

메타데이터를 이용한 기계공학분야 웹 정보자원의 주제 게이트웨이 설계에 관한 연구

A study on the Design of Subject Gateway of Web Information Resource Using Metadata in Mechanical Engineering

송재윤, 이응봉, 충남대학교 대학원 문헌정보학과

Jae-Yun Song, Eung-Bong Lee

Graduate School of Library & Information Science, Chungnam National University

본 고에서는 공학분야 중에서도 그 응용범위가 가장 광범위하다고 할 수 있는 기계공학분야의 이용자에게 수준 높은 웹 정보자원을 효율적으로 제공하기 위하여 검색엔진을 통하여 자동 수집된 웹 정보자원을 기계공학분야 주제전문가의 평가를 통해 핵심정보를 선정하고 이를 더블링크어 메타데이터를 응용한 메타데이터 요소에 따라 조직하여 웹 상의 검색 인터페이스에 의해 이용될 수 있도록 주제 게이트웨이 시스템을 설계하였다. 정보자원의 검색방법으로는 키워드 검색과 주제분류체계에 따른 브라우징 검색 등 두 가지 방법을 이용하였다.

1. 서 론

정보통신기술과 인터넷의 발달은 정보이용환경에 큰 변화를 일으키고 있다. 컴퓨터와 네트워크의 이용이 가능한 곳에서는 누구든지 필요한 정보를 찾을 수 있고 이용할 수 있게 된 것이다. 인터넷을 통한 정보자원의 폭발적인 증가는 이용자로 하여금 정보선택의 범위를 넓혀 주기도 하지만 반대로 핵심정보를 선택하게 할 수 있는 정보선택의 효율성 측면에서는 많은 문제점을 노출하기도 한다.

인터넷의 이용이 보편화되면서 지금은 정보의 생산자와 소비자의 구별이 없어지고 있으며, 이는 누구나 정보의 생산자 또는 소비자가 될 수 있다는 것을 의미한다. 자원의 등록에 있어 공식적인 통제가 이루어지지 않고 있기 때문에 인터넷 상의 정보자원들은 양적인 면은 풍부할지 모르지만 질적인 면에서는 수준 이하의 정보원들이 포함될 가능성 또한 적지 않다.

현재 이렇게 광범위하고 다양한 인터넷 정보자원을 효과적으로 수집, 활용하도록 도와주는 대표적인 기술이 각종 검색엔진들이다. 각 검색

색엔진들은 로봇 에이전트를 이용하여 인터넷 상에 흩어져 있는 정보들을 수집하고 이들 정보의 내용을 자동으로 색인하여, 이용자가 찾고자 하는 정보에 대한 색인정보, 위치정보, 요약정보들, 즉 인터넷 데이터들에 대한 데이터인 메타데이터를 제공해 준다.

최근 들어 인터넷 정보자원의 검색효율성을 높이기 위한 여러 가지 노력으로 메타 검색엔진이나 메타 디렉토리과 같은 웹 탐색 서비스를 통해 이용자의 요구에 적합한 특정 정보의 위치를 식별하도록 도와주고 있다. 그러나 이러한 검색엔진들은 인터넷 정보자원에 대한 단순한 위치정보만을 나타내는데 불과할 뿐 정보의 질적 수준을 전문적으로 차별화해 주지는 못하고 있다. 효율적인 인터넷 정보서비스를 위해서는 인터넷 상의 방대한 양의 정보를 신속하게 접근할 수 있으면서 그 정보의 질적 수준을 보장해 줄 수 있는 또 다른 서비스 체계가 필요하다.

주제 게이트웨이는 수많은 웹 정보자원 중에서 질적으로 우수한 자원의 접근을 위해 특정 주제분야의 주제전문가와 정보전문가가 직접 정보자원을 평가하여 선정한 후 각 자원에 대한 구체적인 기술사항을 메타데이터를 이용하여 작성해 놓은 시스템이라고 할 수 있다.

본 연구의 목적은 공학분야 중에서도 그 응용범위가 가장 광범위하다고 할 수 있는 기계공학분야의 웹 정보자원을 검색로봇에 의해 자동으로 수집을 한 후, 다시 메타데이터를 이용하여 조직하는 주제 게이트웨이를 설계함으로써 기계공학분야의 이용자에게 수준 높은 웹 정보자원을 효율적으로 제공하고자 하는 것이다.

연구의 방법은 첫째, 기계공학분야의 전문 주제분류체계를 선정한 후 검색로봇 에이전트를 이용하여 기계공학관련 웹 정보자원들을 수집하여 색인한다. 둘째, 수집된 웹 정보자원을 대상으로 기계공학을 전공한 주제전문가가 사이트 평가를 통해 핵심 정보자원을 선정한다. 셋

째, 선정된 핵심 정보자원들을 더블링크어의 메타데이터 요소를 응용한 변형 메타데이터에 의해 조직한다. 넷째 조직된 정보자원에 대해 검색 인터페이스를 구축한다. 이용자 검색방법은 키워드 검색과 주제분류체계에 따른 브라우징 검색의 두 가지 방법을 이용한다.

2. 국내·외 연구개발 동향

메타데이터를 이용한 주제별 게이트웨이에 대한 연구는 웹 정보자원이 급격히 증가한 1990년대 후반부터 활발해졌다고 할 수 있다.

국내의 경우 주제 게이트웨이에 관한 연구로 오삼균(1999)은 인터넷 대학강의안의 검색을 위한 Metadata DB 구축에서 Gateway to Faculty Syllabi(GFS)로 명명한 메타데이터 데이터베이스 시스템을 소개하면서 특정 기관 또는 이용자 그룹을 배려한 인터넷 정보의 여과 필요성 및 그 의의를 메타데이터라는 틀 안에서 고찰하였다.

이숙희(2001)는 인터넷상의 주제별 정보 게이트웨이에 관한 논문에서 인터넷상의 학술 정보자원을 주제별로 엄선하여 전문가가 직접 질적 수준을 통제하고 있는 주제별 정보 게이트웨이(Subject-Based Information Gateways)의 서비스 내용과 구조에 대해 연구하였다.

한중엽(2001)은 해양수산분야 전자정보자원 메타데이터 시스템 구축 방안 연구에서 해양수산분야의 웹 정보자원을 포함한 전자정보자원의 효율적인 메타데이터 시스템 구축을 수립하기 위한 방안을 제시하고 있다.

국내의 주제 게이트웨이 구축 사례로는 LG 상남도서관에서 홈페이지를 통해 과학기술분야에 대한 주제별 게이트웨이 서비스를 시행하고 있다. 각 주제별로 메타데이터를 이용해 웹 정보자원을 목록하고 자원유형에 따라 정보를 제공해 주고 있다.

국외의 경우 Traugott Koch(2000)는 질적으로 통제된 주제 게이트웨이에 대한 정의를 기술하면서 현재 서비스되고 있는 주제 게이트웨이의 경험적인 비교를 통해 각 주제 게이트웨이들이 어떠한 형태로 서비스되고 있는지를 분석하였다.

Thomas Fischer 등(2000)은 독일의 지구과학, 수학, 문화 분야의 대표적인 주제 게이트웨이인 SSG-FI 프로젝트에 대한 연구를 수행하였다.

일본의 경우 尾城孝一(2000)은 일본동경공업대학 부속도서관에서 서비스하고 있는 이공학분야의 주제 게이트웨이인 TDL서비스의 구축과 운영에 대한 논문을 발표하였다.

국외의 주제분야별 정보 게이트웨이 구축 사례로는 공학분야의 EEVL과 EELS, 의학분야의 OMNI, 사회과학분야의 SOSIG, 미술, 디자인, 건축, 미디어 분야의 ADAM, 화학분야의

MetaChem 등이 있다.

최근의 연구개발 동향을 살펴보면 대부분의 경우 포괄적인 주제분야를 다루고 있는 것을 알 수 있으며, 보다 세분화된 주제분야에 대한 웹 정보자원의 주제 게이트웨이에 대한 연구는 미약하다고 할 수 있다. 또한 국내의 경우는 국외에 비해 그 연구개발이 아직 초기 단계에 머물고 있는 실정이며 데이터베이스에 수록된 자료의 양이 매우 부족하고 자원평가 방법이나 메타데이터의 작성기준 등이 체계화되어 있지 않다는 것을 알 수 있다.

3. 기계공학분야의 주제분류체계 비교

분류체계를 활용한 웹 정보자원의 조직화는 주제접근 도구로서 이용자에게 정확한 주제접

<표 1> 기계공학분야 주제분류체계 비교

EI Classification	DDC	LC	KDC
600 기계공학, 일반	620 공학과 관련 작용	TJ 기계공학과 기계장치	550 기계공학
610 발전소와 발전기	620.1 공학기계와 재료	170-177 기계응용	551 기계역학, 재료 및 설계
620 원자력기술	620.2 음향과 관련 진동	181-210 기계운동	552 공구와 제작장비
630 유체역학	620.3 기계적 진동	212-225 제어공학,	553 열공학과 원동기
640 열과 열역학	621 응용물리학	자동제어시스템	554 유체공학, 기억학, 진공학
650 항공우주공학	621.1 증기공학	241-254 기계구조, 일반	555 정밀기계
660 자동차공학	621.2 수력기술	255-265 열 엔진	556 자동차공학
670 조선과 해양공학	621.4 열공학과 원동기	267 터빈	557 철도공학, 기관차
680 철도공학	621.5 공압, 진공, 저온기술	268-740 증기공학	558 항공우주공학, 우주항법학
690 재료처리	621.6 송풍기, 팬, 펌프	603-695 기관차	559 기타공학
	621.8 기계공학	751-830 기타 모터와 엔진	
	625 철도공학	840-935 수압기계	
	627 수력공학	940 진공기술	
	629 기타공학	950-1030 공압기계	
	629.04 교통(수송)공학	1040-1119 원동기를 제외한 기계	
	629.1 항공우주공학	1125-1345 기계공장 및 업무	
	629.2 내연육상차량	1350-1418 승강기와 운송기	
	629.3 공기쿠션차량	1425-1475 리프팅과 압력기	
	629.4 우주비행학	1480-1496 농업기계	
	629.8 자동제어공학		

근과 주제어의 제시, 그리고 검색용어의 문맥적인 의미와 브라우징 기능을 제공한다는 점에서 그 의의가 있을 것이다.

기계공학분야는 그 응용범위가 매우 광범위하고 여러 공학분야에 복합적으로 연관되어 있기 때문에 정확한 주제분류가 매우 까다롭다고 할 수 있다.

<표 1>은 공학분야의 대표적인 주제 게이트웨이인 EELS와 EEVL에서 사용하고 있는 Engineering Information사의 주제분류체계와 기존의 도서분류체계 중 가장 보편적으로 사용되고 있는 DDC와 KDC, LC분류체계에서 기계공학분야를 비교한 것이다.

각 분류체계를 비교, 분석한 결과 공학분야의 전문 분류체계인 EI Classification이 기계공학분야의 세부주제를 가장 현실적이고 체계적으로 잘 표현하고 있는 것으로 나타나 본 연구를 위한 주제분류체계로 채택하였다.

주제분류체계에 따른 웹 정보자원의 조직과 분류는 주제별 웹 디렉토리의 기능을 가지면서 이용자로 하여금 웹 정보자원의 검색 효율성을 높이고 브라우징을 용이하게 하는 주제접근 도구로서의 역할을 할 수 있을 것이다.

4. 주제 게이트웨이 구축을 위한 메타데이터 요소 선정

현재 각 주제분야별로 다양한 형태의 메타데이터가 존재하고 있다. 이는 이용자의 수준과 응용분야마다 요구되는 데이터 요소와 수준이 다르기 때문에 각 주제분야에 적합한 메타데이터를 채택하여 사용하자는 요구 때문일 것이다.

본 연구에서 사용되는 메타데이터 요소는 더블링크어 메타데이터의 15가지 요소와 스웨덴의 대표적인 공학분야 주제 게이트웨이인 EELS의 메타데이터 요소를 참고하여 기계공학

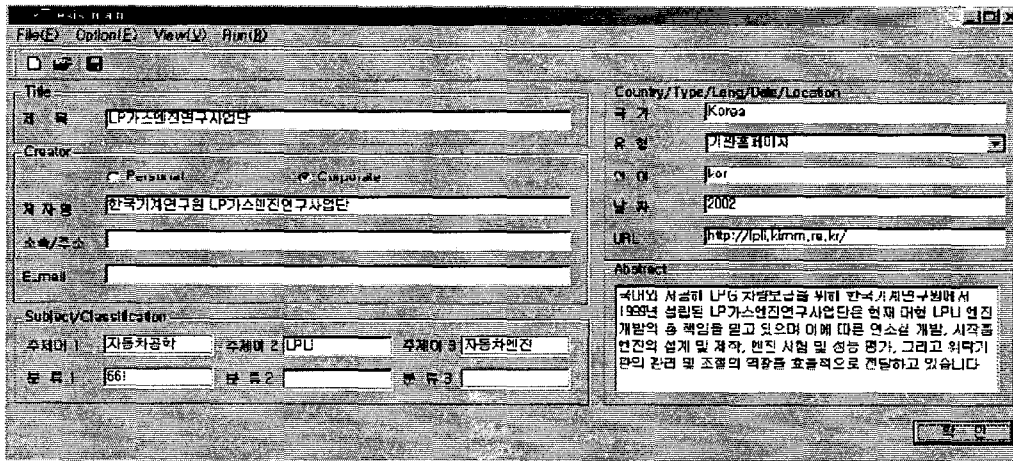
분야의 주제 게이트웨이에 적합한 10개의 메타데이터 요소를 선정하였다.

메타데이터 요소 중 자원의 유형으로는 기관 홈페이지, 전자저널, 전자회의, 국제회의안내, 목차정보, 데이터베이스, 특허규격, 카탈로그, 용어집, 소프트웨어, 메일링리스트, FAQ, 게시판, 디렉토리, 동향정보의 15가지를 지정하기로 하였다.

<표 2>는 기계공학분야 주제 게이트웨이 구축을 위해 선정한 메타데이터 요소와 내용이며, <그림 1>은 채택된 메타데이터 요소에 따라 자원을 기술하기 위한 메타데이터 입력 화면의 예이다.

<표 2> 채택된 메타데이터 요소

요소명	정의	비고
제목 (Title)	제작자나 발행자가 자원에 부여한 제목	DC EELS
제작자 (Creator)	자원의 내용에 책임을 진 개인이나 기관	DC
국가 (Country)	자원의 내용에 책임을 진 국가	EELS
요약 (Description)	자원의 내용에 대한 기술	DC EELS
주제 (Subject)	자원의 주제를 표현한 단어나 구	DC EELS
분류 (Classifications)	자원의 주제에 따른 분류	DC EELS
유형 (Type)	자원의 유형	DC EELS
언어 (Language)	자원의 내용을 기술한 언어	DC EELS
날짜 (Date)	자원이 제작되거나 수정된 날짜	DC
URL	자원의 위치	DC EELS



<그림 1> 메타데이터 입력 화면의 예

5. 웹 정보자원의 선정기준

주제 게이트웨이 시스템의 목적이 질적으로 우수한 웹 정보자원의 검색과 접근에 있다고 보았을 때 그 자원의 질적인 평가를 위해서는 공식적인 선정기준이 마련되어 있어야 하며 주제전문가의 참여도 필수적이라 할 수 있다.

본 연구에서 제시한 기계공학 주제 게이트웨이의 정보자원 선정은 한국기계연구원 소속의 기계공학전공 박사급 학위를 취득한 주제전문가들의 사이트 평가를 통해서 이루어지도록 하였다.

주제 게이트웨이의 데이터베이스에 수록하기 위해 채택되어질 수 있는 정보자원의 선정을 위해서 다음 4가지의 선정기준을 제시하고자 한다.

첫째는 접근성이다. 정보자원은 링크가 되어 있어야 하며 접근이 가능해야 한다.

둘째는 유지보수이다. 사이트는 지속적으로 내용이 갱신되어야 하며 항상 최신성을 유지하고 있어야 한다.

셋째는 문서화(documentation)이다. 정보원 자체에 대해서 식별이 가능하도록 이용 가능한

최소한의 정보를 가지고 있어야 한다.

넷째는 정보자원 생산기관의 책임성이다. 정보원 제작기관에 대한 상세한 정보와 정보원의 유용성, 정보제공에 대한 책임 및 저작권 문제 등이 해결되어야 한다.

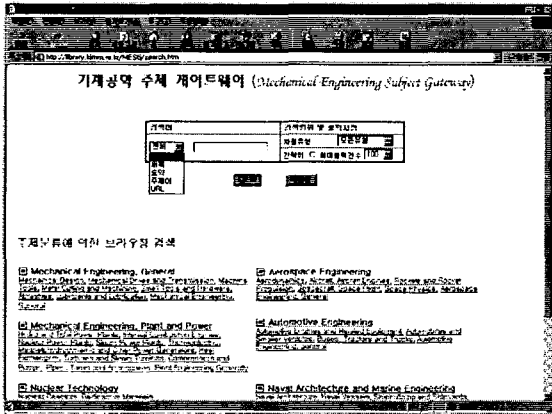
이러한 네 가지 선정기준을 바탕으로 하여 주제전문가들에 의해 웹 정보자원의 평가가 이루어지고 데이터베이스를 위한 정보원으로서 선정이 되어져야 할 것이다.

6. 기계공학분야 주제 게이트웨이를 위한 검색 인터페이스 구축

주제 게이트웨이 설계를 위해 먼저 <그림 1>의 윈도우 상에서 작동하는 전용의 메타데이터 입력 어플리케이션을 이용하여 웹 정보자원의 내용을 각 항목에 입력한다. 등록된 메타데이터는 일단 DBMS에 저장되며 일괄처리에 의해 검색용 서버로 이동이 된 후 웹을 통한 검색 인터페이스에 의해 검색이 이루어지도록 한다.

메타데이터의 검색방법은 키워드를 통한 불리언 검색을 이용하는 방법과 주제분류체계에

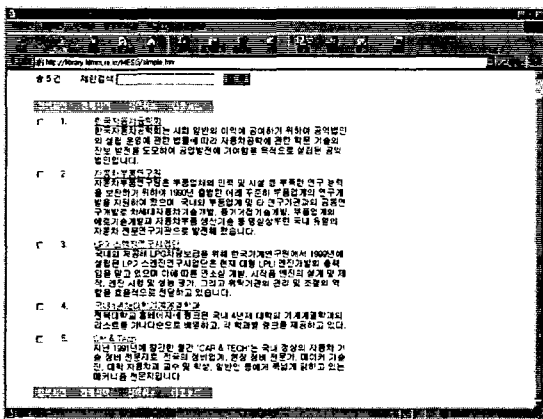
다른 브라우징 검색방법이 있다.



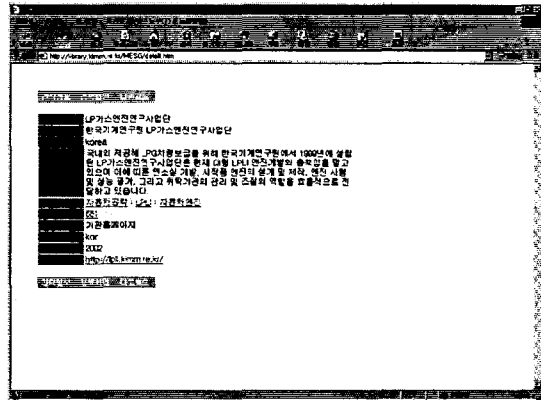
<그림 2> 기본 검색인터페이스 화면의 예

<그림 2>는 기계공학분야 주제 게이트웨이의 기본 검색 인터페이스의 예이다. 키워드 검색은 제목과 요약, 주제, 그리고 URL 등의 필드별 제한검색이 가능하며 자원유형에 따른 검색 또한 가능하다. 이용자가 필드와 관련된 여러 단어를 입력하면 단어와 단어 사이는 자동적으로 'AND' 연산자로 처리가 되며, 불리언 검색식을 통해서 제한적인 검색이 가능하다.

<그림 3>은 검색결과의 간략정보 화면을 보여주고 있다. 간략정보 화면을 통해서 해당 정보원의 제목과 요약을 확인할 수 있다.

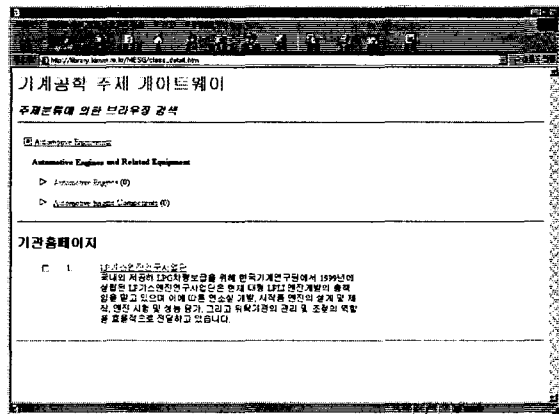


<그림 3> 간략정보 화면의 예



<그림 4> 상세정보 화면의 예

간략정보의 제목을 클릭하면 <그림 4>와 같이 메타데이터의 상세정보 화면이 나타난다. 상세정보 화면에서는 앞에서 선정된 메타데이터 요소에 따라 웹 정보자원이 기술되고 있다. 실제 웹 정보자원으로의 연결을 위해서는 URL을 클릭함으로써 가능하다.



<그림 5> 브라우징 검색 화면의 예

또 다른 검색방법으로는 주제분류에 의한 브라우징 검색방법이 있다. <그림 2>의 검색 인터페이스 초기화면에서는 EI 분류체계에 따른 10개의 대분류와 대분류에 포함된 하위 중분류까지를 기술하고 있는데, 이용자는 관심있는 주제에 대해 계층적으로 따라가다 보면 대분

류, 중분류, 소분류 순으로 세부주제와 관련된 웹 정보자원에 대해 체계적으로 브라우징을 할 수가 있다

<그림 5>는 주제분류에 따른 브라우징 검색의 예이다. 대분류 Automotive Engineering의 중분류인 Automotive Engines and Related Equipment를 브라우징 한 결과를 보여주고 있다. 주제와 관련된 웹 정보자원이 검색되고 결과화면은 키워드 검색과 동일하다는 것을 알 수 있다.

7. 결 론

본 연구는 방대한 웹 정보자원 중에서 기계공학이라는 특정 주제분야의 자원을 주제전문가에 의해 평가하고 선정하여 메타데이터를 이용해 기술해 줌으로써 양질의 웹 정보자원을 검색하고 이용할 수 있는 주제 게이트웨이를 구축하기 위한 것이다. 일반 검색엔진과는 달리 주제전문가 집단에 의해 웹 정보자원의 평가와 분석이 이루어지고 그 결과를 바탕으로 메타데이터를 이용해 각 자원의 접근을 보장하는 기술 사항이 작성되기 때문에 주제 게이트웨이의 이용자는 정보 자체의 정확성과 신뢰성을 보장 받을 수 있다.

최근에 늘어나는 웹 정보자원의 증가량을 고려한다면 특정 전문주제와 관련된 주제 게이트웨이의 구축은 필수적이라 할 것이다.

주제별 정보 게이트웨이가 시험적 운영 단계를 넘어 전문적인 정보시스템으로 발전을 하기 위해서는 늘어나는 웹 정보자원의 양적 규모에 비례해서 주제전문가의 필터링을 통해 과연 얼마만큼의 정보량을 어떻게 관리할 수 있는가 하는 문제가 남아있다. 이것이 해결된다면 주제 게이트웨이가 지니고 있는 질적인 특성을 최대한 활용할 수 있으며 유용한 웹 정보자원의 제공을 위한 인터넷 검색 도구로서 발전을

할 수 있을 것이다. 이를 위해서는 주제전문가의 자발적인 참여가 절실히 요구되며, 필요하다면 정보보상제를 도입하여 양질의 정보자원에 대한 평가를 수행하고 웹 정보자원의 선정에 기여한 주제전문가에게 합당한 혜택이 돌아갈 수 있는 정책이 마련되어야 할 것이다. 사서 또한 양질의 웹 정보자원을 선정하고 조직할 수 있는 주제전문성을 키워 인터넷 상의 새로운 역할을 수행해 나갈 필요가 있다.

참고문헌

- 이숙희. 2001. “인터넷상의 주제별 정보 게이트웨이에 관한 연구”. 『한국문헌정보학회지』. 35(1) : 27-45
- 한종엽. 2001. “해양수산분야 전자정보자원 메타데이터 시스템 구축방안 연구”. 『제8회 한국정보관리학회 학술대회 논문집』. : 117-124
- 오삼균. 1999. “인터넷 대학강의안의 검색을 위한 Metadata DB 구축”. 『정보관리학회지』. 32 : 149-164
- Traugott Koch. 2000. “Quality-controlled subject gateways: definitions, typologies, empirical overview”. 『Online Information Review』. 24(1) : 24-34
- 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
- 尾城孝一. 2000. “サブジェクトゲートウェイの構築と運営 - 理工學分野の高品質なインターネットリソースの提供をめざして-”. 『情報の科學と技術』. 50(5) : 280-289

문헌정보처리연구회. 1998. 『메타데이터의
형식과 구조』. 서울: 문헌정보처리연구
회

LG 상남도서관 <<http://www.lg.or.kr>>

Engineering E-Library, Sweden (EELS)
<<http://eels.lub.lu.se>>

東京工業大學電子圖書館(Titech Digital
Library) <<http://tdl.libra.titech.ac.jp>>