

국내 주요 검색 포털의 디렉터리 서비스 정보자원 선정 및 분류작업 개선방안

Improving Methods for Resources Selection and Classification Practice of Major Korean Directories

김 성 원*
Sung-Won Kim

차 례

- | | |
|-----------------------------|-----------------------|
| 1. 서 론 | 4. 디렉터리 서비스 분류작업 개선방안 |
| 2. 연구방법 및 범위 | 5. 결 론 |
| 3. 디렉터리 서비스 분류작업 분석 및
평가 | • 참고문헌 |

초 록

인터넷을 통한 정보의 유통량은 크게 증가하였으나 정보의 저장, 유통, 그리고 검색의 비효율성이 상존하고 있으며, 이를 보완하기 위한 방안으로 상당수의 검색 포털은 분류체계에 따라 정보자료를 조직화하여 제공하는 디렉터리 서비스를 제공하고 있다. 이 연구에서는 국내의 대표적 인터넷 검색 포털에서 제공되고 있는 디렉터리 서비스의 실제 분류작업을 살펴보고, 이를 기반으로 디렉터리 서비스의 개선안을 제시하였다.

키 워 드

인터넷, 정보검색, 디렉터리 서비스, 분류체계, 분류작업

*연세대학교 대학원 문헌정보학 박사졸업
(Ph. D. of Library & Information Science, Yonsei Graduate School, sungwonk@paran.com)
• 논문접수일자: 2005년 11월 15일
• 게재확정일자: 2005년 12월 5일

ABSTRACT

While the amount of information exchanged through internet has dramatically increased recently, certain inefficiencies still exist with regard to the storage, distribution, and retrieval of information. As a means of improving efficiency in accessing information, many search portals provide directory services to present organized guidance to information, based on the classification schemes. This study examines the classification activities practiced by the major search portals in Korea and makes some suggestions to improve the quality of directory services.

KEYWORDS

Internet, Information Retrieval, Directory Service, Classification Schemes, Classification Practice

1. 서 론

인터넷을 통한 정보의 유통량이 크게 증가함에 따라 인터넷을 통한 유용정보의 취득 가능성이 높아졌다. 인터넷을 통해 유통되는 정보는 양적으로는 폭발적으로 증가하고 있으나 질적인 측면에서는 검색된 정보의 질적 수준을 보장하기 어렵다. 또한 전통적 형태의 정보자원에 비해 가변적이고 불안정하다는 특성을 지닌다.

조직적이지 못한 정보의 저장, 유통 및 키워드 중심의 검색행태를 보완하고 검색효율을 제고하기 위한 노력의 일환으로 대다수의 검색 포털에서는 디렉터리 서비스를 통해 조직화된 정보검색 서비스를 제공하고 있다. 디렉터리 서비스는 키워드 검색결과를 확장, 세분화 해 주고 인터넷 정보자료를 주제별로 조직화할 수 있다는 장점을 지니는 반면, 비논리적인 분류

항목 전개와 실제 분류작업 과정에서의 오류로 인해 관련자료를 분산시킬 수 있다. 이는 키워드 검색에 비해 비교적 양질의 정보들을 선택하여 분류해놓은 디렉터리 서비스가 비논리적인 분류체계나 작업상의 오류로 인해 오히려 여러 가지 부작용과 제한점을 발생시킬 수 있음을 의미한다.

이 연구에서는 국내의 대표적 인터넷 검색 포털에서 제공되고 있는 디렉터리 서비스의 실제 분류작업의 결과를 살펴보고, 이를 기반으로 개선안을 제시하여 디렉터리 서비스를 통한 정보검색 효율을 향상시킬 수 있는 방안을 제시하는 것을 목적으로 한다.

2. 연구방법론 및 선행연구 개관

2.1 연구방법 및 범위

검색 포털에서 제공하는 디렉터리 서비스를 분석해 보고자 하는 본 연구에서는 네이버(<http://www.naver.com>), 야후 코리아(<http://kr.yahoo.com>), 엠파스(<http://www.empas.com>) 등 세 포털에서 제공하고 있는 디렉터리 서비스를 연구대상으로 선정하였다. 연구대상 선정배경은 이들 세 포털이 국내의 주요 포털중 검색기능을 중심으로 하는 3대 포털로 인식되고 있는 것에 기반하였다.

디렉터리 서비스 분류작업 결과의 이론적 분석을 위해 문헌정보학 분야의 관련 이론에 기반하여 연구의 대상으로 선정된 포털 사이트 디렉터리 서비스의 분류작업을 분석 평가하였다. 구체적인 분류작업의 분석은 뉴스, 미디어 분야를 중심으로 수행하였다. 다만, 분석 및 평가항목별로 더 적합하고 빈도가 높은 항목이 있는 경우 뉴스, 미디어 이외의 일부분야를 분석대상에 포함시켰다. 뉴스, 미디어 분야를 선택한 이유는 연구대상으로 선정된 세 포털에서 공통적으로 전개하고 있는 최상위 항목이며, 디렉터리 서비스를 통한 검색의 비율이 가장 높은 항목이라는 선행연구 결과를 반영하였다(이준호, 박소연, 권혁성 2003).

디렉터리 서비스에 대한 이 연구는 다음의 관점에서 중요성과 의의를 찾을 수 있다. 첫째, 이론적인 관점에서 디렉터리 서비스는 키워드 중심의 검색과 비교하여 인간을 통해 조직된 정보를 대상으로 브라우징, 검색의 확장 과 특정화, 문맥파악, 그리고 정보요구의 명확화 기능 등을 제공하므로 검색효율을 향상시

킬 수 있다는 점이다. 둘째, 실제적인 관점에서 이용자의 검색행태를 분석하여, 키워드 검색이 주류를 이루기는 하지만 디렉터리 서비스를 통한 검색비율도 상당한 비중을 차지한다는 점이다.

이 연구는 몇 가지 제한점을 지니고 있다. 우선 연구의 대상을 국내 주요 웹 검색 포털에서 제공하는 세 디렉터리 서비스만으로 제한하였으므로 좀더 많은 포털의 디렉터리 서비스를 포함하는 연구를 수행할 필요가 있다. 주제분야에 있어서도 심층분석은 뉴스, 미디어 분야를 중심으로 하였던 바, 더 많은 주제분야를 포함하는 연구를 수행할 필요가 있다.

2.2 선행연구 개관

디렉터리 서비스 관련 선행연구는 크게 문헌정보학 분야에서 개발된 문헌분류체계를 인터넷 정보자원을 대상으로 하는 검색 서비스에 적용하는 것과 관련된 연구, 특정주제 분야의 인터넷 정보자원의 조직화를 위한 디렉터리 분류체계의 구축에 관련된 연구를 중심으로 이루어졌다. 또한 이와는 별도로 디렉터리 서비스의 전반적인 구축지침을 제시하거나 실제 디렉터리 서비스 이용행태 분석과 관련된 연구들이 수행되고 있다.

문헌정보학 분야에서 도서관 정보자원의 조직화를 위해 개발된 분류체계를 인터넷 정보자원에 적용하는 방안에 관련된 연구로 다음과 같은 주요 연구들이 있다.

스베노니어스(Svenonius 1983)는 온라인 검색 환경에서 도서관 분류체계를 적용하는 것이 검색효율을 높이는 데 도움이 된다는 것을 밝혔다. 마키와 드마이어(Markey and Demeyer 1986)는 DDC를 활용한 온라인 프로젝트에서 도서관 분류체계가 이용자의 주제접근, 브라우징 등에 있어 유용성이 크다는 것을 밝혔다. 달베르그(Dahlberg 1995)는 인터넷 정보자원에 LCC와 DDC를 적용하는 문제를 검토하여 분류체계의 적용이 검색효율 향상에 유용함을 밝혔다. OCLC의 비진 괴에츠(Vizine Goetz 1996, 2002)는 1996년과 2002년의 연구에서 도서관 분류체계인 LCC와 DDC를 인터넷 디렉터리 서비스의 분류체계와 비교하여 계층구조, 자모순 배열, 다언어 자원 등의 적용가능성 등의 상관성이 있음을 밝힌 바 있다. 코흐(Koch 1997)는 DESIRE (Development of a European Service for Information on Research and Education) 프로젝트의 일부로 수행된 연구에서 DDC, LCC, UDC, BC 등 10여개의 현존 전통적 문헌 분류체계가 인터넷 서비스 환경에서 어떻게 적용되어 활용되고 있는지를 고찰하고 그 장단점과 향후 잠재성, 타 시스템과의 호환성, 확장성 등에 대해 논했다. 국내의 연구로 남영준(1998)은 KDC를 기반으로 전 주제분야의 인터넷 정보자원을 효율적으로 검색할 수 있는 새로운 분류체계를 제시하였다. 한국데이터베이스진흥센터(1998)는 연구과제를 통해 인터넷 정보자원의 효율적인 검색을 위한 분류체계 작

성지침을 제시하고 과학기술분야의 분류체계를 제시하였다. 박철완(2002)은 전자도서관을 위한 분류체계 구축을 위해 필요한 연구과제와 특성을 제시하였다.

박창호, 염성숙, 이정모(1999)는 사용자 중심의 분류체계가 검색 효율에 미치는 영향을 분석하였다. 이란주(1999)는 웹 검색 엔진의 기능과 특징에 대해 전반적으로 고찰한 바 있다. 김정현(2001)은 인터넷 정보자원의 조직화 필요성 및 MARC를 이용한 목록문제를 고찰하였다. 신동민(2001)은 디렉터리 분류체계 구성시 적용할 10개 지침을 제시하였다. 이우범(2002)은 사회과학 분야를 중심으로 디렉터리 분류체계가 인터넷 정보자원의 활용도를 제고할 수 있음을 밝혔다. 남영준, 최승순(2002)은 디렉터리 분류체계를 분석하여 설계원칙 및 KDC와의 통합분류체계를 제시하였다.

특정주제 분야를 중심으로 인터넷 정보자원의 조직화를 위한 디렉터리 분류체계의 구축방안이나 지침제시 관련 연구로는 김영보(1997), 최희운(1998), 최재황(1998), 이명희(2000) 등 다수의 연구가 국내에서 수행되었다.

3. 디렉터리 서비스 분류작업 분석 및 평가

이 장에서는 디렉터리 서비스의 분류체계를 활용하여 인터넷 정보자원을 디렉터리 서비스에 등록하는 분류작업 과정 및 결과의 문제점들을 제시하고 평가해 보도록 한다.

3.1 정보자원 등록기준 분석

방대한 인터넷 정보자원을 대상으로 하는 디렉터리 서비스는 높은 지적 능력을 지닌 인간에 의해 선별된 정보자원을 제공한다는 점에서 키워드 검색과 차별화된 효용을 갖는다. 디렉터리 서비스의 이러한 특화된 효용을 고려한다면 등록하여 제공할 정보자원의 선정기준을 정하고 적용할 필요가 있다. 세 디렉터리 서비스에 등록된 정보자원을 살펴보면, 디렉터리 서비스에서 제공할 가치가 떨어지는 정보자원까지 지나치게 광범위하게 포함하고 있다.

등록여부 결정과정에서 적용할 인터넷 자원의 정보자원으로서의 가치 평가기준과 관련하여 상당수의 연구가 수행되었다. 관련 연구들을 종합하면 인터넷을 통해 유통되는 정보자원은 크게 접근성, 내용의 질적 수준, 이용 편의성의 세 가지 측면에서 평가될 수 있다(안인자 1997; 배금표 1999; 이용봉 1999; 이해영, 남태우 1999; 황혜경, 정영미 1999; 남태우, 이해영 2001). 이들의 연구를 종합하여 정보자원의 등록여부를 결정하는 과정에서 참고할 수 있는 자세한 세부기준들을 정리하면 다음과 같다.

- 사이트의 접근성과 가용성
- 정보의 안정성과 유지관리보수의 품질
- 저자/발행기관의 신뢰도와 권위
- 정확성
- 정보의 구조
- 객관성
- 최신성

- 정보자원의 목적
- 정보의 범주 및 범위
- 대상이용자 및 이용자의 요구에 대한 적합성
- 링크의 품질과 유효성
- 이용 편의성
- 정보의 일람성과 조직성
- 경제성

모든 등록후보 정보자원에 대해 완벽하게 등록여부의 선정기준을 적용하는 것이 필요하겠으나, 특히 등록의 필요성이 의심되는 정보자원에 대해서는 신중한 검토를 통해 나름대로의 지표별 선정기준을 정하고 적용함으로써 디렉터리 서비스의 질적 향상을 도모해야 한다.

현재 각 디렉터리 서비스에 등록된 정보자원을 살펴보면, 등록 정보자원으로서의 가치가 의심되는 경우가 상당수 있음은 전술한 바와 같다. 이들 정보자원은 몇 가지 유형으로 나누어 볼 수 있다.

우선 자사 별도 서비스 항목의 개별 페이지들을 디렉터리 서비스에 사이트로 등록한 경우가 많다. 예를 들어, 네이버 <엔터테인먼트>류의 '영화 > 영화작품 > 데드 캠프'에 등록된 사이트인 '데드 캠프 작품정보'는 사이트가 아니라 네이버의 콘텐츠 서비스인 '네이버 영화' 중 한 페이지일 뿐이다. <인물, 사람들>류의 '배우' 종려시' 항목에 등록된 '종려시 이미지 모음'은 사이트가 아니고 네이버의 이미지 검색 서비스 결과 페이지를 등록한 것이며, '종려시 배우소개' 사이트 역시 위에서 예시한 '네이버 영화' 서비스의 한 페이지일 뿐이다. 야후 <음

악)류의 '가수, 음악가 > 국내 가수 > 윤도현 밴드 > 앨범 소개' 항목에 전개된 8개 항목 모두 '야후! 비트 박스' 서비스 아래의 웹 페이지들이다.

항목주제와 관련된 보도기사나 인터넷 게시판의 게시내용을 사이트로 등록한 경우도 상당수 있다. 엠파스의 <뉴스, 미디어>류의 '뉴스 > 뉴스 인물' 아래 전개된 개인별 항목들에 등록된 사이트의 상당수는 해당 인물에 대한 보도기사나 게시판상의 게시내용을 등록하고 있다. 예를 들어, '길지빈' 항목 아래 등록된 '빵떡이는 아가씨 길지빈'은 '엠파스 유행'에 게재되어 있는 기사이며, 그나마 자체 작성된 것이 아니라 외부에서 인용한 내용이다. '뉴스 > 뉴스 인물 > 김영완' 항목에 등록된 6개 사이트는 '엠파스 뉴스' 서비스에서 외부 보도 기사를 클리핑한 페이지를 등록한 경우이다. 이런 페이지 단위까지 디렉터리 서비스의 분류대상으로 포함시킬 필요는 없다. 또한 '김영완'의 예에서는 6개 등록 사이트 중 세 사이트가 접속이 안되는 삭제 사이트로 지나친 세부 페이지를 정보자원으로 등록해서는 안되는 근거가 된다.

다른 한 유형으로 자사의 주요 사업분야관련 정보자원을 과도하게 등록한 경우이다. 네이버 <게임>류 '게임 > 게임명 > 릴 온라인 > 프리뷰'에 등록된 세 정보자원은 한 두 쪽짜리 웹 페이지로서 아무리 게임 분야가 네이버의 주요 사업분야임을 고려하더라도 등록대상으로 적당치 않은 정보자원으로 판단된다.

이상에서 볼 수 있는 바와 같이 세 디렉터리 서비스 모두 광범한 정보자원의 망라적 제공을 위해 디렉터리 서비스 등록 자원의 선정에 있어 엄격한 기준을 적용하고 있지는 않다. 또한 자사의 서비스나 사업분야에 대해서는 우대하고 있음도 볼 수 있다.

디렉터리 서비스 등록대상에 관해 명확한 확립이 필요하다. 사이트가 아니라 사이트내의 한 페이지만을 분류의 대상으로 선정하는 경우는 되도록 지양되어야 한다. 특별히 의미 있는 정보를 제공하는 것으로 판단하여 등록을 결정하는 경우라고 하더라도 최소한 메뉴 이상의 정보단위로 제한하는 것이 필요하다. 한 두 쪽 짜리 웹 정보자원을 사이트로 등록하는 것은 선별된 정보자원을 제공한다는, 달리 말하면 재현율보다는 정확률이 높은 정보검색을 제공한다는 디렉터리 서비스의 특화된 효용을 저하시킬 수 있다.

3.2 사이트 유효성관리 평가

한 번 등록된 사이트의 유효성관리 문제는 정보자원의 접근성, 가용성, 안정성에 관련된 항목으로 디렉터리 서비스의 기본적인 기능에 해당된다. 분류자는 많은 노력과 시간을 투자하여 정보자원을 수집·선정·조직하고, 이용자는 많은 노력을 들여 검색을 수행하고 결과를 확인한 후에만 디렉터리 서비스에 등록된 사이트에 접속할 수 있다. 만약, 어렵게 조직하고 검색한 정보자원에 접속을 시도하였는데,

접속이 되지 않는다면 이용자는 차라리 검색결과에 포함되어 있지 않은 경우보다 더 많은 좌절과 디렉터리 서비스에 대한 불만을 느끼게 될 것이다. 등록과 관련하여 등록 사이트의 유효성과 등록내용의 변동여부를 지속적으로 검토할 장치가 필요하다. 등록 사이트의 유효성과 관련하여 다음의 몇 가지 문제를 발견할 수 있다.

더 이상 유효하지 않은 사이트를 등록하고 있는 경우가 있다. 등록된 사이트의 상당수가 사이트 폐쇄나 서버 이전 등의 이유로 더 이상 접속이 불가능하다. 디렉터리 서비스 관리자에 의해 발견된 경우 “접속이 불안정한 사이트입니다” 혹은 “정상적으로 연결되지 않는 사이트입니다” 라는 안내가 되고 있다. 하지만, 관리자에 의해 파악되지 않은 상당수의 유효하지 않은 등록 사이트를 발견할 수 있다. 이렇게 접속이 불안하거나 정상적인 연결이 어려운 사이트는 시차를 두고 몇 차례 검토하여 사이트 복구 어려운 것으로 판단되는 경우 제적해야 한다. 2003년 10월 네이버에 등록된 사이트 200개를 무작위 추출하여 검토한 결과, 폐쇄되었거나 접속이 불가능한 상태인 사이트는 12개로 무작위 추출한 검토대상 등록 정보자원의 6%에 달하는 것을 발견할 수 있었다.

사이트 자체에 접속은 가능했으나 사이트 성격이 완전히 변화된 경우가 있다. 물론, 디렉터리 서비스의 등록내용은 기존의 내용을 담고 있고 현행화되어 있지 않다. 예를 들어, 엠파스의 경우 기업, 쇼핑몰 > 결혼, 만남 > 결혼

항목에 분류되어 있는 ‘웨딩 TV(www.iweddingtv.co.kr) 웨딩 컨설팅, 웨딩 사진, 드레스, 신혼여행, 가구, 한복, 예물, 웨딩 TV 결혼바람회 제공’ 사이트를 클릭하면 전혀 다른 URL을 가진 엔코코(<http://www.nkoko.com/>)라는 검색 포털 성격의 사이트로 연결된다. 또한 ‘생활, 취미 > 가족, 가정 > 육아 > 육아일기’ 항목에 분류된 ‘노은이네집(web22.bchome.co.kr) 김노은 출생, 육아일기, 앨범, 가족소식’을 클릭하면 역시 URL이 <http://makehome.bcline.com/ems/ed0000p0>인 이메일 마케팅 솔루션 서비스 이용 페이지로 접속된다. 즉, 디렉터리에 소개된 사이트와 접속된 사이트는 URL 및 사이트 성격이 전혀 다르다.

사이트가 폐쇄되어 사이트 호스팅 업체나 쇼핑몰 호스팅 업체, 그리고 다른 사이트로 리디렉션 되는 경우가 있다. 이런 경우는 대부분 자체 서버를 운용하기 어려운 중소기업업체가 호스팅 업체의 서버를 이용하다 사이트 운용을 포기한 경우이며 쇼핑, 기업 등의 항목에서 많이 발견할 수 있다. 예를 들어, 네이버의 경우 ‘기업, 쇼핑 > 의류, 패션 잡화 > 여성 의류’ 항목에 분류된 ‘쉬즈카페’(<http://www.shezcafe.com/>) 사이트에 접속하면 쇼핑몰 호스팅 업체인 ‘Cafe24 EC Hosting’(<http://echosting.cafe24.com/>)의 휴면 쇼핑몰 안내 페이지로 연결된다.

이 외에도 사이트의 서버 및 URL(Uniform Resource Locator)을 유지한 채 다른 서버로

정보자원을 옮긴 후, 기존 등록된 사이트에는 새로이 개설한 사이트로의 링크만을 제공하는 경우도 있다.

이상의 몇 가지 경우에서 알 수 있듯이 웹 정보자원의 불안정성의 정도는 관리의 한계를 벗어나 있는 지도 모른다. 하지만 주기적으로 등록된 사이트에 접속하여 정보자원의 접속가능성과 기존 등록내용과 차이를 파악한 후, 변동이 있는 경우 다시 확인하는 방안과 같은 적합한 방안을 강구하여 사이트의 유효성과 등록정보의 정확성을 지속적으로 검토, 관리해야 한다.

3.3 동일 정보자원의 중복분류 평가

키워드 검색의 등장으로 전산작업을 통한 다수의 접근점을 제공할 수 있게 됨으로써 제한된 수의 접근점 밖에는 제공할 수 없었던 기존 정보조직 상의 문제점은 상당부분 해소되었다고 할 수 있다. 문헌이라는 매체의 물리적인 특성상 서가 상의 특정위치를 지정하는 서가 분류기호는 하나밖에 지정할 수 없었고, 서지 분류의 경우는 제한적으로 부여되어 왔다. 이에 따라 정보자료의 주제가 여러 분류항목에 분류될 수 있는 자료의 경우에는 분류가 아닌 주제명이나 저자명으로 추가적인 접근점을 제공하는 것이 일반적이다.

인터넷 정보자원은 물리적인 특성상 기존 도서관의 도서나 비책자와 같이 특정위치에 존재하는 것이 아니므로 정보원에 적합한 여러 분류항목에 중복분류할 수 있고, 이를 통해 재

현율(Recall Ratio)을 향상시킬 수 있다. 디렉터리 서비스를 제공하는 경우 이러한 장점을 살려 가능한 해당되는 정보자원을 여러 분류항목에 중복분류 할 수 있다.

중복분류는 '분류체계' 및 '분류작업' 모두에 관련된 문제이다. 분류체계의 관점에서 분류항목 간 '상호배타성' (mutually exclusiveness) 이 완전한 상태라면 동일주제는 복수개의 관련 분류항목을 가질 수가 없으며, 중복분류 작업도 고려할 필요가 없다. 동일주제가 분류될 수 있는 복수 개의 분류항목의 전개는 지양되어야 하나, 앞에서 살핀 바와 같이 사용자들에게 좀더 다양한 접근점을 주기 위한 시도에서 동일계층에서 상이한 구분기준을 동시에 적용하여 중복되는 항목이 설치되어 있는 현실이다. 분류작업의 관점에서 특정주제와 정보자원에 대해 적용가능한 복수의 분류항목이 있다면 존재하는 유사 분류항목들에 대해 빠짐없이 분류하여 이용자의 상이한 분류경로 선택 시 해당 자원이 제공되지 못하고 누락되는 경우를 방지하는 것이 더 잘된 분류작업이라고 할 수 있을 것이다.

중복분류와 관련하여 중요한 변수는 '@' 로 표시되는 참조 링크이다. 참조 링크의 활용은 중복분류 작업의 필요성을 상당히 낮춰주기 때문이다. 특정 분류항목에서 주제의 중복이나 관련성이 있는 모든 분야에 빠짐없이 참조 링크를 해주었다면, 그 분류항목 하나에 분류하는 것으로 중복분류 작업의 효과를 발휘할 수 있다. 이는 중복분류를 부여하려 했던 분류항목이 이미 참조 링크로 연계되기 때문이다.

따라서 분류체계의 관점에서는 항목 간 배타성을 확보하면서 중복되는 분류항목을 지양해야 하며, 다양한 접근점을 제공한다는 목적에서 불가피하게 중복성 있는 분류항목을 전개하였다면 항목 간의 참조 링크를 적절하게 처리하여야 한다. 분류작업의 관점에서는 중복분류가 가능한 항목에 포괄적으로 분류작업을 처리하되 항목 간의 참조 링크를 최대한 활용하여야 한다.

2005년 1월 수집한 데이터를 분석한 결과 세 디렉터리 서비스의 중복분류 현황을 <뉴스, 미디어>류를 중심으로 살펴보면, 네이버와 야후는 동일류 내에서의 중복분류는 지양하고 있는 것으로 밝혀졌다. 세 디렉터리 서비스 <뉴스, 미디어>류에 등록된 정보자원 중 참조 링크

된 분류항목에 등록된 정보자원을 제외하면 네이버 3,532건, 야후 2,376건, 엠파스 3,776건이었다. 이 가운데 둘 이상의 분류항목에 중복분류되어 2회 이상 수집된 정보자원은 엠파스의 경우 '한겨레21', 'am7', 'Seven Days' 등 모두 55건이었다. 네이버와 야후는 중복분류된 정보자원이 없었다. 네이버의 경우 데이터 처리과정에서 두건이 중복분류된 것으로 파악되었으나 확인결과 '이너모스트(<http://innermost.org>)'가 '이너모스트'로 동일 분류항목에 중복분류되어 있고, 두개의 URL로 등록된 동일 사이트인 '부천일보(<http://www.puchon.co.kr>)'와 '부천닷컴(<http://www.puchon.com>)'이 별도 사이트로 중복분류되어 분류작업의 오류로 확인되었다. 이를 표로

<표 1> 디렉터리 서비스 <뉴스, 미디어>류 등록 및 중복분류 정보자원

(2005. 1월 수집 데이터 기준)

구분	네이버	야후 코리아	엠파스
등록 정보자원수	3,532	2,376	3,776
중복분류 자원수	0(2*)	0	55

* 네이버의 중복분류는 동일자원임을 파악하지 못한 오류로 발생

정리하면 <표 1>과 같다.

엠파스가 중복분류 처리한 55건의 정보자원을 분석하면, 신문 사회면 관련 자원이 8건, 서울신문 각 섹션 관련 항목이 7건, 연합뉴스 관련 항목이 8건으로 이들 분류항목에서 참조 링크가 제공되어 있다면 분류자의 부담을 크게 줄여줄 수 있음을 알 수 있다.

관점을 달리하여 참조 링크를 포함한 중복분류 정도를 살펴보면, 야후의 참조 링크 활용도가 높은 것으로 파악된다. 배희진 등이 수행한 연구에서 수집한 데이터를 인용하면 세 디렉터리 서비스 <뉴스 미디어>류에 등록된 자원수는 참조 링크 제외 시 네이버 1,970건, 야후 2,070건, 엠파스 3,440건이다(배희진, 이준호,

〈표 2〉 디렉터리 서비스 〈뉴스, 미디어〉류 참조 링크 포함 여부에 따른 중복분류 정보자원

(2004. 수집 데이터 기준)

구 분	네이버	야후 코리아	엠포스
참조 링크 제외 정보자원수	1,970	2,070	3,440
참조 링크 포함 정보자원수	39,341	105,424	45,390
확장배수	20	51	13

박소연 2004). 그런데 참조 링크를 포함하여 등록된 정보자원수는 네이버 3만9,341건, 야후 10만5,424건, 엠포스 4만5,390건으로 야후의 참조 링크 제공이 네이버와 엠포스에 비해 포괄적인 것으로 파악된다. 참조 링크의 포함에 따라 정보자원이 확장되는 정도는 네이버 20배, 야후 51배, 엠포스 13배이다. 이를 표로 정리하면 〈표 2〉와 같다.

위 분석을 통해 네이버와 야후는 동일류항목 내에서는 특정 정보자원을 중복분류하지 않는다는 것을 알 수 있다. 엠포스는 동일류항목 내에서도 참조 링크가 제공되지 않는 항목에 대해서는 중복분류하고 있음을 알 수 있다. 또한 참조 링크의 제공 측면에서 야후가 가장 포괄적으로 참조 링크를 활용하고 있음을 알 수 있다.

일부 정보자원의 경우에는 참조 링크와 분류항목의 중복성과 무관하게 중복작업이 필요한 경우도 있다. 예를 들어, 수학교사의 개인 홈페이지에 심도깊은 수학 학습정보와 개인 정보 및 매뉴가 동시에 제공되어 있다면, 이 사이트는 개인 홈페이지 항목과 수학교사 관련 항목에 중복분류 할 수 있다. 이 경우 분류체계에서 주제간의 관련성으로 참조 링크를 전개할

수 없으므로 중복분류 작업이 필요한 경우이다. 특정주제 분야에 대해 심도깊고 비중있는 정보를 담은 개인 홈페이지가 이러한 정보자원의 대표적인 예라고 할 수 있다.

3.4 유사 정보자원의 분산 분석

유사문헌의 분산문제는 디렉터리 서비스의 단점으로 지적되는 것으로 분류체계 자체의 비논리성으로 인해 발생할 가능성이 있는 것으로 지적된 바 있다. 세 디렉터리 서비스 검토결과 유사문헌의 분산은 분류체계 전개상의 비논리성에 더하여 분류자의 작업오류로 발생하는 경우도 상당수 발견할 수 있다. 분류체계의 문제로 기인될 수 있는 경우는 유사한 개념의 분류체계가 중복으로 존재하는 상황이다. 즉, 분류항목 간 상호배타성 원칙에 위배된 경우이다. 유사한 분류항목이 두 곳에 존재한다면 당연히 유사 정보자원의 분류작업 또한 그 두 곳으로 각각 분산될 수밖에 없는 것이다. 분류작업의 문제로 기인될 수 있는 경우는 분류항목 자체는 차별화된 항목 범위와 분류원칙을 가지고 있으나, 항목의 범위와 분류원칙을 분류자가

충분히 숙지하지 못하여 원칙에 어긋난 분류작업을 하는 경우, 또한 분류대상 정보자원의 성격을 충분히 이해하지 못한 상태에서 잘못 분류하는 경우이다. 이 경우 특정 정보자원은 분류되어야 할 분류항목이 아닌 다른 분류항목으로 분류되고 그에 따라 유사 정보자원이 원래 분류되어야 할 분류항목과 분류되면 안되는 분류항목들로 분산되는 경우이다. 이러한 경우는 명백한 오분류라고 할 수 있다.

유사 정보자원의 분산현상이 가장 두드러진 분야가 기업과 쇼핑 관련 항목이다. 기업과 쇼핑 하위의 상호 관련된 분류항목들에 등록된 정보자원을 검토하면 기업과 쇼핑 하위의 분류항목에 등록된 정보자원간의 차별화가 안되며 유사한 정보자원이 기업과 쇼핑 하위의 유관항목에 광범위하게 분산된 것을 발견할 수 있다. 이러한 현상은 우선 세 디렉터리 서비스 모두에서 경제/기업/쇼핑 관련 정보자원이 차지하는 비중이 네이버 52.8%, 야후 24.2%, 엠파스 34.6%로 야후를 제외하고는 가장 크고, 다수를 차지하는 중소기업의 특성상 기업을 소개하는 자사 홈페이지를 통하여 온라인 판매기능을 포함하고 있기 때문이다. 원칙적으로 기업의 소개 정보를 담고 있는 정보자원은 기업의 하위 항목에, 온라인 판매기능을 담고 있는 정보자원은 쇼핑 하위 항목에 분류하면 되는 간단한 문제이나 실제 정보자원의 구성이 양자를 겸하고 있어 분류작업 과정에서 명확한 구분이 쉽지 않다.

네이버는 2003년 기업과 쇼핑이 각각 <비즈니스, 경제>와 <쇼핑>의 별도의 류에 속해 있던

이유로 이러한 문제들이 광범위하게 발견되었다. 최근 네이버가 <기업, 쇼핑>류를 통합 구성한 것은 광범위하게 발견되는 문제를 해결하기 위하여 차별화하기 어려운 기업과 쇼핑을 묶어 유사 관련 정보자원들의 분산 현상을 최소화하기 위한 고려가 있다.

한 예로 가구 관련 항목을 살펴보면, 야후의 가구 관련 항목인 <비즈니스와 경제>류의 '쇼핑과 서비스(B2C, C2C) > 주거와 생활 > 가구' 항목 아래에는 '메이커'와 '소매' 항목이 전개되어 있다. 제조사와 소매상은 이론적으로 구분이 명확할 수 있으나 실제 두 항목에 분류된 정보자원을 살펴보면, 광범위한 유사 정보자원의 분산현상을 볼 수 있다. 또한 '비즈니스와 경제 > 기업 간 거래(B2B) > 가구 > 주거용 가구 > 제조' 항목에 분류된 정보자원들 역시 앞의 두 분류항목과 유사한 정보자원들이 상당수 포함되어 있다. 인터넷 쇼핑몰의 활성화로 B2C와 B2B의 구분이 무너지고 있는 현실이고 상당수 B2B 업체가 인터넷을 통해 B2C 기능을 겸하고 있으므로 야후 <비즈니스와 경제>류의 '기업 간 거래(B2B)'와 '쇼핑과 서비스(B2C, C2C)' 구분은 유사 정보자원의 분산현상을 심화시키는 결과를 빚고 있다. 엠파스의 아동용 가구 관련 항목으로 <기업, 쇼핑>류 아래 '가구 > 아동용 가구' 항목과 '인터넷 쇼핑몰 > 가구 > 아동, 유아용' 항목이 있다. 이론상 전자는 아동용 가구의 제조사 관련 정보자원을 후자는 온라인 판매 기능을 제공하는 정보자원을 분류하면 될 것이다. 실제

등록된 자원들을 비교하면 두 항목 간에 유사 항목이 분산되어 있음을 발견할 수 있다. 이러한 현상은 비단 '가구' 뿐만이 아니라 전업종에 걸쳐 공통되게 발견된다.

기업, 쇼핑 관련 항목 외에도 상당수의 분류 항목에서 유사한 정보자원들이 분산되어 분류되어 있는 것을 발견할 수 있다. 이러한 분산이 발생한 대표적인 항목으로 대학에 설치된 연구실, 장애인 신문, 생활협동조합 등의 분류항목이 있으며 이외에도 기계공구, 섬유, 공인중개사, 스키, 떡/한과, 화가, 열대어/관상어, 미술품/갤러리, 외식업 분야의 치킨, 피자 등에 이르기까지 다양한 분야의 분류항목에서 유사 정보자원의 분산을 확인할 수 있다.

위 예시들을 통해 알 수 있듯이 이론상으로는 명확한 구분이 가능하나 실제 정보자원의 특성이 이론에 기반한 분류항목의 적용을 어렵게 하는 경우가 상당수 발생한다. 이러한 현상은 단순히 분류작업의 오류라기보다는 분류체계와 항목의 범위가 확립되지 않았거나, 분류자 간에 확실한 원칙이 공유되지 않았기 때문이다.

이러한 문제를 해결하기 위해서는 네이버의 <기업, 쇼핑>류의 경우와 같이 구분이 어려운 항목을 통합하거나 각 분류항목별로 분류되어야 할 정보자원의 정의 및 범위에 대해 확실히 규정하여 분류항목별 차별성을 확보하여야 한다. 그 지침은 분류자에게 충분히 숙지케 하여 분류작업의 오류를 최대한 방지하여야 한다. 또한 비용문제로 인해 실현 가능성은 의문시되

나, 이미 분류된 작업에 대해 재검증하는 품질 관리과정을 일상업무 중 한 과정으로 편입시켜 결과를 서로 검토하게 하는 방법도 고려할 수 있다.

위와 같은 분류항목의 포함범위 명확화와 작업절차 확립의 중요성을 엿볼 수 있는 항목이 있다. 네이버 <뉴스, 미디어>류의 '신문 > 분야별 신문' 아래에는 '전기, 전자', '정보통신, IT', '컴퓨터' 항목이 전개되어 있다. 우선 정보통신 관련 정보자원을 분류해야 하는 분류자의 관점에서 세 항목을 구분하기도 쉽지 않다. 실제 세 항목에 분류되어 등록된 자원을 살펴보면, '전자신문'은 모든 항목에 등록되어 있고 유사 정보자원이 분산되어 있음을 볼 수 있으며, 심지어 정보통신 관련 정보자원이 '분야별신문 > 과학' 아래 분류되어 있는 경우도 있다. 엠파스 <뉴스, 미디어>류의 '신문 > 전문지' 항목 아래에도 '전기, 전자', '정보통신', '컴퓨터' 의 세 정보통신 관련 항목이 전개되어 있다. 그러나 실제 세 항목에 분류되어 등록된 자원을 살펴보면, 상당히 차별화된 정보자원이 등록되어 있음을 볼 수 있다. '전기, 전자' 항목에는 전기, 전력관련 정보자원이, '정보통신' 항목에는 국내의 대표적인 정보통신 전문지들이, '컴퓨터' 항목에는 국내 주요 일간지의 IT, 과학 뉴스' 관련 정보자원이 등록되어 있다. 물론, 이용자들이 항목 간의 차이를 표목어를 통해 직관적으로 이해하고 접근하기를 기대하기는 어려운 분류항목의 전개이나 항목 간의 차별화를 위해 나름대로의 지침을 가

지고 애쓴 흔적을 볼 수 있으며, 내부 처리지침 공유의 중요성을 볼 수 있는 항목이다.

3.5 정보자원 표제 및 설명문 기술 평가

분류작업과 관련하여 정보자원의 표제와 설명문 기술문제를 다루고자 한다. 기본적으로 통제어를 사용하는 분류항목 표목어는 전거 데이터를 구축하고 활용해야 할 것이며, 자연어를 사용하는 정보자원의 표제와 설명문의 경우에는 등록의 대상이 되는 해당 사이트의 원 정보를 최대한 활용하여 작성되어야 한다.

디렉터리 서비스가 분류하여 제공하는 정보자원의 표제 기술과 관련하여 가장 근본적인 어려움은 정보자원의 불균질성에 기인한다. 과거에는 정보의 일방적 수용자였던 다수의 비전문가에 의해 생산되는 정보자원이 대부분을 차지하는 현실에서 정보자원으로부터 적합한 표제를 채기하는 일은 쉽지 않다.

또한 현재 디렉터리 서비스는 네이버, 야후, 엠파스 등 서비스 제공사에서 정보자원을 수집하여 등록하는 경우도 있으나 정보자원의 권리가 정보자원의 등록을 신청할 수도 있다. 후자의 경우 정보자원의 권리가 정보자원의 제목, 설명문, 그리고 분류되어 등록될 분류항목을 신청하게 되어 있다. 디렉터리 서비스 제공사에서 검토하여 최종적으로 등록내용을 결정하지만, 비전문가인 정보자원의 권리자도 디렉터리 서비스 등록과정에 많은 영향을 미치게 되고 표제, 설명문에도 영향을 미치게 된다.

현재 야후가 표제선정의 지침으로 제시하고 있는 것은 회사인 경우 사업자 등록증에 나타나는 공식 회사명, 온라인 업체(쇼핑몰)인 경우 홈페이지에 나타나 있는 공식 서비스 명, 강좌를 제공하는 주제사이트인 경우 홈페이지 제목 및 서비스 명, 개인 홈페이지인 경우 사이트 운영자 이름 등이다

정보자원은 자체에 정보자원을 가장 잘 표현해주는 표제를 발견할 수 있다. 메타 데이터가 있는 경우 이를 참고하거나, HTML 헤더의 타이틀 부분을 채기하는 방법도 검토할 수 있다. 디렉터리 서비스에 등록된 정보자원의 표제는 대부분 이러한 과정을 거쳐 적합한 표제로 표현되고 있으나, 일부의 경우에는 정보자원의 내용을 제대로 반영하지 못하거나 원 정보자원에서 찾을 수 있는 정보를 제대로 담고 있지 못한 경우가 있다. 앞에서 예시한 '이너모스트', '부천닷컴'이나 야후에 등록된 '녹두신문'이 좋은 예이다.

분류작업과 관련된 또 하나의 문제는 정보자원 설명문이다. 동일 분류항목에 등록된 유사 정보자원의 설명문을 살펴보면, 설명문 기술에 있어 일관성을 찾아보기 힘들다. 야후 <비즈니스와 경제>류의 '쇼핑과 서비스(B2C, C2C) > 스포츠 > 볼링 > 소매' 아래 등록된 17개 정보자원의 설명문을 살펴보면, 기술형식이나 용어선정에 있어 아무런 기준이 없음을 알 수 있다. 용품을 나타내는 용어를 살펴보면, '볼', '볼링 공', '볼링 볼', '신발', '볼링 신발', '볼링 슈즈', '볼링 화', '볼링 복', '의

류', '아대', '손목보호대', '장갑', '글러브' 등 동일 품목에 대해 다양한 표현을 일관성 없이 사용하고 있다. 또한 판매기능 관련으로도 '온라인 볼링 용품 쇼핑몰', '볼링 용품 판매', '볼링 용품 온라인 판매' 등 다양한 형태로 일관성 없이 표현하고 있다. 또한 중복 분류된 동일한 정보자원의 설명문을 비교해 봐도 같은 제공사에서 작성된 것이라고 생각하기 어려울 정도로 차이를 볼 수 있다. 네이버에서 동일 분류항목에 중복등록한 '이너모스트', '부천일보'의 설명문을 비교해 들어, 다음과 같이 공통점이나 일관성을 찾기 어렵다.

• 예시 1) 이너모스트

(<http://www.innermost.org>)

이너모스트 설명문: 정치문화 포털 웹진, 매체 비평, 정치, 경제, 국제문제 토론방, 핵심 인터뷰, 사진모음

이너모스트 설명문: 새천년민주당 지지 웹진, 정책 칼럼, 국제 문제, 매체 비평 제공

• 예시 2) 부천일보

(<http://www.puchon.co.kr>)

부천일보 설명문: 부천시 정치, 경제, 사회, 문화 등 관련 뉴스 제공, 부천시내 업소 검색 서비스

부천닷컴 설명문: 종합 일간 인터넷 신문 및 포털 사이트, 지역신문 부천 21과 함께 기사 수록, 주제토론실, 생활정보, 시민 리포트, 구독 안내

위 두 항목은 디렉터리 서비스가 동일 정보 자원임을 파악하지 못하고 중복등록한 오분류

의 예이나 그렇다고 하더라도 동일한 제공사에서 작성된 설명문이라 하기에는 너무 상이하다. 위 네이버의 예에서 볼 수 있듯이 동일자원에 대해 작성된 설명문인 경우라도 일관성을 유지하는 것이 쉬운 일은 아니다. 설명문 작성 시에는 해당 분류항목 내 등록 정보자원의 설명문을 참고하여 기술에 일관성과 통일성을 기해야 한다. 동일 분류항목 내에서는 용어 및 기술순서의 일관성을 맞추려면 완성도 높은 디렉터리 서비스가 될 것이다.

설명문 작성에 관한 자세한 기술(記述)지침이 준비되어 분류자들 사이에 공유되어야 한다. 디렉터리 서비스의 설명문을 비교 검토한 결과 설명문 작성시 갖추어야 할 기술요소들로 아래와 같은 항목을 고려할 수 있으며, 기술하는 순서에 대해서도 일관되게 적용해야 한다.

- 내용: 정보자원이 다루고 있는 내용
- 정체성: 정보자원의 정체성
- 기능: 기업소개외의 경우, 업체의 사업영역(수입, 수출, 생산, 제조, 개발, 판매, 임대 등)
- 변별성: 정보자원을 다른 정보자원과 구별할 수 있는 특징
- 소재지: 정보자원 대상 조직의 지리적 소재지
 설명문은 정보자원의 표제나 분류항목과 관계없이 그 정보자원에 대해 파악할 수 있도록 자체 완결성을 갖도록 작성되어야 한다. 또한 이용자의 입장에서 과연 어떤 정보를 설명문에서 전달받는 것이 가장 도움이 될 것인가를 고려하여 작성되어야 한다.

3.6 기타의 문제점

디렉터리 서비스 분류항목에 분류된 정보자원을 살펴보면, 수많은 분류작업상의 문제점과 오류를 발견할 수 있다. 이런 분류작업상의 오류는 몇 가지 유형으로 나누어 볼 수 있다.

위의 생협외 경우처럼 정보자원의 내용을 정확히 파악하지 못하였거나 파악된 내용을 반영하지 못하여 발생하는 분류 작업상의 오류이다. 분류작업은 정보자원의 내용을 정확히 파악하는 것에서 시작하여 정보자원의 내용이 등록할 가치가 있다면 해당되는 분류항목을 찾아 먼저 등록된 정보자원의 내용을 확인하고 등록 작업을 해야 한다. 네이버 <뉴스, 미디어>류의 '잡지, 웹진 > 정치' 항목에는 정보자원인 '이너모스트(<http://innermost.org>)'가 '이너모스트'로 두 번 등록되어 있다. 동일한 정보자원이 등록이 되었음에도 불구하고 '이너모스트'가 '이너모스트'로 표제부터 잘못되어 등록된 것은 명백한 실수로 파악되며 정보자원의 내용을 제대로 파악하고 해당 항목에 기 등록된 정보자원을 검토하지 않아 발생한 오류이다. 네이버 '신문 > 지역신문 > 경기도' 항목에 등록된 '부천일보'는 URL만을 달리하는 동일 정보자원이 동일 분류항목 내에 '부천닷컴(<http://www.puchon.com>)'으로도 등록되어 있다.

해당 항목에 적합한 정보자원이 존재함에도 불구하고 부적합한 관련 정보자원을 등록한 경우이다. 야후 <인문, 사회과학>류의 '문헌정보

학 > 문헌정보학과' 항목은 대학교 문헌정보학과의 홈페이지 정보자원을 등록하고 있다. 이 가운데 연세대학교 문헌정보학과로 등록된 정보자원은 학과의 홈페이지가 아니고 졸업동문들이 운용중인 동문회 홈페이지이다. 야후의 통합검색 기능을 사용하여 검색하면 검색결과에 연세대학교 문헌정보학과 공식 홈페이지가 검색됨에도 불구하고 정작 해당항목에는 동문회 관련 정보자원이 등록되어 있는 것이다. 등록된 정보자원의 설명문에는 동문회 관련 정보자원임을 기술하고 있어 정보자원에 대한 내용 파악은 완결된 것으로 판단된다. 만약 동문회 관련 자원의 등록을 검토하였더라도 항목에 더 적합한 정보자원이 있다면 그것을 찾아 등록해야 하며 관련 항목으로 등록하기로 결정하였다면 부가적인 정보자원으로 등록해야 한다.

정보자원을 분류할 수 있는 세부항목이 있음에도 상위 항목에 분류한 경우이다. 특정 정보자원을 분류할 때 정보자원을 분류할 수 있는 세부항목이 있음에도 불구하고 상위 분류항목에 분류한 경우를 상당수 발견할 수 있다. 이러한 경우는 세부항목이 설치되어 있는 모든 분류항목에서 공히 발견된다. 엠파스 <기업, 쇼핑>류의 '스포츠 > 스키' 항목의 하위 항목으로 '대여, 렌탈' 항목이 전개되어 있으나 상위 '스키' 항목의 정보자원 가운데 대여, 렌탈 관련 업체의 정보자원이 포함되어 있다. 또한 하위의 '대여, 렌탈' 항목은 스키장별로 하위 항목을 구성하고 있으나 상위항목에 분류된 정보자원 가운데 특정 스키장 인근에 소재하여

하위 세부항목에 등록하는 것이 적합한 정보자원이 다수 포함되어 있다. 분류대상의 주제를 가장 특정하게 포함해주는 분류항목이 있을 때에는 가능한 가장 세분된 분류항목에 소속시켜야 하며, 상위의 분류항목으로 분류해서는 안 된다.

이와는 반대로 상위 항목에 분류되어야 할 정보자원이 하위 세부항목에 분류된 경우도 있다. 분류대상 정보자원의 주제가 여러 개념이나 요소들을 포괄하고 있을 때, 한 개념이나 요소를 위해 전개된 하위분류항목으로 분류하는 사례들이 다수 발견된다. 위의 경우와 마찬가지로 세부항목이 전개된 모든 분류항목에서 공히 발견된다. 엠파스 <기업, 쇼핑몰>류의 '인터넷 쇼핑몰 > 의류, 신발 > 해외명품 > 폴로' 항목에 분류된 16개 정보자원 가운데 폴로 전문업체가 아닌 다수 브랜드를 취급하는 업체가 있다. 이런 정보자원들은 상위 항목으로 분류되어야 한다.

원인을 짐작할 수 없는 명백한 오분류 작업도 흔히 눈에 띈다. 대표적인 예로 네이버는 '지저스 인터넷 TV(<http://www.jesus.or.kr/>) 기독교 전문인터넷 방송, 설교방송, 강좌, 선교21C, 가스펠 음악, 프로그램 안내 등 수록'을 '뉴스, 미디어 > 인터넷방송 > 경제, 재테크' 아래에 분류했으며 '한국여성상담센터(<http://www.iffeminist.or.kr/>) 가정폭력, 성폭력, 가족문제, 청소년, 성문제 등 광범위한 여성문제 상담 프로그램을 지원'을 '라이프스타일 > 여성 > 기관, 단체 > 여성예술단

체'로 분류했다. 엠파스의 '건강, 의학 > 건강 > 금연 > 협회, 단체' 아래에는 주제와 상관 없는 '대한민속그네협회중앙회(www.minsokgune.co.kr) 민속그네, 단오축제 소개, 그네뛰기, 제기차기, 널뛰기, 율놀이 등 보급 및 대여, 행사실적 수록', '한국레크리에이션 프로그램지도센터(www.korrec.com) 기업체, 일반 행사와 레크리에이션 연구, 연수, 체육대회, 송년행사, 야유회, 이벤트 등 안내', '한국여가레크리에이션학회(web.sportskorea.net/kslr) 한국여가레크리에이션학회 학술자료, 논문검색, 관련서적 정보 제공' 등이 소속되어 있다.

이외에도 주제전문가가 아닌 분류자에 의해 분류작업이 이루어지는 현실에 따라 정보자원이 다루고 있는 주제에 대한 이해부족으로 발생하는 분류작업의 오류 등 수많은 작업상의 오류를 찾을 수 있었다. 정보자원을 정확히 파악하고 최적의 해당 분류항목을 확인하여 일관성 있는 분류작업이 이루어져야 하며, 지속적인 수정과 검토를 통해 분류작업의 완성도를 높여야 한다. 또한, 분류자는 해당 주제분야에 대한 이해도를 증진시키는 노력도 병행해야 할 것이다.

4. 디렉터리 서비스 분류작업 개선방안

앞 장에서는 디렉터리 서비스의 분류체계를 활용하여 인터넷 정보자원을 디렉터리 서비스에 등록하는 분류작업 과정 및 결과의 문제점

들을 제시하고 평가해 보았다. 이 장에서는 이를 기반으로 하여 디렉터리 서비스의 개선안에 대해 제시해 보도록 한다.

4.1 분류대상 선정 및 등록 정보자원 관리

디렉터리 서비스는 양적으로 방대하고 지수 함수적으로 증가하는 인터넷 정보원을 대상으로 한다. 효율적인 디렉터리 서비스를 제공하기 위해서는 평가과정을 거쳐 방대한 자원 중에 정보자원으로서의 가치있는 자원만을 선별하여 제공하는 것이 필수적이다. 특히 인터넷 상의 정보는 전술한 바와 같이 비구조적이며 동적인 동시에 정보자원의 질적 가치에 있어서 큰 차이가 나므로 등록 전에 엄격한 선정작업을 거쳐 가치있는 정보를 제공해야 디렉터리 서비스의 근본적인 효용성을 확보할 수 있다.

이러한 선정작업을 위해서는 디렉터리 서비스 정보자원으로 등록 여부를 결정하기 위한 평가와 선정기준이 필요하다. 인터넷 정보자원의 선정기준에 대해서는 상당수의 연구가 있었으며 제시된 다양한 평가기준에 대해서는 전술한 바 있다.

이 연구에서 고찰한 세 디렉터리 서비스도 나름의 등록대상 정보자원 선정기준을 적용하고 있을 것으로 판단되나, 선정기준의 확립 및 공유는 관리자의 입장에서 매우 중요한 문제이므로 검토와 일관된 적용노력이 기울여져야 한다.

분류대상의 선정에 있어 기본적으로는 독립

된 사이트를 대상으로 함을 원칙으로 할 것이 권장되며, 여기에 포함된 하위 웹 페이지를 등록하고자 하는 경우는 최소한 웹사이트에서 제공하는 메뉴 중에 포함된 내용으로 제한하는 것이 권장된다. 현재의 등록관행도 웹 사이트를 기본대상으로 하고 있는 것으로 파악되나 적지 않은 수의 부적절한 웹 페이지를 등록해 놓은 것을 볼 수 있다. 만약, 특정 웹 페이지를 분류의 대상으로 채택하는 경우라면 상응하는 이유가 있어야 하며, 합리적인 기준에 맞추어 일관되게 처리해야 한다. 특히, 개인 홈페이지의 등록문제는 위의 평가기준에 맞추어 검토되어야 할 필요가 있는 대표적인 정보자원이다. 등록된 상당수의 개인 홈페이지가 정보량 및 가치에 있어 등록 필요성이 의심스러운 것으로 판단된다.

일단 등록된 사이트도 그 유효성을 관리하는 절차가 있어야 한다. 연구자가 2003년 네이버에 등록된 정보자원 200개를 검토한 결과, 폐쇄되었거나 접속이 불가능하여 유효하지 않은 사이트가 12개에 달했음은 전술한 바와 같다. 또한 접속은 가능하더라도 내용이 완전히 변경되어 등록내용이 부적합 하거나 등록된 사이트는 운용이 중단되어 웹 호스팅 업체의 사이트로 접속되는 경우도 온라인 쇼핑 관련 항목에서 다수 발견할 수 있었다. 자동화된 점검절차를 마련하여 사이트의 유효성과 내용수정에 대한 지속적인 관리절차의 운용이 필요하다.

등록 및 검색 사이트의 유효성확인 문제는 웹 정보자원의 특징을 논할 때 반드시 언급되

는 휘발성 또는 가변성과 관련된 것으로 이의 해결을 위해 OCLC는 PURL(Persistent URL)을 구축하고 있고 또 다른 방안으로 URI(Uniform Resource Identifier)의 도입이 연구되고 있기도 하다. 세 디렉터리 서비스의 경우 등록된 사이트가 접속이 불안정하거나 정상적인 접속이 되지 않는 경우 사이트 명 옆에 별도의 기호를 덧붙여 표시하고 있다. 이를 통해 어느 정도는 이용자에게 편리함을 제공하고 있으나, 지속적인 점검과 더불어 유효성을 감시할 수 있는 방안이 수립되어야 한다. 예를 들어, 도메인 등록 사이트 등록현황의 참조 등 외부기관의 자원을 활용하는 방법도 고려할 수 있으며, 여러 방법을 통해 폐쇄된 것으로 확인된 사이트의 경우라면 등록된 레코드를 삭제해야 한다.

4.2 분류체계의 정비

국내 주요 포털에서 제공하는 디렉터리 서비스의 분류체계의 여러 가지 문제점에 대해서는 앞 절에서 살펴본 바와 같다. 키워드 검색에 비해 정확한 검색결과를 제공하는 디렉터리 서비스를 개선하여 더욱 많은 이용을 유도하며 서비스의 효율성을 높이기 위해서는 기반이 되는 분류체계의 정비가 필수적이라 할 수 있다.

디렉터리 서비스의 분류체계 정비와 관련하여 지속적인 개선작업이 필요한 항목으로 다음과 같은 것들이 있다.

- 비논리적이고 일관성을 상실한 분류항목의

인용순

- 주제의 특성을 무시한 자모순 배열
- 모호한 항목 간의 계층성
- 항목 간 배타성을 상실한 중복 분류항목
- 망라성과 포괄성을 상실한 분류항목의 전개
- 정보자원의 문헌적 근거와 무관하게 설치되는 분류항목 간의 불균형
- 일관성을 상실한 분류항목 전개
- 부적절하고 일관성을 잃은 표목어 선정
- 분류항목의 성격 및 범위의 불명확성

모든 분류체계는 완벽하지 않다. 다양한 사고체계를 가지는 분류자와 이용자가 이용하며, 아무리 신중한 검토로 구축된 분류체계라 하더라도 주제의 분화나 학제적(interdisciplinary) 주제의 출현을 모두 포괄할 수 없기 때문이다. 그렇다고 완벽할 수 없다는 것이 보편적이고 상식적인 논리에 어긋나는 분류체계 구성에 대한 면죄부가 될 수는 없다. 완벽한 분류체계를 구축하겠다는 의지로 최선의 분류체계를 구축하기 위한 지속적인 노력을 기울여야 한다.

디렉터리 서비스는 상이한 연령, 계층, 교육 수준의 이용자들이 다양한 정보요구를 가지고 이용하는 서비스이다. 이를 사용하는 이용자의 사고체계 또한 통일되어 있거나 획일적이지 않다. 따라서 가장 보편적이고 상식적인 논리체계를 기반으로 분류체계가 구성될 수 있도록 많은 노력을 경주해야 한다. 또한, 신중한 검토를 거쳐 채택된 전개방식이라면 관련 항목에서도 일관성 있게 적용되어야 한다. 한편으로 인터넷 정보자원의 특성에 따라 새로이 등장하는

정보자원을 적절히 포용할 수 있는 유연성과 융통성도 요구된다.

디렉터리 분류체계에서 발견되는 또 하나의 문제는 분류체계의 깊이와 세밀도의 편차이다. 분야에 따라서 분류체계의 상세성에 차이가 있다. 어떤 분류항목은 너무나 세부되어 있어 예를 들면, 네이버 <엔터테인먼트>류의 '방송프로그램 > TV프로그램 > 볼세 > 이서진 장세훈역@' 항목과 같이 특정 TV 드라마의 특정 배역이나 특정 게임의 리뷰/프리뷰까지도 개별 분류항목이 될 정도이다. 이와는 반대로 네이버 '한방병원', '치과병원' 처럼 너무 포괄적인 단계를 끝으로 더 이상 세부된 항목이 전개되지 않는다. 대중문화의 비중이 확대되고 다른 분야에 비해 많은 관심과 비중을 차지하는 현상 등 인터넷이라는 매체의 사용자들이 지니는 속성과 욕구를 반영한 것일 수 있다.

네이버 '치과병원' 항목은 2005년 1월 현재 590여 건에 이르는 정보자원이 등록되어 있어 디렉터리 서비스로 접근한 이용자라고 하더라도 특정 치과병원을 찾기 위해서는 다시 키워드 검색을 하던지 아니면 전체 자료를 다 확인하기 위해 50개 이상의 페이지를 열람해야 한다. 네이버 '한의원, 한방병원' 항목도 이와 마찬가지로 500개 이상의 정보자원이 등록되어 있다. 한의학이라는 분야가 서양의학처럼 보편적으로 통용되는 분류체계를 지니고 있지 못하다는 것을 고려하더라도 이 분류항목을 유용하기 만들기 위해서는 어떤 형태로든 하위 분류항목의 전개가 필요하다. 실제로 엠파스는 한

의학의 분류체계에 기반하여 분류항목을 전개하고 있으며 네이버에 등록된 많은 한방병원관련 정보자원들은 사이트 설명문에서 전문분야에 대해 언급하고 있다. 따라서 그러한 정보들과 문헌분류체계를 토대로 한방병원의 세부분류체계를 고안할 필요가 있다.

4.3 분류체계 일람도구

디렉터리 서비스는 키워드 검색기능을 제공하므로 분류표 본표에 해당하는 별도의 카테고리 목록을 제공하지 않더라도 통합검색이나 키워드 검색 등 검색기능을 통해 원하는 카테고리 및 특정 카테고리의 디렉터리 분류체계 내의 위치를 확인할 수 있다. 하지만 전개된 모든 동위 항목을 찾고자 하는 경우나 특정 동위항목으로의 이동을 원하는 경우 상위항목으로 이동하여 확인하고 다시 동위항목으로 이동해야만 하는 등 이용상 불편함이 따른다.

이런 문제점을 해결하여 이용자의 편리성을 향상시키기 위해서는 디렉터리 분류체계의 전체적 구조, 각 대분류별 계층구조와 상하위 카테고리 전개, 관련 카테고리 경로 등을 입체적으로 편리하고 쉽게 검색해볼 수 있는 카테고리 탐색기의 제공이 필요하다. 최근 네이버의 디렉터리 서비스에 도입된 '카테고리 탐색기'는 분류표에 해당하는 장치로서 분류체계의 구조에 따라 상하위 항목과 동위 항목으로 쉽게 이동할 수 있는 방법을 제공한다. 다른 디렉터리 서비스에서도 적극적인 검토와 도입이 필요

하다.

이런 카테고리 탐색기는 이용자들에게 현재 자신이 이용하고 있는 분류항목이 디렉터리 서비스 전체 분류체계에서 어느 위치에 있는지를 파악할 수 있게 해준다. 또한 검토 결과에 따라 자신의 정보요구에 맞게 상하위, 동위 항목 간의 이동을 용이하게 해주는 효용이 있다. 분류자의 관점에서도 카테고리 탐색기는 전체적인 분류체계를 일목요연하게 일별할 수 있게 해주어, 주제간의 관계를 파악한 상태에서 최적의 분류항목을 선정하는 과정에 도움을 줄 수 있다.

디렉터리 일람도구는 이용자들의 이용 편의성을 증진시킬뿐 아니라 분류자의 입장에서조차 전체 분류체계를 쉽게 검토할 수 있는 도구를 제공함으로써 분류작업의 효율을 향상시켜 작업을 돕는 도구로서 기능한다.

4.4 내부작업 도구의 정비

디렉터리 서비스는 이용자들의 이용을 목적으로 한다. 디렉터리 서비스의 개선을 위해서는 이용자들의 효과적인 이용만큼이나 구축된 분류체계에 따라 분류자가 효율적이고 일관성있게 분류작업을 수행할 수 있도록 각종의 작업 보조도구들을 마련하고 정비할 필요가 있다.

부적절한 분류작업이나 유사항목 간 차별성 없는 분류작업으로 인해 비슷한 사이트들이 유사 분류항목들에 혼재되는 현상이 발생하게 된 원인 중 하나는 잘 정비된 카테고리 해설집이

없기 때문으로 추정된다. 각 분류항목별로 범위주기(Scope Note)를 활용하여 이 카테고리에는 어떤 문서들이 분류되어야 하는지에 대해 원칙을 자세히 기술함으로써 유사 카테고리 간 분류작업의 혼동을 막고, 유사 사이트들이 분산되는 것을 방지해야 한다. 각 카테고리 항목별로 해당 카테고리에 분류 등록하는 정보자원의 범위를 규정하는 해설집을 구비한다면 관련 정보원이 유사 카테고리로 나뉘어 등록되는 경우를 방지할 수 있을 것이며 분류자가 바뀐다 하더라도 일관성 있게 분류 및 등록작업이 진행될 수 있을 것이다.

또한 내부용 작업 매뉴얼을 작성하고 해당 분야의 분류자가 숙지할 필요가 있다. 분류작업 중에 어려움을 겪었던 사례에 대해 검토내용과 결정한 분류방안을 반드시 내부 작업 매뉴얼에 기록으로 남겨 분류자 간 등록의 일관성이 유지될 수 있도록 해야 한다. 매뉴얼에 관련한 제언을 추가하자면 일람이 용이한 형태의 통일된 서식을 마련하고, 그 서식대로 모든 분류자가 담당 분야에 대해 매뉴얼을 작성함으로써 공유도(interoperability), 정보습득의 신속성, 가독성, 열독성을 높이고, 분류자변경 시 후임 분류자가 작업요령을 신속히 습득하여 실무에 적용시키기 쉬워야 한다. 이런 작업 매뉴얼은 분류자변경 시 후임자가 어려움이나 실수 없이 작업할 수 있도록, 심지어 구두 오리엔테이션이 필요없을 정도의 수준으로 정밀하고 자세하게 기술되어야 한다.

카테고리 해설집, 작업 매뉴얼 등에 더하여

한 가지 분류작업의 보조도구로 구비되어야 하는 것이 용어사전이다. 그것이 주제명 표목표가 되었던 시소러스가 되었던 디렉터리 서비스가 표목어로 선정된 통제어휘와 동의어, 관련어 등 자연어를 연결해주는 용어사전이 필요하다. 이를 통해 일관성 있는 표목어의 제공이 가능하다.

3.6절의 분류작업 검토를 통해 정보자원 표제어 및 설명문 기술에 일관성이 없음을 지적하였다. 이는 기술대상이 되는 정보자원의 어느 부분에서 무엇을 어떻게 기술해야 하는가의 기술에 관련된 지침의 부재로 인해 발생한다. 각 기술항목별 정보원과 기술항목, 기술방법 등에 대해 지침을 마련해야 한다.

4.5 외부자원 활용

이상에서 언급하였던 여러 가지 문제점들은 디렉터리 서비스 제공자들의 자체적인 노력으로 해결될 수도 있을 것이나, 이 과정에서 소요되는 비용의 문제를 고려한다면 외부자원의 활용 방안도 고려할 수 있다.

차별화된 자체 디렉터리 분류체계를 개발, 유지하고 제공하는 것은 이용자를 유인할 수 있는 특화된 서비스가 될 수 있다. 이용자를 유인할 수 있고 한 번 이용해본 이용자가 지속적으로 이용할 수 있도록 하려면, 지속적인 이용의 이유가 이용자가 특정 디렉터리의 분류체계에 익숙해졌기 때문이 아니라 다른 디렉터리 서비스와 비교했을 때 상대적으로 이용이 쉽고

분류체계가 우수하기 때문이어야 한다. 이렇게 자체 디렉터리 서비스의 상대적 우위를 유지하기 위해서는 많은 자본과 인력의 투입이 전제되어야 한다.

야후의 경우 2003년 현재 분류체계의 구성과 실제 정보자원의 분류작업을 위해 20여명의 분류전담 직원이 참여하고 있는 것으로 알려지고 있다. 자체 분류체계 구축을 위한 노력과 투자를 통해 신속한 주제의 추가와 분류의 변경이 가능한 것 또한 사실일 것이나, 한편으로는 폐쇄된 범위의 자체인력만이 참여하는 경우, 주제전개 및 아이템 분류에 있어 편향, 비논리성, 일관성의 부족으로 인해 보편적 논리나 상식과 괴리가 발생할 가능성도 있을 수 있다.

이러한 편향의 문제를 해결할 수 있는 방안으로 적정한 외부자원의 활용이 검토되어야 한다. 구체적인 방안으로 각 디렉터리 서비스의 하위 주제구분 및 전개에 있어서는 표준화된 외부 분류표를 참조하거나 주기적으로 외부 전문가를 동원하여 분류체계를 점검토록 하는 것도 외부자원 활용의 한 방안이 될 수 있다. 또한 표목어의 선정을 위한 선거통제 문제에 있어서도 기 구축된 외부 선거통제 시스템의 활용을 검토할 수 있을 것이다.

또한, 자체 디렉터리 서비스만이 아니라 외부 디렉터리 서비스를 참조하는 노력도 필요하다. 미국 야후의 디렉터리 서비스를 접속하여 들어, 새로 설치된 분류항목이 소개되어 있고, 웹 브라우저의 크기를 조절하면 그에 따라 화면에 표시되는 내용이 재배열되는 것을 볼 수

있다. 세 디렉터리 서비스 중에서 야후 코리아가 새로 등록된 정보자원에 '새내기' 표시를 부가하고 있는 것이나 엠파스가 브라우저 크기 조절에 따라 화면내용이 재배열 되는 것도 같은 기능을 한다. 네이버의 경우 디렉터리 서비스 메인 페이지에 새로 추가된 정보자원을 소개하고 있으나, 각 분류항목 화면이 아닌 첫 화면에서 제공되는 이유로 정보량은 제한적이다. 이런 기능은 참조하여 도입할 수 있는 일례에 불과하다. 관련 서비스를 참고하면 더욱 편리한 서비스를 제공할 수 있다.

4.6 사용자 인터페이스의 개선

디렉터리 서비스는 웹 서비스를 통해 제공되므로 이용자와 서비스 간의 상호작용은 이용자 인터페이스를 통해 이루어진다. 이용자 인터페이스가 디렉터리 서비스의 효율성에 미치는 영향은 지대하다.

4.3절에서 언급한 분류체계 일람도구 역시 이용자 인터페이스의 하나로 간주할 수 있고, 네이버의 '카테고리 탐색기' 구축은 전체 분류체계를 일별할 수 있는 도구를 제공함으로써 이용자뿐만 아니라 분류자에게도 유용한 도구이다. 또한 동일 상위항목 아래 전개된 동위 하부 분류항목 간을 이동하고자 하는 경우 '카테고리 탐색기'가 없는 상태에서는 상위의 항목으로 이동한 후 다시 하위 항목으로 이동하는 2단계의 이동이 불가피 했으나, '카테고리 탐색기'가 제공되면 한 단계의 이동만으로 동위

항목으로의 이동이 가능하다. 여타의 디렉터리 서비스에서도 적극적인 도입이 필요하다.

등록된 정보자원의 배열순과 관련하여 도입을 검토할 만한 이용자 인터페이스는 엠파스의 '적합도순', 네이버의 '네이버 랭킹' 순, '관련도순' 그리고 야후의 '관련순' 등 배열방법의 선택권이다. 디렉터리 서비스는 아니지만 키워드 검색 결과 화면에서 세 검색 포털은 항목에 따라 디렉터리 서비스에서 제공하고 있는 자포순, 인기순, 등록일순 이외의 배열방식 선택권을 제공하고 있다. 이러한 배열순 선택권을 제공한다면 이용자의 다양한 요구에 더 나은 대응이 가능하다. 또한 일부 항목에서 제공하고 있는 것과 같이 동위 항목의 배열기준을 구분기준별로 선택하여 배열할 수 있도록 이용자 인터페이스를 제공하는 것도 논리적인 주제접근을 가능하게 해주는 개선사항이다.

5. 결 론

이상의 연구를 통해 디렉터리 서비스 정보자원의 선정 및 분류작업의 문제점을 정보자원 등록기준 분석, 사이트 유효성관리 평가, 동일 정보자원의 중복분류 평가, 유사 정보자원의 분산 분석, 정보자원 표제 및 설명문 기술 평가, 그리고 기타의 문제점 등의 관점에서 살펴 보았다. 또한 이러한 문제인식에 기반하여 디렉터리 서비스의 개선방안으로서 분류대상 정보자원의 선정 및 관리, 분류체계의 정비, 분류체계 일람도구, 내부작업 도구의 정비, 외부자

원 활용, 그리고 이용자 인터페이스의 개선 등을 제시하였다.

각 개선방안별 구체사항을 살펴보면, 분류 대상 정보자원의 선정 및 관리와 관련하여 선정기준의 수립 및 최소 등록단위의 일관된 적용과 등록자원 유효성 검토의 필요성을 개선방안으로 제시하였다. 분류체계의 정비와 관련하여 분류체계의 논리성과 유연성 그리고 항목 내 등록자원의 계량적 균형유지를 개선안으로 제시하였다. 분류체계 일람도구 도입과 관련하여 이용자가 이용중 언제라도 편리하게 전체적인 분류체계를 일별해 볼 수 있는 분류체계 일람도구의 도입 필요성을 제시하였다. 내부작업도구의 정비와 관련 카테고리 해설집, 작업 매뉴얼, 용어사전 등 분류작업의 일관성을 유지하는 데 도움이 될 수 있는 각종의 보조도구의 구비 필요성을 언급하였다. 외부자원 활용과 관련하여 외부 전거 시스템의 활용이나 벤치마킹의 필요성에 대해 언급하였다. 마지막으로 이용자 인터페이스의 개선과 관련하여는 분류체계 일람도구의 도입과 더불어 다양한 배열순의 선택장치의 도입을 제시하였다.

전술한 바와 같이 이제 인터넷 정보자원은 일상적인 생활에서뿐만 아니라 학술연구활동을 위한 정보수집의 과정에서도 중요한 비중을 차지하고 있다. 이렇게 중요성이 증가하고 있는 인터넷 정보자원의 효율적인 검색 및 유통을 위해 제공되고 있는 디렉터리 서비스의 개선을 위해 많은 연구와 노력들이 경주되어야 할 것이다.

참고문헌

- 곽철완. 2002. 전자도서관을 위한 분류체계에 대한 소고. 『국회도서관보』, 39(2): 3-14.
- 김영보. 1997. 『인터넷 탐색엔진의 분류체계에 대한 연구: 컴퓨터 인터넷분야를 중심으로』. 석사학위논문, 성균관대학교 대학원, 문헌정보학과.
- 김정현. 2001. 도서관에 있어 인터넷 자원의 목록문제. 『국회도서관보』, 38(4). [인용 2004. 03. 14].
<http://www.nanet.go.kr/libinfo/data/library/nal2001_082.htm>.
- 김태수. 2000. 분류의 이해. 서울: 문헌정보처리연구회.
- 남영준. 1998. 웹 문서분류체계의 분석 및 새로운 설계. 『한국문헌정보학회지』, 32(3): 207-230.
- 남영준, 최승순. 2002. 한국십진분류체계와 웹 문서의 통합 분류체계 개발. 『국회도서관보』, 39(1): 25-43.
- 남태우, 이해영. 2001. 우수한 웹 자원 선정을 위한 웹 자원의 질적 평가에 관한 연구. 『국회도서관보』, 38(4). [인용 2004. 3. 14].
<http://www.nanet.go.kr/libinfo/data/library/nal2001_081.htm>.
- 박창호, 염성숙, 이정모. 1999. 사용자 중심의 홈페이지 분류체계가 분류 검색에 미치

- 는 효과. 『한국 인지과학회 논문지』, 11(1): 47 65
- 배금표. 1999. 인터넷 정보자원 평가요소. 『제6회 한국정보관리학회 학술대회 논문집』, 89 94.
- 배희진, 이준호, 박소연. 2004. 로그 분석을 통한 네이버 이용자의 디렉터리 접근 행태에 관한 연구. 『한국도서관·정보학회지』, 35(1): 1 17.
- 신동민. 2001. 인터넷 검색엔진의 디렉토리 구성에 관한 연구. 『정보관리학회지』, 18(2): 143 163.
- 안인자. 1997. 전자도서관의 WWW 데이터 평가기준에 대한 연구. 『정보관리학회지』, 14(2): 249 268.
- 이관주. 1999. 웹 검색엔진의 기능과 특징에 관한 연구. 『도서관』, 54(3): 3 27.
- 이명희. 2000. 국내 디렉터리 검색엔진의 교육 분야 분류체계 연구. 『국회도서관보』, 37(2): 25 46.
- 이명희. 2000. 인터넷 학술연구정보의 포털사이트 구축에 관한 연구. 『한국도서관·정보학회지』, 31(4): 313 339.
- 이우범. 2002. 인터넷 검색엔진의 분류체계 분석 및 이용증대 방안 연구. 『민족문화』, 13:211 233.
- 이용봉. 1999. 인터넷 웹 사이트 문서의 평가기준 및 방법. 『한국도서관·정보학회지』, 30(1): 151 169.
- 이준호, 박소연, 권혁성. 2003. 질의 로그 분석을 통한 네이버 이용자의 검색행태 연구. 『정보관리학회지』, 20(2): 27 41.
- 이혜영, 남태우. 1999. 인터넷 자원의 질적 평가에 대한 연구. 『제6회 한국정보관리학회 학술대회 논문집』, 99 104.
- 한국데이터베이스진흥센터. 1998. 데이터베이스의 효율적인 검색을 위한 분류체계 및 초록 작성 지침. 서울: 한국과학기술원 부설 연구개발정보센터.
- 황혜경, 정영미. 1999. 정보자원으로서의 웹사이트 평가에 관한 연구. 『제6회 한국정보관리학회 학술대회 논문집』, 95 98.
- Bergman, Michael K. 2001. "The Deep Web : Surfacing Hidden Value." *The Journal of Electronic Publishing*, [cited 2004. 7. 10]. <http://www.press.umich.edu/jep/07_01/bergman.html>.
- Dahlberg, Ingtraut. 1995. "The Future of Classification in Library and Networks, a Theoretical Point of View." *Cataloging and Classification Quarterly*, 21(2): 23 35.
- Hudon, M. 2000. "Innovation and Tradition in Knowledge Organization Schemes on the Internet." In *Dynamism and Stability in Knowledge Organization: Proceedings of the Sixth International Isko Conference*, edited by Beghtol, C.,

- Howarth, L, and Williamson, N.J, Wurzburg, Germany: Ergon Verlag. Quoted in Zins, Chaim, 2002. "Models for Classifying Internet Resources." *Knowledge Organization*, 29(1) : 20-28.
- Lyman, Peter, 2002. "Archiving the World Wide Web." In *Building a National Strategy for Digital Preservation Issues in Digital Media Archiving*. Washinton, D.C. : CLIR. 재인용: 윤희운. 2003. 디지털 장서관리의 패러독스 분석. 『한국비블리아』, 14(1): 5-24.
- Markey, Korea and Demeyer, Ahn N, 1986. *Dewey Decimal Classification Online Project: Evaluation of Library Schedule and Index Integrated into the Subject Searing Capabilities of an Online Catalogue*. Dublin, Ohio: OCLC Online Computer Library Center Inc. Quoted in Vizine Goetz, Diane, 1996. "Using Library Classification Schemes for Internet Resources." In *Proceedings of The OCLC Internet Cataloging Project Colloquium*. [cited 2003, 09, 03]. <<http://www.lub.lu.se/desire/radar/reports/D3.2.3/>>.
- Koch, Traugott., et al, 1997. *The Role of Classification Schemes in Internet Resource Description and Discovery*. [cited 2003, 9, 2]. <<http://www.lub.lu.se/desire/radar/reports/D3.2.3/>>.
- Svenonius, Elaine, 1983. "Use of Classification in Online Retrieval." *Library Resources & Technical Services*, 27(1): 76-80.
- Van der Walt, Martin, 1998. "The Structure of Classification Schemes used in Internet Search Engines." *Advances in Knowledge Organization*, 6: 379-387.
- Vizine Goetz, Diane, 2002. "Classification Schemes for Internet Resources Revised." *Journal of Internet Cataloging*, 5(4): 5-18.
- Zins, Chaim, 2002. "Models for Classifying Internet Resources." *Knowledge Organization*, 29(1): 20-28.