

# 국내 학술지 논문에 인용된 인터넷 참고문헌의 특성 및 활동성 연구

- 국내 문헌정보학분야 4개 학술지를 중심으로 -

## A Study on Internet Reference Characteristics and Activity of Four Major Journals in Library & Information Science in Korea

김 규 환(Gyu-Hwan Kim)\*

### < 목 차 >

<p>I. 서론</p> <p>II. 연구 설계</p> <ol style="list-style-type: none"> <li>1. 분석 데이터</li> <li>2. 용어의 정의</li> <li>3. 분석내용</li> <li>4. 분석 프로세스</li> </ol> <p>III. 인터넷 참고문헌의 수/비율 및 변화추이</p> <ol style="list-style-type: none"> <li>1. 참고문헌 중 인터넷 참고문헌의 수/비율, 연도별 변화 추이</li> <li>2. 기사 1건당 평균 인터넷 참고문헌의 수, 연도별 변화 추이</li> <li>3. 인터넷 참고문헌을 1개 이상 포함한 기사 수/비율, 연도별 변화추이</li> </ol>	<ol style="list-style-type: none"> <li>4. 기사 1건당 인터넷 참고문헌의 평균 인용률 범위</li> </ol> <p>IV. 인터넷 참고문헌의 특성</p> <ol style="list-style-type: none"> <li>1. 도메인 유형</li> <li>2. 자료 유형</li> <li>3. 언어 유형</li> </ol> <p>V. 인터넷 참고문헌의 활동성 및 변화 추이</p> <ol style="list-style-type: none"> <li>1. 인터넷 참고문헌의 활동성</li> <li>2. 인용년도 이후 인터넷 참고문헌의 비활동성 비율의 변화추이</li> <li>3. 인터넷 참고문헌의 비활동성에 대한 오류 메시지</li> </ol> <p>VI. 결론 및 제언</p>
--	---

### 초 록

본 연구의 목적은 국내 학술지 논문에 인용된 인터넷 참고문헌의 특징과 활동성을 조사하는 것이다. 분석대상은 국내 문헌정보학 분야 4개 학술지(한국정보관리학회지, 한국도서관·정보학회지, 한국문헌정보학회지, 한국비블리아학회지)의 지난 8년(2002년부터 2009년) 동안 게재된 논문에 포함된 참고문헌들이다. 주요 분석결과를 정리하면 다음과 같다. 첫째, 참고문헌 중 인터넷 참고문헌의 비율은 15.5%이었으며 2002년 이후 인터넷 참고문헌의 비율은 매년 꾸준히 증가하였다. 둘째, 인터넷 참고문헌의 자료유형은 일반 웹문서(html, htm 등)가 84.3%로 가장 많았다. 셋째, 인터넷 참고문헌의 도메인 유형은 비영리기관이 35%, 대학/교육기관이 20%로 가장 많았다. 넷째, 인터넷 참고문헌의 비활동성은 44.8%로 높게 나타났고, 인용된 지 10년 후에는 최대 53.3%까지 증가하였다.

키워드: 인터넷 참고문헌, 인터넷 참고문헌 특성, 인터넷 참고문헌과 활동성

### ABSTRACT

This study examined the frequency, format, and activity of Internet references of four major journals in Library and Information Science in Korea. Sampled articles of each journal were published during 2002-2009. The results show five implications. 1) Overall 15.5% of all references were Internet references. 2) The number of Internet references has continuously increased since 2002. 3) The type of Internet references were most commonly "web document(84.3%), followed by" PDF(14%). 4) Nearly 55% of Internet reference were ".or(org)" and ".ac(edu)". 5) The percentage of inactive Internet references increased from 34.4% at 3 years to 44.1% at 5 years and to 53.3% at 10 years after publication. Although this study may not be representative of the entire field, it shows that Internet references occurred frequently and were often inaccessible within years after publication.

Keywords: Internet references, Characteristics of internet references, Activity of internet references

\* 한국과학기술정보연구원 국내정보팀 선임연구원(gyuhwan@kisti.re.kr)

• 접수일: 2012년 8월 28일 • 최초심사일: 2012년 9월 14일 • 최종심사일: 2012년 9월 21일

## I. 서론

오늘날 점점 더 많은 연구자들이 연구결과물을 보급하는데 인터넷을 사용하고 있으며 처음부터 인터넷으로 출판하는 경우도 점점 증가하고 있다. 새로운 방식으로 출판된 학술자료들은 웹사이트 등 여러 정보 출처에 링크되고, 때로는 전체 참고문헌 시스템에 링크되기도 한다. 또한 인터넷 기술 발달 덕분에 과거에는 서로 연계시켜 상상하기 힘들었던 많은 연구들의 관계를 시각화하여 보는 것도 가능해지게 되었다. 그 결과 학술논문이나 학술서적들이 인터넷 자료를 인용하는 비율이 점점 증가하고 있는데, 연구자들은 전통적 인쇄본 자료만큼이나 인터넷 자료도 다양하고 광범위하게 인용하고 참고하고 있다.<sup>1)</sup>

인터넷 자료를 포함한 모든 참고문헌은 이용자가 실제로 내용을 확인할 수 있을 때만 그 효과가 제대로 발휘될 수 있다. 그러나 웹 자료 중 상당 부분을 보존하기 위해 설립된 인터넷 아카이브(Internet Archive)의 조사결과에 따르면, 오늘날 웹상에 나타난 자료의 평균 수명이 겨우 100일 밖에 안 되는 것으로 조사되었다.<sup>2)</sup> Lawrence 등은 참고문헌으로 제시한 인터넷 자료가 사라지는 비율이 논문 출판 2년 후에는 23%, 7년 후에는 53%까지 증가한다는 연구결과를 제시하였다.<sup>3)</sup> 또한, Dellavalle 등은 영향력이 매우 큰 의학, 과학 학술지 중 일부를 선정하여 그 학술지에 실린 논문들이 어느 정도의 인터넷 자료를 참고했으며 그렇게 인용된 인터넷 참고자료의 평균 수명이 어느 정도 되는지 조사했다. 이 연구 조사 대상 학술지들에 실린 논문 중 30%가 적어도 1개 이상의 인터넷 자료를 참고문헌으로 인용하였고, 인용된 인터넷 자료가 사라지는 비율은 학술지가 출판된 지 3개월 후 3.8%에서 15개월 후에는 10%, 27개월 후에는 13%까지 증가하는 것으로 확인되었다<sup>4)</sup>. 한편, 인용하거나 참고했던 인터넷 자료들은 굳이 찾으려고 노력하면 찾아내는 것이 불가능하지 않은 것도 사실이다. 그러나 해당 자료를 찾고 보면 원래 사이트가 아닌 다른 사이트에 있는 경우가 대부분이며, 더 큰 문제는 그렇게 찾아낸 인터넷 자료가 원래의 그 자료인지를 확인할 길이 없고, 원래 자료라고 해도 내용이 바뀌지 않았다고 보장할 수 없다는 것이다.

본 연구는 이러한 모든 사실을 감안하여 국내 학술지 논문에 인용된 인터넷 참고문헌의 현황은 어떠한지를 조사하고자 하였다. 이를 위해 국내 문헌정보학분야 학술지에 인용된 참고문헌을 연구

1) Clifford A. Lynch, *Preserving Digital Information to Support Scholarship, In the Internet & the University*, 2003, <<http://www.educause.edu/ir/library/pdf/ffpiu029.pdf>> [Cited 2012, 6. 13].

2) Rick Weiss, *On the Web, Research Work Proves Ephemeral : Electronic Archivists Are Playing Catch-up in Trying to Keep Documents From in History's Dustbin*, 2003, <[http://stevereads.com/cache/ephemeral\\_web\\_pages.html](http://stevereads.com/cache/ephemeral_web_pages.html)> [Cited 2012, 6. 13].

3) Steve Lawrence et al., "Persistent of Web References in Scientific Research," *IEEE Computers*, Vol.34, No.2(2001), pp.26-31.

4) Robert Dellavalle et al., "Going, Going, Gone : Lost Internet References," *Science*, Vol.302(October 2003), pp.787-788.

대상으로 인터넷 참고문헌의 비율과 특성, 그리고 활동성을 조사하였다. 국내 문헌정보학분야 학술지만을 연구대상으로 한 것이 본 연구의 제한점일 수 있으나, 현재까지 국내 학술지 논문에 인용된 인터넷 참고문헌의 현황을 조사한 연구가 없었다는 점에서 본 연구의 의의가 있다고 하겠다. 향후 국내 학술지 논문에 인용된 인터넷 참고문헌의 손실을 최소화할 수 있는 다양한 방안들을 마련하는데 본 연구결과가 기초자료로 활용되었으면 한다.

## II. 연구 설계

KISTI는 매년 국내 학술지를 수집하고 이에 수록된 논문의 서지사항과 해당 논문에 인용된 참고문헌의 서지사항을 조직한 KSCI(Korea Science Citation Index) 데이터베이스를 제작하여 서비스하고 있다. 이에 본 연구에서는 국내 문헌정보학분야 학술지 논문에 인용된 인터넷 참고문헌의 비율과 특성, 그리고 활동성을 조사하기 위해 KISTI의 KSCI 데이터베이스를 활용하였다.

### 1. 분석 데이터

분석데이터는 KSCI DB에 구축된 2002년부터 2009년까지 국내 문헌정보학분야 핵심 학술지 4종에 포함된 1,666건의 기사와 참고문헌 42,032건을 대상으로 하였다. 2012년 6월 현재 기준으로 KSCI 데이터베이스의 구축사업의 진행상황을 보면, 국내 문헌정보학분야 학술지의 2010년부터 2011년까지의 참고문헌들은 아직 구축 중에 있는 관계로 본 연구에서는 2002년부터 2009년까지의 참고문헌만을 분석대상으로 제한하였다.

〈표 1〉 분석데이터

학회명	학술지명	기간	기사수	참고문헌수
한국정보관리학회	한국정보관리학회지	2002-2009	464	11,056
한국도서관·정보학회	한국도서관·정보학회지	2002-2009	488	12,882
한국문헌정보학회	한국문헌정보학회지	2002-2009	482	12,698
한국비블리아학회	한국비블리아학회지	2002-2009	232	5,396
합계			1,666	42,032

### 2. 용어의 정의

가. 참고문헌(References)은 학술지 논문의 말미에 게재된 모든 문헌 리스트를 의미한다. 단, 한국도서관·정보학회의 경우와 같이 참고문헌을 각주로 대신하는 경우는 이들을 참고문헌에 포

#### 4 한국도서관·정보학회지(제43권 제3호)

합하였다.

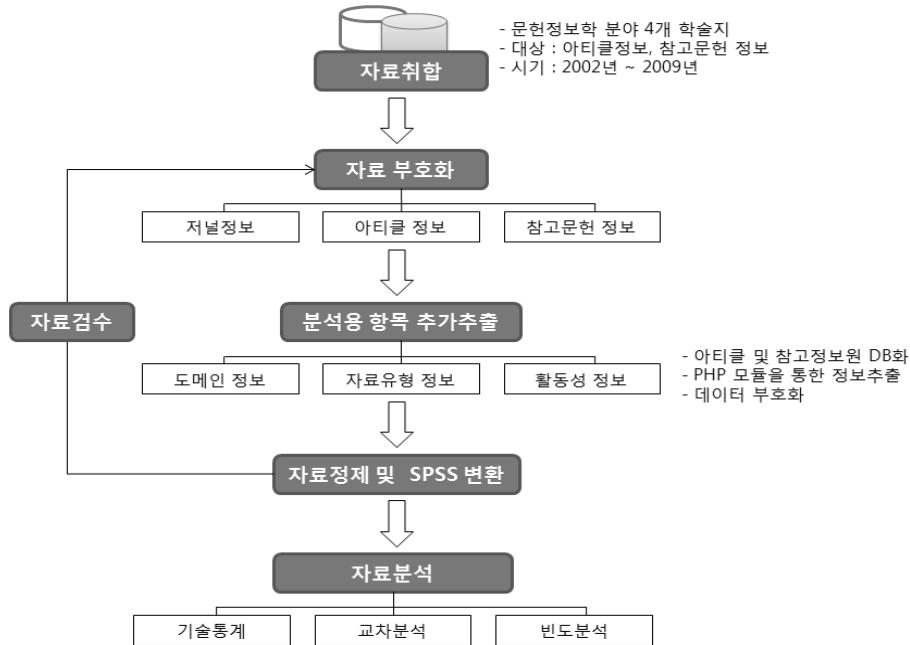
- 나. 인터넷 참고문헌(Internet References)은 참고문헌 중 원형정보에 URL 정보를 포함하고 있는 모든 참고문헌 리스트를 의미한다.
- 다. 활성 인터넷 참고문헌(Active Internet References)은 웹 브라우저로 접근할 때 에러 메시지가 없이 정보가 보이는 URL을 가진 모든 인터넷 참고문헌을 의미한다.
- 라. 비활성 인터넷 참고문헌(Inactive Internet References)은 웹 브라우저로 접근할 때 에러 메시지가 나타나는 URL을 가진 모든 인터넷 참고문헌을 의미한다.
- 마. 활성 인터넷 참고문헌일 경우 활동성(Activity)이 있다고 하였고, 비활성 인터넷 참고문헌일 경우 비활동성(Inactivity)이 있다고 하였다.

### 3. 분석내용

본 연구에서는 다음과 같은 세 가지 측면으로 분석내용을 조사하였다. 첫 번째는 인터넷 참고문헌의 특성을 알아보기 위해 인터넷 참고문헌의 언어 유형, 도메인 유형, 자료 유형을 조사하였다. 두 번째는 인터넷 참고문헌의 수/비율과 연도별 변화추이를 알아보기 위해 인터넷 참고문헌의 수/비율, 인터넷 참고문헌 수의 연도별 변화추이, 기사 1건당 평균 인터넷 참고문헌의 수, 기사 1건당 평균 인터넷 참고문헌의 수의 연도별 변화추이, 인터넷 참고문헌을 1개 이상 포함한 기사 수/비율, 인터넷 참고문헌을 1개 이상 포함한 기사 수의 연도별 변화추이, 기사 1건당 인터넷 참고문헌의 평균 인용확률 범위를 조사하였다. 셋째는 인터넷 참고문헌의 활동성을 알아보기 위해 인터넷 참고문헌의 비활동성 비율, 도메인 유형별/자료 유형별 인터넷 참고문헌의 비활동성 비율, 인용년도 이후 비활동 인터넷 참고문헌 비율의 변화추이, 비활동 인터넷 참고문헌의 오류 메시지 유형, 도메인 유형별/언어 유형별 비활동 인터넷 참고문헌의 오류 메시지 유형을 조사하였다.

### 4. 분석 프로세스

본 연구에서 분석 프로세스는 <그림 1>과 같다. 단계별로 보면 원데이터 추출 및 데이터 정제, 인터넷 참고문헌의 분석항목 도출 및 자료 부호화, 최종 통계처리의 3단계로 진행하였다.



〈그림 1〉 분석 프로세스

가. 원데이터 추출 및 데이터 정제

KSCI 데이터베이스로부터 국내 문헌정보학분야의 핵심 4개 학술지에 인용된 참고문헌의 1) 저널 정보, 2) 기사정보, 3) 참고문헌 원형정보, 4) 언어정보, 5) 출판년도, 6) 인터넷 참고문헌 URL 정보를 포함한 6개 필드의 데이터값을 추출하여 엑셀파일에 저장하였다. 추출된 6개 필드의 데이터 값에 대한 정제 작업을 실시하였다. 특히 본 연구에서 중요한 정보인 인터넷 참고문헌 URL에 대한 저자들의 기재 오류(http://, hwp://, htt://, http://, http : //, htt :/// 등)가 많이 발견되었다. 이에 URL 해독의 기계처리를 위해 저자들의 URL에 대한 기재 오류들을 찾아서 일일이 정제하여 반영하였다.

6 한국도서관·정보학회지 (제43권 제3호)

A	B	C	D	E	F	G	H	M
Journal name	Journal N	Journal PubYear_Vol_Issue_Ai	Article number	DataTyp	Lang	Citation_YEA	URL(Internet Reference)	RawData
한국정보관리학회	JBGRBQ	JBGRBQ_1995_v12n1_45_002	JAKO199511920383018	o	eng	1995	http://www.halsoft.com/html-tk	
한국정보관리학회	JBGRBQ	JBGRBQ_1995_v12n1_45_005	JAKO199511920383018	o	eng	1995	http://info.cern.ch/hypertext/www/Tools/Word-proc/filters.html	
한국정보관리학회	JBGRBQ	JBGRBQ_1995_v12n1_45_001	JAKO199511920383018	o	eng	1995	http://cs.dal.ca/htmlasst/htmlfaq.txt	
한국정보관리학회	JBGRBQ	JBGRBQ_1995_v12n1_45_003	JAKO199511920383018	o	eng	1995	ftp://ftp.cs.dal.ca/htmlasst/usrtools.exe	
한국정보관리학회	JBGRBQ	JBGRBQ_1995_v12n1_45_004	JAKO199511920383018	o	eng	1995	ftp://ftp.cuhk.hk/pub/www/windows/util/cu_html.zip	
한국정보관리학회	JBGRBQ	JBGRBQ_1995_v12n2_171_036	JAKO199511920383134	o	eng	1995	http://hoohoo.ncsa.uiuc.edu/cgi/primer.html	
한국정보관리학회	JBGRBQ	JBGRBQ_1995_v12n2_171_016	JAKO199511920383134	o	eng	1995	http://www.cc.gatech.edu/gvu/stats/NSF/Extrap.GIF	
한국정보관리학회	JBGRBQ	JBGRBQ_1995_v12n2_171_031	JAKO199511920383134	o	eng	1995	http://www.oac.ucl.edu/indiv/MIME/MIME.html	
한국정보관리학회	JBGRBQ	JBGRBQ_1996_v13n2_39_012	JAKO199611920383295	o	eng	1996	http://cs.wpi.edu/~nemleem/mpegcontents.html	
한국정보관리학회	JBGRBQ	JBGRBQ_1996_v13n2_39_016	JAKO199611920383295	o	eng	1996	http://www.mpl.jp/nasa.gov/~ndr/tiff/#w	
한국정보관리학회	JBGRBQ	JBGRBQ_1996_v13n2_39_001	JAKO199611920383295	b	eng	1996	http://www.arbortext.com/wp.html	
한국정보관리학회	JBGRBQ	JBGRBQ_1996_v13n2_189_038	JAKO199611920383345	o	eng	1996	http://seamless.com/rcf/article.html	
한국정보관리학회	JBGRBQ	JBGRBQ_1996_v13n2_189_045	JAKO199611920383345	o	eng	1996	http://www.hotwired.com/wired/whitepaper.html	
한국정보관리학회	JBGRBQ	JBGRBQ_1997_v14n1_47_036	JAKO199711920383383	o	eng	1997	http://www.nlc-bnc.ca/rla/documents/libraries/resource-sharing/nil-protocol/z-ill-1a.doc	
한국정보관리학회	JBGRBQ	JBGRBQ_1997_v14n1_47_041	JAKO199711920383383	o	eng	1997	http://www.nlc-bnc.ca/rla/documents/libraries/resource-sharing/nil-protocol/z-ill-2a.doc	
한국정보관리학회	JBGRBQ	JBGRBQ_1997_v14n1_47_039	JAKO199711920383383	o	eng	1997	http://lcweb.loc.gov/z3950/agency/profiles.html	
한국정보관리학회	JBGRBQ	JBGRBQ_1997_v14n1_47_020	JAKO199711920383383	o	eng	1997	http://lcweb.loc.gov/z3950/lcserver.html	
한국정보관리학회	JBGRBQ	JBGRBQ_1997_v14n1_47_008	JAKO199711920383383	o	eng	1997	http://europagate.dlv.dk/summary.htm	
한국정보관리학회	JBGRBQ	JBGRBQ_1997_v14n1_47_038	JAKO199711920383383	o	eng	1997	http://www.dstc.edu.au/RDU/reports/zclients.html	
한국정보관리학회	JBGRBQ	JBGRBQ_1997_v14n1_47_032	JAKO199711920383383	o	eng	1997	http://www.vtls.com/gateway.htm	
한국정보관리학회	JBGRBQ	JBGRBQ_1997_v14n1_47_011	JAKO199711920383383	o	eng	1997	http://www.evos.be/lib/gsr.htm	
한국정보관리학회	JBGRBQ	JBGRBQ_1997_v14n1_47_010	JAKO199711920383383	o	eng	1997	http://www.geac.com/products/librany/	
한국정보관리학회	JBGRBQ	JBGRBQ_1997_v14n1_47_006	JAKO199711920383383	o	eng	1997	http://lcweb.loc.gov/z3950/agency/brussels/4br.html	
한국정보관리학회	JBGRBQ	JBGRBQ_1997_v14n1_47_027	JAKO199711920383383	o	eng	1997	http://www.bibsys.no/one-report	
한국정보관리학회	JBGRBQ	JBGRBQ_1997_v14n1_47_037	JAKO199711920383383	o	eng	1997	http://www.vtls.com/zclient.htm	
한국정보관리학회	JBGRBQ	JBGRBQ_1997_v14n1_47_013	JAKO199711920383383	o	eng	1997	http://www.iii.com/screens/opacintro.html	
한국정보관리학회	JBGRBQ	JBGRBQ_1997_v14n1_47_018	JAKO199711920383383	o	eng	1997	http://www.nlm.nih.gov/publications/staff_publications/rogers/z3950/z3950.html	
한국정보관리학회	JBGRBQ	JBGRBQ_1997_v14n1_47_031	JAKO199711920383383	o	eng	1997	http://vinca.cndr.org/software/site/guide.html	

〈그림 2〉 분석필드와 데이터값

통계분석을 위해 텍스트 중심의 자료를 부호화하였다. 세부적으로 저널 정보 부호화(저널유형, 저널ID 등), 기사 정보 부호화(기사ID, 기사출판년도 등), 참고문헌 정보 부호화(참고문헌 ID, 언어, 연도, URL 등), 언어정보 부호화(영어, 프랑스어 등), 출판년도 및 인용년도 부호화를 수행하였다.

나. 인터넷 참고문헌의 분석항목 도출 및 자료 부호화

인터넷 참고문헌의 특성과 비활동성을 조사하기 위한 분석항목은 1) 도메인 유형, 2) 자료유형, 3) 활동성 유무로 구분하였다. 모든 분석항목은 PHP 모듈을 개발하여 인터넷 참고문헌의 URL을 해독하여 추출하였다. 첫째, 도메인 유형은 인터넷 참고문헌의 URL 정보에서 최상위 도메인(1단계 도메인)과 한국 도메인의 경우는 지역도메인을 제외한 2단계 도메인으로부터 추출하였다. 추출된 도메인 유형 및 부호화 작업결과는 〈표 2〉와 같다.

〈표 2〉 도메인 유형(1단계/2단계) 및 부호화

도메인 유형	부호	1단계 도메인	2단계 도메인
국제적인 기구/기관	1	int	
미연방군사기관	2	mil	
네트워크 관리기관	3	net	ne
대학기관	4	edu	ac
비영리기관	5	org	or
연구기관	6	-	re
영리기관 및 기업	7	com	co
정부기관	8	gov	go
기타	9	-	-

둘째, 자료 유형은 인터넷 참고문헌의 URL 정보에서 파일확장자를 기반으로 추출하였다. jsp, asp, php의 웹언어 또는 별도의 세부페이지 유형이 존재하지 않는 경우에는 일반 웹문서로 간주하여 부호화하였다. 추출된 자료 유형 및 부호화 작업결과는 <표 3>과 같다.

<표 3> 자료 유형 및 파일확장자

자료 유형	부호	파일확장자
텍스트	1	txt, dat
MS 엑셀	2	xls, xlsx
MS 파워포인트	3	ppt, pptx
MS 워드	4	doc, docx
PDF	5	pdf
한글	6	hwp
일반웹문서	7	html, htm, jsp, asp, php 등

셋째, 활동성 유무는 인터넷 참고문헌의 URL이 현재 정상적으로 접속되는지 또한 접속이 되지 않을 경우 원인이 무엇인지를 확인하는 과정이다. 다량의 인터넷 참고문헌을 수작업으로 확인할 경우 인력 및 시간적 비용이 많이 소요될 수 있다. 이에 본 연구에서는 인터넷 참고문헌의 URL 정보를 데이터베이스화하여 시스템을 통해 웹사이트 활동성 정보를 자동적으로 체크할 수 있는 모듈을 개발하여 운용하였다. 시스템 환경은 웹서버(Apache), DB(Mysql), 프로그램 언어(PHP)를 기반으로 하였고 추출 프로세스는 다음과 같다.

첫째, 인터넷 참고문헌의 고유ID와 URL값을 데이터베이스에 입력하였다.

둘째, 인터넷 참고문헌의 URL을 PHP 언어로 호출하여, cURL 모듈을 통해 각 인터넷 참고문헌의 URL 접속 후 HTTP CODE를 회신받아 출력하였다.

셋째, HTTP CODE(그림 3 참조)를 해석하였다.

넷째, 출력된 결과를 부호화하여 통계분석용 자료로 변환하였다.

아이디	HTTP CODE	URL
1	404	http://www.halsoft.com/html-tk
2	404	http://info.cern.ch/hypertext/www/Tools/Word-proc.filters.h
3	0	ftp://cs.dal.ca/htmlasst/htmlafaq.bt
4	0	ftp://ftp.cs.dal.ca/htmlasst/usrttools.exe
5	550	ftp://ftp.cuhk.hk/pub/www/windows/util/cu_html.zip
6	0	http://hoohoo.ncsa.uiuc.edu/cgi/primer.html
7	200	http://www.cc.gatech.edu/gvu/stats/NSF/Extrap.GIF
8	404	http://www.oac.uci.edu/indiv/MIME/MIME.html
9	0	http://cs.wpi.edu/~nemleem/mpegcontents.html
10	0	http://www.mpl.jpl.nasa.gov/~ndr/tiff/#w
11	0	http://www.arbortext.um/wp.html
12	0	http://seamless.com/rci/article.html
13	200	http://www.hotwired.com/wired/whitepaper.html
14	404	http://www.nic-bnc.ca/ifa/documents/libraries/resource-s
15	404	http://www.nic-bnc.ca/ifa/documents/libraries/resource-s
16	200	http://icweb.loc.gov/z3950/agency/profiles.html
17	404	http://icweb.loc.gov/z39.50/icserver.html
18	0	http://europagate.dtv.dk/summary.htm
19	0	http://www.dstc.edu.au/RDU/reports/zclients.html
20	404	http://www.vtis.com/gateway/htm
21	200	http://www.ewos.be/lib/gsr.htm

<그림 3> 인터넷 참고문헌의 고유ID와 회신 받은 HTTP CODE 값

다. 통계처리

SPSS 프로그램(SPSS Korea 18 version)을 이용한 통계 처리를 수행하였다. 사용된 통계분석 기법은 기술통계 분석, 교차분석, 빈도분석이다.

### Ⅲ. 인터넷 참고문헌의 수/비율 및 변화 추이

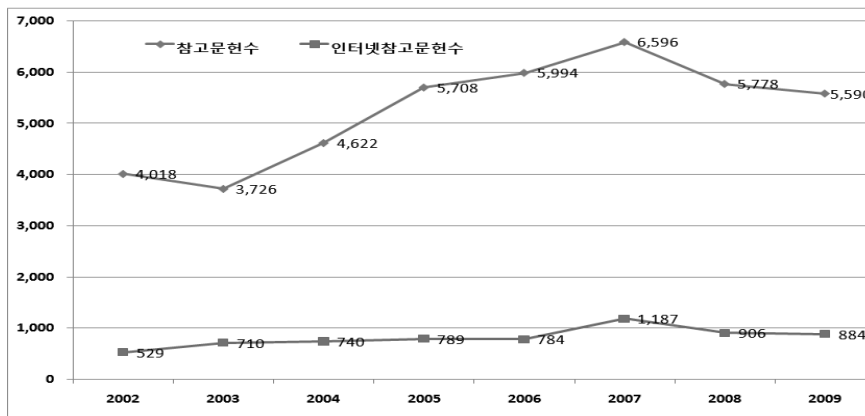
1. 참고문헌 중 인터넷 참고문헌의 수/비율, 연도별 변화 추이

<표 4>와 같이 국내 문헌정보학분야의 기사 수는 총 1,666건이며 여기에 인용된 참고문헌 수는 42,032건이었다. 인용된 참고문헌 중 인터넷 참고문헌은 15.5%(6,529건)인 것으로 나타났다. 그리고 대부분의 학술지에서 참고문헌 대비 인터넷 참고문헌 수의 비율이 15%에서 17%로 유사하게 나타났다.

<표 4> 참고문헌 중 인터넷 참고문헌의 수/비율

학술지명	기사 수	참고문헌 수	인터넷 참고문헌 수	인터넷 참고문헌 비율 (%)
한국정보관리학회지	464	11,056	1,697	15.3%
한국도서관·정보학회지	488	12,882	1,991	15.5%
한국문헌정보학회지	482	12,698	1,913	15.1%
한국비블리아학회지	232	5,396	928	17.2%
전체	1,666	42,032	6,529	15.5%

기사에 인용된 참고문헌 수의 연도별 변화추이(<그림 4>)를 보면 참고문헌 수는 2002년에 4,018건을 시작으로 급격히 그 수가 증가하였고 2007년에 최대 6,596건까지 증가하였다. 반면 기사에 인



<그림 4> 참고문헌 및 인터넷 참고문헌 수의 연도별 변화추이



용된 인터넷 참고문헌 수의 연도별 변화추이를 보면 인터넷 참고문헌은 참고문헌의 급격한 변동과 상관없이 2002년에 529건에서 2009년에 884건까지 완만히 증가하는 모습을 보였다. 연도별로 보면, 2002년에 인터넷 참고문헌이 529건에서 2008년 710건으로 34%정도 증가하였고, 그 이후 2006년까지 710건에서 789건까지 소폭 상승세를 보였다. 특히 2007년에 51%의 높은 증가율(1,187건)을 보인 이후, 2008년과 2009년에는 평균 900건 내외를 유지하였다.

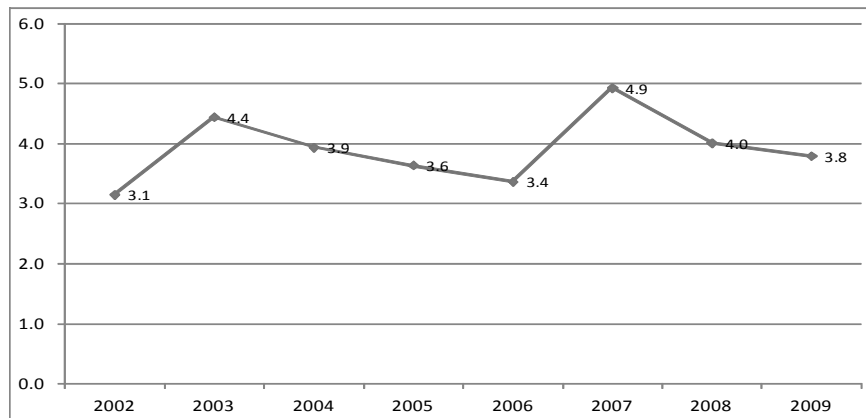
## 2. 기사 1건당 평균 인터넷 참고문헌의 수, 연도별 변화 추이

<표 5>와 같이 기사 1건당 평균 참고문헌 수는 25.2건이었고, 기사 1건당 평균 인터넷 참고문헌의 수는 3.9건인 것으로 나타났다.

<표 5> 기사 1건당 평균 인터넷 참고문헌 수

기사 수	참고문헌 수	인터넷 참고문헌 수	기사 1건당 평균 참고문헌 수	기사 1건당 평균 인터넷 참고문헌 수
1,666	42,032	6,529	25.2	3.9

기사 1건당 평균 인터넷 참고문헌수의 연도별 변화추이(<그림 5>)를 보면 전체 기사 1건 당 평균 인터넷 참고문헌 수는 2002년부터 2009년까지 일정한 비율은 아니지만, 전반적으로 증가하는 모습을 보였다. 기사 1건당 평균 인터넷 참고문헌의 증가 현상은 다음에 제시될 인터넷 참고문헌을 1개 이상 포함한 기사 수의 연도별 변화추이에서 보다 뚜렷하게 나타났다.



<그림 5> 기사 1건당 평균 인터넷 참고문헌 수의 연도별 변화추이

3. 인터넷 참고문헌을 1개 이상 포함한 기사 수/비율, 연도별 변화추이

〈표 6〉과 같이 기사 1,666건 중 인터넷 참고문헌을 1개 이상 포함하고 있는 기사의 수는 1,022건으로 전체 기사 중에서 61.3%로 높게 나타났다.

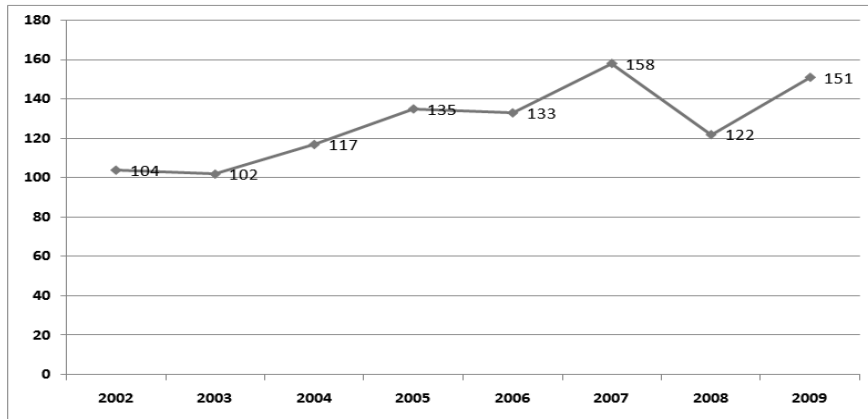
〈표 6〉 인터넷 참고문헌을 1개 이상 포함한 기사 수/비율

기사 수	인터넷 참고문헌을 1개 이상 포함한 기사 수	인터넷 참고문헌을 1개 이상 포함한 기사 비율(%)
1,666	1,022	61.3%

기사 1,666건 중 인터넷 참고문헌을 2개에서 10개 이상 포함한 기사의 수와 비율을 순차적으로 제시하면 아래와 같다.

- 인터넷 참고문헌 2개 이상 포함한 기사 수 : 806개(48%)
- 인터넷 참고문헌 3개 이상 포함한 기사 수 : 659개(40%)
- 인터넷 참고문헌 4개 이상 포함한 기사 수 : 534개(32%)
- 인터넷 참고문헌 5개 이상 포함한 기사 수 : 448개(27%)
- 인터넷 참고문헌 6개 이상 포함한 기사 수 : 393개(24%)
- 인터넷 참고문헌 7개 이상 포함한 기사 수 : 338개(20%)
- 인터넷 참고문헌 8개 이상 포함한 기사 수 : 280개(17%)
- 인터넷 참고문헌 9개 이상 포함한 기사 수 : 244개(14%)
- 인터넷 참고문헌 10개 이상 포함한 기사 수 : 203개(12%)

다음으로 인터넷 참고문헌을 1개 이상 포함한 기사 수의 연도별 변화추이(〈그림 6〉)를 보면, 인터넷 참고문헌을 1개 이상 포함한 기사 수는 매년 증가하였다. 연도별로 보면, 인터넷 참고문헌을 1개 이상 포함한 기사 수는 2002년 104건을 시작으로 2005년에 135건까지 꾸준히 증가하였고, 2007년에 최대 158건까지 증가하였다.



〈그림 6〉 인터넷 참고문헌을 1개 이상 포함한 기사 수의 연도별 변화추이

#### 4. 기사 1건당 인터넷 참고문헌의 평균 인용률 범위

연구자가 기사 1건당 인터넷 참고문헌의 인용할 평균 범위를 조사하였다. 조사내용은 연구자가 기사를 작성할 경우 일반적으로 인터넷 참고문헌을 인용하는 평균 비율과 인터넷 참고문헌을 1건 이상 인용했다는 조건하에 인터넷 참고문헌을 인용할 평균 비율을 조사하였다.

〈표 7〉 기사 1건당 인터넷 참고문헌의 평균 인용확률 범위

기사 1건당 평균 참고문헌 수	기사 1건당 평균 인터넷 참고문헌 수	인터넷 참고문헌을 1개 이상 포함한 기사 1건당 평균 인터넷 참고문헌 수	기사 1건당 인터넷 참고문헌 평균 인용 확률 범위
25.2	3.9	6.4	15.5%~25.4%

〈표 7〉와 같이 기사 1건당 평균 참고문헌 수는 25.2건이고, 기사 1건당 평균 인터넷 참고문헌의 수는 3.9건이다. 이는 참고문헌을 1개 이상 포함한 기사의 경우, 인용한 인터넷 참고문헌의 비율이 평균 15.5%라는 것을 의미한다. 여기서 인터넷 참고문헌을 1개 이상 포함한 기사만으로 조사범위를 좁혀서 보면, 인터넷 참고문헌을 1개 이상 포함한 기사는 1건당 평균 6.4건의 인터넷 참고문헌을 포함한다. 이는 인용한 인터넷 참고문헌의 비율이 평균 25.4%라는 것을 의미한다. 이는 어떤 연구자가 기사를 작성할 경우, 기사 1건당 평균적으로 인터넷 참고문헌을 인용할 확률이 평균 15.5%에서 25.4%라는 것을 의미한다. 특히 연구자가 기사를 작성할 때 참고문헌 중 인터넷 참고문헌을 1개라도 인용하는 경우라면 인터넷 참고문헌을 인용할 확률이 25.4%로 더 높아진다는 것을 의미한다.

## IV. 인터넷 참고문헌의 특성

### 1. 도메인 유형

도메인 유형(<표 8>)은 인터넷 참고문헌이 기재된 URL 정보에서 최상위 도메인(1단계 도메인)과 한국 도메인의 경우는 지역도메인을 제외한 2단계 도메인으로부터 추출하였다. 인터넷 참고문헌의 도메인 유형을 보면, 비영리기관(or, org)에서 발행한 인터넷 자료가 35%로 가장 많이 인용되었다. 다음으로 대학/교육기관(ac, edu)의 인터넷 자료가 20%, 영리기관의 인터넷 자료가 14%, 정부기관(go, gov)의 인터넷 자료가 13.7%순으로 인용되었다.

<표 8> 도메인 유형

구분	빈도	비율(%)
비영리기관(or, org)	2,288	35%
대학/교육기관(ac, edu)	1,308	20%
정부기관(go, gov)	896	13.7%
영리기관(co, com)	913	14%
네트워크 관리기관(ne, net)	237	3.6%
연구기관(re)	72	1.1%
국제적인 기구/기관(int)	25	0.4%
미연방군사기관(mil)	8	0.1%
기타	782	12%
전체	6,529	100%

### 2. 자료 유형

자료 유형(<표 9>)은 인터넷 참고문헌의 URL 정보에서 파일확장자를 기반으로 추출하였고, jsp, asp, php의 웹언어 또는 별도의 세부페이지 유형이 존재하지 않는 경우에는 일반 웹문서로 간주하였다. 인터넷 참고문헌의 자료 유형을 보면, 일반웹문서가 전체의 84.3%로 압도적으로 많이 인용되었다. 다음으로 PDF가 14%로 인용되었다. 그 외 DOC, HWP, PPT, TXT 문서는 거의 인용되지 않았다.

〈표 9〉 자료 유형

구 분	빈도	비율(%)
일반웹문서	5,501	84.3%
PDF	916	14%
DOC	51	0.8%
HWP	13	0.2%
PPT	43	0.7%
TXT	5	0.1%
전체	6,529	100%

### 3. 언어 유형

언어 유형(〈표 10〉)은 참고문헌 원형정보에 기재된 언어를 기준으로 하였다. 예를 들어, 참고문헌의 원형정보의 논문 제목이 'L'Exception au droit d'auteur en faveur des personnes handicapees'일 경우 프랑스어 자료로 판단하였다. 그러나 언어 유형이 해당 언어를 사용하는 국가와 반드시 일치하는 것은 아니다. 인터넷 참고문헌에 사용된 언어 유형을 보면 영어로 표기된 인터넷 자료가 81.8%로 가장 많이 인용되었고 다음으로 한국어 자료가 16.5%로 인용되었다. 그 외 일본어 자료나 중국어 자료, 프랑스어 자료는 거의 인용되지 않았다.

〈표 10〉 언어 유형

구 분	빈도	비율(%)
영어	5,345	81.8%
한국어	1,072	16.5%
일본어	96	1.5%
기타(중국어, 프랑스어 등)	16	0.2%
전체	6,529	100%

언어별 도메인 유형(〈표 11〉)을 보면, 영어 자료는 비영리기관에서 생산된 인터넷 자료가 36.9%로 가장 많이 인용되었고 다음으로 대학에서 생산된 인터넷 자료가 21.8%로 인용되었다. 일본어 자료는 정부기관에서 생산된 인터넷 자료(37.5%)와 대학에서 생산된 인터넷 자료(36.5%)가 비슷한 비율로 가장 인용되었다. 한국어 자료의 경우는 전반적으로 도메인 유형이 고른 분포를 보였으며, 그 중에서 비영리기관(28.5%)과 정부기관(21.2%)에서 생산된 인터넷 자료가 가장 많이 인용되었다. 기타 언어(중국어, 프랑스어 등)의 경우는 대부분 정부기관에서 생산된 인터넷 자료(37.5%)가 많이 인용되었다.

〈표 11〉 언어별 도메인 유형

구 분	국제기구/기관	미연방군사기관	네트워크관리기관	대학/교육기관	비영리기관	연구기관	영리기관 및기업	정부기관	기타
영어	24 (0.4%)	8 (0.1%)	116 (2.2%)	1,164 (21.8%)	1,970 (36.9%)	18 (0.3%)	730 (13.7%)	627 (11.7%)	688 (12.9%)
일본어	0 (0%)	0 (0%)	1 (1%)	35 (36.5%)	10 (10.4%)	0 (0%)	0 (0%)	36 (37.5%)	14 (14.6%)
한국어	1 (0.1%)	0 (0%)	120 (11.2%)	108 (10.1%)	306 (28.5%)	54 (5%)	183 (17.1%)	227 (21.2%)	73 (6.8%)
기타	0 (0%)	0 (0%)	0 (0%)	1 (6.3%)	2 (12.5%)	0 (0%)	0 (0%)	6 (37.5%)	7 (43.8%)
전체	25 (0.4%)	8 (0.1%)	237 (3.6%)	1,308 (20%)	2,288 (35%)	72 (1.1%)	913 (14%)	896 (13.7%)	782 (12%)

언어별 자료 유형(〈표 12〉)을 보면, 영어 자료는 대부분 82.2%가 일반웹문서이고 16%가 PDF로 나타났다. 일본어 자료는 일반웹문서가 66.7%로 가장 많았고 다음으로 PDF가 32.3%를 차지하였다. 한국어 자료는 일반웹문서가 96%로 압도적으로 많았다.

〈표 12〉 언어별 자료 유형

구 분	일반웹문서	PDF	DOC	HWP	PPT	TXT
영어	4,392 (82.2%)	855 (16%)	51 (1%)	1 (0%)	41 (0.8%)	5 (0.1%)
일본어	64 (66.7%)	31 (32.3%)	0 (0%)	0 (0%)	1 (1%)	0 (0%)
한국어	1,029 (96%)	30 (2.8%)	0 (0%)	12 (1.1%)	1 (0.1%)	0 (0%)
기타	16 (100%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
전체	5,501 (84.3%)	916 (14%)	51 (0.8%)	13 (0.2%)	43 (0.7%)	5 (0.1%)

## V. 인터넷 참고문헌의 활동성 및 변화추이

### 1. 인터넷 참고문헌의 활동성

2012년 6월 13일을 조사 기준으로 하여 인터넷 참고문헌의 활동성(〈표 13〉)을 조사한 결과, 인터넷 참고문헌 6,529건 중 44.8%(2,922건)의 인터넷 참고문헌이 비활동성을 보이는 것으로 나타났다. 다시 말하면 현재 시점에서 국내 문헌정보학분야 학술지 논문에 인용된 인터넷 참고문헌의 상당수가 더 이상 접근하여 이용할 수 없는 자료가 되었다는 것을 의미한다. 참고문헌은 이용자가 실제로 내용을 확인할 수 있을 때만 그 효과가 제대로 발휘될 수 있다. 그러한 점에서 이미 비활성화된 인터넷 참고문헌을 포함한 기사를 찾아 읽는 연구자들은 해당 논문의 논리적 근거로 사용된 인터넷 참고문헌을 신뢰할 수 없고, 결국 논문 전체에 대한 신뢰성을 확보할 수 없게 된다. 따라서 인터넷 참고

문헌의 손실을 최소화할 수 있는 효율적인 인터넷 참고문헌 관리방안 마련이 필요하다고 하겠다.

<표 13> 인터넷 참고문헌의 활동성

구분	빈도	비율(%)
인터넷 참고문헌의 활동성	3,607	55.2%
인터넷 참고문헌의 비활동성	2,922	44.8%
전체	6,529	100%

학술지별로 인터넷 참고문헌의 활동성(<표 14>)을 보면, 대부분의 학술지들에서 인터넷 참고문헌의 비활동성은 40%에서 50%사이로 매우 높게 나타났다.

<표 14> 학술지별 인터넷 참고문헌의 활동성

학술지명	인터넷 참고문헌 활동성		전체
	비활동	활동	
한국정보관리학회지	710 (41.8%)	987 (58.2%)	1,697 (100%)
한국도서관·정보학회지	977 (49.1%)	1,014 (50.9%)	1,991 (100%)
한국문헌정보학회지	833 (43.5%)	1,080 (56.9%)	1,913 (100%)
한국비블리아학회지	402 (43.3%)	526 (56.7%)	928 (100%)
전체	2,922 (44.8%)	3,607 (55.2%)	6,529 (100%)

도메인 유형별 인터넷 참고문헌의 활동성(<표 15>)을 보면, 영리기관(co, com)에서 생산된 인터넷 참고문헌의 비활동성이 36.8%로 가장 낮은 것으로 나타났다. 반면 연구기관(re), 국제기구/기관(int), 대학/교육기관(ac, edu), 정부기관(go, gov)에서 생산된 인터넷 참고문헌의 비활동성은 41%에서 최대 62.5%로 영리기관(co, com)에 비해 상대적으로 높게 나타났다.

<표 15> 도메인 유형별 인터넷 참고문헌의 활동성

도메인 유형	인터넷 참고문헌 활동성		전체
	비활동	활동	
미연방군사기관(mil)	5 (62.5%)	3 (37.5%)	8 (100%)
연구기관(re)	35 (48.6%)	37 (51.4%)	72 (100%)
국제적인 기구/기관(int)	12 (48%)	13 (52%)	25 (100%)
네트워크 관리기관(ne, net)	110 (46.4%)	127 (53.6%)	327 (100%)
대학/교육기관(ac, edu)	613 (46.9%)	695 (53.1%)	1,308 (100%)
비영리기관(or, org)	1,039 (45.4%)	1,249 (54.6%)	2,288 (100%)
정부기관(go, gov)	367 (41%)	529 (59%)	896 (100%)
영리기관 및 기관(co, com)	336 (36.8%)	577 (63.2%)	913 (100%)
기타	405 (51.8%)	377 (48.2%)	782 (100%)
전체	2,922 (44.8%)	3,607 (55.2%)	6,529 (100%)

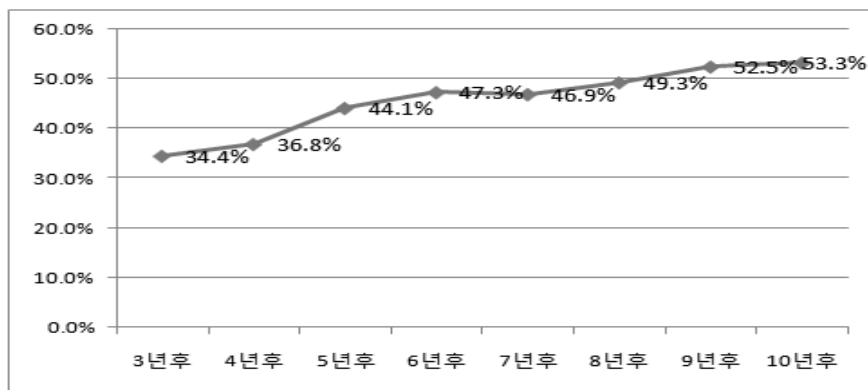
자료 유형별 인터넷 참고문헌의 활동성(<표 16>)을 보면, TXT 포맷의 인터넷 참고문헌의 비활동성이 100%로 가장 높았다. 다음으로 HWP과 DOC 포맷의 인터넷 참고문헌의 비활동성이 각각 61.5%, 51%로 높게 나타났고 PDF 포맷이 47.9%, PPT 포맷이 44.2%, 일반 웹문서가 44.1% 순으로 나타났다. 자료 유형별 인터넷 참고문헌의 활동성의 특징은 일반 웹문서에 비해 HWP, DOC, PDF, PPT 포맷 자료의 비활동성이 상대적으로 높다는 것이다.

<표 16> 자료 유형별 인터넷 참고문헌의 활동성

자료유형	인터넷 참고문헌 활동성		전체
	비활동	활동	
TXT	5 (100%)	0 (0%)	5 (100%)
HWP	8 (61.5%)	5 (38.5%)	13 (100%)
DOC	26 (51%)	28 (49%)	51 (100%)
PDF	439 (47.9%)	477 (52.1%)	916 (100%)
PPT	19 (44.2%)	24 (55.8%)	43 (100%)
일반웹문서	2,425 (44.1%)	3,076 (55.9%)	5,501 (100%)
전체	2,922 (44.8%)	3,607 (55.2%)	6,529 (100%)

2. 인용년도 이후 인터넷 참고문헌의 비활동성 비율의 변화추이

2012년 6월 13일 현재를 기준으로 인용년도 이후 인터넷 참고문헌의 비활동성 비율의 변화 추이(<그림 7>)를 보면, 인용된 지 3년이 지나면 34.4% 정도의 인터넷 참고문헌이 비활동성을 보였다. 그리고 인용연도가 증가할수록 인터넷 참고문헌의 비활동성이 지속적으로 증가하였는데 인용된 지 6년 후에는 47.3%까지 비활동성이 증가하였고, 10년 후에는 인터넷 참고문헌의 비활동성이 최대 53.3%까지 증가하였다.



<그림 7> 인용연도 이후 인터넷 참고문헌의 비활동성 비율의 변화추이



이런 결과는 앞에서 제시된 Lawrence의 연구결과와 유사하다고 하겠다. Lawrence와 본 연구의 연구대상 분야가 서로 다르고, Lawrence의 연구결과가 2001년 이전의 인터넷 참고문헌의 비활동성에 대한 변화추이를 보여주고 본 연구가 2002년부터 2009년까지의 인터넷 참고문헌의 비활동성에 대한 변화추이를 보여준다고 할 때 앞으로 대부분의 학문분야에서 인터넷 참고문헌의 비활동성은 계속적으로 증가할 것으로 예상된다.

### 3. 인터넷 참고문헌의 비활동성에 대한 오류 메시지

인터넷 참고문헌의 비활동성과 관련된 오류 메시지(<표 17>)를 보면, 대부분 [404 not found(문서를 찾을 수 없음)]의 오류 메시지가 전체 55.9%로 가장 높았고, 다음으로 [0 사이트 없음]이 31.7%를 차지하였다.

<표 17> 인터넷 참고문헌의 비활동성에 대한 오류 메시지

HTTP 오류	내용	빈도	비율
404	문서를 찾을 수 없음	1,632	55.9%
0	사이트 없음	926	31.7%
403	접근 금지	115	3.9%
400	요청 실패	59	2.0%
302	요구한 데이터가 변경된 URL에 있음을 명시	51	1.7%
301	요구한 데이터를 변경된 URL에서 찾았음	42	1.4%
500	서버 내부 오류	30	1%
405	메서드 허용 안됨.	17	0.6%
205	새 문서 없음	13	0.4%
기타	기타 오류	37	1.4%
합계		2,922	100%

학술지별로 인터넷 참고문헌의 비활동성에 대한 오류 메시지(<표 18>)를 보면, 모든 학술지에서 [404 not found(문서를 찾을 수 없음)]과 [0 사이트 없음]의 오류 메시지가 가장 높게 나타났다.

<표 18> 학술지별 인터넷 참고문헌의 비활동성에 대한 오류 메시지

학술지명	인터넷 참고문헌의 비활동성에 대한 오류 메시지			
	[0] 사이트없음	[404] 문서를 찾을 수 없음	기타	합계
한국정보관리학회지	229 (32.3%)	392 (55.2%)	89 (12.5%)	710 (100%)
한국도서관·정보학회지	319 (32.7%)	550 (56.3%)	108 (11%)	977 (100%)
한국문헌정보학회지	239 (28.7%)	479 (57.5%)	115 (13.8%)	833 (100%)
한국비블리아학회지	139 (34.6%)	211 (52.5%)	52 (12.9%)	402 (100%)
합 계	926 (31.7%)	1,632 (55.9%)	364 (12.4%)	2,922 (100%)

도메인 유형별 인터넷 참고문헌의 비활동성에 대한 오류 메시지(<표 19>)를 보면, 국제적인 기구/기관(int), 미연방군사기관, 네트워크관리기관(mil), 연구기관(re)들은 [0 사이트 없음]의 오류 메시지의 비율이 [404 not found(문서를 찾을 수 없음)]보다 상대적으로 높게 나타났다. 반면, 대학/교육기관(ac.edu), 비영리기관(or.org), 영리기관(co.com), 정부 기관(go.gov) 등은 [404 not found(문서를 찾을 수 없음)]의 오류 메시지의 비율이 [0 사이트 없음]보다 상대적으로 높게 나타났다.

<표 19> 도메인 유형별 인터넷 참고문헌의 비활동성에 대한 오류 메시지

도메인 유형	인터넷 참고문헌의 비활동성에 대한 오류 메시지			
	[0] 사이트없음	[404] 문서를 찾을 수 없음	기타	합계
국제적인 기구/기관(int)	5 (41.7%)	2 (16.7%)	5 (41.6%)	12 (100%)
미연방군사기관(mil)	5 (100%)	0 (0%)	0 (0%)	5 (100%)
네트워크 관리기관(ne, net)	66 (60%)	38 (34.5%)	6 (5.5%)	110 (100%)
대학/교육기관(ac. edu)	209 (34.1%)	347 (56.6%)	57 (9.3%)	613 (100%)
비영리기관(or, org)	218 (21%)	678 (65.3%)	143 (13.7%)	1,039 (100%)
연구기관(re)	23 (65.7%)	6 (17.1%)	6 (17.2%)	35 (100%)
영리기관 및 기관(co, com)	109 (32.4%)	161 (47.9%)	66 (19.7%)	336 (100%)
정부기관(go, gov)	129 (35.1%)	186 (50.7%)	52 (14.2%)	367 (100%)
기타	162 (40%)	214 (52.8%)	29 (7.2%)	405 (100%)
합계	926 (31.7%)	1,632 (55.9%)	364 (12.4%)	2,922 (100%)

언어 유형별로 인터넷 참고문헌의 비활동성에 대한 오류 메시지(<표 20>)를 보면, 영어 자료와 일본어 자료는 [404 not found(문서를 찾을 수 없음)] 오류 메시지의 비율이 각각 61.5%, 79.3%로 [0 사이트 없음] 보다 상대적으로 매우 높게 나타났다. 반면 한국어 자료의 경우는 [0 사이트 없음] 오류 메시지가 [404 not found(문서를 찾을 수 없음)]보다 상대적으로 높게 나타났다.

<표 20> 언어 유형별 인터넷 참고문헌의 비활동성에 대한 오류 메시지

언어 유형	[0] 사이트없음	[404] 문서를 찾을 수 없음	기타	합계
영어	631 (26.8%)	1,449 (61.5%)	276 (11.7%)	2,356 (100%)
일본어	6 (20.7%)	23 (79.3%)	0 (0%)	29 (100%)
한국어	285 (54%)	155 (29.4%)	88 (16.6%)	528 (100%)
기타	4 (44.4%)	5 (55.6%)	0 (0%)	9 (100%)
합계	926 (31.7%)	1,632 (55.9%)	364 (12.4%)	2,922 (100%)

## VI. 결론 및 제언

본 연구에서 지금까지 수행된 연구의 결과를 요약하면 다음과 같다.

첫째, 국내 문헌정보학분야 학술지 기사 1,666건을 대상으로 분석한 결과, 참고문헌은 총 42,032건 이었고, 이 중 인터넷 참고문헌은 6,529건으로 전체 참고문헌 중 15.5%가 인터넷 참고문헌인 것으로 나타났다. 연도별 인터넷 참고문헌의 수의 변화를 보면 조사 시점인 2002년 이후 매년 꾸준히 증가한 것으로 나타났다.

둘째, 기사 1건 단위로 분석한 결과, 기사 1건당 평균 25.6건의 참고문헌이 포함되었고, 기사 1건당 인터넷 참고문헌은 평균 3.9건인 것으로 나타났다. 연도별 기사 1건당 평균 인터넷 참고문헌의 수는 조사 시점인 2002년 이후 일정하지는 않지만 증가하는 양상을 보였다.

셋째, 인터넷 참고문헌을 1개 이상 포함한 기사 단위로 분석한 결과, 전체 기사 1,666건 중 인터넷 참고문헌을 1개 이상 포함한 기사는 1,022건으로 전체 기사 중 61.3%로 매우 높게 나타났다. 또한, 연도별 인터넷 참고문헌을 1개 이상 포함한 기사의 수는 조사 시점인 2002년 이후 증가하는 양상을 보였다.

넷째, 어떤 연구자가 기사를 작성한다고 가정할 때, 해당 연구자가 인터넷 참고문헌을 인용할 확률은 최소 15.5%에서 최대 25.4%인 것으로 나타났다.

다섯째, 인터넷 참고문헌에 언어 유형, 도메인 유형, 자료 유형을 보면, 영어 자료가 81.8%로 가장 많이 인용되었고 다음으로 한국어 자료가 16.5%로 많이 인용되었다. 도메인 유형은 비영리기관(.or, .org)이 35%, 대학/교육기관(.ac, .edu)이 20%, 영리기관(.com, .co)이 14%, 정부기관(.gov, .go)이 13.7%순으로 많이 인용되었다. 자료 유형은 일반웹문서(.html, .htm, .jsp, .asp, .php 등)가 84.3%로 가장 많이 인용되었고 다음으로 PDF가 14%인 것으로 나타났다.

여섯째, 인터넷 참고문헌의 비활동성은 44.8%로 매우 높게 나타났다. 인터넷 참고문헌의 비활동성 비율은 인용된 지 3년 후에는 34.4%, 5년 후에는 44.1%, 10년 후에는 53.3%까지 증가하였다. 인터넷 참고문헌의 주요 비활동성 원인은 55.9%가 “[404] 문서를 찾을 수 없음”이었고, 31.7%가 “[0] 사이트 없음”으로 나타났다.

이상의 연구결과를 종합해 볼 때, 본 연구의 분석대상인 국내 문헌정보학분야 학술지가 국내 모든 학문분야를 대표하지는 못하지만 국내 학술지 기사에서 인터넷 참고문헌을 인용하는 비율이 증가할 것이다. 특히 인터넷 참고문헌의 인용년도 이후로부터 인용기간이 늘어날수록 인터넷 참고문헌의 비활동성 비율은 더 심각하게 증가할 것으로 예상된다. 향후 국내 학술지 기사에 인용된 인터넷 참고문헌의 비활동성을 최소화하기 위한 방안들이 마련되어야 할 것이다.

본 연구에서는 인터넷 참고문헌의 비활동성을 최소화하기 위한 방안을 마련하는데 고려되어야

할 몇 가지 사항을 제안하고자 한다.

첫째, 이미 인용된 인터넷 참고문헌에 대한 비활동성을 최소화하기 위한 현실적인 대책을 수립하는 것은 어렵다고 보인다. 그 이유는 비활동성을 보이는 인터넷 참고문헌은 이미 해당 자료가 손실되어 존재하지 않기 때문에 접근할 수가 없고, 아직까지 활동성을 보이는 인터넷 참고문헌의 경우에도 실제로 해당 URL로 접근하여 찾은 자료가 원래의 그 자료인지를 확인할 길이 없고, 원래 자료라고 해도 내용이 바뀌지 않았다고 보장할 수 없다. 따라서 현실적으로 대책 수립이 가능한 대상은 현재 시점부터 학회에 투고되는 기사에 인용되는 인터넷 참고문헌들일 것이다.

둘째, 저자는 인용하고자 하는 인터넷 자료가 PDF 포맷의 자료일 경우, 해당 자료가 인쇄본으로 출판되었는지를 확인할 필요가 있다. 만약 PDF 자료와 인쇄본으로 동시에 출판되고 있다면 논문을 작성할 때 인터넷 자료의 URL 주소와 함께 인쇄본에 대한 출처정보도 함께 제시해 주는 것이 필요할 것이다. 이를 위해서는 저자들의 자발적인 노력이 요구될 수밖에 없다. 또는 저자가 해당 PDF 파일을 다운로드하여 저장하여 관리하고, 이용자가 요구할 경우 해당 자료를 제공하는 것도 가능할 것이다. 그러나 이 경우는 해당 자료가 오픈 액세스 자료가 아닌 경우 저작권 문제가 발생할 수 있다는 한계점이 있다.

셋째, 학회에서 저자로 하여금 논문을 투고할 때, 인용한 모든 인터넷 참고문헌에 대한 캡처 화면을 제출하도록 요구할 수 있다. 이럴 경우 투고 되는 논문에 인용된 인터넷 참고문헌을 효율적으로 학회가 수집하여 관리할 수 있을 것이다. 그러나 이 경우도 해당 인터넷 참고문헌들이 오픈 액세스 정책 하에 이용할 수 있는 자료가 아닌 경우, 저자들은 저작권자의 이용허락을 받은 후에 캡처 화면을 제출해야 한다. 이는 저자들에게 적지 않은 부담을 줄 수 있다.

넷째, 이용자에게 해외에서 운용되고 있는 Internet Archive([www.archive.org](http://www.archive.org)), Google([www.google.com](http://www.google.com)), OAIS(Open Archival Information System)를 통해 해당 인터넷 참고문헌을 확인하고 참조하도록 안내하는 것도 가능한 방법이다. 그러나 Internet Archive([www.archive.org](http://www.archive.org)), Google([www.google.com](http://www.google.com)), OAIS(Open Archival Information System)가 국내 학술지 기사에 인용된 모든 인터넷 참고문헌을 포함하고 있지 못하며, 만약 안내된 해당 사이트에서 인터넷 참고문헌을 찾아다 하더라도 찾은 자료가 원래의 그 자료인지를 확인할 길이 없고 원래 자료라고 해도 내용이 바뀌지 않았다고 보장할 수 없다는 문제점은 여전히 존재한다.

이상과 같이 인터넷 참고문헌의 비활동성을 최소화하기 위해서는 저자와 학회의 자발적인 노력뿐만 아니라 저작권 문제의 해결 등 많은 제약사항들이 극복되어야 한다. 향후 저자와 학회, 그리고 정부기관을 포함한 유관 기관들이 모여 이와 같은 국내의 상황을 이해하고 보다 합리적인 방안도 도출하기 위한 논의가 이루어지길 기대해 본다.

## 참고문헌

- 최선희 등. 국내 학술정보 가공 지침서 2010. 대전 : 한국과학기술정보연구원, 2010.
- Dellavalle, Robert et al., "Going, Going, Gone : Lost Internet References," *Science*, Vol.302 (October 2003), pp.787-788.
- Lawrence, Steve et al., "Persistent of Web References in Scientific Research," *IEEE Computers*, Vol.34, No.2(2001), pp.26-31.
- Lynch, Clifford. Preserving Digital Information to Support Scholarship, In *The Internet & the University*, 2003. <<http://www.educause.edu/ir/library/pdf/ffpiu029.pdf>> [Cited 2012. 6. 13].
- Lawrence, Steve and C. Lee Giles. "Accessibility of information on the web," *Nature*, Vol.23, No.3(1998), pp.33-39.
- McCown, Frank et al., *The Availability and Persistence of Web References in D-Lib Magazine*, Old Dominion University Department of Computer Science, Norfolk, VA 23529 USA, 2005. <<http://iwaw.europarchive.org/05/papers/iwaw05-mccown1.pdf>> [Cited 2012. 6. 13].
- Ostrom, Elinor, Charlotte Hess. 지식의 공유. 김민주, 송희령 역. 서울 : 타임북스, 2010.
- Weiss, Rick. *On the Web, Research Work Proves Ephemeral : Electronic Archivists Are Playing Catch-up in Trying to Keep Documents From in History's Dustbin*, 2003, <[http://stevereads.com/cache/ephemeral\\_\\_web\\_\\_pages.html](http://stevereads.com/cache/ephemeral__web__pages.html)> [Cited 2012. 6. 13].