 대한 정보나 메타데이터 없이는 자원의 관리와 검색에 많은 어려움이 있기 때문에 IEEE는 이러한 문제점을 해결하고자 LOM을 만들게 되었다.

3) GEM(Gateway to Educational Materials)

GEM 메타데이터는 1996년 미국 교육부가 인터넷을 통해 서비스할 수 있는 교육자료, 강

<표 1> 참고한 메타데이터

스키마 이름	국가	분야
Dublin Core	세계적	일반 웹 자원
IEEE Learning Object Metadata(3.6)	세계적	교육자료
GEM Element Set and Profile(s)	미국	교육자료
FIU DL metadata	미국	일반 웹 자원
KERIS 메타데이터	한국	학술연구자료

의계획서, 커리큘럼 등에 대한 접근성을 높이기 위하여 더블린코어의 기반위에 교육분야의 특수성을 고려한 몇 개의 요소를 추가하여 개발한 새로운 교육 자료 메타데이터이다. GEM의 메타데이터는 실제 자료에 대한 분석 결과와 이용자의 실제 요구사항을 토대로 개발된 것으로 먼저, 다양한 교수 자료들의 샘플을 내용 분석하여 교육자료의 유형과, 정보 요소를 규정하였다. 다음으로 교사들과의 인터뷰를 통하여 정보를 찾기 위하여 어떠한 탐색을 하는지, 정보 탐색 과정에서 중요하게 생각하는 요소는 무엇인지, 많이 이용하는 자료의 유형은 무엇인지 등 GEM이 내용분석을 통하여 도출해낸 요소들에 대하여 상세한 조사를 하였다. GEM은 개발된 메타데이터를 토대로 검색시스템을 구축하여 이를 기반으로 다양한 교육 자료에 대한 메타데이터 서비스를 제공하고 있으며, 인터넷 상의 모든 교육 자료들에 대한 원스톱 접근을 실현하는 것을 목표로 하고 있다. GEM의 역할은 더블린코어가 모든 유형의 웹 자원에 대한 일반적인 범위의 기술적인 요소를 제공한다면, GEM은 교육 분야에 대한 전문적이고 확장된 정보와 저자, 분야별 전문가들의 데이터 등을 첨가하여 도큐멘테이션을 지원할 수 있도록 하는 것이라 주장하였다(Sutton 1998).

4) FIU(Florida International University)
Digital Library

FIU의 디지털도서관 프로젝트에서는 Dublin Core를 기반으로 하되 몇 개의 자체적인 요소와 한정어를 추가하는 방식으로 새로운 메타데이터 세트를 만들었다.

5) KERIS 메타데이터

한국교육학술정보원에서 교육자료에 대한 효율적인 메타데이터 구축을 위해 개발한 메타데이터 세트로서 학술지 메타데이터, 학술 논문 메타데이터, 멀티미디어 교육 자료 메타데이터의 세 가지 세트를 구분하여 개발하였다. 본 연구에서는 학술지 메타데이터와 학술 논문 메타데이터 세트를 참조하였다.

2. 3 참고한 메타데이터 세트의 비교 분석

해외 및 국내에서 통용되고 있는 관련 메타데이터 스키마들을 조사하여 각각의 요소를 정리하였다. KERIS, IEEE LOM, GEM, FIU DL 등의 메타데이터 스키마 각각의 요소들을 Dublin Core 요소를 기준으로 매핑하여 정리한 결과는 <표 2>와 같다.

2. 4 메타데이터 요소의 추출과정

더블린 코어를 기준으로 타 메타데이터 요소를 매핑하여 공통의 의미를 가진 요소를 도출하고 그룹화 하여 정리한 <표 2>를 기준으로 청소년 대상의 메타데이터 요소를 도출하였으며, 그 과정은 다음과 같다.

첫째, 청소년의 자원 식별 및 선택에 필수적인 요소와 선택적인 요소 그리고 불필요한 요소가 무엇인가를 고려하였다. 둘째, <표 2>를 기준으로 필수요소, 부가요소, 불필요한 요소를 구분하였다. 셋째, <표 2>에 반영되지 않은 요소 가운데 추가할 필요가 있다고 판단되는 요소를 추가하였다. 위의 과정을 통해 도출된 메타데이터의 요소는 <표 3>과 같다.

〈표 2〉 메타데이터 요소의 매핑과정

Dublin Core	LOM	GEM	FIU DL	KERIS	
				학술논문	학술지
Title	Title	Title	Title	Title	Title
Creator	Contribute	Creator	Creator	Creator	Creator
Subject	Classification	Subject	Subject	Subject	Subject
Description	Description	Description	Description	Description	Description
Publisher	Contribute	Publisher	Publisher	Publisher	Publisher
Contributor		Contributor	Contributor	Contributor	Contributor
Type		Type	Resource Type	Type	Type
Date		Date		Date	Date
Format	Format	Format	Format	Format	Format
Identifier	Identifier	Identifier		Identifier	Identifier
Source		Source	Source	Source	Source
Language	Language	Language	Language	Language	Language
Relation	Relation	Relation	Relation		Relation
Coverage	Coverage	Coverage	Coverage place Coverage time		Coverage
Rights	Rights	Rights	Access Restrictions		

〈표 3〉 메타데이터 도출과정

	구분	내용	참고메타데이터	비고
필수요소	Title	제목	DC, LOM, GEM, FIU, KERIS	
	Identifier	식별자	DC, LOM, GEM, FIU, KERIS	URL
	Subject	주제	DC, LOM, GEM, FIU, KERIS	
	Audience levels	이용대상자 - 초등학생 - 중학생 - 고등학생 - 일반	LOM, GEM	더블린 코어 기본요소에는 해당하지 않으나 필요하다고 판단.
	Format	자원 형식	DC, LOM, GEM, FIU, KERIS	
	Date	날짜	DC, GEM, KERIS	
부가요소	Creator	저자	DC, LOM, GEM, FIU, KERIS	
	Description	설명	DC, LOM, GEM, FIU, KERIS	
	Type	자원 유형 - 참고자료 - 실험자료 - 교수학습자료 - 시험자료 - 대중매체	DC, GEM, FIU, KERIS	
	Language	언어	DC, LOM, GEM, FIU, KERIS	

2. 5 도출된 메타데이터 요소

도출된 메타데이터 세트는 크게 필수요소와 부가요소로 구성된다. 필수요소로 Title, Identifier, Subject, Audience, Format, Date의 요소는 더블린 코어를 비롯하여 참고한 전 메타데이터 세트에서 두루 사용하고 있는 요소일 뿐만 아니라 해당 자원을 식별하기 위한 가장 근본이 되는 요소이므로 필수요소로 정하였다. Audience 요소는 공통적인 요소로 도출되지는 않았지만, 청소년이라는 한정된 이용자 집단을 고려하고 자원선택에 도움을 주기 위해 필수요소로 선정하였으며, 학부모와 교사 등의 이용자를 고려한 일반 이용자 항목을 추가하였다.

Creator, Description, Type, Language와 같은 부가요소는 더블린 코어를 비롯하여 참

고한 대부분의 메타데이터 세트에서 사용하고 있는 요소이므로 포함했다. 그러나 웹 자원의 특성상 자원의 Creator(저자)나 Type(자원 유형)은 불명확한 경우가 있고, Format(자원 형식)과 Language(언어)는 자원의 선택에 도움은 되나 필수 요소로 선정한 요소들에 비해 자원선택의 기준을 제공하는 데 있어 자원의 내용에 해당하는 결정적인 요소는 아니므로 부가 요소로 선정하였다.

Contributor, Publisher, Relation, Source, Rights등과 같은 요소는 더블린 코어의 기본 요소지만 참고한 타 메타데이터 세트에서 반영하지 않은 경우가 있으며 자원의 내용과 직접적인 연관이 없으므로 반영하지 않았다. 도출된 메타데이터의 요소 및 속성과 내용을 정리하면 <표 4>와 같다.

<표 4> 청소년 대상 과학정보 메타데이터 셋

구분	요소	속성	설명	
필수요소	제목		자원에 부여된 제목	
	식별자		자원의 URL	
	주제	주제요소참조	자원의 주제에 관한 사항	
	이용 대상자		초등학생	
			중학생	
			고등학생	
			일반	
자원 형식		웹사이트	자원의 주요 이용 대상자	
		웹 문서		
		멀티미디어자료		
	날짜		메타데이터 작성일자	
부가요소	저자		자료를 작성한 사람 또는 기관의 이름	
	설명		자원의 내용과 관련된 설명	
	자원 유형		참고자료	자료의 내용적 성격
			실험자료	
			학습자료	
			시험자료	
			대중매체	
언어		한국어	자원 표현 언어 형식	
		영어		

3. 청소년 과학정보 메타데이터 DB 구축

3.1 웹 자원 정보 수집

메타데이터의 정보 수집을 위해서 먼저 수집대상사이트를 선정한 후 사이트 내에서 불필요한 콘텐츠를 제외하고, 과학과 관련된 콘텐츠만을 선별하는 작업을 수행하였다. 실제적인 웹 자원의 수집은 선별된 사이트의 세부적인 URL을 추출하여 SWING 시스템을 통해 자동으로 수집하였으며, 필요에 따라 수동 입력을 병행하였다. SWING 시스템은 K연구원이 서비스하는 대표적인 정보서비스의 하나인 웹디렉토리 서비스의 기본이 되는 시스템으로, 국내외 과학기술 관련 사이트의 콘텐츠를 통합검색하고 그 결과를 각각의 기술분류별로 재분류하여 제공하는 서비스이다.

메타데이터 구축을 위해 사용될 사이트의 선정은 주로 청소년 대상 과학정보 콘텐츠를 다루는 포털사이트 및 교사 및 교수 홈페이지, 대학 및 연구기관의 사이트 등을 중심으로 선별하였다. 웹사이트 선정에 사용된 내부 기준은 다음과 같다.

- 교과과정에 중심을 둔 분류체계라는 것을 감안하여 초, 중, 고등학교 교사의 개인홈페이지들 중 질적으로 우수한 사이트
- 포털사이트의 과학 분류에서 상위에 랭크되어 있는 사이트들
- 질적으로 우수한 과학 포털사이트
- 과학관련 도감, 사전
- 과학관련 대중매체

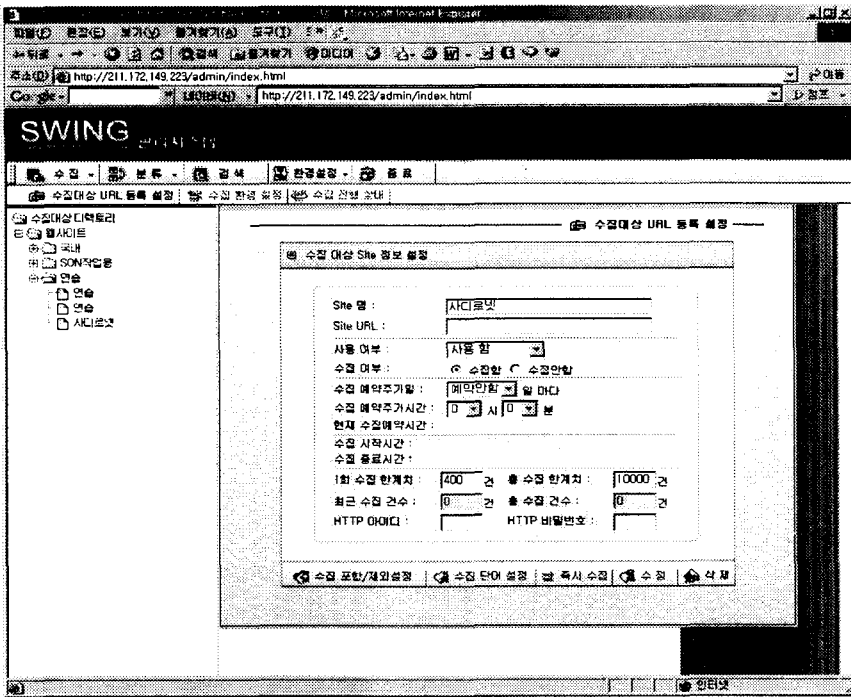
메타데이터 수집 시스템인 SWING을 이용하여 자동으로 웹 자원을 수집하는 과정은 다음과 같다. 먼저, SWING 시스템의 상위 메뉴 가운데 '수집' 과 하위 메뉴인 '수집대상 URL 등록 설정' 버튼을 이용한다. 다음으로 수집 대상 디렉토리를 선택하여, 수집된 웹 자원들이 저장될 디렉토리를 지정해 준다. '수집 대상 SITE 정보 설정' 창이 실행되면 수집 대상 사이트에 대한 상세한 정보를 입력하고 '즉시 수집' 버튼을 클릭한다. <그림 1>에서 보는 바와 같이 사이트의 URL과 수집 한계치 등을 설정하여 수집한다.

3.2 메타데이터 작성 절차

수집된 홈페이지 및 웹문서 그리고 멀티미디어 자료를 대상으로 SWING 시스템의 자동 및 수동 분류 방식과 입력기를 사용하여 메타데이터를 구축하였다. 입력하는 방법은 자동 입력과 수동 입력의 방식을 사용하였다.

메타데이터의 자동 입력은 구축 대상이 되는 사이트로부터 메타데이터 수집기가 자동으로 수집해 온 웹 자원들을 대상으로 메타데이터 작성자가 적절한 분류 코드를 할당하고 내용에 대한 가공 작업 등을 통하여 이루어졌다. 입력의 세부과정은 다음과 같이 설명할 수 있다.

첫째, SWING 시스템의 '분류' 메뉴 하위의 '수동/자동 분류' 메뉴를 이용하여 작성하고자 하는 웹 자원을 선택한다. 이 때, SWING 시스템의 여러 기능 가운데 '분류' 메뉴와 '수동/자동 분류'의 메뉴를 이용하되, 해당 자원의 수집 기준이 되는 날짜를 설정하여 구축 대상 자



〈그림 1〉 수집대상 SITE 정보 설정

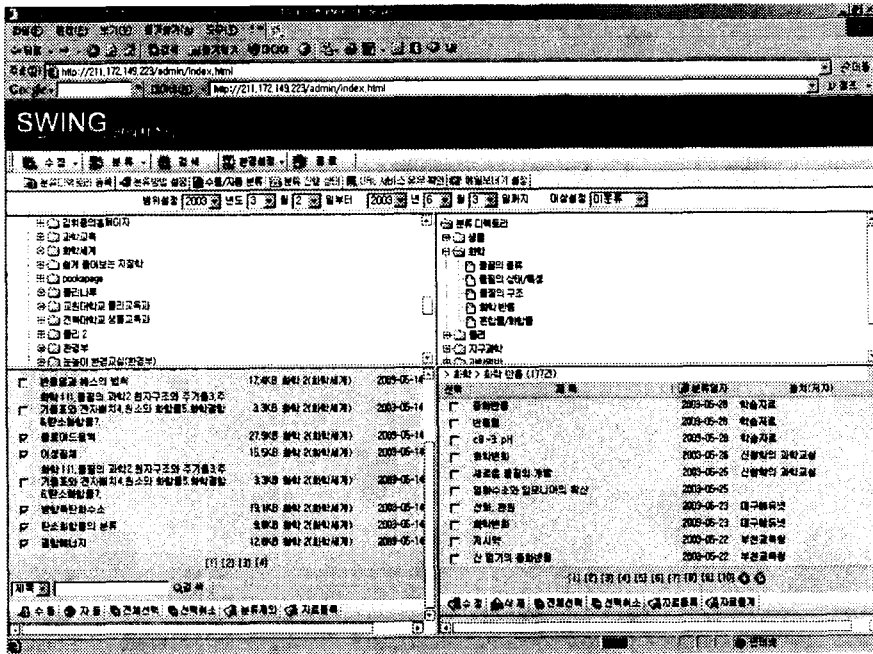
원을 적절히 선택하도록 한다. 수집대상 디렉토리로부터 수집된 여러 사이트들 가운데 입력하고자 하는 웹사이트 혹은 웹사이트의 세부 항목을 선택한다.

둘째, 수집대상 디렉토리로부터 구축 대상이 되는 웹 사이트를 선택한 후 실제 입력 대상이 되는 웹 자원을 선별한다.

셋째, 선택한 웹 자원들을 분류 디렉토리 가운데 적절한 분류 코드를 선택하여 해당 디렉토리로 할당한다. 〈그림 2〉에서 보는 바와 같이 그림의 우측 상단에 위치한 분류 디렉토리 가운데 입력 대상 웹 자원의 세부 주제를 선택한 후, 좌측 하단의 '수동' 버튼을 클릭하면 좌측 분류 대상 웹 자원이 우측 상단의 분류 디렉토리에 해당하는 주제로 분류가 완료되었음을 우측 하단의 결과를 통해 알 수 있다.

넷째, SWING 시스템을 통해 분류작업을 완료한 메타데이터를 작성자가 확인하고 수정한다. 분류된 웹 자원을 수정하기 위해 해당 웹 자원을 체크박스를 이용하여 선택하고 '수정' 버튼을 클릭한다. 다섯째, '수정하기' 창에서 작성자가 메타데이터의 내용을 수정한 후 '확인' 버튼을 클릭하여 입력작업을 완료한다.

수동 입력은 자동 입력을 통하여 메타데이터를 구축할 수 없는 경우에 한하여 사용하였다. SWING 시스템을 통하여 웹 자원을 수집할 경우 해당 사이트의 구조나 URL의 구조 그리고, 웹 자원의 형태에 따라 자동 수집이 어렵거나 불가능한 경우 이러한 문제점을 보완하기 위한 방안으로 수동 입력 방법을 선택하였다.



〈그림 2〉 분류코드 선택

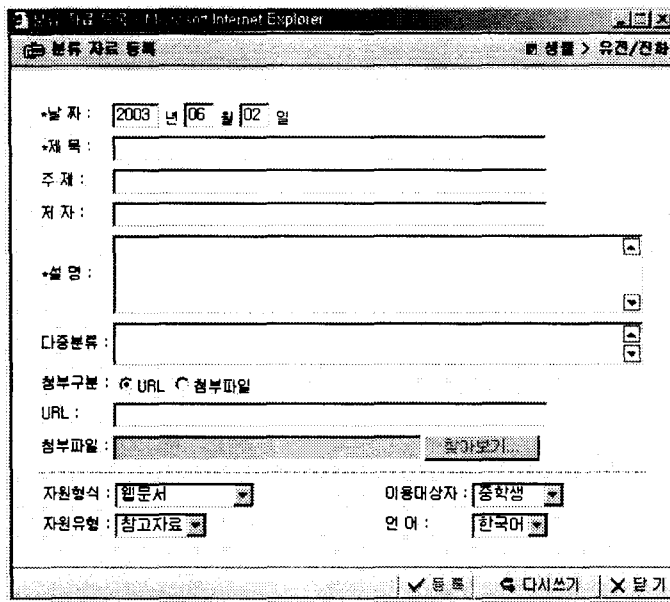
SWING 시스템을 이용한 수동 입력의 세부과정은 다음과 같다. 먼저 분류 디렉토리에서 수동 입력을 하고자 하는 웹 자원의 주제를 선택하고, 분류 디렉토리로부터 입력하고자 하는 웹 자원의 세부 주제를 선택한다. 다음으로 우측 하단의 '자료등록' 버튼을 클릭한다. '자료등록' 버튼을 클릭하면 〈그림 3〉에서 보는 바와 같이 수동으로 메타데이터를 입력할 수 있는 입력창이 실행된다. 각각의 메타데이터 요소에 해당하는 내용을 텍스트와 콤보박스에 설정된 값을 적절히 선택하여 입력하고, '등록' 버튼을 클릭하여 수동 입력을 완료한다.

3. 3 메타데이터 데이터베이스 구축

본 연구에서 구축한 청소년 과학정보 메타

데이터 데이터베이스는 주제별로 물리, 화학, 생물, 지구과학 등 6개 대주제 아래 데이터를 수집하였다. 데이터 수집 기간은 2003년 4월 1일부터 2003년 6월 30일까지 3개월간 데이터 수집 작업을 하였다. 수집된 데이터의 현황은 〈표 5〉와 같이 총 8,000개로 주제별로 비교적 균등하게 데이터를 수집하였으며, 이용대상별로 각 대상 간에 중복될 수 있도록 하였다.

청소년 과학정보 메타데이터 데이터베이스의 검색 화면은 〈그림 4〉와 같다. 상단의 검색창에 검색어를 입력하여 필요한 과학정보를 찾을 수 있으며, 검색창 우측에 도움말 기능을 주어 처음 시스템을 이용하는 사람도 쉽게 기능을 파악할 수 있도록 하였다. 검색창 아래 정보의 소재를 쉽게 찾을 수 있도록 주제별로 접근할 수 있는 주제별 디렉토리를 검색



〈그림 3〉 메타데이터 수동 입력 화면

〈표 5〉 주제별 메타데이터 DB 현황

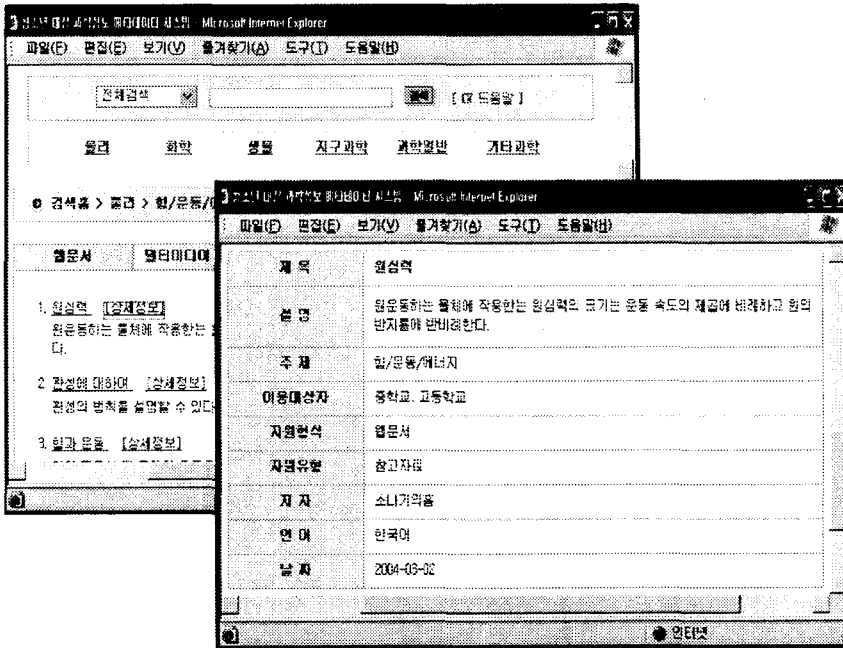
주제분야	데이터 수
물리	1,323
화학	1,202
생물	1,533
지구과학	1,487
과학일반	1,215
기타과학	1,240
합계	8,000

할 수 있다.

검색된 자료의 리스트를 보면서 찾고자 하는 자료를 클릭하면 해당하는 사이트로 바로 이동하도록 링크를 설정하였으며, 상세정보를 클릭하면 해당 사이트의 보다 상세한 정보를 볼 수 있다. 식별자인 자료의 URL은 따로 표기하지 않고 자료의 제목에 링크를 걸어 필요 시 바로 접근 할 수 있게 하였다.

본 연구에서는 청소년 대상의 과학정보 디

렉토리 분류체계는 설계하지 않고 기 개발된 과학분야 디렉토리 분류체계를 활용하였다. 이 분류체계는 이용자의 수준을 고려하여 6개의 대분류와 24개의 중분류로 간결하고 이해하기 쉬운 분류 체계를 도출하는 데 중점을 두었으며, 청소년들의 가장 효율적인 접근점이 교과과정임에 착안하여 7차 교육과정을 중심으로 분류를 전개하였다(곽승진 2003).



〈그림 4〉 청소년 과학정보 메타데이터 데이터베이스 검색 화면

4. 메타데이터 시스템의 타당성 평가

4.1 평가의 개요 및 절차

본 연구에서 개발한 청소년 과학정보 메타데이터 세트와 컬렉션 DB 시스템에 대하여 시스템의 실제 이용자인 과학교사와 청소년을 대상으로 이용자 중심의 타당성 평가를 실시하였다.

본 연구에서 선정된 피실험자들은 청소년 과학포털사이트인 LG사이언스랜드 회원을 대상으로 청소년에 해당하는 중·고등학교 학생과 과학교사로 구성하였다. 이 과학사이트는 2003년 12월 현재 중고등학교 학생 회원이 8천명, 교사 2천명, 기타 학부모 등이 8천명으

로 총 1만 8천 여명이다. 이 중에서 실험에 참여할 의사가 있는 학생과 과학교사 회원 약 1%에 해당하는 100명이 실험에 참여 하였다. 이들은 메타데이터 시스템을 통해 찾고자 하는 과학분야 탐색 질의 두 개를 가지고 실험에 참가하며, 실험이 끝난 후 시스템에 대한 타당성을 설문조사하여 정량적으로 평가하였다. 실험기간은 2004년 1월 5일부터 1월 17일 까지 총 13일간 진행되었다.

4.2 피실험자의 인구통계학적 특성

본 연구의 실험에 참가한 피실험자들은 100명으로 <표 6>과 같으며, 과학교사가 20명(20%), 학생이 80명(80%)이었다. 학생들의 학년은 중학생이 51명(64%)이고 고등학생이

〈표 6〉 피실험자의 학년 현황

학년		과학교사		학생		합계	
		인원(명)	%	인원	%	계	%
중학교	1학년	4	20	11	14	15	15
	2학년	6	30	22	28	28	28
	3학년	2	10	18	22	20	20
고등학교	1학년	5	25	16	20	21	21
	2학년	3	15	12	15	15	15
	3학년	0	0	1	1	1	1
합계		20	100	80	100	100	100

29명(37%)이었다.

피실험자들의 검색 행태를 알아보기 위하여 피실험자들이 정보를 찾을 때 주로 사용하는 검색 방법을 조사한 결과는 〈표 7〉과 같다. 가장 많은 응답으로 메타검색을 포함한 키워드형이라는 응답이 전체의 73%였으며, 다음으로 디렉토리 검색방법이었다. 정보를 찾을 때 검색방법이 일정하지 않다고 응답한 피실험자도 12%였다. 대부분의 피실험자들은 본인이 잘 알고 있는 정보를 찾을 때에는 자주 이용하는 인터넷 검색서비스에 접속하여 검색창에 키워드를 입력하여 검색을 실시한다고 하였다. 그리고 검색결과가 너무 많거나 관련된 정보를 추가로 알아보기 위해서, 또한 잘 모르는 정보를 찾을 때에는 주로 디렉토리 검색서비스를 주로 이용한다고 하였다.

4. 3 타당성 평가 결과

청소년 과학정보 메타데이터 시스템에 대한 과학교사와 학생들로 구성된 피실험자들의 타당성 평가는 과학문제 해결에 도움이 되는 정도, 메타데이터 시스템에 대한 만족도에 대한 것으로 평가 결과는 다음과 같다.

1) 과학문제 해결에 도움이 되는 정도

피실험자들이 실험에 참가하기 전에 본인들이 찾아보고자 하는 과학관련 질의 두 개를 가지고 청소년 과학정보 메타데이터 시스템을 이용한 후, 과학문제 해결에 어느 정도 도움이 되었는지에 대하여 설문문항을 리커트식 5단계 척도(5 point Likert Scale)로 설계하여 조사하였다. 조사 결과는 〈표 8〉과 같이 전체 76%

〈표 7〉 정보를 찾을 때 주로 사용하는 방법

내용	과학교사		학생		합계	
	인원수	%	인원수	%	인원수	%
키워드(메타검색)	16	80	57	71	73	73
디렉토리(주제별 검색)	3	15	12	15	15	15
일정하지 않음	1	5	11	14	12	12
합계	20	100	80	100	100	100

〈표 8〉 과학문제 해결에 도움이 되는 정도

내용	과학교사		학생		합계	
	인원수	%	인원수	%	인원수	%
많은 도움이 됨	3	15	20	25	23	23
도움이 되는 편임	12	60	41	51	53	53
보통	4	20	16	20	20	20
별 도움이 안됨	1	5	2	3	3	3
전혀 도움이 안됨	0	0	1	1	1	1
합계	20	100	80	100	100	100

가 “도움이 되는 편”이라고 응답하였다. 교사와 학생의 차이는 거의 없으나 “많은 도움이 된다”는 교사(15%)보다 학생(25%)이 10% 더 높았다.

2) 메타데이터 시스템에 대한 만족도

청소년을 위한 과학정보 메타데이터 시스템에 대하여 어느 정도 만족하고 있는지 설문항을 리커트식 5단계 척도(5)로 설계하여 조사하였다. 조사 결과는 〈표 9〉와 같이 전체 73% 이상이 “만족”이라고 응답하였다. 교사보다는 학생의 만족도가 조금 더 높았으며, “매우 만족”의 경우는 교사(5%)보다 학생(19%)이 15% 더 높았다.

청소년을 대상으로 한 과학정보 메타데이터

시스템이 만족스럽다면 어떤 부분인지 추가적으로 조사하였다. 조사 결과는 〈표 10〉과 같이 “참고할만한 유용한 자료가 많다”가 36%로 가장 높았다. 다음으로 “분류가 알아보기 쉽고, 멀티미디어 자료가 많다”가 28%, “정보 찾기가 편리하며, 검색결과가 만족스럽다”는 20%의 순으로 나타났다. 과학교사들은 대부분 참고할만한 유용한 정보가 많고, 분류가 편리하고 멀티미디어 정보가 많은 것에 만족하는 것으로 나타났지만, 학생들이 시스템에 대해 만족하는 부분은 과학교사들 보다 비교적 다양한 것이 차이점이라 할 수 있다.

본 연구에서 개발된 청소년 과학정보 메타데이터 시스템은 청소년들의 과학문제 해결에 도움이 되었으며, 또한 시스템의 이용에 대한

〈표 9〉 메타데이터 시스템에 대한 만족도

내용	과학교사		학생		합계	
	인원수	%	인원수	%	인원수	%
매우 만족	1	5	15	19	16	16
만족	13	65	44	55	57	57
보통	6	30	16	20	22	22
불만족	0	0	4	5	4	4
매우 불만족	0	0	1	1	1	1
합계	20	100	80	100	100	100

〈표 10〉 메타데이터 시스템의 만족스러운 부분

내용	과학교사		학생		합계	
	인원수	%	인원수	%	인원수	%
참고할만한 유용한 자료가 많다	10	50	26	33	36	36
분류가 알아보기 쉽고, 멀티미디어 정보가 많다	8	40	20	25	28	28
정보 찾기가 편리하며, 검색결과가 만족스럽다	2	10	18	22	20	20
디자인이 깔끔하고 인터페이스가 편리하다	0	0	16	20	16	16
합계	20	100	80	100	100	100

만족도도 높은 것으로 나타났다. 따라서 본 연구에서 제시한 청소년 과학정보 메타데이터 기술요소 및 데이터베이스 구축 등의 서비스 시스템에 대한 타당성이 있다고 할 수 있다.

5. 결론

본 연구는 국내외 청소년 대상 과학정보 사이트의 내용이 양적인 측면에서는 종류가 다양하고 또 콘텐츠의 수량도 많은 것으로 나타나고 있으나, 서비스와 콘텐츠의 질적 수준이 낮아 우수한 정보를 정확하게 찾기 어려운 문제점을 극복하기 위해 수행되었다. 이를 위해 본 연구에서는 청소년을 대상으로 하는 과학정보 메타데이터 시스템을 지원하는 메타데이터 기술요소를 개발하였으며, 이를 토대로 메타데이터 컬렉션 DB를 구축하여 이용자 중심의 평가를 하였다.

메타데이터 기술요소는 더블링크어 형식 및 IEEE, FIU DL, GEM, KERIS 프로젝트 연구 자료를 비교하여 관련 요소들을 매핑한 다음 필수 요소와 부가 요소로 도출하였다. 필수

요소는 제목, 식별자, 주제, 이용대상자, 자원 형식, 날짜로 구성되며 부가 요소는 저자, 설명, 자원유형, 언어가 선정되었다. 메타데이터 요소는 청소년을 위한 정보의 메타데이터임을 감안하여 필수적인 최소한의 요소와 함께 보다 개인화된 이용자의 접근점을 제공해 주기 위해 이용자 대상을 구분하기 위한 이용대상자(audience) 요소를 반영하고, 자원의 선택에 도움을 줄 수 있는 부가 요소를 반영하였다.

본 연구에서 구축한 청소년 과학정보 메타데이터 시스템은 물리, 화학, 생물 등 6개의 대주제 아래 8,000개의 데이터를 자동 및 수동 수집 시스템을 이용하여 데이터베이스를 구축하였다. 실험을 통해 이용자들 대부분이 본 연구에서 구축된 메타데이터 시스템이 본인의 과학문제 해결에 도움이 되고, 시스템 이용에 대한 만족도도 높은 것으로 나타났다.

이런 결과는 본 연구에서 제시하는 청소년 과학정보 메타데이터 요소와 서비스 시스템이 청소년을 대상으로 하는 과학정보 서비스의 질적 수준을 높일 수 있을 것이며, 청소년들의 과학정보에 대한 접근성을 향상시킬 수 있다는 것을 보여주는 것이라 할 수 있다.

참 고 문 헌

- 고영만, 오삼균. 1999. 인터넷 공공도서관 구축 모형 연구. 『정보관리학회지』, 16(4): 109-123.
- 곽승진. 2003. 『청소년 대상 과학분야 메타검색 시스템과 교과별 디렉토리시스템의 구축 및 사용성 평가 연구』, 박사학위논문, 성균관대학교 대학원, 문헌정보학과.
- 류범중, 이응봉. 2002. 해외 기술정보 메타 데이터베이스 구축 및 관리 시스템에 관한 연구. 『한국문헌정보학회지』, 36(3): 201-214.
- 이은경, 오삼균. 2001. 인터넷 대학강의안 전산학분야 메타데이터 시스템 구축 및 평가. 『정보관리학회지』, 18(1): 65-84.
- 최한석. 2001. 메타데이터 기반 학술지 논문 종합목록시스템 설계 및 구현. 『정보관리학회지』, 18(2): 57-76.
- 한국교육학술정보원. 2002. 『교육용 콘텐츠 메타데이터 작성 및 활용 안내서』, 서울: 한국교육학술정보원, 사업보고 PR 2002-3.
- Draft Standard for Learning Object Metadata. 2002. [cited 2004.1.11]. <http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf>.
- Dublin Core Metadata Initiative(DCMI). 1995. [cited 2004.2.10]. <<http://dublincore.org/index.shtml>>.
- Dublin Core Metadata Initiative(DCMI) Education Working Group. 1999. [cited 2004.2.10]. <<http://dublincore.org/groups/education/>>.
- Gateway to Educational Materials. 2001. [cited 2004.2.8]. <<http://www.geminfo.org/>>.
- Hill, Linda L. 1999. "Collection Metadata Solution for Digital Library Applications." *Journal of the American Society for Information Science*, 50(13): 1169-1181.
- Hoffman, J., 1999. *Information Seeking Strategies and Science Content Understandings of Sixth Grade Students Using On-line Learning Environments*. Ph.D. diss., University of Michigan.
- IEEE P1484.12 Learning Object Metadata Working Group. 2001. [cited 2003.10.11]. <<http://ltsc.ieee.org/wg12/indea.html>>.
- Koch, T., et. al 1997. "The Role of Classification Schemes in Internet Resource Description and Discovery." [cited 2003.12.8]. <<http://www.ub2.lu.se/desire/radar/reports/D3.2.3/>>.
- Sutton, Stuart A., et al. "Networked Information Discovery and Retrieval for Educational Materials on the Internet: Metadata Development, Deployment, and Evolution." *ASIS*

- '98(1998): 54-62.
- Thomas Fischer and Heike N. 2000. "SSG-FI -Special Subject Gateways to High Quality Internet Resources for Scientific Users." *Online Information Review*, 24(1): 64-68.
- Vizine-Goetz, D. 1996. "Using Library Classification Schemes for Internet OCLC Internet Cataloging Colloquium." [cited 2003.12.10]. <<http://www.oclc.org/oclc/man/colloq/v-g.htm>>.