

## 웹 아카이브 OASIS 개선방안에 관한 연구

남재우<sup>1</sup>, 이수영<sup>2\*</sup>

<sup>1</sup>건국대학교 문헌정보학과 조교수, <sup>2</sup>대림대학교 도서관미디어정보과 조교수

### A Study on the Improvement of Web Archive OASIS

Jae-Woo Nam<sup>1</sup>, Su-Young Lee<sup>2\*</sup>

<sup>1</sup>Assistant professor, Department of Library and Information Science, Konkuk University

<sup>2</sup>Assistant professor, Department of Library Media and Information Science, DaeLim University

**요약** OASIS는 온라인상의 디지털 정보자원을 체계적으로 수집·관리·보존하기 위해 2004년도부터 시작된 국립중앙도서관의 웹 아카이빙 프로젝트이다. 본 연구에서는 OASIS 웹 사이트의 접속 로그 분석과 이용자 설문조사를 실시하여 다음과 같은 문제점을 도출하였다. 첫째, OASIS 프로젝트에 대한 사람들의 인지도는 매우 낮았고, 웹 사이트의 첫 접속자가 많았다. 둘째, OASIS 이용률 제고를 위한 적극적인 홍보나 서비스 개편 등이 미흡했다. 연구에서는 이에 대한 개선점으로 자체적인 직접 홍보와 타 기관과 연계한 간접홍보를 강화할 것을 제안하였다. 또한 이용자 맞춤형 서비스를 통해 서비스를 고도화하고, 흥미와 재미를 유발하는 콘텐츠 강화할 것을 제안하였다.

**주제어** : 웹 아카이빙, 웹아카이브, 오아시스, 국립중앙도서관, 로그 분석

**Abstract** OASIS(Online Archiving & Searching Internet Sources) is a web archiving project of the National Library of Korea started in 2004 to systematically collect, manage, and preserve online digital information resources. In this study, the following problems were derived by analyzing the access log of the OASIS website and conducting a user survey. First, people's awareness of the OASIS project was very low, and there were many first-time visitors to the website. Second, active promotion and service reorganization to improve the use of OASIS was insufficient. The study suggested that the improvement point of this was to strengthen its own direct promotion and indirect promotion in connection with other agencies. In addition, it was proposed to enhance the service through user-customized services and to reinforce content that induces interest and fun.

**Key Words** : Web Archiving, Web Archives, OASIS, National Library, Log Analysis

## 1. 서론

### 1.1 연구목적 및 필요성

웹은 1993년 첫 웹 사이트가 생성된 이후 현재까지 정보 전달, 공유, 소통, 협력 등 다양한 역할을 수행하고 있다. 특히 공공기관의 웹 자원은 관리하는 시스템이 기관마다 다르고, 웹 기록물이 여러 곳에 분산되어 있을 뿐만 아니라, 웹 사이트의 수명이 짧기 때문에 체

계적 수집과 보존관리가 쉽지 않다. 이에 국립중앙도서관에서 2004년도부터 온라인상의 디지털 정보자원을 체계적으로 수집·관리·보존하기 위해 시작된 웹 아카이빙 사업이 바로 OASIS(Online Archiving & Searching Internet Sources, 이하 OASIS)이다[1].

웹 자원 아카이빙(이하 웹 아카이빙)에 대한 연구는 OASIS의 운영에 발맞추어 확대되고 있다. 2004년 웹 아카이빙 필요성에 대한 연구를 시작으로 정책 제안 및

\*Corresponding Author : Soo-Young Lee(sulibrary@daelim.ac.kr)

법제도 분석, 선정 지침 및 개선안, 콘텐츠 특성, 인터페이스 유형 및 특성 분석 등 서비스 사용성으로 연구의 범위가 확장되고 있다. 이와 같은 연구는 크게 세 가지로 나누어 살펴볼 수 있다.

첫째 OASIS의 수집자원 및 시스템 운영 방안에 대한 연구이다. 이는 OASIS의 전반적인 개선 방안을 시작으로 수집자원에 대한 선정 지침 및 인터페이스에 이어 재난기록의 아카이브 운영에 대한 연구로 확대된다. 다수의 연구를 통해 수집자원의 양적 성장에 초점을 맞추는 것이 아니라 국가적 디지털자산으로 적합한 콘텐츠 가치를 가진 웹 자원이 수집되기 위한 지속성, 최신성, 정확성에 대한 개선 방안을 확인할 수 있다[2-5]. 둘째, 이용자를 중심으로 하는 아카이브의 관리 및 이용자 서비스에 대한 연구이다. 이용자의 '유사기관과 연계된 통합 검색서비스' 등 요구사항과 이용자 참여 유형을 분석한 연구 등을 통해 웹 아카이브 구축을 위해서는 이용자가 쉽고 자신만의 서비스로 활용할 수 있도록 유연성, 안정성, 표준성을 제공하여야 하는 필요성을 확인할 수 있다. 이는 독점적 기관 중심 관리에서 개방적 이용자 활용 중심으로 전환되는 패러다임의 반영을 기반으로 하고 있다[6-7]. 셋째, 웹 아카이빙을 기반으로 공공데이터, 정보공개, 웹 아카이브 등에 대한 법적·정책적 고찰로 구분된다[8]. 이렇듯 OASIS의 수집정책이나 콘텐츠에 대한 분석 및 정책 연구는 간헐적으로 수행되었고, 국가기록원 등 타 아카이브 이용에 대한 연구는 수행되었으나, OASIS를 대상으로 사용성을 파악하고 이에 대한 개선방안을 제안한 연구는 아직까지 부재한 상태이다.

이에 본 연구는 선행 연구를 토대로 오아시스에 대한 사용성과 활용성을 확대하기 위한 방안을 고찰해 보고자 한다. 특히 웹 로그분석을 통해 오아시스 이용자의 실질적인 이용행태를 파악하고, 별도의 설문조사를 통해 웹 아카이빙과 오아시스 이용에 대한 인식조사를 실시하였다. 이는 실제 이용에서 나타나는 사실적 현상을 분석하고 이에 대한 이용자의 심리적 근거를 발견하고자 하는 상호보완적인 방법으로 기존 이론적 연구와의 차별점이라 할 수 있다. 연구 목적은 다음과 같다.

- OASIS 웹 사이트의 로그 분석을 통해 이용자의 이용패턴을 확인하고 문제점을 파악한다.
- OASIS에 대한 인식 조사를 통해 사용자의 구체적인 이용과 웹 아카이빙에 대한 인식 등을 조사한다.

· 로그 분석과 인식 조사 결과를 분석해 문제점을 진단하여 개선점을 제시한다.

## 1.2. OASIS 개요

OASIS는 국립디지털도서관 건립사업 전까지 온라인상의 디지털 정보자원을 체계적으로 수집, 관리, 보존하기 위한 국립중앙도서관의 대표적인 사업으로, 2008년 국제 인터넷 보존 컨소시엄(International Internet Preservation Consortium, 이하 IIPC) 정회원 가입 전까지 국내 기술을 활용한 독자적인 프로젝트로 추진되었다[10]. 초기에는 보존 가치가 있는 국내 웹 사이트, 웹 문서들을 기계적인 방식으로 자동 크롤링(crawling)하는 방식과 정부 기관 등이 웹 상의 게시판을 통해 공개하고 있는 중요 문서들을 수집 후 목록 작업을 거쳐 등록 및 서비스를 수행하였다. 또한 초기 OASIS 파일 형태의 디지털 정보자원을 관리하고 장기 보존하기 위해 주요 정보통신기술 관련 제조사, 기술 자원, 파일 포맷, 소프트웨어에 대한 정보를 수집하고 관리하였다. 이러한 정보들은 디지털 정보 자원의 장기 보존을 위해 보존 대상 디지털 객체와 함께 메타데이터 형태로 패키징 되어 2009년 이후 디지털 장서관리시스템에 관리 기능이 흡수되었다.

이후 OASIS 관리시스템의 가장 큰 변화는 IIPC에서 제공하고 있는 오픈 소스 소프트웨어의 도입과 저장 파일 포맷을 ISO 28500 표준인 WARC(Web Archive Data Format)으로 전환하며 디지털 장서관리시스템으로 이관한 것이다. IIPC의 오픈 소스 소프트웨어는 크롤링과 하베스팅(Crawling and harvesting), 웹 아카이브 디스플레이(Display/view web archive), 분석(Analytics)으로 이루어진다. IIPC 정회원 가입 후 국립중앙도서관에 도입된 소프트웨어는 웹 크롤러인 Heritrix와 아카이빙된 파일을 보기 위한 Wayback이다.

OASIS는 2004년도부터 추진하고 있는 웹 자원 아카이빙 사업의 수집 방향을 전환하고 선택적 수집을 체계화하기 위해 2013부터 2014년까지 OASIS 시스템 기능개선 사업을 수행하였다. 2015년에는 국가 재난·이슈·주제 분야에 대한 아카이브를 구축하면서 전면적으로 개편하였다. 이를 위해 재해·재난 분류체계를 확립하고 구축 목적과 수집 범위를 명확히 하기 위한 아카이브 구축 방향을 설정하였으며, 재난 유형별 주관기관으로 자료 수집을 시작하였다. 선정 기준에 따라 국

가 재난, 선거, 올림픽 등 국가적으로 중요한 이슈, 인문·사회·과학 주제 분야별 주요한 테마와 컬렉션, 전문가 추천 이슈와 테마 등을 구축하고 있다.

## 2. 연구 방법

### 2.1 연구 설계

연구목적을 달성하기 위해 웹 로그 분석을 통해 이용자의 이용 패턴을 확인하고, 설문조사를 실시하여 OASIS의 접속과 이용, 인식에 대한 조사를 실시하였다.

### 2.2 웹 로그 분석 방법 및 내용

로그 분석이란 로그데이터를 이용하여 트래픽을 파악하고, 이 트래픽이 지닌 의미를 분석해 나가는 것이다[11]. 로그 분석은 크게 웹 서버 로그파일 방식, 페이지 태깅방식, 패킷 스니핑 방식의 세 가지로 분류할 수 있다[11]. OASIS 웹 사이트 로그 분석은 웹 서버 로그 파일 방식으로 수행되었다. 이는 일정 기간 기록된 데이터를 분석하는 배치형 분석에 많이 사용한다. 또한, 서버 수가 적고 서버 관리자와 분석가가 일치하거나 동일한 조직에 속한 경우 많이 사용된다[11].

Table 1. OASIS log analysis content

item	analysis content
funnel	funnel analysis such as search engines
new visit	analysis of the number of new visitors
page view	number of web pages viewed by visitors
popular page	pages visited by many visitors
duration time	analysis of dwell time by page
frequency of visits	How often users visit the website

본 연구에서는 실시간 웹 로그가 아닌 OASIS 단일 서버에 일정 기간 누적된 로그 파일을 분석하였으며, 로그파일의 형식에 따라 Table 1과 같이 유입 경로, 신규 방문, 페이지 뷰, 인기 페이지, 체류 시간, 방문 빈도의 7개 항목으로 구분하여 분석하였다. 분석 기간은 2018년 7월부터 2019년 9월까지 15개월이다.

### 2.3 이용자 인식조사 방법 및 내용

OASIS에 대한 일반적 인식을 조사하기 위해 일반 이용자를 대상으로 설문조사를 실시하였다. 설문조사 기간은 2020년 5월 4일부터 22일까지 총 18일이며, 온라인 설문을 통해 수행하였다. 설문지에는 웹 아카이

빙의 개념과 OASIS 프로젝트에 대한 개요, 용어 설명 등을 함께 작성하여 응답자가 설문 내용에 대해 잘 인지할 수 있도록 조치하였다. 조사내용은 Table 2와 같이 OASIS의 이용과 인식 등 10개의 문항이다.

Table 2. Survey content

division	survey item
connect	connection frequency connection path
use	purpose of use content of interest Intention to use again willingness to recommend to acquaintances
recognition	OASIS recognition OASIS cognitive pathway the importance of web archiving web resource types to be collected and preserved

## 3. 연구 결과

### 3.1 로그 기록 분석 결과

#### 3.1.1 유입 경로와 신규 방문

이용자가 OASIS에 접속하는 경로를 나타내는 유입 경로 분석 결과 Table 3에서와 같이 포털사이트를 통한 검색 유입이 43.9%를 나타내고 있었다.

Table 3. OASIS website funnel

item	%	specific funnel
search engine	43.9	portal site search
external domain	32.9	links to other websites
direct inflow	17.3	URL input, bookmark
internal inflow	5.7	inside the website
SNS	0.2	blog, Facebook, etc.

가장 많이 이용하는 포털사이트는 네이버(64.6%)였으며, 뒤를 이어 다음(30.3%), 구글(4.7%), 기타(0.4%)의 순으로 나타났다. 외부도메인 유입은 타 웹 사이트에서 OASIS의 연결을 제공하기 위한 링크 등으로 32.9%의 유입률을 나타냈고, 이용자가 브라우저 주소창에 OASIS 웹 사이트 주소를 직접 입력하거나 즐겨찾기 등을 통해 유입되는 직접 유입은 17.3%를 나타냈다. 내부 유입은 유입을 추적하는 스크립트가 없는 웹페이지에 처음 방문 후 스크립트가 있는 페이지로 이동한 경우가거나 이미 열어 놓았던 페이지의 방문 세션이 일정 시간 이후 끊겼다가 다시 연결되는 경우이며 5.7%의 유입률을 나타내고 있었다. 마지막으로 블로그나 카페 등을 SNS를 통해 유입되는 경우는 0.2%로 네이버 52.2%, 다음 26.7%, 페이스북 15% 순으로 유입이 많았다.

Table 4. New visit to the OASIS website

quarter	all visitor	new visitors	new visit rate
2018. Q3	5,226	4,522	86.53%
2018. Q4	13,525	11,179	82.65%
2019. Q1	15,461	13,322	86.17%
2019. Q2	32,325	28,333	87.65%
2019. Q3	58,434	41,229	70.56%
sum/average	124,971	98,585	78.89%

방문 수를 살펴보면 Table 4에서와 같이 전체 방문 수가 전반적으로 증가추세를 보이고 있었다. 신규 방문 수도 전체 방문 수와 비슷한 추이로 증가 추세를 보이고 있었는데 5개 분기 평균 신규 방문율은 78.89%로 높은 값을 나타내고 있었다.

### 3.1.2 페이지 뷰

페이지뷰는 1회 방문 시 열람한 페이지를 나타내며, 방문당 페이지 뷰는 페이지뷰를 방문 수로 나눈 결과 값이다.

Table 5에서와 같이 접속 수의 꾸준한 증가와 함께 페이지뷰는 증가세를 나타냈으나, 방문당 페이지뷰는 2018년 3분기부터 일시적인 상승을 제외하고 지속해서 감소하는 추세를 나타내고 있었다. 최종 5개 분기를 종합하면 이용자들은 OASIS에 접속해서 평균 1.79개의 페이지를 열람하고 있었다.

Table 5. OASIS website page view

quarter	all visitor	page view	page views
			per visit
2018. Q3	5,226	12,760	2.44
2018. Q4	13,525	38,705	2.86
2019. Q1	15,461	34,163	2.21
2019. Q2	32,325	50,921	1.58
2019. Q3	58,434	87,543	1.50
sum/average	124,971	224,092	1.79

### 3.1.3 선호페이지와 체류시간

선호페이지는 방문자의 접속이 높은 페이지를 의미하며 현황은 Table 6과 같다. 분석 결과 웹 사이트 보존파일 검색, 주제검색, 메인화면, 자료검색, 컬렉션, 가나다 검색, OASIS 시스템 소개, 재난 아카이브(시간순)의 순서로 접속률이 높았다. 방문당 페이지뷰는 웹 사이트 보존파일 검색이 2.99페이지로 가장 높았으며, OASIS 소개, 재난 아카이브 순으로 높게 나타났다.

Table 6. OASIS preference page and time of stay

ranking	page title	per visit	
		page view(%)	duration time
1	search archived website	2.99(8.06)	6초~15s
2	subject search	1.31(6.50)	5s under
3	main screen	1.69(5.43)	5s under
4	data search	1.90(4.23)	5s under
5	collection-time order	1.73(3.19)	5s under
6	srarch-alphabetically	1.39(2.74)	5s under
7	OASIS system Introduction	2.58(1.96)	5s under
8	disaster Archive-time order	2.25(1.82)	5s under

선호페이지의 방문당 체류시간(페이지별 이용자가 체류한 시간의 평균값)을 살펴보면 자료검색이 최대 15초 가장 높게 나타났으며, 그 외 5초 이하로 나타났다.

### 3.1.4 방문 빈도

방문 빈도는 이용자의 OASIS 방문 빈도를 나타낸다. 분석 결과 Table 7에서와 같이 1회 방문이 83.9%, 2회 방문 5.4%, 3~5회 방문 1.8%, 10회 이상 방문이 8.2%를 차지하고 있었다. 2회 이상 방문 시부터 페이지뷰와 체류시간이 소폭 증가하고 있으며, 10회 이상 방문하는 경우 페이지뷰와 (2.39페이지), 체류시간(6.44초)이 가장 높게 나타났다.

Table 7. Frequency of visits to the OASIS website

frequency	all visitor	visit rate	per visit	
			page view	duration time
1 time	9,285	83.94%	1.23	17s
2 time	605	5.44%	1.67	45s
3~5 time	202	1.83%	1.69	34s
6~9 time	61	0.55%	1.49	54s
10 times over	912	8.24%	2.39	6:44s
sum/average	11,062	100%	1.36	51s

## 3.2 로그 분석 결과의 시사점

OASIS 웹 로그 분석을 통한 시사점은 다음과 같다. 첫째, 이용자의 OASIS 유입 경로는 포털사이트의 검색유입이 43.9%로 가장 많았다. 포털사이트 중 네이버(64.6%)와 다음(30.3%)이 가장 많이 유입되는 경로이며 구글(4.7%)을 통한 유입은 적은 수치를 나타내고 있었다. 그 이유는 포털사이트에서 'OASIS' 또는 '오아시스'를 키워드로 검색할 경우 네이버에서는 OASIS 웹 사이트가 상위 순위 안에 노출되며, 다음에서는 후 순

위에, 구글에서는 다음보다 더욱 후 순위에 노출되기 때문이다. 포털사이트의 검색 결과에 노출되기 위해서는 OASIS 웹 사이트를 각 포털사이트에 등록하거나, 검색엔진의 크롤링을 위해 OASIS 웹 사이트의 메타태그에 키워드를 작성하는 방법이 있다. 이를 검색엔진 최적화(Search Engine Optimization, SEO)라고 한다. 현재 OASIS 웹 사이트의 메타태그는 '국립중앙도서관' 'OASIS', '도서관', '웹 검색', '아카이빙'으로 작성되어 있으나 여기에 '오아시스' 한글 키워드를 포함시키고 '웹검색'과 같은 불필요한 키워드를 '웹 트랜드'와 같은 연관 키워드로 대체하는 등 개선이 필요하다. OASIS의 주체인 국립중앙도서관은 다양한 방법을 통해 주요 포털사이트의 상위 검색 결과에 나타날 수 있도록 적극적으로 노력해야 할 것이다.

둘째, 신규 방문 수 증가에 따른 접속 수 증가로 페이지 뷰는 지속해서 증가하고 있으나, 방문 당 페이지 뷰는 1.79로 감소세를 나타내고 있다. 이는 OASIS에 접속해서 2페이지 이상을 열람하지 않는다는 것으로 이용자가 콘텐츠에 흥미를 느끼지 못하고 웹 사이트를 바로 이탈하는 것으로 해석할 수 있다. 신규 방문자의 증가는 웹 사이트나 콘텐츠에 대한 이용자의 요구가 있는 것으로 볼 수 있으므로, OASIS 사이트의 디자인을 개선하거나 이용자가 흥미를 느낄 수 있는 콘텐츠를 메인화면에 배치하는 등 이탈 방지를 위한 노력이 필요하다.

셋째, OASIS의 웹 사이트 보존파일 검색은 이용자가 가장 선호하고 체류시간이 높은 페이지로 분석되었다. 그 외 주제 검색, 메인 화면, 자료 검색, 시간 순 컬렉션, 가나다 검색 등이 방문당 페이지 뷰가 높은 것으로 분석되었다. 따라서 이용자가 선호하고 비교적 오래 체류하는 페이지는 대체로 검색 관련 페이지로 볼 수 있다. 이용자가 검색어를 입력하고 결과를 얻는 것은 OASIS와 이용자 간의 상호작용이라고 할 수 있다. 온라인 시스템의 상호작용성은 이용자의 체류 시간을 증가시키고 지속적 방문을 유발하는 것으로 연구되었다 [12-13]. 온라인 쇼핑몰의 경우 높은 체류시간은 구매에 긍정적 영향을 미치고 있었으므로 [14-15], 검색관련 페이지는 OASIS 웹 사이트에 긍정적 역할을 할 것으로 추측된다. 그러나 원하는 검색 결과가 빠르게 제공되지 않을 때도 검색화면에 체류하는 시간이 길어질 수 있으므로 좀 더 면밀한 분석이 필요할 것이다.

넷째, 방문 빈도를 살펴봤을 때 OASIS에 1회 방문

자가 83.9%로 이용자의 대부분을 차지하고 있었고, 10회 이상 방문자(8.24%), 2회 방문자(5.4%)의 순으로 집계되었다. 1회 방문자는 첫 방문자를 나타내며 방문당 체류시간은 17초로 매우 낮게 나타났으므로, 이들 첫 방문자는 OASIS 웹 사이트 접속 후 곧바로 사이트를 이탈한다고 볼 수 있다. 따라서 OASIS 웹 사이트는 이용자가 재방문할 수 있는 콘텐츠를 마련하는 등 대중적인 활용도를 높일 방안을 마련해야 할 것이다.

### 3.3 설문조사 결과

#### 3.3.1 모집단의 구성

설문조사를 통해 OASIS 이용 현황과 인식을 분석하였고 응답자의 인구통계학적 구성은 Table 8과 같다.

전체 1,283명의 설문 응답자 중 여성이 58.5%, 남성이 41.5%를 나타내고 있으며, 연령대는 30대(47.8%), 20대(21.3%), 40대(21.2%), 50대(6.9%), 60대 이상(2.9%)의 순으로 나타났다. 직업별로는 회사원이 64.9%로 가장 많았으며, 뒤를 이어 대학생(21.3%), 연구원(4.8%), 대학원생(2.2%), 기타(6.6%) 순으로 조사되었다. 기타 응답자는 구직자, 주부 등 무직자로 구성되어 있다.

Table 8. Population composition (n=1,283)

	division	n	%
gender	female	750	58.5
	male	533	41.5
age	20s	273	21.3
	30s	613	47.8
	40s	272	21.2
	50s	88	6.9
	over 60s	37	2.9
job	office worker	762	59.4
	college student	273	21.3
	educational facilities	71	5.5
	researcher	62	4.8
	postgraduate student	28	2.2
	etc	87	6.8

#### 3.3.2 웹 아카이빙에 대한 인식

웹 아카이빙에 대한 인식 조사를 위해 Table 9과 같이 웹 아카이빙 개념에 대한 사전 인지 여부, 인지 경로, 웹 아카이빙 중요성, 보존 대상 웹 자원에 대해 질문하였다.

**Table 9. Awareness of web archiving**

	division	n	%
awareness of web archiving	know	734	57.2
	do not know	549	42.8
web archiving awareness path (n=734)	News, newspapers, etc.	238	32.4
	while web surfing	192	26.2
	acquaintance	158	21.5
	books, journals, etc.	82	11.2
importance of web archiving (n=1,283)	national library	64	8.7
	important	1,109	86.4
	not important	30	2.3
web resources to be collected (n=3,272, multiple response)	do not know	144	11.2
	public institutions	889	27.2
	National issues, events	847	25.9
	Major companies, media, etc.	403	12.3
	non-profit organization	398	12.2
	by academic subject	356	10.9
	education, research institute	221	6.8
	all Korean websites	158	4.8

웹 아카이빙에 대해서는 57.2%(734명)가 사전에 그 개념을 알고 있었으며, 42.8%는 잘 모르고 있는 것으로 조사되었다. 다음으로 웹 아카이빙 사전 인지 응답자 734명을 대상으로 인지 경로에 대해 파악하였고, 그 결과 뉴스, 신문 등 미디어를 통한 인지(32.4%)가 가장 높았으며, 다음으로 웹 서핑 중 우연히(26.2%), 친구, 동료 등 지인(26.2%), 도서, 학술지 등(11.2%), 국립중앙도서관을 통해서 알게 된 경우가 8.7%로 나타났다. 웹 아카이빙의 중요성에 대해서는 중요하다(86.4), 중요하지 않다(2.3%), 잘 모르겠다(11.2%) 순으로 조사되었다.

수집해야 할 웹 자원에 대한 인식(복수응답)에서는 국가·공공기관의 웹 자원이 27.2%로 가장 높게 나타났으며, 선거, 재난, 행사 등 국가적 이슈(25.9%), 국내 주요기업·언론(12.3%), 민간·시민단체 등 비영리조직 웹 자원 (12.2), 인문·사회·과학 등 주제별 웹 자원(10.9%), 교육·연구기관 웹자원(6.8%), 대한민국 웹 사이트 전체 (4.8%) 순으로 조사되었다.

3.3.3 OASIS의 접속

OASIS 웹 사이트의 접속 빈도를 분석한 결과 Table 10과 같이, 응답자의 75.6%가 첫 방문이었으며, 월 1~2회(15.4%), 주 1~2회(8.1%), 매일 접속(0.9%)의 순으로 조사되었다.

**Table 10. Access the OASIS website**

	division	n	%
connection frequency (n=1,283)	first access	970	75.6
	1-2 times a month	197	15.4
	1-2 times a week	104	8.1
	almost everyday	12	0.9
connection method (n=313)	search engine	213	68.1
	link to National Library	86	27.5
	URL input, bookmark	12	3.8
	SNS	2	0.6

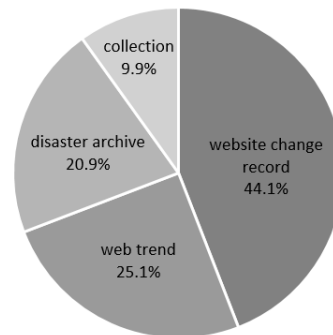
처음 접속자를 제외한 313명의 응답자를 대상으로 OASIS 접속 방법에 대해 질문하였다. 그 결과 포털사이트 검색을 통한 접속이 62.7%로 가장 높았으며, 국립중앙도서관 홈페이지 내 OASIS 링크를 통한 접속이 30.6%, 도메인 직접 입력 또는 즐겨찾기를 통한 직접접속이 5.8%, 국립중앙도서관의 블로그나 페이스북 등 SNS를 통한 접속이 0.9%로 나타났다.

3.3.4 OASIS 이용에 대한 인식

OASIS 콘텐츠의 사용 목적을 조사하기 위해, OASIS 사용 경험이 있는 응답자에게는 실제 목적을, 처음 접속자에게는 예상되는 사용 목적을 응답받았다. 그 결과, Table 11과 같이 최신 정보 습득(32.0%), 업무 활용(28.4%), 유희적 목적(22.3%), 학업 및 학술 목적(17.4%) 순으로 조사되었다.

**Table 11. Use of OASIS (n=1,283)**

	division	n	%
purpose	information acquisition	410	32.0
	business use	364	28.4
	playful purpose	286	22.3
	academic purpose	223	17.4
future use	Will use	1,121	87.4
	Will not use	162	12.6
recommended	recommendation	1,030	80.3
	Not recommended	253	19.7



**Fig 1. Most Interest Content in OASIS**

OASIS에서 가장 관심 있는 콘텐츠는 Fig. 1에서와 같이 웹 사이트 변경기록(44.1%), 웹 트렌드(25.1%), 재난 아카이브(20.9%), 재난 컬렉션(9.9%) 순으로 조사되었다. 그리고 87.4%의 응답자가 향후 재 이용 의사를 밝혔고, 19.7%의 응답자가 주변 지인에게 OASIS를 추천할 의사가 있다고 응답하였다.

### 3.4 설문조사 결과의 시사점

일반 이용자의 OASIS 접속과 콘텐츠 이용분석은 현재 이용자들의 사용 현황을 진단하고 OASIS의 문제점을 파악하여 사용성 증가 방안을 마련할 수 있는 기초 자료로 활용될 수 있을 것이다. 설문조사 결과 다음과 같은 시사점을 도출할 수 있다.

첫째, 웹 아카이브의 개념과 필요성 등에 대해 많은 사람이 인식하고 있지 못하고 있었다. 웹 아카이빙 개념에 대해서는 응답자의 57.2%만 사전에 인지하고 있었고, 43.8%의 응답자는 알고 있지 못했다. 그러나 웹 아카이빙의 중요성에 대해서는 응답자의 86.4%가 중요하다고 응답하였다(설문지에 웹 아카이빙의 개념을 명시). 이를 종합하면 아직까지 많은 사람들이 웹 아카이빙에 대해 잘 모르고 있었지만, 그 개념을 알게 되면 중요성을 인식할 가능성이 높을 것으로 예상해 볼 수 있다. 웹 아카이브의 개념과 필요성 등을 사람들에게 인식시키는 일은 OASIS 프로젝트에 대한 근거가 되며 OASIS 웹 사이트의 실질적 이용으로 이어질 것이다.

둘째, OASIS 프로젝트에 대한 적극적인 홍보가 필요하다. 설문 응답자의 76.5%가 설문조사를 통해 OASIS를 처음 접속하였으므로 실제 많은 이용이 없는 것으로 볼 수 있다. 그러나 87.4%의 응답자가 향후 이용 의사를 보였고, 80.3%의 응답자가 주변에 추천할 것이라고 응답하였다. 따라서 많은 사람이 아직까지 OASIS 프로젝트에 대해 모르고 있지만, 그 중요성과 사이트의 기능을 알게 되면 이용과 방문이 늘어날 것으로 예상된다. 따라서 국립중앙도서관은 홈페이지 및 SNS를 통해 오아시스를 소개하고 참여형 이벤트 등을 통해 많은 사람이 직접 이용해 볼 수 있도록 노력해야 할 것이다.

셋째, OASIS의 수집대상 웹 자원은 OASIS 자원 수집지침에 근거하되, 우선 순위를 고려할 필요가 있다. 수집대상 웹 자원에 대한 중요도는 국가-공공기관의 웹 자원(27.2%)과 재난-선거 등 국가적 이슈 관련 웹 자원(25.9%) 등 국가기관이나 국가적 이슈 관련 웹 자원에

대한 아카이빙이 50% 이상을 차지하였다. 따라서 사람들은 대체로 국가기관이나 공공기관의 업무에서 생산된 웹 자원을 보존대상으로 중요하게 생각하고 있음을 알 수가 있다.

넷째, OASIS 웹 사이트는 유희적 요인을 내포하고 있으며 이는 이용자의 관심 요소이다. OASIS 이용에 대해서는 정보습득, 업무 활용, 학업 및 학술 목적이 각각 25~32% 구간에서 비슷한 분포로 나타났으며 유희적 목적도 22.3%로 조사되었다. 유희적 요인은 웹 사이트에서 느끼는 흥미와 재미로 감성적 요인이며 이에 대한 만족도가 높을 경우 체류시간이 늘어나고[16], 브랜드 태도에 긍정적 영향을 미치는 것으로 연구되었다[17]. 또한, 웹 사이트 만족도에 영향을 미치고[18], 웹 사이트 사용의 동기유발 요인으로 작용하기도 한다[19].

이용자가 OASIS에서 흥미-재미를 느낄 수 있는 요인은 아카이빙 된 과거의 웹 사이트를 볼 수 있거나 웹 트렌드를 확인해볼 수 있기 때문으로 추측된다. 이는 OASIS의 관심 콘텐츠에 대한 조사에서 확인해 볼 수 있는데 웹 사이트 변경기록(44.1%), 웹 트렌드(25.1%) 등 과거부터 현재까지의 모습과 경향을 알 수 있는 콘텐츠들에 대해 많은 69.2% 응답자들이 관심을 가지고 있었다. 따라서 웹 아카이브는 다양한 목적으로 사용될 가능성이 있으며 특히 흥미와 재미 요소를 적절히 유지한다면, 좀 더 많은 사람이 관심을 가질 수 있게 될 것이다.

## 4. 결론 및 제언

OASIS 접속 로그 분석과 설문조사를 종합하여 다음과 같은 문제점을 도출하였다. 첫째, 웹 아카이빙과 OASIS 프로젝트에 대한 사람들의 인지도가 매우 낮았다. OASIS 웹 사이트는 첫 접속자가 많았고, 이탈률이 높은 문제점이 있다. 둘째, OASIS 이용률 제고를 위한 적극적인 홍보 활동이나 서비스 개편 등이 이루어지지 않고 있다. 이에 대한 개선점을 제안하면 다음과 같다.

### 4.1 OASIS 프로젝트 홍보

OASIS 사이트는 로그 분석과 설문조사에서 신규접속자가 75% 이상이었으며, 체류 시간이 매우 적었다. 또한, OASIS의 접속은 포털사이트 검색을 통한 접속이 가장 많지만, 실제 포털사이트 검색 시 사이트가 잘 검색되지 않는 등 문제점도 존재한다. 따라서 대내적으로는 검색 엔진 최적화(SEO)와 같은 웹 마케팅 기법을 통해 인

터넷에서 웹 사이트가 잘 인식되도록 노력해야 할 것이다.

한편 설문조사 결과에서는 50% 이상의 응답자가 웹 아카이빙과 OASIS 프로젝트에 대해 알고 있지 못했다. 이는 2004년도에 시작된 OASIS 프로젝트 역사와 비교해 매우 약한 인지도와 이용률이다. 따라서 대외적으로 OASIS의 목적과 활용성에 대한 적극적인 홍보가 필요한 시점이며, 그 방법으로 국립중앙도서관의 SNS 홍보, 참여형 이벤트 등을 통한 직접 홍보 방안을 제안해 볼 수 있다. 또한, 공공기관의 기관소개(연혁) 페이지에 OASIS에 아카이빙된 과거 웹 사이트의 링크를 제공하거나, 국가기록원 기록정보서비스로 웹 사이트 변경기록 콘텐츠를 제공한다면 훌륭한 간접홍보가 될 것이다.

#### 4.2 OASIS 서비스 고도화

로그 분석 결과 OASIS 웹 사이트에서 사람들이 가장 많이 방문하는 페이지는 주로 검색 관련 페이지였다. 설문조사 결과에서 이용자들은 최신 정보 습득(32.0%) 및 업무(28.4%)와 학업·학술(17.4%) 목적으로 OASIS를 가장 많이 이용할 것으로 분석되었다. 이 두 가지 결과를 연결하면 이용자는 결국 정보검색(최신정보, 업무정보, 학업·학술정보) 목적으로 OASIS를 이용하는 것으로 볼 수 있다. 한편, 대부분의 설문 응답자는 OASIS를 재이용하고 주변 지인에게 추천할 의향을 가지고 있지만, 추천 의향이 없다는 응답자도 20% 가까이 존재하고 있었다. 이는 개인적으로 OASIS의 필요성이 인정되고 다시 방문해서 좀 더 세부적인 내용을 탐색할 수는 있지만, 다른 사람에게 추천할 파급효과까지는 연결되지 못하는 OASIS의 한계점을 나타내고 있다.

OASIS 프로젝트가 시작된 지 15년이 지난 시점이다. OASIS는 그동안 서버 증설 및 수집기 도입, 표준화 등 웹 자원의 수집과 보존이라는 일차적 목적 달성을 위해 노력해 왔다. 이제는 업무 및 연구 활용성, 과거 기록의 열람 등을 초점으로 이용자가 원하는 서비스를 제공하고 있는지 검토할 시점이다. 현재 OASIS는 크게 검색, 컬렉션, 웹 트렌드의 세 가지 서비스를 제공하고 있으나, 서비스의 기능적인 부분을 포함하여 이용자 중심의 개방적인 서비스가 될 수 있도록 서비스 고도화 작업이 필요하다.

#### 4.3 유해적 요소 강화

웹 서비스에서 이용자가 흥미를 느끼는 요인은 웹 사이트의 체류 시간, 만족도, 동기 유발 등에 있어 긍정

적인 영향을 가져온다[16-19]. 설문조사에서 이용자가 가장 관심 있는 콘텐츠는 웹 사이트 변경 기록(44.1%)과 웹 트렌드(25.1%), 재난 아카이브(20.9%) 등 이었다. 특히 웹 사이트 변경 기록은 로그 분석에서도 이용자가 가장 많이 접속하고 오래 체류하는 페이지로 분석되었다. 웹 사이트 변경 기록은 현재와 다른 과거 웹 사이트를 열람할 수 있는 서비스이며, 웹 트렌드는 Fig. 2에서와 같이 키워드의 관심도 변화를 시간의 흐름에 따라 검색해 볼 수 있는 메뉴이다.

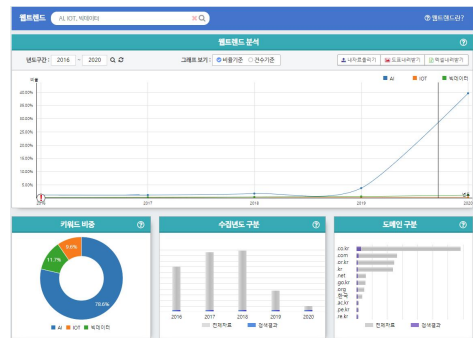


Fig 2. Keyword trend analysis of Oasis

이용자에게 관심 있는 모든 콘텐츠가 유해적 요소를 내포한다고 볼 수는 없지만, 웹 사이트 및 키워드 트렌드 변화를 직접검색하고 열람하는 것은 알지 못하는 과거를 돌이켜본다는 점에서 타 콘텐츠에 비해 흥미 요소가 강할 것이다. 따라서 이와 같은 '재미있는' 콘텐츠를 좀 더 부각하고 타 콘텐츠와 연계하거나 이용자가 참여할 수 있는 서비스를 개발한다면 웹 사이트의 사용성에 긍정적 영향을 가져올 것이다. 예를 들면 OASIS에 로그인 시스템을 구축하고 이용자의 학교나 직장 등 정보를 입력하여 '나와 연관 있는 웹 사이트 기록보기' 서비스 제공한다면 과거를 회상해 볼 수 있는 개인화 서비스가 될 것이다. 또한 개인의 관심 주제를 등록하고 로그인 시 관심 주제와 연관된 키워드의 트렌드 분석 화면을 제공하는 것도 이용자가 흥미를 느낄 수 있는 맞춤형 서비스의 좋은 방법이 될 것이다.

## REFERENCES

- [1] OASIS. (n.d.), *Online Archiving & Searching Internet Sources*, National Library of Korea. (Online). <http://www.oasis.go.kr>



[2] C. O. Yoon. (2010). Research on the OASIS, a web archive in Korea. *Journal of the Korean Society for Library and Information Science*, 44(2), 5-27. DOI : 10.4275/KSLIS.2010.44.2.005

[3] C. O. Yoon. (2014). An analysis of the contents of OASIS, the national web archive in Korea. *Journal of Korean Library and Information Science Society*, 45(4), 45-65. DOI : 10.16981/kliss.45.201412.45

[4] H. J. Kim. (2010). The analysis of interface types and characteristics of web archiving: with a focus on the web archiving project initiated by the nation. *Journal of Korean Society of Archives and Records Management*, 10(2), 147-170.

[5] H. J. Kim. (2017). A study on the web archiving of disaster records: focusing on the roles and functions of the OASIS disaster archive. *Journal of Korean Society of Archives and Records Management*, 17(2), 27-56. DOI : 10.14404/JKSARM.2017.17.2.027

[6] J. W. Nam & S. H. Kim. (2009). A study on the application of web 2.0 for archival information services. *Journal of the Korean Society for Library and Information Science*, 43(2), 123-146. DOI : 10.4275/KSLIS.2009.43.2.123

[7] Y. S. Kim. (2010). A theoretical study on establishing archive 2.0. *Journal of Korean Society of Archives and Records Management*, 10(2), 31-52.

[8] Y. M. Jung. (2017). A comparative study of Korea and Japan in their web archiving policies. *The Japanese Modern Association of Korea*, 57, 433-454.

[9] M. R. Kim. (2003). *System log analysis technique using fuzzy c-means clustering*. Doctoral dissertation Maste thesis. EwHa Womans University, Seoul

[10] S. W. Lee. (2019). A Study on the Evaluation of Serviceability of Web Resource Archiving System. *Journal of D-Culture Archives*, 2(2), 15-28.

[11] J. H. OH, J. H. Kim & J. W. Kim. (2011). A study on the development of realtime online maketing system using web log analytics. *The Journal of Society for e-Business Studies*, 16(3), 249-261. DOI : 10.7838/jsebs.2011.16.3.249

[12] J. Alba et al. (1997). Interactive home shopping: consumer, retailer, and manufacturer incentives to participate in electronic marketplaces. *Journal of Marketing*, 61(3), 38-53. DOI : 10.2307/1251788

[13] S. Y. Kim & Y. H. Ju. (2001). Perceived

interactivity and web site loyalty/on the role of flow as a mediating variable. *Journal of Consumer Studies*, 12(4), 185-208.

[14] J. H. Choi, S. M. Han & Y. S. Lee. (2004). The relationship between stickiness and purchasing behavior in internet shopping mall: the moderating. *Korea Marketing Review*, 19(2), 3-22.

[15] L. Lin, P. J. H. Hu, O. R. L. Sheng, & J. Lee. (2010). Is stickiness profitable for electronic retailers?. *Communications of the ACM*, 53(3), 132-136. DOI : 10.1145/1666420.1666454

[16] S. Y. Hwang & J. H. Seo. (2003). A study on the increase of intuitive play in web site : especially on the internet shopping mall design using tactile user experience. *Archives of Design Research*, 53(3), 22-32.

[17] D. Y. Yon & H. J. Lee. (2004). Analysis of the key factors that influence emotional web design and its effects on brand attitude : focus on fastfood websites. *Archives of Design Research*, 55(1), 153-162.

[18] D. G. Kim, S. S. Jang & J. E. Kang. (2006). The research on evaluating the quality for a website of providing tourism information. *Journal of Tourism and Leisure Research*, 18(3), 311-325.

[19] K. S. Jung. (2003). An empirical study on web sites evaluation by industry type. *Information System Review*, 5(2), 91-108.

남 재 우 (Jae-Woo Nam)

[정회원]



- 2009년 2월 : 중앙대학교 기록관리학과 (기록관리학석사)
- 2013년 8월 : 중앙대학교 문헌정보학과 (문헌정보학박사)
- 2017년 3월~현재 : 건국대학교 문헌정보학과 교수

· 관심분야 : Archival Service, Information Science  
· E-Mail : lisnam@kku.ac.kr

이 수 영 (Su-Young Lee)

[정회원]



- 2008년 8월 : 중앙대학교 교육대학원(문헌정보교육 석사)
- 2017년 2월 : 중앙대학교 문헌정보학과 (문헌정보학박사)
- 2011년 3월~현재 : 대림대학교 도서관미디어정보과 교수

· 관심분야 : Information Library service  
· E-Mail : sulibrary@daelim.ac.kr