

인용정보와 연구자 홈페이지를 이용한 디지털 도서관 장서개발 방안 연구

Developing Digital Library Collection Using Citation and Homepage Information

이지연(Jee Yeon Lee)*

초록

오늘날 정보환경은 정보이용자로 하여금 도서관 장서뿐만 아니라 다양한 디지털 정보자원으로의 접근을 가능하게 하였다. 변화하는 환경에서 이용률이 증가하는 디지털 정보자원으로 웹사이트, 각종 데이터, 홈페이지 등을 들 수 있다. 이 연구는 ISI Web of Knowledge의 Essential Science Indicators에서 제공하는 인용정보를 이용하여 임상의학 분야의 피인용 횟수가 높은 저자들과 이 저자들의 논문을 인용한 저자들의 홈페이지에 수록된 연구 성과물을 추출한 후, 의학대학 디지털 도서관 및 학술포털과 비교 분석하였다. 10,000여명의 저자 중 연구 성과물 정보를 제시하는 146명의 홈페이지 연구정보를 분석하였는데, 홈페이지에서만 제공하는 연구정보는 학술지 논문, 단행본, 학술대회 발표논문, 강의노트를 포함하였다. 이 중 학술지 논문은 약 15%, 단행본은 약 32%, 학술대회 발표논문은 약 48%, 강의노트는 100% 모두 홈페이지를 통해서만 접근할 수 있는 연구 성과물로 분석되었다. 실제로 의학대학 도서관과 학술포털에서 제공하는 연구정보가 대부분 일치했던 점을 고려할 때, 저명한 연구자들의 홈페이지를 통한 연구정보의 확충이 디지털 도서관 장서개발의 양적, 질적 향상을 위한 방안이 될 수 있다.

ABSTRACT

Nowadays the information environment enables users to access the traditional library collection as well as various digital information resources. In this rapidly changing environment, the use of digital information resources such as web sites, data, and homepages has increased. In this research, highly-cited authors' research outcomes as well as the research outcomes of the people, who cited the highly-cited authors' works, were extracted then compared with information stored in the medical colleges' digital libraries and the academic information portals in the clinical medicine area by using the citation information provided by Essential Science Indicators from ISI Web of Knowledge. Out of 10,000 authors, 146 people's homepages, which present research outcomes, were analyzed. The research outcomes listed in the homepages included journal papers, monographs, conference proceedings, and lecture notes. About 15% of the journal papers, 32%

* 연세대학교 문헌정보학과 조교수 (jlee01@yonsei.ac.kr)

of the monographs, 48% of the conference proceedings, and 100% of lecture notes were accessible only through the homepages. The research outcomes accessible from the homepages were almost analogous to the ones available through the medical college's digital libraries and the academic information portals. Therefore, the digital library collection will be improved and expanded quantitatively and qualitatively by collecting and using the information in the homepages of the prestigious researchers.

키워드: 디지털 도서관, 장서관발, 웹 자원, 인용정보, 저자 홈페이지
digital library, collection development, web resources, citation,
researcher homepages

1 서론

도서관은 건물이나 시설(의 유무), 주된 이용자 계층, 서비스의 종류와 제공 형태 등 다양한 요소에 의해 그 성격이 구분될 수 있지만, 무엇보다도 도서관의 내용과 특징을 드러내는 부분이 장서이다. 장서개발은 이러한 점에서 지속적으로 연구 분야 및 현장에서 관심대상이 되어왔고 장서개발 정책 및 기준, 장서개발 시스템, 장서개발 요소 등 다양한 측면에서 연구가 진행되었다. 장서개발 요소로는 수집 및 평가, 정리, 구축, 유지보수 등 장서관리의 일련의 과정이 포함되는데 이 중에서도 도서관에서 필요한 양질의 정보자원을 수집하는 부분은 장서개발에서 중요한 단계인 동시에 재정과 밀접하게 연관된 어려운 문제이기도 하다.

장서 수집 시의 고려사항은 국내외의 논문을 통하여 여러 차례에 걸쳐서 소개되었고 지속적으로 증보되었는데, 이용자의 요구를 만족시켜야 한다는 대명제와 이를 실현하는데 필요한 재정적 확보 사이에서 최적의 절충점을 찾기 위해 고려해야 하는 문제들이 제시되었다. 또한 장서 수집은 정보화 시대, 디지털 정보자원의 증가와 더불어 고려사항들이 더욱 복잡해지고 추가되었는데, 외국의 경우 디지털 시대의 장서개발에 대한 여러 연구 자료가 출판되었다(Vignau and Quesada 2006; Joint 2006; Kiondo 2004; Ramirez and Gyeszly 2001; Kovacs and Elkorday 2000). 또한 외국의 국립을 포함한 관중별 도서관의 경우, 장서수집에 대한 정책과 방향을 구체적으로 제시하고 있으며, 미국, 영국, 캐나다, 호주 등 여러 나라의 도서관에서 전자자원에 대한 수집 및 관리방안이 웹사이트를 통해서 구체적으로 명시되어 있다. 반면 국내의 경우는 소수의 연구 논문이 출판되었으며(이지연, 김성연 2003; 강숙희 1999a; 강숙희 1999b), CD-ROM이나 온라인 서비스 등 초기의 전자정보 서비스를 다루고 있거나 개념적인 소개에 그치고 있어서 빠르게 변모하는 디지털 정보자원의 수집 및 개발에 대한 후속적인 연구가 필요한 시점이다. 또한 국내의 도서관에서 자관에 맞는 구체적이고 적절한 장서개발 정책을 수립하여 시행하는 경우가 드문 상황이며, 장서개발 정책이 수립되어 있는 경우도 모두 내부 참고용으로 사용하고 있어서 외부적으로 도서관 홈페이지나 간행물을 통하여 장서개발 정책을 홍보하고 있는 경우는 찾을 수 없다.

본 연구는 디지털 장서의 수집에 초점을 두고, 디지털 도서관에서 특히 웹 자원을 수집하여 장서개발에 적용할 수 있는 방안을 제시하고자 한다. 인터넷 정보자원의 경우, 광범위하고 다양한 유형의 정보를 제공하는 장점이 있는 반면에 제공하는 정보의 신뢰성과 권위성 측면에서 문제점이 지적되어 왔다. 인터넷 정보자원과 대조적으로 디지털 도서관의 경우 전문적인 장서관리가 이루어지는 장점이 있는 대신에 제공되는 장서의 종류 면에서 한계를 가지고 있다. 인터넷을 통하여 제공되는 정보자원 중 아직 디지털 도서관에서 제공되지 않고 있으나 이용자에게 유용한 정보를 찾아서 도서관 장서개발에 고려하는 것은 의미 있는 것으로 판단된다. 이 연구는 웹 자원 중 연구자 홈페이지에 수록된 연구 정보를 기존의 디지털 도서관 장서와 비교하여 어떤 부분이 겹치고 어떤 부분이 홈페이지를 통해서만 제공되고 있는지를 파악하여, 현재 디지털 도서관에서 제공되고 있는 않는 연구 정보에 대하여 분석하고 이를 디지털 도서관의 장서 수집에 적용할 수 있는 방안을 제시하고자 한다. 웹 자원 중 신뢰성 있는 연구 정보를 수집하기 위하여 연구자 홈페이지 선정을 위하여 인용정보를 이용하였다.

웹 자원은 웹을 통하여서만 존재하는 정보를 제공하는 정보원으로 웹사이트, 홈페이지, 블로그, 커뮤니티 등을 통하여 존재하는 정보이다. 이 연구는 웹 자원 중 연구자 홈페이지

에 수록된 연구 성과물을 대상으로 하여 기존의 디지털 도서관 장서에 포함되지 않은 부분은 어떤 정보자원인지 파악하고 이를 디지털 도서관 장서에 포함시킬 수 있는 방안을 제시하고자 하였다. 최근에 많은 연구자가 홈페이지를 통하여 연구물을 게시하고 있으며 이 중에 원문을 제공하는 경우도 적지 않다. 또한 홈페이지에 수록된 연구정보에는 출판된 단행본이나 학술지 논문 이외에도 학술대회 발표자료, 강의노트, 연구 데이터 등 디지털 도서관 사이트를 통해서 접할 수 없는 다양한 웹 자원이 포함된다. 이러한 웹 자원은 연구자, 강연자 집단에 유용한 정보가 될 수 있고, 실제로 이용자 연구를 통하여 많은 연구자 집단이 동료, 비공식 커뮤니케이션, 저명한 연구자의 홈페이지를 통하여 정보를 획득한다는 사실이 밝혀졌다. 예를 들어, 공학자 집단은 출판된 연구물에서 필요한 정보를 획득하기보다는 자신들이 가지고 있는 지식, 동료, 혹은 소속 기관 내에 존재하는 정보원을 더 많이 이용한다는 사실이 여러 연구를 통하여 밝혀졌다(Holland and Powell 1995; Pinelli 1991; Case, Borgman, & Meadow 1986; Raitt 1985; Shuchman 1981). 또한 과학자 집단은 정보추구에서 오래 전부터 알려진 'invisible college'의 가장 왕성한 이용자 집단이고, 사회과학자는 자연과학자 집단과 더불어 직장이나 학술대회에서 동료들과 비공식 커뮤니케이션을 통하여 필요한 정보의 상당 부분을 획득하는 것으로 밝혀졌다. 또한 동료를 통한 이러한 정보추구 행태는 모든 학문 분야에 있어서 점점 보편화되는 추세를 보이고 있다(Case 2002).

이 연구는 위와 같은 연구자 간의 정보교류가 학술정보의 접근에 중요함을 인식하고, 이러한 정보를 디지털 도서관을 통하여 접근할 수 있는 방안으로, 연구자 홈페이지에 있는 정보를 활용하는 방안을 다루었다. 연구자 홈페이지는 공식적으로 출판된 연구물 이외에 개인적으로 혹은 비공식적으로 생산된 연구물과 연구정보를 포함하는 경우가 많이 있다. 예를 들어, 연구자 홈페이지에는 본인이 출판한 단행본, 학술지 논문 이외에 발표자료, 강의자료, 경우에 따라서 자신의 관심 연구 분야 추천자료 등 다양한 연구정보를 담게 되는 경우가 종종 있다. 그러나 지금까지 거의 모든 디지털 도서관은 장서의 권위성과 질적 우수성을 보장하기 위하여 공식적 경로를 거쳐서 출판된 단행본이나 논문만을 포함하였다. 그러나 실제로 미출판자료에 대한 연구자 집단의 이용 요구가 지속적으로 증가하고 있고, 인터넷이나 웹을 통한 정보접근의 기회가 확대되면서 웹 자원을 이용하는 추세는 가속화되었다. 따라서 디지털 도서관 장서에서 기존에 접근하던 자료 이외에 연구자 홈페이지를 통해 접근할 수 있는 정보의 종류와 양은 어느 정도 되는지 조사하고, 이러한 자료를 디지털 도서관에서 흡수하여 이용자에게 제공할 필요성에 대하여 진단할 필요가 있다.

실제로 이 연구를 시작하는 초기에 물리학 분야의 외국대학 연구자와 정보획득에 대하여 비공식적이지만 심도 있는 면접을 수행할 기회가 있었는데, 피면접자는 자신이 필요한 대부분의 정보를 물리학 분야, 특히 자신이 연구하고 있는 세부 분야의 저명한 학자들의 홈페이지를 방문하여 획득한다고 대답하였다. 동료들과의 대화를 통하여 필요한 정보를 직접 입수할 때도 있지만, 연구 정보를 수집할 때 홈페이지를 이용하면 시간적, 공간적 제약 없이 손쉽게 원하는 정보를 얻을 수 있다고 하였다. 면접을 통하여 알게 된 저명한 학자의 홈페이지에서 양질의 연구 정보를 획득할 수 있다는 답변에 기반을 두어, 이를 디지털 도서관 장서개발 대상으로 선정할 필요성에 대한 진단 및 방안 연구를 시작하였다.

2 디지털 정보자원의 장서개발 관련 연구

21세기에 들어서면서 전자정보원 그리고 더 나아가 디지털 정보자원에 대한 관심이 늘어 나게 되었고, 이러한 동향은 도서관 장서 수집에도 영향을 미치게 되었다. 전자정보의 수집은 인쇄 출판되는 정보에 비해 기술적 측면, 비용적 측면, 저작권 측면에서 복잡한 문제를 수반하게 된다. 따라서 기존의 인쇄매체를 통한 출판물을 다루는 장서수집 전략을 그대로 전자정보원에 적용하는 것은 무리이며, 더구나 인터넷을 통하여 전달되는 전자정보원의 경우는 전통적인 장서 수집 방식으로는 그 한계가 많다(Gessesse 2000). 이러한 한계에 대한 인식을 공유하는 연구물이 21세기에 들어서면서 증가되었는데, 이러한 연구물들의 경향은 크게 두 가지로 나누어 볼 수 있다.

첫째는 전자정보원의 수집 기준 및 정책에 관련된 부분이다. 국립도서관의 경우는 미국 의회도서관이나 호주 국립도서관을 예로 들 수 있다. 미의회도서관은 전자정보원의 장서선정 기준을 홈페이지(<http://www.loc.gov/acq/devpol/electronicselectionguidelines.html>)를 통해 명시하고 있는데, 전자정보원을 입수자원(acquire), 수집자원(collect), 연결자원(link)으로 나누어서 정의를 내리고 있으며 장서 선정 기준을 단계별로 제시하고 있다(표 1).

[표 1] 미의회도서관 전자정보원 장서선정 요약

정의	<ul style="list-style-type: none"> ◆ 전자정보원: 다양한 형태로 코드화시켜서 컴퓨터를 통하여 접근할 수 있도록 만든 정보자원 ◆ 입수자원(acquire): 계약, 기증, 구매, 교환 등 정식 입수절차를 거쳐 원거리에서 혹은 기관 내부에서 직접 연결할 수 있는 전자자원 ◆ 수집자원(collect): 도서관에서 장서선정 절차를 거쳐 선택한 소유권을 가지고 있는 전자자원 ◆ 연결자원(link): 다른 곳에서 제공하는 정보를 도서관 웹사이트 상에서 포인터로 연결하는 경우 	
선정 기준	일반적 기준	<ul style="list-style-type: none"> ◆ 정보의 유용성(특히 미의회 및 연구자 집단에 유용한 정보) ◆ 정보제공자의 신뢰성, 평판 ◆ 제공되는 정보의 독창성 ◆ 학문적 내용(scholarly content) ◆ 디지털 형태로만 제공되는 내용 ◆ 소멸 위험성이 높은 정보(사고, 고의적 삭제, 회색문헌 등)
	세부 기준	<ul style="list-style-type: none"> ◆ 정보 내용의 질적 우수성, 학문적 가치 ◆ 컴퓨터에서 작업하는 연구자나 스태프의 업무를 지원 ◆ 접근성(서버 안정성, 멀티미디어 정보의 적절한 이용, 저작권 준수, 이용자 등록 정보 관리, 플러그인 소프트웨어 조향) ◆ 이용자 인터페이스, 네비게이션의 편리함 ◆ 기술적 기준과 장비의 필요성 ◆ 영구 보유 자원으로서의 가치

호주 국립도서관은 장서부분을 두 부류로 나누어 개발 정책을 제시하고 있는데, 하나는

일반적인 모든 정보자원을 다루고 있으며 다른 하나는 전자정보원을 따로 뽑아서 구체적이고 자세하게 다루고 있다. 그리고 그림 1에서 보는 바와 같이 전자정보원의 장서개발 정책에서 수서기준(acquisition guidelines), 디지털화 정책(digitisation policy), 전자정보원 선정, 보존을 위한 온라인 출판물 선정 등 장서 수집/선정과 관련된 항목이 다수를 차지하고 있다.



[그림 1] 호주 국립도서관 장서개발 정책(<http://www.nla.gov.au/policy/colldevt.html>)

공공도서관의 경우는 이란주(2005)의 논문에서 모범적인 장서개발 정책을 가지고 있는 것으로 평가하였던 퀸스 도서관을 예로 들 수 있다. 장서개발에 적용되는 자료의 형태를 ‘인쇄물과 시청각 미디어’ 및 ‘전자미디어’로 나누고 전자미디어의 사용과 관련하여 인터넷 이용 정책을 상세하게 기술하고 있다(퀸스 도서관 2004). 다만, 공공도서관이 앞의 국립도서관과 비교하여 장서 수집 자체보다는 디지털 장서의 이용에 중점을 두고 장서 개발 정책을 진행하고 있다는 것을 알 수 있다. 뉴욕 공공도서관(www.nypl.org/branch/)이나 시카고 공공도서관(www.chipublib.org)에서도 홈페이지 최상위 항목 중 하나로 ‘디지털 장서’ 항목을 포함시킨 점에서 디지털 장서의 이용 측면을 강조하고 있음을 알 수 있다.

디지털 정보자원 수집과 관련된 연구물의 두 번째 특징은 웹 자원을 다루고 있다는 점이다. 웹 자원은 그 범위가 다양하고 유동적이어서 온라인 데이터베이스나 CD-ROM과 같은 보편화된 전자정보원과는 차이가 있다. Kovacs와 Elkordy(2000)는 이미 수년전에 엄선된 웹 자원의 필요성을 언급하면서, 우수한 웹 자원(high-quality web resources)으로 디지털 도서관 장서를 구성하는 것은 중요하지만 동시에 다양한 형태의 이질적인(unfamiliar) 정보자원을 도서관 장서로 포함시키는 것은 복잡한 문제임을 지적하였다.

인터넷이나 웹이 있기 전에는 얻을 수 없었던 정보들을 현재는 인터넷이나 웹을 통하여 이용자들이 접근하고 있으며, 이는 도서관에 도전인 동시에 새로운 기회이기도 하다. 실제로 최근에 도서관 웹페이지를 통하여 이러한 인터넷 정보자원에 접근할 수 있도록 서비스를 제공하고 있는 예가 생겨나고 있다. 시카고 공공도서관은 최상위 메뉴에 ‘디지털 장서’ 뿐

만 아니라 ‘엄선된 인터넷 정보자원’ 항목을 넣음으로써 인터넷 상의 정보자원에 이용자가 접근할 수 있도록 하고 있다(그림 2). 시카고 공공도서관에서 제공하는 인터넷 정보자원은 시카고 근처의 이용자들에게 유용하다고 판단되는 웹 자원을 다양한 주제에 의해서 접근하도록 하고 있다(그림 3). 예를 들어, 통계의 경우 센서스 데이터, 범죄 관련 통계자료, 경제 분야 통계자료, 빈곤, 의료 등의 사회 관련 통계자료들을 제공하고 있다(그림 4).



[그림 2] 시카고 공공도서관에서 제공하는 인터넷 정보자원
(<http://www.chicagopubliclibrary.org/cpl.html>)

Selected Internet Resources

The Internet sites listed in the categories below are recommended by librarians at the Chicago Public Library and are selected based on their quality and utility in answering common questions asked at CPL reference desks. In addition to these individual sites, librarians at CPL refer patrons to larger collections of sites, such as [Illinois Clicks](#), the [Librarians' Index to the Internet](#), and the [Internet Public Library](#). Like the sites referenced in the categories below, the sites included in these larger collections are selected by information professionals based on their usability and the accuracy and timeliness of their content.

Almanacs, Encyclopedias, Dictionaries, and Directories	History
Architecture, Art and Music	History Fair
Biographies	Holidays
Books, Literature, and Poetry	Homework Help
Business & Company Information	Law
Careers & Jobs	Parenting Resources
Chicago Area Directory of Organizations	Reader Resources
Children's Resources	Reference Shelf
Consumer Information	Science Fair, Science and Technology
Copyright, Patents & Trademarks	Senior Citizens Internet Resources
Education & Scholarship Information	Small Business & Start-up Information
Gay, Lesbian, Bisexual & Transgender Concerns	Statistics
Genealogy	Stocks & Investments
Government	Taxes
Grants and Funding	Teacher Resources
Health & Medicine	Teens

[그림 3] 시카고 공공도서관 인터넷 정보자원의 주제영역
(<http://www.chicagopubliclibrary.org/008subject/008subject.html>)

Selected Internet Resources

Home > Selected Internet Resources > Government Information

Statistics

Below you'll find general statistical sites at various governmental levels. In the subsections following, you'll find specific sites for the [census](#), [crime statistics](#), [economic statistics](#), and [social statistics](#).

- [FedStats](#)
The Federal Interagency Council on Statistical Policy maintains this Web site showcasing over 70 Federal agencies' statistics on the Web. Check the [A to Z index](#) for your topic.
- [Statistical Abstract of the United States](#)
This almanac is a fixture at reference desks everywhere because it offers statistics on almost every measurable topic from population to manufacturing. To save time, check [U.S.A. Statistics in Brief](#). There's also a [State Rankings](#) section.
- [Illinois Statistical Abstract](#)
A comprehensive compilation of economic and demographic data relating to Illinois and to political subdivisions within the state.
- [Illinois County Statistical Abstract](#)
A comprehensive compilation of economic and demographic data relating to specific Illinois counties.
- [Chicago Fact Book](#)
Chicago's most recent official demographic, labor force, and education statistics.

Census

The [Census Bureau](#) offers statistics on all kinds of topics, from population to manufacturing, housing to employment. It's an enormous Web site offering thousands of statistics and studies. The following links are meant to help you get started with your research:

- [American FactFinder](#)
The central resource for new Census data. A useful place to start is the [data sets](#) page.
- [Subjects Index](#)
A handy way to find information on the Census Web site.
- [Illinois QuickFacts](#)
Demographic and economic data for Illinois and its counties (or other [states](#)).
- [Census 2000 Results for Northeastern Illinois](#) from the Northeastern Illinois Planning

[그림 4] 시카고 공공도서관에서 제공하는 인터넷 정보자원 예 - 통계부분
(<http://www.chicagopubliclibrary.org/008subject/005genref/statistics.html>)

이상에서 설명한 외국의 국립 및 공공도서관은 전자정보원이나 인터넷 정보자원을 장서개발 및 정보자원 제공에 있어 항목을 분리시켜 집중적으로 관리하고 있음을 알 수 있었다. 장서개발 정책 및 기준의 제시, 장서 선정 범위 확대, 웹페이지 상의 이용자 접근 방식 개선 등 장서개발 및 관리의 전 과정이 철저하게 이루어지고 있으며 웹페이지를 통하여 적극적으로 홍보하면서 이용을 유도하고 있었다.

웹 자원은 웹페이지를 통하여 제공되는 각종 데이터를 의미하며 특히 웹을 통해서만 제공되는 자원을 말한다. 반면 웹 기반 자원은 인쇄매체나 다른 매체를 통하여 제공되는 정보가 동시에 웹을 통하여서도 제공되는 것을 가리킨다. 이러한 차이점을 볼 때 웹 자원은 다른 매체를 통해서 접근하기 어려운 정보를 제공하므로 정보 획득의 중요한 채널이 될 수 있다. 웹 자원의 중요함을 인식하면서 생겨난 용어가 'webliographies' 혹은 'web directories'라고 불리는 웹 서지이다. 특정 주제에 대하여 도서관 검색시스템을 통하여 얻게 되는 자료에 대한 정보를 서지(bibliographies)라고 한다면, 이러한 개념을 웹 자원에

적용시킨 것이 웹 서지이다. 웹 서지는 도서관에서 얻게 되는 정보와 같이 이용자에게 유용하고 우수한 웹 자원을 의미한다. Kovacs와 Elkordy(2000)는 논문의 부록에 장서개발 도구에 대한 웹 서지를 제공하고 있다. 여기에는 100여개의 장서개발 도구에 대한 정보 및 웹사이트 주소를 제공하고 있는데, 웹을 통해서만 제공받을 수 있는 유용한 정보를 리스트 형태로 제공하고 있다.

디지털 장서, 그중에서 특히 웹 자원은 웹이 활성화되고 웹 2.0과 같은 형식보다는 내용에 치중하는 웹 환경에서 도서관 장서로도 중요한 의미를 지니고 있으며, 이에 대해 심도 있게 고려할 필요가 있다.

3. 연구자 홈페이지 정보를 이용한 장서개발 방안 연구

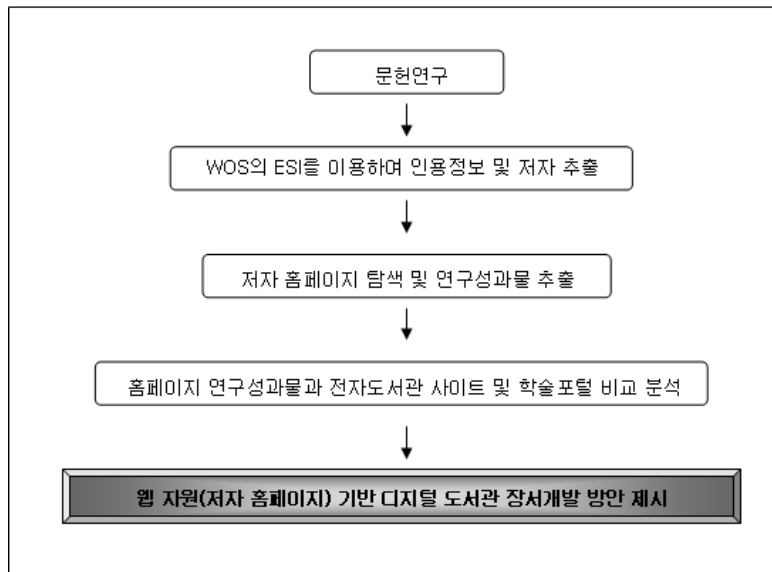
3.1 연구과정

연구를 위하여 분야를 선정하였는데, 연구와 현장이 같이 연결된 분야로 임상의학(clinical medicine) 분야를 택하였다. 임상의학 분야의 저명한 연구자를 선정하기 위하여 연구자들의 인용정보를 이용하였다. 우선 선정된 분야에서 인용이 가장 많이 된 상위 피인용 저자를 ISI사의 Web of Science에서 제공하는 ESI(Essential Science Indicators)를 이용하여 추출하였다. ESI는 저자별, 기관별, 국가별, 학술지별로 인용순위를 검색할 수 있는 응용도구이다. ESI를 이용하여 1996년 1월 1일부터 2006년 12월 31일까지 임상의학 분야에서 출판된 논문을 대상으로 하여 가장 많이 인용된 저자 30명을 추출하였다. 각 저자의 논문을 추출하여 총 17,625개의 논문 리스트를 작성하였다(표 2). 이 중 인용빈도가 높은 상위 120개의 논문을 대상으로 하여 이 논문을 인용한 총 63,774명의 인용저자 리스트를 추출하였다. 이상의 과정을 통하여 추출한 30명의 상위 피인용 저자와 10,236명의 인용저자의 홈페이지를 인터넷 상에서 탐색하여 피인용 저자명 중 중복된 이름을 제외하고 총 146개의 홈페이지에 수록된 논문 및 기타 연구 성과물 리스트를 작성하였다. 홈페이지에 수록된 연구 성과물을 디지털 도서관 사이트와 학술포털에서 검색하여 비교, 분석하였다(그림 5).

[그림 5] 연구과정 개요

3.2 데이터 수집

ESI에서 분야를 임상의학으로 지정하여 인용된 횟수에 의하여 순위를 추출한 과정은 그림 6과 같으며, 결과는 표 2와 같다. 그림 6에 의하면, ISI Web of Knowledge (<http://portal.isiknowledge.com>)의 ESI 메인화면에서 인용순위 중 'scientist' 항목을 선



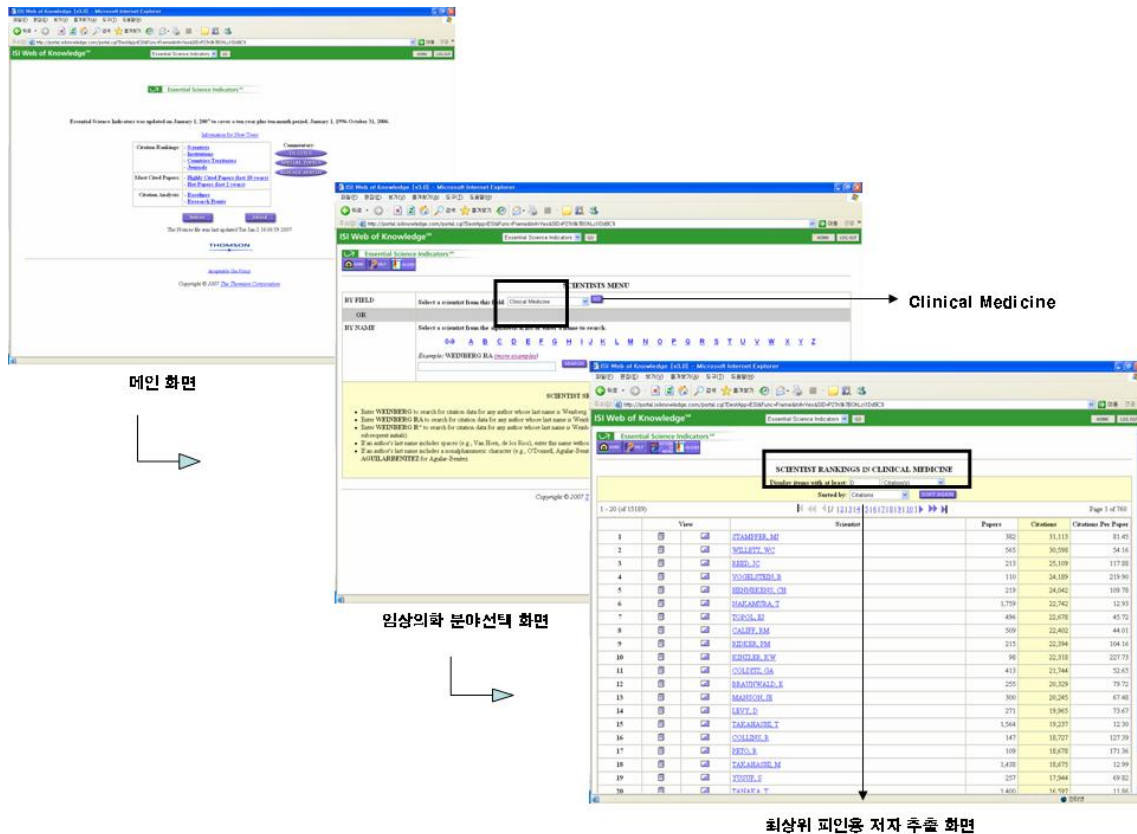
택한 후 다음 화면에서 분야를 '임상의학'으로 선택한다. 그리고 나면 임상의학 분야의 최상위 피인용 저자의 리스트가 화면에 출력된다. 다음 과정은 이 중 상위 30명의 피인용 저자 및 이들의 논문을 인용한 저자를 찾는 과정으로 그림 7과 같다. 그림 6의 마지막 화면에서 피인용 저자의 이름을 선택하면 그 저자의 논문 리스트로 이동한다. 이 중 한 논문을 선택하면 이 논문에 대한 구체적인 정보가 화면에 나타나고, 이 정보 중 'cited references' 항목을 선택하면 이 논문을 인용한 저자와 논문의 서지사항을 보여 주는 화면으로 이동한다.

상위 30명의 피인용 저자의 논문 중 피인용 횟수가 높은 상위 120개의 논문을 선택하여 이 논문을 인용한 횟수가 높은 10,000명의 인용저자 이름을 대상으로 홈페이지를 탐색하였다. 이 때 30명의 상위 피인용 저자도 홈페이지 탐색 대상에 포함하였다.

ISI Web of Knowledge의 ESI로부터 데이터를 수집한 과정을 정리하면 다음과 같다.

- 1) 임상의학 분야에서 피인용 횟수가 높은 상위 30명의 저자 추출
- 2) 상위 30명의 피인용 저자의 논문 중에서 인용빈도가 높은 120개의 논문 추출
- 3) 120개의 논문을 인용한 인용논문의 저자 추출(63,774명)
- 4) 3번에서 추출한 저자 중 120개의 논문에 대하여 인용빈도가 높은 저자 10,000명 추출
- 5) 1번에서 추출한 30명과 4번에서 추출한 10,000명을 홈페이지 탐색 대상으로 선정

위에서 설명한 과정에 의하여 ESI에서 추출된 연구자의 홈페이지를 탐색한 결과 232개의 홈페이지가 검색되었고, 중복제거 과정을 거친 결과 총 146개의 홈페이지가 검색되었다. 임상의학 분야에서 홈페이지 탐색 대상으로 한 의학자 중 중복률($146/232=0.629$)을 고려하면 약 2.32%($146/6290$) 가량이 홈페이지를 운영하고 있는 것으로 조사되어 예상보다 적은 홈페이지 운영률을 보였는데, 이는 의학자 집단이 다른 연구자 집단에 비해서 직업의 특성상 홈페이지를 직접 운영하기보다는 홈페이지 제작 업체나 보조 인력에 의해서 운영하는 경우가 많기 때문인 것으로 분석되었다.



[그림 6] 상위 피인용 저자 추출 과정

TOP PAPERS FOR STAMPER, MJ IN CLINICAL MEDICINE

피인용 저자의 상위 인용 논문 리스트

논문 샘플

앞의 논문을 인용한 저자 리스트

The screenshot shows the following content:

- Author List:** A list of authors including Stamper, M.J., and others, with their affiliations and publication counts.
- Full Record:** A detailed view of a paper titled "Inflammation, aspirin, and the risk of cardiovascular disease in apparently healthy men" by Stamper et al., published in the *New England Journal of Medicine*.
- Citing Articles—Summary:** A list of 7 articles that cite the Stamper et al. paper, including works by Young et al., Danese et al., Chipman, Stenberg et al., Garcia et al., Camero et al., and Ridker et al.

[그림 7] 피인용 저자의 논문을 인용한 저자 추출 과정

[표 2] 임상의학 분야 상위 30위 피인용 저자

순위	저자명	ESI에서 검색된 논문 수	총 피인용 횟수	논문 1개당 피인용 횟수
1	STAMPFER, MJ	382	31,113	81.45
2	WILLETT, WC	565	30,598	54.16
3	REED, JC	213	25,109	117.88
4	VOGELSTEIN, B	110	24,189	219.90
5	HENNEKENS, CH	219	24,042	109.78
6	NAKAMURA, T	1,759	22,742	12.93
7	TOPOL, EJ	496	22,678	45.72
8	CLIFF, RM	509	22,402	44.01
9	RIDKER, PM	215	22,394	104.16
10	KINZLER, KW	98	22,318	227.73
11	COLDTIZ, GA	413	21,744	52.65
12	BRAUNWALD, E	255	20,329	79.72
13	MANSON, JE	300	20,245	67.48
14	LEVY, D	271	19,965	73.67
15	TAKAHASHI, T	1,564	19,237	12.30
16	COLLINS, R	147	18,727	127.39
17	PETO, R	109	18,678	171.36
18	TAKAHASHI, M	1,438	18,675	12.99
19	YUSUF, S	257	17,944	69.82
20	TANAKA, T	1,400	16,597	11.86
21	BARNES, PJ	330	16,392	49.67
22	NAKAMURA, Y	1,331	15,756	11.84
23	WANG, Y	1,029	15,306	14.87
24	TUOMILEHTO, J	344	15,061	43.78
25	LIBBY, P	197	15,036	76.32
26	TANAKA, M	1,367	14,759	10.80
27	SUZUKI, K	1,647	14,481	8.79
28	MATSUZAWA, Y	334	14,312	42.85
29	RIFAI, N	204	14,234	69.77
30	HERMAN, JG	122	14,224	116.59

3.3 연구결과 분석

홈페이지 탐색과정을 통하여 검색된 146개의 홈페이지에 수록된 연구 성과물을 다음과 같이 비교 분석하였다. 우선 146개에 수록된 연구 성과물은 학술지 논문 6,163편, 단행본 45권, 학술대회 발표논문(proceedings) 914편, 강의노트 78편으로 구성되었다. 강의노트의 경우 연구보다는 교육에 더 관련이 있는 자료이나 교육자료 또한 연구정보를 바탕으로 작성되는 경우가 일반적이어서 분석 대상 자료로 포함하였다. 연구자 홈페이지를 통하여 추출한 이상의 연구 성과물을 디지털 도서관 및 학술포털을 통하여 제공되고 있는 연구 성과물과 유형과 특징 면에서 비교 분석하기 위하여 연세대학교 디지털 도서관(의과대학 도서관 및 중앙도서관)과 구글 스칼라를 이용하였다. 의과대학 도서관인 경우 미국의학도서관에서 제공하는 생의학 관련 의학 데이터베이스인 PubMed에서 연구 성과물을 탐색하였다. 비교

분석한 결과는 표 3과 같다.

[표 3] 시스템별 연구 성과물 비교

연구 성과물 검색시스템	학술지 논문	단행본	학술대회 발표논문	강의노트	합
1. 연구자 홈페이지	6,126 (85.5%)	45 (0.6%)	914 (12.8%)	78 (1.1%)	7,163 (100.0%)
2. 의과대학 디지털 도서관 사이트	5,072 (91.0%)	31 (0.6%)	469 (8.4%)	0 (0.0%)	5,572 (100.0%)
3. 일반대학교 디지털 도서관 사이트	182 (91.9%)	13 (6.6%)	3 (1.5%)	0 (0.0%)	198 (100.0%)
4. 구글 스칼라	5,200 (91.1%)	31 (0.5%)	478 (8.4%)	0 (0.0%)	5,709 (100.0%)

표 3에 나타난 시스템별로 검색된 연구 성과물을 살펴보면, 연구자 홈페이지에 나타난 학술지 논문의 상당 부분이 디지털 도서관 사이트나 학술포털을 통해서 제공되는 것으로 나타났다. 특히 의과대학 디지털 도서관 사이트와 학술포털인 구글 스칼라는 상당수 링크로 연결되어 구글 스칼라에서 검색된 논문은 의과대학 디지털 도서관에서도 제공하는 데이터베이스인 PubMed로 연결되고 있었다. 따라서 이 두 시스템 간에 상당히 일치되는 연구 성과물 정보를 제공하고 있음을 알 수 있었다. 다만, 일반대학교 중앙도서관 사이트인 경우는 의학분야가 전체 장서에서 차지하는 비중에 따라 장서비율이 달라질 수 있어서 나머지 시스템에 비하여 전체적인 연구 성과물 수가 적은 것으로 나타났다.

연구자 홈페이지의 경우는 학술지 논문이 전체 연구 성과물의 85.5%, 학술대회 발표논문이 12.8%, 단행본이 0.6%, 강의노트가 1.1%를 차지했다. 반면 의과대학 디지털 도서관과 구글 스칼라의 경우 학술지 논문의 비율이 91% 정도 차지했으나, 실제로 검색된 논문 수 자체는 홈페이지에서 제공하는 논문 수의 82.8%(5,072/6,126)와 84.9%(5,200/6,126)를 의과대학 디지털도서관과 구글 스칼라가 각각 제공하고 있었다. 학술대회 발표논문도 시스템별로 학술지 논문 다음으로 많이 검색되는 연구 성과물 유형이었으나 의과대학 디지털 도서관이나 구글 스칼라에 비해 홈페이지에서 약 2배가량 많은 수의 발표논문이 검색되었다. 단행본의 경우도 홈페이지가 더 많은 양의 정보를 제공하고 있었으며, 특히 강의노트의 경우에는 홈페이지를 통해서만 제공되고 있었다(표 4).

[표 4] 자료 유형별 시스템 연구 성과물 비교

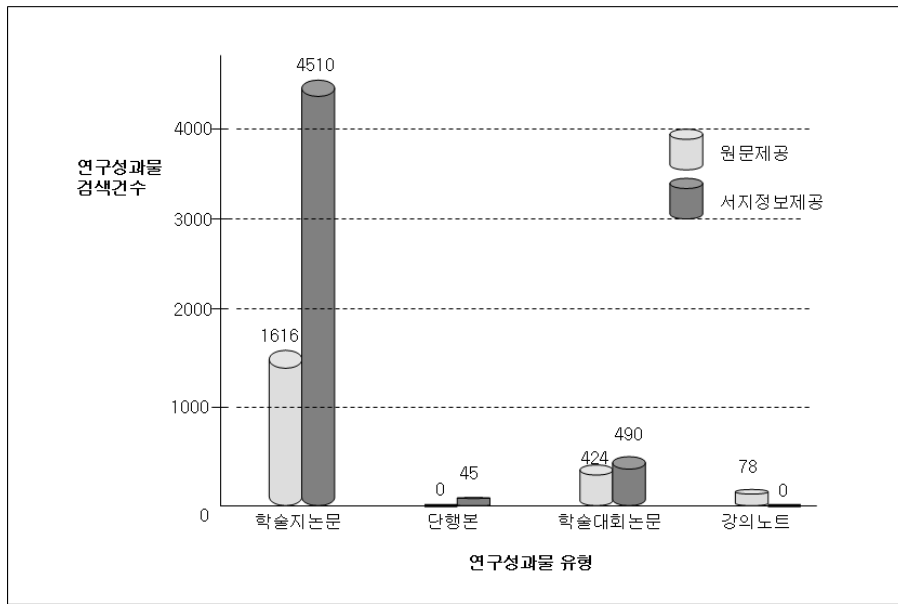
자료 유형	의과대학 디지털 도서관 사이트	일반대학교 디지털 도서관 사이트	구글 스칼라	비교 기준: 연구자 홈페이지
1. 학술지 논문	5,072 (82.8%)	182 (3.0%)	5,200 (84.9%)	6,126
2. 단행본	31 (68.9%)	13 (28.9%)	31 (68.9%)	45
3. 학술대회 발표논문	469 (51.3%)	3 (0.3%)	378 (41.4%)	914
4. 강의노트	0 (0.0%)	0 (0.0%)	0 (0.0%)	78

다음은 각각의 시스템을 통하여 제공된 연구 성과물의 원문제공 여부를 조사하였다(표 5). 원문제공은 이용자의 접근권한에 따라서 달라질 수 있는데, 의과대학 디지털 도서관의 경우는 대학 내 구성원인 경우 PubMed 데이터베이스의 원문제공이 가능하여서 모든 논문을 원문으로 받아볼 수 있으나 대학 내 구성원이 아닌 경우는 서지정보를 입수하는 것이 그치게 된다. 본 연구의 경우 원문접근이 가능한 이용자 아이디를 사용하였으므로 모든 검색된 학술지 논문 및 학술대회 발표논문의 원문을 추출하는 것이 가능하였다. 단행본의 경우도 전자책을 통하여 상당수를 원문으로 제공하고 있었다. 구글 스칼라는 자체적으로 제공하는 원문정보도 있었지만, 상당수의 연구 성과물이 웹 데이터베이스, 도서관 시스템, 그리고 저자 홈페이지로 연결하여 원문을 제공하고 있어서 학술포털로서의 성격을 나타내고 있었다. 연구자 홈페이지의 경우는 의과대학 디지털 도서관이나 구글 스칼라보다 원문 제공 건수 및 비율에 있어서 낮은 것으로 나타났다. 그러나 주목할 만한 부분은 연구자들이 홈페이지를 통해서 점점 더 많은 양의 원문을 제공하고 있다는 점이다. 홈페이지는 서지정보만을 얻을 수 있을 것이라는 전통적인 관념이 바뀌고 있음을 나타냈다(그림 8).

[표 5] 시스템별 검색된 연구 성과물의 원문제공 건수

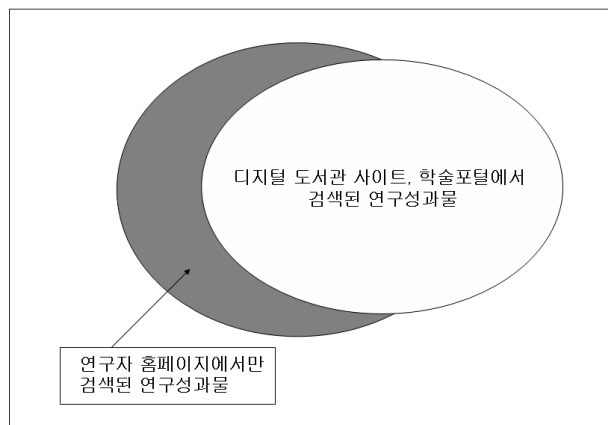
연구 성과물 검색시스템	학술지 논문		단행본		학술대회 발표논문		강의노트	
	원문	비율* (%)	원문	비율 (%)	원문	비율 (%)	원문	비율 (%)
1. 연구자 홈페이지	1,616	26.4	0	0.0	424	46.1	78	100.0
2. 의과대학 디지털 도서관 사이트	5,072	100.0	29	93.5	469	100.0	0	0.0
3. 일반대학교 디지털 도서관 사이트	182	100.0	1	7.7	3	100.0	0	0.0
4. 구글 스칼라	5,111	98.3	29	93.5	465	97.3	0	0.0

* 비율은 각 유형의 정보자원별로 검색된 전체 정보자원 수 중 원문을 제공하는 정보자원 수를 의미함



[그림 8] 홈페이지에 수록된 연구 성과물의 원문제공 건수

이 연구에서 비교 분석의 관심이 되는 부분은 연구자 홈페이지를 통해서만 제공되지만, 디지털 도서관 사이트나 학술포털에 의해서는 제공되지 않는 연구 성과물 정보의 특성에 대하여 알아보는 것이다(그림 8의 회색으로 표시된 부분). 연구자 홈페이지를 통해서만 제공되는 정보는 학술지 논문의 일부, 학술대회 발표논문의 상당수, 단행본의 일부 및 강의노트 전체이다. 임상의학 분야의 경우, 학술지 논문과 단행본은 연구자 홈페이지나 디지털 도서관, 학술포털에서 모두 정보를 찾을 수 있었으나 학술대회 발표논문이나 강의노트와 같이 상대적으로 덜 공식적인 과정을 통하여 배포되는 정보는 홈페이지를 통해서 상당수 혹은 전부를 추출할 수 있는 경우였다. 즉 홈페이지를 탐색하는 경우에만 검색할 수 있는 연구 성과물이나 연구 관련 자료가 있음을 알 수 있었다. 다만 원문제공에 있어, 홈페이지의 경우는 강의노트를 제외한 모든 유형의 연구 성과물에서 디지털 도서관이나 학술포털보다 낮은 비율을 나타냈다.



[그림 8] 장서개발 대상 연구 성과물

4. 저자 홈페이지를 이용한 디지털 도서관 장서개발 방안

저자 홈페이지에 수록된 연구정보를 분석한 결과, 기존의 디지털 도서관이나 학술포털을 통해서 접근할 수 없는 연구정보가 홈페이지에 포함되어 있음을 확인하였다. 홈페이지를 통해서만 접근할 수 있는 정보는 학술지 논문의 약 15%, 단행본의 약 32%, 학술대회 발표 논문의 약 48%, 강의노트는 100% 전체에 해당하였다. 이는 그림 8의 회색으로 표시한 부분에 해당한다. 이 부분에 해당하는 연구정보의 유용성을 검증하기 위하여 임상의학 분야 연구자에게 유용성 평가를 하도록 하였다. 회색부분에 해당하는 연구 성과물의 리스트를 보여 주고 이에 대한 전체적인 평을 하도록 하는 오픈된 방식의 평가를 수행하였는데, 80% 이상의 연구물에 대하여 긍정적인 평가를 내렸다. 강의노트에 대한 의견은 교육현장의 교수자와 학습자 모두에게 유용한 정보임을 언급하였다. 홈페이지 연구정보의 유용성 검증은 임상의학뿐만 아니라 다양한 분야에서 체계적으로 수행할 만한 가치가 있는 후속 연구임을 알 수 있었다.

저자 홈페이지에 수록된 정보를 분석한 결과, 장점과 단점은 다음과 같다.

장점:

- 1) 디지털 도서관이나 학술포털을 통하여 제공받을 수 없는 연구정보에 접근할 수 있다.
- 2) 저명한 연구자의 홈페이지에서 추출한 연구정보이므로 신뢰성 있는 정보를 제공할 수 있다.
- 3) 최신 연구물이나 현재 출판 중인 연구물에 접근하기 쉬워 정보의 신속성이 보장된다.
- 4) 원문으로 제공되는 정보가 계속 증가하고 있다.
- 5) 웹 로봇 등을 이용하여 정보의 주기적인 갱신이 가능하다.
- 6) 정보추구 방식에서 중요하게 인식된 동료나 연구자 간의 비공식 커뮤니케이션의 대체 역할을 기대할 수 있다.

단점:

- 1) 저자 홈페이지에 수록된 원문정보의 경우, 저작권에 대한 요구사항이 공식적인 출판경로나 상업적인 업체를 통하는 경우보다 복잡할 수 있다.
- 2) 연구자의 홈페이지 운영률에 따라 제공하는 장서의 양이 한정된다.

저자 홈페이지 정보를 디지털 도서관 장서개발에 적용하기 위하여 다음과 같은 사항이 선행되어야 한다.

- 1) 저자 홈페이지를 탐색하고 연구정보를 추출하는 웹 로봇이 개발되어야 하고 추출된 정보의 정확성이 요구된다. 참고로 본 연구에서는 탐색된 홈페이지와 추출된 정보의 정확성을 위하여 수작업으로 데이터를 수집하였는데, 상당히 많은 시간과 노력이 필요하였다. 따라서 데이터 수집 과정을 자동으로 수행할 수 있는 시스템이 필요하고, 시스템의 정확성도 보장되어야 한다.
- 2) 디지털 도서관에서 준수할 수 있는 저자 홈페이지 연구정보 이용에 대한 저작권 기준이 마련되어야 한다.

- 3) 임상의학 분야뿐만 아니라 다양한 주제 분야에서의 연구자 홈페이지 정보에 대한 연구가 수행되어야 한다.
- 4) 홈페이지에서 추출되는 정보에 대한 유용성을 검증하기 위하여 체계적인 이용자 연구가 수행되어야 한다.

이상의 선행조건이 만족되어 저자 홈페이지 연구정보의 유용성이 검증되고, 이를 디지털 도서관 장서개발에 적용할 경우 그 기대효과는 무엇보다도 현재 제공되는 디지털 도서관 장서의 양적, 질적 향상을 가져올 수 있다. 이 연구에서 비교 분석의 대상이 되었던 의학대학 디지털 도서관과 학술포털은 거의 대부분 동일한 연구정보를 제공하고 있었다. 다시 말하자면, 의학대학 디지털 도서관을 통하여 제공되는 정보가 학술포털에서 제공되는 정보와 차별성이 거의 없는 것으로 파악되었다. 이는 도서관 측면에서 학술포털이나 기타 웹 자원 시스템으로 서비스 대상 이용자의 유실을 의미한다. 인터넷이나 웹을 통하여 정보에 접근하는 연구자가 점점 늘고 있는데, 이는 네트워크를 통한 디지털 정보자원의 손쉬운 접근과 더불어 차별성 없는 도서관 장서에도 영향이 있을 것으로 판단된다. 따라서 현재 학술포털에서 제공하지 못하고 있는 연구정보를 디지털 도서관에서 제공할 수 있는 방안으로 웹 자원 중 하나인 연구자 홈페이지 정보를 활용할 필요성이 있음을 확인할 수 있었다.

5. 결론

도서관에서 장서개발은 물체에 비유하면 핵(core)이고 성분(content)을 결정짓는 매우 중요한 부분이다. 인터넷과 웹으로 대표되는 정보환경은 도서관으로 하여금 제공되는 디지털 정보자원의 양적 증가와 다양성, 질적 우수함을 요구하게 되었다. 학술포털이나 인터넷 검색엔진의 보편화 추세는 도서관에 위기 상황을 불러왔다. 그 결과 이를 극복할 방안을 모색해야 하는 것이 오늘날 도서관의 과제가 되고 있다.

도서관의 위기 극복과 제공되는 정보자원의 향상을 위하여 본 연구는 인용정보와 연구자 홈페이지에 있는 연구 성과물 정보를 활용해야 하는 필요성에 대하여 진단하고, 이를 디지털 도서관 장서개발에 적용할 때 고려되고 선행되어야 하는 조건들을 방안으로 제시하였다. 임상의학 분야로 한정하여 ESI에서 추출한 인용정보와 연구자 홈페이지에서 추출한 연구정보를 분석한 결과를 요약하면 다음과 같다.

첫째, 연구자 홈페이지는 학술지나 단행본과 같이 공식적으로 유통되는 정보자원에 비하여 학술대회 발표논문이나 강의노트와 같이 상대적으로 비공식 혹은 덜 공식적으로 유통되는 정보자원을 제공받을 수 있는 장소로 확인되었다.

둘째, 공식적으로 유통되는 학술지 논문이나 단행본의 경우도 연구자 홈페이지만을 통해서 접근할 수 있는 자료가 있었던 점을 고려하면, 제공되는 출판물의 포괄성도 우수한 것으로 나타났다.

셋째, 인용정보 중 피인용 횟수를 분석하여 저명한 연구자를 파악하고, 이들의 홈페이지에서 수집된 자료를 대상으로 하였으므로 제공되는 정보의 신뢰성 및 유용성이 있는 것으로

평가되었다.

넷째, 출판 중 혹은 연구진행 중인 자료가 포함되어 있는 점을 고려할 때 최신성 있는 자료의 제공이 가능한 것으로 판단되었다.

연구자 홈페이지 정보를 디지털 장서관리에 이용하기 위하여, 자동적인 연구데이터 추출 시스템의 개발이 필요하고 다양한 분야에서 그 필요성을 진단하는 연구와 더불어 홈페이지에서 추출된 연구정보의 유용성을 검증하는 후속연구가 수행되어야 할 것이다.

참고문헌

- 강속희. 1999a. 전자정보서비스가 대학도서관 장서관리에 미치는 영향. 『한국문헌정보학회 학술발표논문집 제 7집』, 1999년 4월.
- 강속희. 1999b. 정보기술이 대학도서관 장서관리에 미치는 영향과 전자정보서비스의 관계. 『한국문헌정보학회지』, 33(1): 23-48.
- 이란주. 2005. 도서관 장서관리정책 수립에 관한 연구. 『한국비블리아학회지』, 16(2): 203-224.
- 이지연, 김성연. 2003. 디지털 도서관의 웹자원 기반 장서관리. 『한국기록관리학회지』, 3(1): 111-128.
- Case, D. 2002. *Looking for Information: A Survey of Research on Information Seeking, Needs, and Behavior*. San Diego: Academic Press.
- Gessesse, K. 2000. "Collection development and management in the twenty-first century with special reference to academic libraries: an overview." *Library Management*, 21(7): 365-372.
- Joint, Nicholas. 2006. "Digital library futures: collection development or collection preservation?" *Library Review*, 55(5): 285-290.
- Kiondo, E. 2004. "Around the world to: The University of Dar es Salaam Library: collection development in the electronic information environment." *Library Hi Tech News*, 21(6): 19-24.
- Kovacs, D. and A. Elkordy. 2000. "Collection development in cyberspace: building an electronic library collection." *Library Hi Tech*, 18(4): 335-359.
- Queens Library. 2004. "Internet Policy." [cited 2007. 2. 13]. <http://www.queenslibrary.org/index.aspx?section_id=2&page_id=12>.
- Ramirez, D. and S. Gyeszly. 2001. "netLibrary: a new direction in collection development." *Collection Building*, 20(4): 154-164.
- Vignau, B. and I. Quesada. 2006. "Collection development in a digital environment: an imperative for information organizations in the twenty-first century." *Collection Building*, 25(4): 139-144.