

학술 커뮤니케이션을 위한 인터넷 정보자원의 안정성에 대한 추적연구*

- 국내 문헌정보학 분야를 중심으로 -

A Longitudinal Study on Stability of Internet Information Resources for Academic Communications in Korean LIS Domain

김 성 원 (Sungwon Kim)**

목 차

- | | |
|---------------------|-----------------------------|
| 1. 서론 | 4. 인터넷 정보자원의 안정성 개선을 위한 시사점 |
| 2. 이론적 배경 및 선행연구 개관 | 5. 결론 |
| 3. 인터넷 정보자원의 안정성 분석 | |

초 록

학술 커뮤니케이션 과정에서 인터넷 정보자원의 활용이 보편화 되었다. 인터넷 정보자원의 안정성을 확보하는 일은 학술 커뮤니케이션의 효율성에 영향을 미치는 요인이다. 인터넷 정보자원의 안정성을 확보하기 위한 노력의 일환으로 학술 커뮤니케이션 과정에서 참고문헌으로 인용된 인터넷 정보자원을 대상으로 안정성(stability)을 검증하기 위한 추적연구를 수행하였다. 인터넷 자원への 접속가능성(accessibility)을 기준으로 정량적 분석을 수행하였고, 접속이 가능한 자원을 대상으로 동일성(uniformity)을 기준으로 정성적 분석을 수행하였다. 정량적 분석결과 접속가능성은 발간후 5년간까지 26.1%(2013년 자원) - 27.4%(2019년 자원), 5-10년 기간에서 28.3%, 10-15년 기간에서 26.5% 저하되는 것으로 분석되었다. 정성적 분석결과 최초 인용자료와 동일성이 확인된 자료의 비율은 발간후 5년 시점에서 39.4%, 10년 경과시 16.8%, 15년 경과시 14.8%인 것으로 분석되었다. 정량적, 정성적 분석 모두에서 시간의 경과에 따른 안정성 저하를 확인할 수 있었다. 연구의 과정에서 도출된 인터넷 정보자원의 안정성 제고를 위한 시사점과 제언을 제시하였다.

ABSTRACT

Internet Information Resources are commonly utilized in the Academic Communication process. Securing the stability of Internet Information Resources is a critical factor in the efficiency of Academic Communication. In an effort to secure the stability of Internet Information Resources, this study conducts longitudinal studies to verify the stability of Internet Information Resources cited as references in the Academic Communication process. Furthermore, this study performs quantitative analysis on the accessibility of Internet Resources and qualitative analysis on the authenticity and uniformity of accessible resources. Quantitative analysis shows that accessibility decreased by 26.1% (2013 data set) - 27.4% (2019 data set) - in the initial 5 years since publication, 28.3% in the 5-10 years span, and 26.5% in the 10-15 years span. Qualitative analysis shows the percentage of citations found to be uniform with the initially cited resource 39.4% after five years of publication, 16.8% after 10 years, and 14.8% after 15 years. In both quantitative and qualitative analysis, we observe a decrease in stability over time. The researcher presents implications and suggestions for improving the stability of Internet Information Resources.

키워드: 인터넷 정보자원, 안정성, 접속가능성, 동일성, 학술 커뮤니케이션

Internet, Internet Information Resources, Stability, Accessibility, Uniformity, Academic Communication

* 이 연구는 충남대학교(교육·연구 및 학생지도비)에 의해 지원되었음.

** 충남대학교 문헌정보학과 교수(sungwonk@cnu.ac.kr / ISNI 0000 0004 7490 2525)

논문접수일자: 2024년 10월 27일 최초심사일자: 2024년 11월 1일 게재확정일자: 2024년 11월 11일

한국문헌정보학회지, 58(4): 377-394, 2024. <http://dx.doi.org/10.4275/KSLIS.2024.58.4.377>

© Copyright © 2024 Korean Society for Library and Information Science

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

1. 서론

학술논문은 기본적인 연구과정의 윤리성에 더하여 정확성과 객관성, 검증성 및 재현성 등의 요건을 갖추어야 한다. 이러한 요건을 갖추었음을 상징적으로 보여줄 수 있는 논문의 구성요소가 참고문헌이다. 해당 분야 연구자들은 특정 연구자의 연구결과에 포함된 참고문헌을 통해 해당 연구가 적합한 선행연구를 참고하여 객관적으로 수행되었음을 확인할 수 있으며, 원하는 경우 결과를 스스로 재현해보기 위해 좀더 구체적인 정보를 담고 있는 인용자료에도 접근할 수 있다. 이러한 의미에서 학술 커뮤니케이션 과정에서 참고문헌의 역할은 지대하다고 할 수 있다. 인터넷 등 정보통신기술의 발전은 일상생활에서의 소통문화뿐 아니라 학술커뮤니케이션 과정에도 영향을 미쳤으며 그 가운데 하나가 인터넷 주소를 포함한 인터넷 자원을 참고문헌으로 활용하게 된 것이다.

인터넷 자원을 참고문헌으로 수용함에 따라 연구자들은 클릭 한번 혹은 약간의 추가적인 노력을 통해 참고문헌의 원문에 접근할 수 있게 되었고, 기존에 참고문헌에 접근하기 위해 도서관을 방문하여 원문을 찾아 복사하는 수고를 줄일 수 있게 되어 연구 효율성을 크게 개선할 수 있었다. 이러한 장점과 더불어 자원의 휘발성이라는 단점도 제기되고 있는 바, 인터넷 정보자원의 경우 여러가지 이유로 시간의 경과에 따라 접속이 되지 않거나 접속이 되더라도 내용이 갱신되어 있어 연구자가 인용한 시점의 내용과는 상이한 자원에 접근하게 된다는 안정성(stability)의 문제가 그것이다. 이러한 문제의식에 기반하여 인터넷 정보자원의 안정성을

검증하고자 시도하였으며, 연구결과를 기반으로 안정성을 제고시킬 수 있는 시사점과 제안을 정리하여 제시하였다.

이 연구는 학술 커뮤니케이션 과정에서 활용된 인터넷 정보자원의 안정성을 검증해보고자 하는 목적에서 수행하였다. 연구에서 포함한 연구문제 및 범위는 다음과 같다. 첫째, 2019년에 인용된 인터넷 정보자원의 안정성을 5년 경과 시점인 2024년에 검증한 후, 결과를 2013년 인용 인터넷 정보자원을 대상으로 5년 경과 시점인 2019년에 수행한 안정성 검증결과와 비교한다. 이를 통해 환경과 시간의 변화에 따라 5년 경과시점에서 인터넷 정보자원의 안정성에 차이가 발생하였는지 확인할 수 있다. 둘째, 2013년 인용된 인터넷 정보자원의 안정성을 10년 경과 시점인 2024년에 검증한 후, 시간 경과에 따른 안정성의 변화를 확인한다. 셋째, 위에서 2013년 인용 인터넷 자원을 대상으로 수행한 10년 경과시점 안정성을 2019년에 2009년 인용 인터넷 자원을 대상으로 수행한 10년 경과시점의 안정성 검증결과와 비교한다. 이를 통해 환경과 시간의 변화에 따라 10년 경과시점에서 인터넷 정보자원의 안정성에 차이가 발생하였는지 확인할 수 있다. 넷째, 2009년에 인용된 인터넷 정보자원을 대상으로 15년 경과시점의 안정성을 검증한다.

위와 같은 분석에 더하여 정성적인 분석을 추가적으로 수행하였다. 안정성 분석은 기본적으로 인용된 참고문헌에 포함된 인터넷 주소(URL: Uniform Resource Locator)의 유효성과 접속가능성(accessibility)에 기반한 정량적 분석이다. 분석의 과정에서 인용된 정보자원의 URL이 유효하고 접속이 되더라도 최초

인용시점에서 연구자가 인용한 내용과는 차이가 있는 경우가 상당수 있음을 발견하였다. 인터넷 주소가 유효하고 자료에 접근할 수 있다고 하더라도 원래 인용한 내용과 차이가 있다면 기술적으로는 안정성이 있다고 할 수 있을 지라도 내용적으로는 안정성이 없는 것이라 판단했다. 이러한 문제의식에 기반하여 기술적으로 접속이 가능한 자원을 대상으로 내용의 동일성(uniformity)이라는 관점에서 정성적인 분석을 수행하였다.

인터넷 정보자원의 안정성 검증을 위한 인용 정보는 한국연구재단 학문분류표 복합학에 속하는 문헌정보학 분야의 학술지인 한국문헌정보학회지를 선정하여 수집하였다. 문헌정보학 분야를 선택한 이유는 인터넷 정보자원을 상대적으로 조기에 활용하기 시작하였고 활용이 활발한 분야이기 때문이다(김성원, 2017). 한국 문헌정보학회지를 선택한 이유는 국내 문헌정보학에서 가장 전통있는 학술지로서 문헌정보학 분야의 다양한 연구를 균형있게 포괄하여 문헌정보학 전반의 활용현황을 파악하기에 적합하다고 판단하였기 때문이다.

한편 분석과정에서 안정성에 영향을 미치는 것으로 보이는 제반 요소들을 통제한다면 안정성을 제고할 수 있을 것으로 판단하였던 바, 안정성에 영향을 미치는 것으로 파악된 제반 요소들을 중심으로 인터넷 정보자원의 안정성 제고를 위한 시사점과 제안점들을 정리하여 제시하였다.

인터넷 정보자원의 안정성을 다룬 다수의 연구가 수행된 바 있다. 이 연구는 동일한 데이터를 대상으로 시간경과에 따른 안정성의 변화를 추적연구(longitudinal)하였다는 점과, 시기의

변화에 따른 안정성의 변화를 검증하였다는 점에서 기존의 연구와 차별성이 있다. 연구결과 시간 구간별 저하의 구체적인 정도를 수치로 파악할 수 있었다. 다만 구체적인 저하의 정도는 이 연구를 위해 수집한 특정 자원군을 분석한 결과로서 일반화하려면 더 많은 데이터로 추가 검증할 필요가 있다고 판단된다. 한편 선행연구에서 수행하였던 자원의 제반 특성(도메인, 생산국, 생산기관의 성격 등)과 안정성과의 상관관계 분석은 선행연구에서 통계적으로 유의미한 차이를 보이지 않았던 이유로 연구에서 제외하였다.

연구를 위한 데이터 중 발간 10년(2013년), 15년(2009년) 경과시점의 안정성 검증을 위한 데이터는 기존연구(김성원, 2019)에서 분석하였던 데이터를 활용하였으며, 5년 경과시점의 안정성 분석을 위해 추가로 필요했던 발간 5년(2019년) 참고문헌 데이터는 한국연구재단의 한국학술지 인용색인(KCI: Korea Citation Index)과 한국과학기술정보연구원의 디지털 큐레이션센터(관련 과제번호: K24L1M1C1)로부터 제공받았다.

2. 이론적 배경 및 선행연구 개관

이 장에서는 본격적인 분석과 논의에 앞서 학술 커뮤니케이션 과정에서의 인터넷 정보자원의 전반적인 활용현황에 대해 살펴보고 선행연구에 대해 정리해보고자 한다.

2.1 인터넷 정보자원 활용현황

인터넷 정보자원이 학술 커뮤니케이션에 활용되기 시작한 것은 1990년대 정보통신기술의 확

산되고 인터넷이 등장한 이후이다. 그 이전에도 제한된 분야의 연구자들이 연구정보의 공유를 위해 컴퓨터 네트워크를 활용하였으나, 네트워크 자체가 권한을 부여받은 일부에게만 접근을 허용하는 폐쇄된 것이었다. 인터넷의 등장으로 네트워크를 통한 정보의 접근 및 교환이 학계뿐 아니라 일반대중으로 확산되었고, 일상에서의 소통방식에도 변화를 주었지만 학술 커뮤니케이션에도 활용되어 학술 커뮤니케이션 방식을 변화시켰다. 대표적으로는 학술지 원문이 디지털화되어 컴퓨터에 저장되었고 저장된 학술논문에 대한 접근과 유통이 인터넷을 통해 이루어지게 되었다. 이런 변화는 학술논문에 포함된 참고문헌에 인터넷으로 접근 가능한 인터넷 자원을 포함하여 인터넷을 통해 참고문헌 원문으로 편리하게 접근할 수 있게 해주었다. 이후 인터넷을 통해 접근 가능한 자원의 증가에 따라 참고문헌에서 인터넷 정보자원의 비중도 증가하였으며 참고문헌의 상당부분을 차지하고 있는 것으로 분석된 바 있다(김성원, 2017).

인터넷 정보자원의 활용현황을 살펴보도록 한다. <표 1>은 국내 문헌정보학 영역의 학술지인 한국문헌정보학회지와 정보관리학회지에 게재된 학술논문에서 인용한 참고문헌과 참고문

헌에 포함된 인터넷 정보자원의 현황을 정리한 것이다. 국내의 학문 대분류별로 참고문헌에 인용된 인터넷 정보자원의 비율을 조사한 연구에서 2010~2016년간 사회과학이 2.95-3.46%, 자연과학이 1.74-2.44%의 범위로 조사된 것(김성원, 2017)에 비교하면 문헌정보학 영역의 활용이 활성화되어 있음을 확인할 수 있다. 아래의 분석은 한국이라는 지역과 문헌정보학이라는 학문분야의 제한점을 갖는다. 문헌정보학 분야가 복합한 영역에 속하여 정보통신 기술의 적용에 적극적이며 한국이라는 국가 또한 정보통신기술의 선도국임을 감안하더라도 국내 문헌정보학 분야에서 인터넷 정보자원이 타 학문분야 대비 높은 비율로 활용되고 있음을 확인할 수 있다.

인터넷 정보자원의 활용현황은 2002-2024년간 국내 문헌정보학 분야 학술지의 참고문헌을 5년 단위로 집계하였다. 년도별 분석을 통해서도 인터넷 정보자원의 활용현황과 추이의 변화를 살펴볼 수 있으나, 년도별 집계는 일목요연한 파악이 쉽지 않을 수 있고 년도별 편차가 발생할 수 있어 5년 단위로 집계하였다. <표 1>을 보면 2002년 이후 현재까지 게재된 논문의 참고문헌 중 인터넷 정보자원의 비중은 25.5%-

<표 1> 국내 문헌정보학 분야의 인터넷 정보자원 활용현황

| 기간 | 한국문헌정보학회지 | | | 정보관리학회지 | | |
|-------------|-----------|--------|-------|---------|--------|-------|
| | 참고문헌 | 인터넷 자원 | 비율 | 참고문헌 | 인터넷 자원 | 비율 |
| 2002년-2005년 | 5,551 | 818 | 14.7% | 5,377 | 1,544 | 28.7% |
| 2006년-2010년 | 9,539 | 2,625 | 27.5% | 7,180 | 3,020 | 42.1% |
| 2011년-2015년 | 8,431 | 2,524 | 29.9% | 7,639 | 3,173 | 41.5% |
| 2016년-2020년 | 7,778 | 1,368 | 17.6% | 7,289 | 2,169 | 29.8% |
| 2021년-2024년 | 7,490 | 2,538 | 33.9% | 6,127 | 2,545 | 41.5% |
| 합계 | 38,789 | 9,873 | 25.5% | 33,612 | 12,451 | 37.0% |

37%를 접하며 최대 42.1%를 접하는 경우도 있음을 볼 수 있다. 학술지에 따라 시기에 따라 차이는 있으나, 인터넷 정보자원의 활용이 시작된 이후 인터넷 정보자원의 꾸준히 활용되고 있음을 볼 수 있다. 가장 최근인 2021년 이후 현재까지의 활용현황을 보면 한국문헌정보학회지의 경우 33.9%, 정보관리학회지의 경우 41.5%의 참고문헌이 인터넷 정보자원인 것으로 집계된다. 이런 차이는 양 학술지의 연구분야 차이에 영향을 받은 것으로 추정된다. 시계열관점에서 보면 최초 활용되기 시작한 이후 그 비중이 증가하여 왔고 일정 수준이상에서는 그 증가추이가 정체되는 것을 볼 수 있다. 변화의 추이와 원인에 대해서는 추가적인 분석이 필요할 것이나, 본 연구에서는 인터넷 정보자원의 활용이 학술 커뮤니케이션 과정에서 중요한 비중을 차지하고 지속적으로 활용되고 있음을 확인하는 것으로 가름하고자 한다.

2.2 안정성관련 선행연구

인터넷 정보자원이 학술 커뮤니케이션에 활용이 보편화 되면서 인터넷 자원의 안정성이 학술 커뮤니케이션의 효율성에 영향을 미칠 수 있음이 인식되었다. 이러한 인식에 따라 인터넷 정보자원 관련 연구들이 수행된 바 있다. 선행연구의 주요 결과로는 학술커뮤니케이션 과정에 인터넷 자원 활용이 보편화 되었다는 점, 활용의 정도는 학문분야별로 차이가 있으나 일정 수준에 이르면 더 이상 증가하지 않으며 인터넷 자원의 안정성은 기존 인쇄매체에 비해 취약하다는 점 등이다.

선행연구들은 크게 인터넷 정보자원의 활용현

황을 분석한 연구와 활용된 인터넷 정보자원의 안정성을 검증한 연구로 대별해 볼 수 있다. 인터넷 정보자원의 활용과 관련한 연구로는 사회과학영역 연구자들을 중심으로 연구에 활용하는 정보원 활용행태를 분석한 박준식 외(2000)의 연구, 문헌정보학 연구자를 중심으로 인터넷 자원의 활용현황을 분석한 장선영, 정영미(2012)의 연구, 연구자의 제반 특성과 인터넷 정보자원 활용과의 관계를 분석한 김성원(2017)의 연구 등을 들 수 있다.

인터넷 정보자원의 안정성이 취약하다는 연구결과에 기반하여 인터넷 정보자원의 안정성에 대한 연구가 수행되었다. 국내의 관련 연구로는 인터넷 정보자원의 활용이 증가하고 있으나 안정성은 취약하다는 결과를 제시한 김홍렬(2006)의 연구, 국내 문헌정보학 분야에서 인용된 인터넷 정보자원을 대상으로 시간경과에 따른 안정성의 변화를 분석한 김성원(2019)의 연구 등을 들 수 있다. 국외에서 수행된 관련 연구로 광범한 관련 자료를 개관한 후 인터넷 정보자원의 불안정성은 'Internet Archive'에 저장하는 등의 방안으로 극복될 수 있음을 제시한 Teixeira da Silva와 Nazarovets(2023)의 연구, 인용된 인터넷 정보자원의 링크 유효성과 대체 경로 접속가능성을 검증하여 접속인 안되는 자원의 58.8%는 대체경로로 접근이 가능함을 확인한 Kumar와 Kumar(2017)의 연구, 인터넷 정보자원의 안정성과 그에 영향을 미치는 요인을 분석하여 최대 요인은 시간이며 자원의 중위 생명주기는 9.3년임을 밝힌 Hennessey와 Ge(2013)의 연구, 인용된 인터넷 정보자원의 링크 유효성과 대체경로를 분석하여 대체경로까지를 포함하면 접속가능성이 95%으로 높아짐

을 확인한 Oranus et al.(2011)의 연구, 그리고 문헌정보학 논문을 대상으로 발간 3-5년 시점에서 안정성을 검증하여 결과를 제시한 Riahinia et al.(2011)의 연구 등을 들 수 있다.

3. 인터넷 정보자원의 안정성 분석

이 장에서는 인터넷 정보자원의 안정성 분석을 위한 데이터의 수집과 정제 과정을 요약하고, 분석결과를 제시하고자 한다. 분석결과는 접속 가능성을 기준으로 검증한 정량적 분석과 동일성을 기준으로 검증한 정성적 분석으로 나누어 시간 경과에 따른 변화를 중심으로 정리하였다. 또한 분석의 과정에 발견한 내용을 기반으로 인터넷 정보자원의 안정성 향상을 위한 시사점과 제안점을 정리하여 4장에 제시하였다.

인터넷 정보자원의 안정성은 학술논문에서 인용한 참고문헌 중 인터넷 주소(URL: Uniform Resource Locator)를 포함한 인터넷 정보자원을 대상으로 발간 5년, 10년, 15년 시점으로 나누어 검증하였다. 분석 시점을 이와 같이 나누는 것은 Hennessey & Ge(2013)의 연구에 따르면 안정성은 발간 5년과 10년 시점에서 의미있는 변화를 보이는 것으로 연구되었기 때문이다.

3.1 분석대상 데이터의 수집 및 정제

인터넷 정보자원의 안정성 검증을 위한 학술 분야로 연구재단 학문분류표의 복합학에 속하는 문헌정보학을 선정하였다. 문헌정보학을 선택한 이유는 문헌정보학이 인터넷 정보자원과 같은 새로운 정보자원을 비교적 선도적으로 연

구에 적용하는 분야로 판명되었기 때문이다(김성원, 2017). 검증대상 학술지로는 한국문헌정보학회를 선정하였는데, 이는 2019년에 수행한 선행연구에서 한국문헌정보학회지 게재논문에 포함된 인터넷 정보자원을 대상으로 안정성을 검증한 연구결과와의 비교를 통해 시기에 따른 안정성의 변화여부를 확인하고자 하였기 때문이다(김성원, 2019).

구체적인 검증을 위한 데이터의 수집 및 정제과정은 다음과 같다. 발간시점 5년 경과 인터넷 정보자원의 안정성 검증을 위해 한국연구재단과 한국과학기술정보연구원으로부터 한국문헌정보학회지 2019년 게재논문 66편에서 인용한 참고문헌 1,311건의 서지데이터를 제공받았다. 제공받은 1,311건의 참고문헌 데이터 가운데 인터넷 주소(URL)가 포함된 213건을 인터넷 정보자원으로 선별하였다. 수집된 인터넷 정보자원 중 중복으로 인용된 자원 5건은 중복집계에 따른 통계상의 왜곡을 방지하고자 1건씩만 포함시켜 208건을 대상으로 안정성 검증을 수행하였다.

발간시점 10년, 15년 경과 인터넷 정보자원의 안정성 검증을 위한 데이터는 선행연구에서 수집한 데이터를 활용하였다. 다만 선행연구에서 발간 5년, 10년 시점에서 안정성 검증한 결과, 이미 접속이 불가능하였던 2009년 108건, 2013년 59건의 자원을 제외하였다. 이는 이미 선행연구 수행시점에서 접속이 불가하였으므로 다시 검증하여도 접속이 불가할 것으로 판단하였기 때문이다. 이렇게 중복자원 및 선행연구에서 접속이 불가능한 것으로 확인된 인터넷 정보자원을 제외한 2009년 175건, 2013년 167건, 2019년 208건을 대상으로 검증작업을 수행하

였다. 수집된 인터넷 정보자원 및 검증대상 정보자원 내역은 <표 2>에 제시하였다.

3.2 인터넷 정보자원 안정성의 정량적 분석

이 절에서는 연구에서 수행한 인터넷 정보자원의 안정성을 정량적 관점에서 검증한 결과를 제시하고자 한다. 기본적으로 인터넷 정보자원의 안정성은 참고문헌에 포함된 인터넷 자원을 대상으로 접속가능여부(링크의 유효성)를 기준으로 검증하였다. 접속가능성을 기준으로 하는 인터넷 정보자원의 안정성은 기본적으로 시간의 경과에 영향을 받는 것으로 연구(Hennessey & Ge, 2013)된 바 있어, 발간시점 5년, 10년, 15년으로 나누어 검증결과를 정리하였다.

3.2.1 발간 5년 경과 인터넷 정보자원의 안정성

이 항에서는 발간 이후 5년이 경과된 인터넷

정보자원의 안정성 분석결과를 정리하였다. 발간 이후 5년이 경과된 2019년 한국문헌정보학회지 게재논문에 포함된 인터넷 정보자원 208건의 접속가능성을 기준으로 안정성을 검증한 결과는 <표 3>과 같다. 검증시점의 차이에 따라 5년 경과시점의 안정성이 차이가 있는지를 비교하기 위해 2013년에 인용된 인터넷 정보자원의 안정성을 5년 경과시점인 2019년에 분석한 결과를 아래 행에 추가하였다.

<표 3>에서 볼 수 있는 바와 같이 2019년에 인용되어 5년이 경과한 인터넷 정보자원 208건 가운데 접속이 가능한 자원이 151건(72.6%), 불가능한 자원은 56건(27.4%)으로 분석되었다. 시기의 따른 안정성의 차이를 비교하기 위해 2013년 자원과 2019년 자원의 5년 경과시점 안정성을 비교하였다. 조사시점이 2019년과 2024년으로 5년의 간격이 있었으나 5년 경과시점에서 분석한 접속가능성은 2013년 인용된 인터넷

<표 2> 데이터 수집 및 정제

| 발간년 | 최초 수집자원수 | 중복자원수 | 기접속 불가자원* | 분석대상 자원수 | 비고: 발간후 경과년 |
|------|----------|-------|-----------|----------|-------------|
| 2009 | 297 | 14 | 108 | 175 | 15년 |
| 2013 | 271 | 45 | 59 | 167 | 10년 |
| 2019 | 213 | 5 | - | 208 | 5년 |
| 합계 | 794 | 64 | 167 | 550 | |

* 2019년 수행한 연구에서 접속이 불가능하였던 자원

<표 3> 5년 경과시 접속가능성 검증결과(건수, 백분율 %)

| 구분 | 검증대상 | 접속가능 자원수 | | 접속불가 자원수 | | 검증시점 |
|------------|-------|----------|--------|----------|--------|--------------|
| | 건수 | 건수 | 백분율(%) | 건수 | 백분율(%) | |
| 2019년 인용자원 | 208* | 151 | 72.6% | 57 | 27.4% | 2024년 발간후 5년 |
| 2013년 인용자원 | 226** | 167 | 73.9% | 59 | 26.1% | 2019년 발간후 5년 |
| 전체 | 434 | 318 | 73.3% | 116 | 26.7% | |

* 2019년 최초수집 213건 중 중복자원 5건 제외 분석대상 자원수

** 2013년 최초수집 271건 중 중복자원 45건 제외 분석대상 자원수

정보자원(73.9%)과 2019년 인용된 인터넷 정보자원(72.6%)간에 큰 차이가 없는 것으로 분석되었다. 시기의 변화에 따라 인터넷과 관련된 기술이 발전하고 안정성에 대한 인식 개선 등 활용환경의 변화에 따라 안정성도 향상되었을 것으로 기대하였으나 2019년 인용된 인터넷 자원을 대상으로 분석한 결과는 5년전인 2013년에 인용된 인터넷 자원의 5년후 시점에서 조사한 안정성과 차이를 보이지 않는다. 2006년 수행된 연구(김홍렬, 2006)에서 5년 경과자료의 57%는 접속이 불가하고 43%만이 접속이 가능한 것으로 조사된 결과에 비교하면 두 검증표본 모두에서 전반적인 안정성이 개선된 것을 볼 수 있다. 즉 인터넷 정보자원의 안정성은 활용 초기시점보다는 개선되었으나, 그 이후 안정성의 관점에서 크게 개선되고 있지 않은 것으로 추정할 수 있다.

한편 2019년 자원가운데 접속불가한 것으로 파악된 자료 57건에 대해 인터넷 포털이나 해당 사이트내의 검색기능 등 대체경로를 통해 최초 인용했던 자료에 접근할 수 있는지의 여부를 검토한 결과 가능한 경우가 29건(50.9%), 접근이 불가능한 경우가 28건(49.1%)으로 분석되었다. 이는 5년 경과시점에서 접속불가 자원을 대상으로 대체경로 접속가능성을 검증한 결과, 가능(45.8%), 불가능(54.2%)으로 분석되

었던 선행연구(김성원, 2019)와 유사한 결과를 보인다.

3.2.2 발간 10년 경과 인터넷 정보자원의 안정성

이 항에서는 발간 이후 10년이 경과된 인터넷 정보자원의 안정성 분석결과를 정리하고자 한다. 발간 이후 10년이 경과된 2013년 한국문헌정보학회지 게재논문에 포함된 인터넷 정보자원 271건 가운데 중복자원 45건, 발간 5년 시점(2019년)에서 접속이 불가한 것으로 판명된 59건을 제외한 167건을 대상으로 접속가능성을 기준으로 안정성을 분석하였다. 분석결과는 <표 4>와 같다. 발간 5년 이후의 안정성 변화추이를 확인하기위해 첫 행에 2019년 분석된 5년 경과시점 검증결과를 추가하였다.

<표 4>에서 볼 수 있는 바와 같이 5년 시점에서 접속이 가능했던 167건 가운데 10년 시점에서 접속이 가능한 자원이 103건(61.7%), 불가능한 자원이 64건(38.3%)인 것으로 분석되었다. 5년 시점에서 접속불가 했던 자원 59건과 10년 시점에서 접속불가 했던 자원 64건을 합산하면 누적 123건(54.4%)이 접속불가한 것으로 분석되었다. 접속불가 자료 123건 가운데 발간 5년 시점에서 접속이 불가능하였던 자원이 59건(26.1%), 5년 이후 10년 경과 시점까지 추가

<표 4> 10년 경과시 접속가능성 검증결과(건수, 백분율 %)

| 구분 | 검증기간 | 검증대상 | 접속가능 자원수 | | 접속불가 자원수 | | 검증시점 |
|------------|-------|------|----------|--------|----------|--------|---------|
| | | 건수 | 건수 | 백분율(%) | 건수 | 백분율(%) | |
| 1-5년 경과기간 | 1-5년 | 226 | 167 | 73.9% | 59 | 26.1% | 발간후 5년 |
| 5-10년 경과기간 | 5-10년 | 167* | 103 | 61.7% | 64 | 38.3% | 발간후 10년 |
| 10년 시점 누적 | 10년 | 226 | 103 | 45.6% | 123 | 54.4% | 발간후 10년 |

* 2013년 최초수집 271건 중 중복자원 45건, 5년 경과시 접속불가 59건 제외 분석대상 자원수 167건

적으로 접속이 불가능해진 자원이 64건(28.3%, 최초 검증대상자원 226건 대비 비율)로 분석되었다. 이는 일반적으로 시간의 경과에 따라 한계접속불가자원도 증가할 것이라는 예상에 부합하는 결과라 판단된다. 또한 10년 시점에서 접속불가 자료가 과반을 넘어 123건(54.4%)로 분석된 결과는 Hennessey & Ge(2013)이 '자원의 중위 생명주기는 9.3년'인 것으로 분석한 연구결과(생명주기 중위값이 9.3년이라 함은 9.3년이 경과하면 자원의 50%가 휘발된다는 의미)와도 부합한다.

3.2.3 발간 15년 경과 인터넷 정보자원의 안정성

이 항에서는 발간 이후 15년이 경과된 인터넷 정보자원의 안정성 분석결과를 정리하고자 한다. 발간 이후 15년이 경과된 2009년 한국문헌정보학회지 게재논문에 포함된 인터넷 정보자원 297건 가운데 중복자원 14건, 발간 10년 시점(2019년)에서의 검토결과 접속이 불가능 것으로 판명되었던 108건을 제외한 175건을 대상으로 접속가능성을 기준으로 안정성을 분석하였다. 분석결과는 <표 5>와 같다. 발간 10년 이후의 안정성 변화추이를 확인하기위해 첫 행에 2019년 분석된 10년 경과시점 검증결과를 추가하였다.

<표 5>에서 볼 수 있는 바와 같이 2009년에 인용되어 15년이 경과한 인터넷 정보자원 297건 중 중복자원 14건, 10년 시점에서 접속불가로 확인된 108건을 제외한 175건 가운데 접속이 가능한 자원이 100건(35.3%), 불가능한 자원은 10년 시점에서 접속불가했던 자원과 누적으로 183건(64.7%)으로 분석되었다. 접속불가자료 183건 가운데 발간 10년 시점에서 접속이 불가능했던 자원이 108건(38.2%), 10년 이후 15년 경과 시점까지 추가적으로 접속이 불가능해진 자원이 75건(26.5%, 최초 검증대상자원 283건 대비 비율)로 분석되었다. 이는 일반적으로 시간의 경과에 따라 한계접속불가자원도 증가할 것이라는 예상과 선행연구 결과와도 부합한다.

발간시점 15년 경과 인터넷 자원을 대상으로 접속가능성을 분석한 연구결과는 비교할 선행연구가 없어 그 자체의 분석에 그치고자 한다. 다만 이러한 분석결과는 추후 인터넷 자원의 안정성 분석관련 연구에서 참조 데이터로 활용될 수 있을 것으로 기대한다. 한편 15년 경과시점의 최종적인 접속가능 자원의 비율이 100건(35.3%)으로 조사된 것은 기존의 연구에서 10년 경과시 모든 자원이 접속불가한 것으로 분석되었던 선행연구(김홍렬, 2006) 결과에 비교하면 안정성이 개선된 것으로 판단할 수 있다.

<표 5> 15년 경과시 접속가능성 검증결과(건수, 백분율 %)

| 구분 | 검증기간 | 검증대상 | | 접속불가 자원수 | | 검증시점 |
|-------------|--------|------|--------|----------|--------|---------|
| | | 건수 | 백분율(%) | 건수 | 백분율(%) | |
| 1-10년 경과기간 | 1-10년 | 283 | 61.8% | 108 | 38.2% | 발간후 10년 |
| 10-15년 경과기간 | 10-15년 | 175* | 57.1% | 75 | 42.9% | 발간후 15년 |
| 15년 시점 누적 | 15년 | 283 | 35.3% | 183 | 64.7% | 발간후 15년 |

* 2009년 최초수집 297건 중 중복자원 14건, 10년 경과시 접속불가 108건 제외 분석대상 자원수 175건

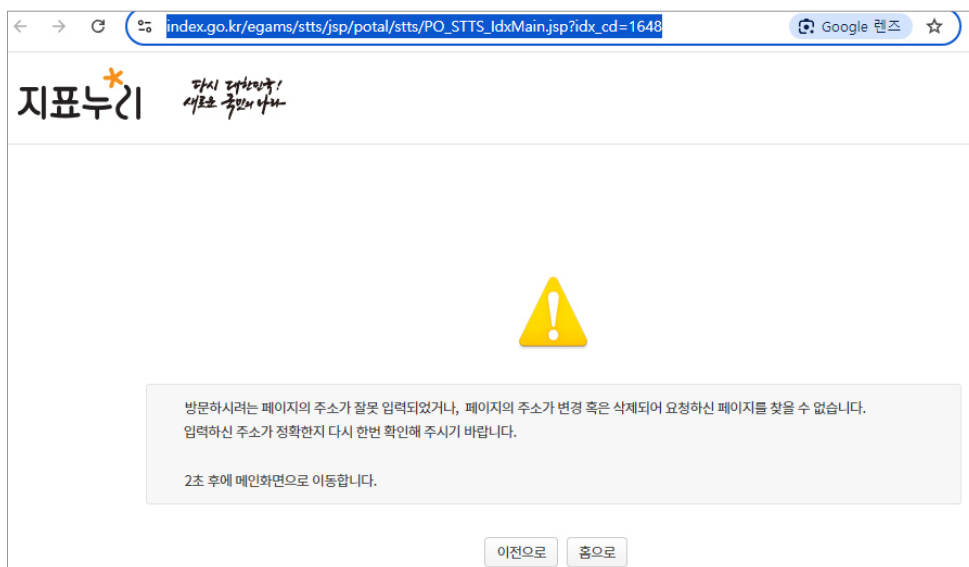
3.2.4 접속불가 인터넷 정보자원 유형

이 항에서는 접속이 불가하였던 인터넷 정보 자원의 접속불가 유형을 정리해 보고자 한다. 이렇게 유형을 구분하는 이유는 본 연구가 지향하는 궁극적인 목표가 인터넷 정보자원의 안정성을 개선하기 위한 것으로 접속이 불가한 경우는 어떤 유형이 있는지를 확인하고 각 유형별로 개선방안을 마련하는 것이 필요하다 판단하였기 때문이다.

검증과정에서 접속이 불가하였던 자원에 대한 추가적인 분석을 통해 몇 개의 접속불가 유형으로 구분해 볼 수 있었다. 첫째 유형은 해당 사이트로의 접속도 되지 않고 오류 메시지 '404 not found'가 뜨는 유형이다(접속불가 유형 1). 둘째 유형은 해당 사이트로는 접속은 되지만 사이트내에 해당 자원이 없음을 안내하는 오류 화면에 정지되는 유형이다(접속불가 유형 2). 위 두 유형이 접속불가의 경우에 만나게 되는

대부분을 차지한다. 이 두 유형 외에도 1) 자원에 연결할 수 없음을 안내하고 해당 사이트내의 검색창을 제시하는 경우, 2) 아래 예시 <그림 1>과 같이 일정시간 해당 자원이 없음을 안내하고 임의적으로 해당 사이트의 최상위 메뉴를 연결하는 경우, 3) 해당 URL이 매도되어 판매 중이라는 안내로 접속되거나 이미 판매되어 완전히 상이한 사이트로 접속되는 경우와 같은 예외적인 경우도 있었다. 접속불가한 인터넷 정보자원을 접속불가 유형별로 구분하면 <표 6>과 같다.

<표 6>에서 볼 수 있는 바와 같이 접속불가한 경우 대부분은 유형 1, 2의 화면을 만나게 된다. 유형 1의 경우 접속을 시도한 이용자는 완전한 접속오류 화면을 만나게 되므로 더 이상 자료에 접근하고자 하는 시도를 중지할 가능성이 크다고 보인다. 유형 2의 경우 해당 사이트의 메뉴화면 안에서 특정자료에의 접근이



<그림 1> 접속불가 안내 후 최상위 메뉴로 이동하는 기타 유형의 화면 예시

〈표 6〉 접속불가 유형별 자원수

| | 2009 | | 2013 | | 2019 | | 합계 | |
|----------|------|-------|------|-------|------|-------|-----|-------|
| | 자원수 | 백분율 | 자원수 | 백분율 | 자원수 | 백분율 | 자원수 | 백분율 |
| 접속불가유형 1 | 35 | 46.7% | 28 | 43.8% | 25 | 41.7% | 88 | 44.2% |
| 접속불가유형 2 | 40 | 53.3% | 30 | 46.9% | 30 | 50.0% | 100 | 50.3% |
| 기타 유형 | - | - | 6 | 9.4% | 5 | 8.3% | 11 | 5.5% |
| 합계 | 75 | 100% | 64 | 100% | 60 | 100% | 199 | 100% |

불가함을 안내하므로 최상위 메뉴로 이동하여 자료를 검색할 동기를 얻을 가능성이 있다. 유형 1 보다는 유형 2가 더 나은 접속불가 대응방안이라 여겨진다. 따라서 사이트 관리자들은 해당 사이트로 유입되는 접속시도가 유효하지 않은 주소를 향해 이루어지는 경우라도 해당 사이트내에서 수용하고 안내를 해주는 장치를 마련하는 것이 권고된다.

한가지 추가적으로 언급하고자 하는 점은 접속불가로 검증된 2019년 자원 가운데 인용표기 시 인터넷 주소가 아닌 연구자 컴퓨터내의 저장위치와 화일명을 기재한 경우 3건이 있었으며, 이 경우 인터넷을 통한 접근은 인용시점부터 원초적으로 불가능하였다고 할 수 있다. 연구자는 인용시 인용방법을 확인하고 학회는 심의과정에서의 검토를 강화할 필요가 있다.

3.3 인터넷 정보자원 안정성의 정성적 분석

앞 절에서는 접속가능여부를 기준으로 인터넷 정보자원의 안정성을 정량적으로 검증해보았다. 이 절에서는 접속이 가능한 것으로 검증된 인터넷 정보자원에 대해 내용적인 동일성(uniformity)을 기준으로 평가하는 정성적 관점에서의 안정성을 검증해보고자 한다. 정성적 검증이 필요한 이유는 비록 인터넷 자원에 기

재된 인터넷 주소로 접속이 되더라도 실제 접속되는 자료의 내용이 연구자가 최초에 인용한 내용과 차이가 있을 수 있기 때문이다. 이렇게 내용상의 차이와 변화가 있다면 인용한 연구자가 참고한 원자료를 확인하고자 하는 후속연구자의 목적에는 부적합한 자료이며, 안정성이 없다고 할 수 있다.

정성적 검증은 최초 수집된 자원에서 중복자원을 제외한 후, 다시 2019년 수행한 선행연구와 이번 연구의 과정에서 접속불가한 것으로 판정된 자료를 제외한 접속가능 자료를 대상으로 수행하였다. 검증 자원은 정량적 검증과 동일하게 발간시점 5년, 10년, 15년 경과 집합으로 나누어 검증하였다. 각 발간년도별 검증대상 자원은 〈표 7〉과 같다.

접속이 가능하였던 인터넷 정보자원에 대한 정성적 분석의 과정에서 인터넷 주소(URL)의 변경여부와 내용적인 동일성의 두 관점에서 분석을 수행하였다. 인터넷 주소의 변경 여부를 포함한 이유는 인터넷 주소가 변경된 경우 내용상의 변화도 수반하는 경우가 많았기 때문이다. 다음 단계로 인터넷 주소의 변경여부와 무관하게 내용상의 변화가 있는지를 확인하였다. 인터넷 주소가 변경된 경우라고 하더라도 내용상의 변화가 없는 경우가 있었기 때문이다. 각 발간년도별 접속이 가능한 자원을 대상으로 인

〈표 7〉 정성분석 대상 자원

| 발간년 | 최초 자원수 | 중복 자원수 | 접속불가 자원수* | 분석대상 자원수 | 비고: 발간후 경과년 |
|------|--------|--------|-----------|----------|-------------|
| 2009 | 297 | 14 | 183 | 100 | 15년 |
| 2013 | 271 | 45 | 123 | 103 | 10년 |
| 2019 | 226 | 18 | 57 | 151 | 5년 |
| 합계 | 794 | 77 | 363 | 354 | - |

* 2019년 수행한 연구와 이번 연구에서 접속이 불가능한 것으로 파악된 자원을 더한 자원수

〈표 8〉 접속가능 자원의 동일성 분석결과

| 구분 | | 내용동일 | 내용갱신 | 합계 | 비고: 최초 자원수 대비 내용 동일 자원수 및 비율 |
|------|--------|------------|------------|-----|------------------------------|
| 2009 | URL 유지 | 42 | 42 | 84 | 42(내용동일) / 283(최초자원), 14.8% |
| | URL 변경 | 0 | 16 | 16 | |
| | 소계 | 42(42.0%) | 58(58.0%) | 100 | |
| 2013 | URL 유지 | 36 | 56 | 92 | 38(내용동일) / 226(최초자원), 16.8% |
| | URL 변경 | 2 | 9 | 11 | |
| | 소계 | 38(36.9%) | 65(63.1%) | 103 | |
| 2019 | URL 유지 | 79 | 53 | 132 | 82(내용동일) / 208(최초자원), 39.4% |
| | URL 변경 | 3 | 16 | 19 | |
| | 소계 | 82(54.3%) | 69(35.7%) | 151 | |
| 총계 | | 162(45.8%) | 192(54.2%) | 354 | |

터넷 주소의 변경여부 및 내용의 동일성 유지라는 정성적인 관점에서 분석한 결과는 〈표 8〉과 같다.

〈표 8〉에 제시된 분석결과를 보면 인터넷 주소(URL)가 변경된 자원의 대다수가 내용상의 변화와 갱신을 수반하고 있음을 볼 수 있다. 예를 들어 2009년에 인용된 접속가능 자원 가운데 URL이 변경된 16건은 내용상으로 모두 수정, 갱신된 것으로 분석되었다. 이러한 인터넷 주소의 변경은 사이트내의 디렉토리 구조와 파일명의 변화에 따라 새로운 주소로 재접속(redirect)되는 경우에서 많이 발생한다. URL의 변화가 내용상의 변화를 야기한다기보다는, 내용상 갱신이나 변화가 있는 경우 수정과정에서 저장장소의 주소인 URL도 변화되는 것으로 이해하는

것이 더 타당한 해석이 될 것이다. URL은 내용 변화 여부를 판단하는 지표로서 활용될 수 있을 것으로 분석된다.

표에 정리된 내용에 포함된 2009년 자료를 중심으로 기술하도록 한다. 2009년 인용된 인터넷 정보자원의 경우, 접속가능 100건 중 URL이 유지된 경우는 84건, 변경되어 새로운 URL로 재접속되는 경우가 16건이다. 새로운 URL로 재접속되는 16건의 경우 정도의 차이는 있으나 모두 내용상 수정, 갱신된 것으로 파악된다. URL을 유지한 84건 가운데 내용이 인용시점과 동일한 경우가 42건이며, URL은 유지되었으나 내용이 갱신된 경우가 42건으로 분석되었다. 내용이 동일하게 유지되었던 42건은 대부분이 특정 업무와 관련된 기관의 보고서나 신문기사로 문

서저장고에 PDF나 HTML의 형식으로 저장된 경우였다.

URL은 유지되었으나 내용은 갱신된 경우는 대부분 특정 서비스나 기관의 최상위 메뉴를 인용한 자원이었다. 특정 서비스를 인용하는 경우 해당 서비스가 중단되거나 변경되는 경우 해당 정보도 삭제되거나 수정될 수밖에 없었을 것으로, 많은 서비스가 시도되고 퇴출되는 상황을 감안하면 많은 변동이 있었음을 이해할 수 있다. 다만 특정 서비스가 중단되는 경우에도 해당 서비스를 대신해 최상위 메뉴로의 접속을 제공하는 경우가 많았다. 특정 기관의 최상위 메뉴를 인용한 경우도 시간 경과에 따라 최상위 메뉴에 포함된 내용은 주기적으로 갱신하는 관행을 고려하면 갱신되어 최초 인용한 내용과 동일성을 유지하기 어려움을 이해할 수 있다. 특정 서비스나 기관의 메인 페이지를 인용하는 경우 장기적으로는 동일성을 기대하기 어려운 것으로 분석되었다. 15년이 경과된 시점에서 인터넷주소와 내용이 완전하게 동일하게 유지되어 접속이 가능한 경우는 42건으로 인용된 인터넷 자원 283건의 14.8%에 불과한 것으로 분석되었다.

검증대상 모두를 포함하여 정리하면 다음과 같다. 인터넷 주소의 변경여부와 무관하게 내용상의 동일성 관점에서 분석해보면 접속가능 자원의 45.8% 정도가 최초에 인용한 자원과 동일한 자원에 접속되는 것으로 분석되었다. 다만 이러한 동일성 분석결과는 접속이 가능한 자원에 대비한 결과로서 만약 최초 인용자원수에 대비한다면 접속이 가능하며 내용 역시 동일하여 최종적으로 최초 인용한 원본자료에 접근할 수 있는 자원의 비율은 5년 경과시 39.4%,

10년 경과시 16.8%, 15년 경과시 14.8%인 것으로 분석되었다. 정성적 분석인 동일성의 관점에서 5년 경과시 39.4%의 자원만이 안정성을 보이는 것은 낮은 수준으로 평가할 수 있다고 판단된다. 또한 최초 인용자원 대비 동일성을 기준으로 하는 정성적 관점의 안정성도 발간 10년 16.8%, 발간 15년 14.8%로 시간의 경과에 따라 저하됨을 볼 수 있다.

4. 인터넷 정보자원의 안정성 개선을 위한 시사점

인터넷 정보자원에 대한 정량적, 정성적 분석의 과정에서 인용문헌이 저장된 사이트 성격, 자료의 유형, 인터넷 주소의 종류에 따라 안정성에 차이가 있음을 발견할 수 있었다. 안정성에 영향을 미치는 것으로 보이는 제반 요소들을 통제한다면 안정성을 제고할 수 있을 것으로 판단되었고 인터넷 정보자원의 안정성 개선을 위한 제안점을 도출할 수 있었다. 이 절에서는 분석의 과정에서 도출된 개선방안에 대해 기술하고자 한다.

우선 접속불가 자원으로부터 얻을 수 있었던 사이트 관리자 측면에서의 개선방안이다. 접속불가 자원은 완전접속 불가와 사이트로 접속된 후 사이트내에서 접속이 불가함을 안내해주는 경우로 대별할 수 있었다. 일부 접속불가 사이트의 경우 사이트내에서 접속불가함을 안내한 후, 5초내외의 시간이 경과하면 최상위 메뉴로 접속해주는 경우를 볼 수 있었다. 완전접속이 불가한 화면을 만나는 경우 이용자는 좌절하고 자료접근을 포기할 가능성이 높아질 것이다. 따

라서 사이트 관리자는 사이트로 유입되는 접속 요청의 주소가 유효하지 않은 경우라도 해당 사이트내에서 안내화면을 제공하고 메인 메뉴로 이동할 수 있는 방안을 제공하는 것이 바람직하다. 이를 통해 이용자는 해당자료에 직접 접근은 못하였으나 해당 사이트내에서 다시 한번 검색할 동기를 부여받을 수 있을 것이다.

관리자 측면에서의 또 다른 개선방안은 사이트 개편시 기존 인터넷 주소에 대한 관리가 필요하다는 점이다. 접속불가한 자원 중 절반 정도는 다른 대체 경로를 통해 접근할 수 있었으며, 원래 저장되었던 사이트에서 인터넷 주소만 변경된 채 저장되어 있는 경우도 있었다. 대표적인 경우로 유럽연합(EU)나 국제도서관협회연맹(IFLA) 등 국제기구 자원의 경우 인용 자원에 포함된 URL로는 접속이 안되지만 해당 사이트의 다른 경로를 통해 발견할 수 있는 경우가 많았다. 인용된 자원이 최초 인용된 사이트내에 존재하나 주소가 변경되어 접속이 안되는 경우가 없도록 최초 부여된 주소가 변경되는 경우 변경된 경우로 재접속 될 수 있도록 관리를 강화하는 것이 권고된다.

인터넷 정보자원을 인용하는 연구자 관점에서의 개선방안이다. 우선 연구자는 인터넷자원 인용시 해당 자원의 성격을 이해하고 자원의 성격에 적합한 방식으로 인용하도록 해야 한다. 정성적 분석결과 최종적으로 원본에 접근할 수 있는 경우는 대부분 기관에서 정리하여 저장해 놓은 PDF 파일을 인용한 경우였다. 만약 동일한 문서가 각각 HTML과 PDF로 저장되어 2개의 버전이 있다면 저장소에 저장되어 있는 PDF(다른 문서편집프로그램이나 이미지 파일 포함) 자료를 인용하는 것이 추후의 접속가능

성을 높일 수 있는 방안이다. 한편 특정 기관에서 제공하고 있는 서비스를 인용하는 경우 해당 서비스의 주요내용이 담긴 이미지를 본문중에 제공하는 것이 바람직하다. 해당 정보원을 'Internet Archives'와 같은 외부 저장소에 저장해 놓는 것도 또 다른 보완책이 될 수 있다. 서비스를 인용한 경우에 시간 경과에 따라 상당수의 서비스가 중단되거나 변경되는 경우가 빈번하며 접속이 불가하거나 접속이 되더라도 인용한 서비스가 아닌 최상위 메뉴로 접속되는 경우가 많기 때문이다. 예를 들어 특정 도서관의 키워드 클라우드서비스를 인용하였는데 시간 경과후 해당 서비스가 중지되어 해당 페이지가 삭제되었다면 접속이 안 될 것이다. 다양한 서비스가 시도되고 또 그 만큼의 서비스가 중지되는 현실을 염두하면 웹사이트에서 특정 서비스를 인용하는 경우 해당 화면을 연구결과물에 포함시켜 두거나 외부 저장소에 저장해두는 것이 바람직하다.

연구자 관점에서의 안정성 개선방안으로 기본적인 인용방식을 준수하는 것과 디지털 객체 식별자(DOI: Digital Object Identifier)를 활용하는 것도 언급하고자 한다. 접속불가 자원의 일부에는 인용자가 인용시 인터넷 주소가 아닌 연구자의 전산장비내 저장위치를 포함하였던 경우나 명백히 오류인 형식으로 인용한 경우도 있었다. 이렇게 잘못된 방식으로 인용한다면 자료에의 접근은 원천적으로 불가능할 수밖에 없으므로 인용요소와 방식을 정확히 준수할 수 있도록 해야 한다. 또한 2020년 전후로 디지털 객체식별자의 적용이 확산되고 있는 바(한국 DOI 센터, 2024) 인용시 DOI를 포함한다면 안정성을 크게 향상시킬 수 있을 것으로 기대된다. 실

계 2019년 인용된 자원 중 DOI가 포함된 자원 모두에 접근할 수 있었다.

마지막으로 연구결과물을 생산하는 학회나 출판사 측면에서의 개선방안이다. 인터넷 자원의 안정성을 제고하기 위해 해당 기관에서 발간하는 연구 결과물에 포함된 인터넷 자원이 정확한 기술요소와 방식으로 인용되었는지 검증을 강화하는 노력이 필요하며, DOI와 같이 안정성을 개선시킬 수 있는 도구를 도입하여 적용하는 노력이 필요하다고 보인다.

이상 사이트 관리자, 연구자, 학술기관 등 학술 커뮤니케이션과 관련된 주체별로 인터넷 정보자원의 안정성을 개선시킬 방안에 대해 제시해보았다. 본 연구의 궁극적인 목적이 인터넷 정보자원의 안정성 제고를 통한 학술 커뮤니케이션의 효율화에 있는 만큼 관련자들의 관심과 노력을 기대해본다.

5. 결론

이 연구를 통해 인터넷 등장 이후 학술 커뮤니케이션에서 지속적으로 활용되고 있는 인터넷 정보자원의 안정성을 정량적, 정성적으로 검증해 보았다. 정량적인 관점에서의 안정성은 인용된 인터넷 자원에 포함된 인터넷 주소를 통한 접속가능여부를 기준으로 분석하였다. 접속가능성에 영향을 미치는 가장 중요한 요인은 시간의 경과임을 반영하여 각각의 검증 자원군(2009년, 2013년, 2019년 인용자료)을 발간후 5년, 10년, 15년으로 나누어 분석하였다.

접속이 되더라도 인터넷 주소의 변경이나 내용의 갱신 등이 발생한 경우 인용한 연구자가

참고하였던 인용시점의 자료와는 차이가 있을 수 있다는 점에 기반하여, 정성적인 관점에서의 안정성을 동일성 기준으로 분석하였다. 이러한 정성적 분석을 통해 인용한 연구자가 인용하였던 자원과 현재 접근이 가능한 자료의 동일성을 기준으로 안정성을 검증할 수 있었다. 정보자원의 안정성을 검증해보고자 하는 본 연구를 통해 밝혀진 내용을 요약하면 다음과 같다.

- 1) 접속가능성을 기준으로 하는 정량적 관점에서의 안정성 분석결과는 다음과 같다. 인터넷 정보자원의 안정성은 시간의 경과에 따라 저하된다. 접속가능성을 기준으로 하는 정량적 관점의 안정성은 발간 5년 시점까지 26.1%(2013년 자원)-27.4%(2019년 자원)의 자료가 접속불가해지는 것으로 분석되었으며, 5-10년 기간에는 5년 경과시점까지의 접속불가자료에 더하여 28.3%(누적 54.4%, 2013년 자원) 정도의 자원에 접근할 수 없게 되는 것으로 분석되었다. 또한 10-15년 기간에는 추가적으로 26.5%(누적 64.7%, 2009년 자원)의 자원이 접속불가한 상황으로 안정성이 저하되는 것으로 분석되었다. 15년 경과시 최초 인용된 인터넷 자원의 65% 내외의 자원에는 접속이 불가능 상태가 되는 것이다. 이는 인터넷 자원 인용 10년후 100% 접속불가한 것으로 조사되었던 선행연구나 자원의 생명주기상 9.3년이 중간값(9.3년이 되면 자원의 50%는 접속이 불가)이라는 선행연구에 비하면 안정성이 개선되었음을 보여준다. 기간별 안정성의 저하 정도는 시간의 경과와 어느정도 비례하는 것으로 분석되었다. 접속이 불가능해진 자원은 최초 인용자원수 대비 1-5년 기간에서 26%, 5-10년

기간에서 28.3%, 10-15년 기간에서 26% 정도 발생하여 안정성이 저하되는 것으로 분석되었다. 5년의 기간이 경과할 때마다 26-28%의 자원이 접속불가해지는 것으로 분석되어 시간 경과에 따라 일정한 정도로 안정성이 저하되는 모습을 보였다.

- 2) 접속이 가능한 자원을 대상으로 최초 인용된 자원과의 내용상 동일성을 기준으로 분석한 정성적 관점에서의 안정성 분석결과는 다음과 같다. 정성적 관점의 안정성은 5년 경과시 39.4%, 10년 경과시 16.8%, 15년 경과시 14.8%인 것으로 분석되었다. 이러한 분석결과에 따르면 정성적 안정성 역시 시간의 경과에 영향을 받는 것으로 판단할 수 있으며, 최초 자원대비 안정성이 10년이 경과된 시점 16.8%, 15년 경과시점 14.8%으로 분석된 결과는 인터넷 자원의 안정성이 여전히 낮은 수준임을 보여준다.
- 3) 이러한 분석의 과정에서 얻을 수 있었던 인터넷 정보자원의 안정성 개선을 위한 제안은 다음과 같다. 자원이 저장된 사이트 관리자는 사이트의 구성변경시 기존 주소에 대한 접속 대안을 강구할 것, 연구자 측면에서는 인용시

정확한 요소와 방식으로 인용할 것, 학술논문을 인용할 경우 안정적인 접속을 제공하는 디지털객체식별자(DOI)를 활용할 것, 인터넷에서만 구할 수 있는 자원을 인용할 경우 기관의 문서저장공간에 저장되어 있는 문서(주로 PDF 형태)를 인용할 것, 그리고 서비스를 이용하는 경우 시간의 경과에 따라 내용이 갱신될 가능성이 크므로 인용목적에 부합하는 화면을 캡처하여 제공할 것 등의 방안을 제시하였다.

본 연구는 학술 커뮤니케이션 과정에서 꾸준히 활용되고 있고 이제는 불가결한 학술정보원으로 위상을 확보한 인터넷 정보자원의 안정성을 정량적, 정성적 관점에서 분석하였다. 5년전 수행된 연구에서 활용하였던 데이터를 사용하여 일정 시간 경과 후 반복적으로 분석하여 시간의 경과에 따른 변화를 추적 분석하였다는 점에서 그 의의가 크다고 할 수 있다. 향후 후속 추적연구도 의미가 있을 것이라 판단된다. 이러한 연구를 통해 인터넷 정보자원의 안정성을 개선하여 궁극적으로 학술 커뮤니케이션 과정의 효율성 제고에 기여할 수 있을 것으로 기대한다.

참 고 문 헌

- 김성원 (2017). 연구자의 특성과 연구 분야가 인터넷 정보자원의 인용에 미치는 영향. 한국비블리아학회지, 28(3), 197-215.
- 김성원 (2019). 학술커뮤니케이션을 위한 인터넷 정보자원의 안정성에 대한 연구. 한국문헌정보학회지, 53(3), 45-62.
- 김홍렬 (2006). 인용문헌에 나타난 웹 자원의 접근성에 관한 연구. 정보관리학회지, 23(1), 181-200.

- 박준식, 박일중, 신상헌 (2000). 대학교수들의 학술정보원 이용행태분석: K대학의 사회과학교수집단을 중심으로. 정보관리학회지, 17(3), 21-43.
- 장선영, 정영미 (2012). 인용 웹 자원의 접근성과 유용성에 관한 연구: 한국문헌정보학회지를 중심으로. 한국정보관리학회 학술대회 논문집, 19, 137-142.
- 한국 DOI 센터 (2024). 한국 DOI 센터(KDC) 소개. 출처: <https://doi.or.kr/kdc/intro>
- Hennessey, J. & Ge, S. X. (2013). A cross disciplinary study and the effectiveness of mitigation techniques. BMC BIOINFORMATICS, 14. In Proceedings of the Tenth Annual MCBIOS Conference 2013. Columbia, MO, USA.
- Kumar, V. D. & Kumar, B. T. S. (2017). Finding the unfound: recovery of missing URLs through Internet Archive. Annals of Library and Information Studies, 64(3), 165-171.
- Oranus, T., Ali, A., Ali, S., & Hadi, S. (2011). Death of web citations: a serious alarm for authors. Malaysian Journal of Library & Information Science, 16(3), 17-29.
- Riahinia, N., Zandian, F., & Azimi, A. (2011). Web citation persistence over time: a retrospective study. Electronic Library, 29(5), 609-620.
- Teixeira da Silva, T. A. & Nazarovets, M. (2023). Archiving website-based references in academic papers: problems caused by reference rot, potential solutions and limitations. Learned Publishing, 36(3), 477-487.

• 국문 참고자료의 영어 표기

(English translation / romanization of references originally written in Korean)

- Jang, Sun-Young & Jung, Young-Mi (2012). A study on the accessibility and usefulness of cited web resource: based on journal of the Korean library and information science society. In Proceeding of the 19th Conference of Korean Society for Information Management, 19, 137-142.
- Kim, Hong-Ryul (2006). A Study on accessibility of web resources via citation analysis. Journal of the Korean Society for Information Management, 23(1), 181-200.
- Kim, Sungwon (2017). The impacts of author's demographic properties and research fields on citing Internet based resources in LIS field. Journal of the Biblia Korean Society for Library and Information Science, 28(3), 197-215.
- Kim, Sungwon (2019). A study on stability of internet information resources for academic communications in LIS domain of Korea. Journal of the Korean Society for Library and Information Science, 53(3), 45-62.

Korea DOI Center (2024). Korea DOI center (KDC) introduction. Available:
<https://doi.or.kr/kdc/intro>

Park, Joon-Shik, Park, Il-Jong, & Shin, Sang-Heun (2000). An analysis on the academic information seeking behavior of college professors: a social scientist group in the K-university. *Journal of the Korean Society for Information Management*, 17(3), 21-43.