

웹애널리틱스를 이용한 아카이브 이용자 분석 사례 연구*

이 효 은** · 임 진 희***

1. 서론
2. 아카이브 이용자 분석과 세분화 필요성
 - 1) 이용자 세분화 개념의 등장과 의미
 - 2) 아카이브 분야의 이용자 세분화 연구동향
 - 3) 아카이브 웹이용자 세분화의 필요성과 요건
3. 이용자분석을 위한 웹애널리틱스 기능 검토
 - 1) 웹애널리틱스의 종류와 특성
 - 2) 웹애널리틱스의 주요 기능
 - 3) 기록관 이용자 변수 적용 가능성 검토
4. 아카이브 웹이용자 분석 사례 : 민주화운동기념사업회를 중심으로
 - 1) 민주화운동기념사업회 오픈아카이브 웹로그분석 현황
 - 2) 민주화운동기념사업회 오픈아카이브 기본지표 분석
5. 결론

* 이 논문은 2015년 2월 명지대학교 기록정보과학전문대학원 기록정보학 석사논문을 요약·수정한 것임.

** 이효은, 명지대학교 대학사료실(제1저자).

*** 임진희, 명지대학교 기록정보과학전문대학원 교수(교신저자).

■투고일 : 2015년 6월 20일 ■최초심사일 : 2015년 6월 30일 ■게재확정일 : 2015년 7월 3일.

[국문초록]

기록정보서비스는 이용자들의 행태를 관찰하고 그들의 정보요구를 파악하여 그에 맞는 기록정보에 연결하고 안내해주는 적극적인 행위이다. 그러나 국내 기록정보서비스에 대한 연구는 이용자를 세분화하여 개인화된 서비스를 제공해야 한다는 당위성을 인정하지만 구체적인 사례나 방안에 대해서는 논의되지 않고 있다. 웹의 대중화와 알 권리의 확산으로 웹기록정보서비스의 중요성이 대두되고 있지만, 현재 우리나라 대부분의 기록관리기관에서는 실무 현장에서의 인식이 부족하고 인력과 예산의 불충분으로 웹로그분석을 실시하지 못하고 있다. 그래서 기관의 기록정보서비스 담당자의 업무 효율성을 높여주면서 예산이 부족한 기관에서 활용될 수 있는 분석 도구의 도입이 필요하다.

웹애널리틱스는 사이트에 방문한 웹이용자가 남긴 웹로그를 분석하여 행태를 분석하는 것이다. 그 종류에는 여러 가지가 있는데 그 중에서도 구글애널리틱스는 무료로 다양한 분석 항목을 제공하고 있어 전 세계 많은 인구가 이미 사용하고 있다.

본 연구에서는 구글애널리틱스를 중심으로 웹애널리틱스를 소개하고 이를 실제 기관에 도입한 민주화운동기념사업회 사료관의 사례를 분석하여 구체적인 웹이용자 세분화와 서비스 개선 방안을 제시하고자 한다.

주제어 : 웹로그분석, 웹애널리틱스, 구글애널리틱스, 아카이브, 이용자분석, 이용자세분화, 기록정보서비스

1. 서론

최근 기록관리기관의 역할이 전통적인 기록물의 보존에서 기록물의 적극적인 활용으로 확산되면서 기록정보서비스에 대한 관심이 증가하고 있다. 웹의 확산으로 이용자들의 기록정보요구의 형태 또한 기존의 비전자기록물에서 전자기록물로 바뀌고 있고 기록관을 직접 방문하기 보다 해당 기관의 웹사이트에서 기록물을 검색하는 이용자가 증가하고 있다. 그만큼 웹이용자에 대한 서비스가 중요해졌고 그를 위해서는 웹이용자의 행태를 파악하기 위한 분석이 이루어져야 한다.

Mary Pugh and Christopher Prom(2011)은 일리노이대학교의 아카이브 웹사이트를 대상으로 구글애널리틱스를 이용하여 측정하고 분석하는 파일럿 프로젝트를 실행하였다. 이 논문은 아키비스트가 이용자 행태를 측정하고, 이용자 행태의 측면을 이해하며 프로그램을 개시하기 위해 온라인 서비스를 개선하기 위한 방법으로써 사용할 수 있는 웹애널리틱스를 소개한다. 기록관 목표와 이용에 관련한 정보에 비춘 웹애널리틱스 데이터를 해석함으로써, 기록관은 접근성을 개선하고, 이용을 증진하며, 이용자 만족도를 향상시킬 수 있다고 역설한다.

하지만 현재 우리나라 기록관리기관에서 웹이용자의 행태에 대한 분석은 거의 이루어지지 않고 있으며 웹사이트의 방문자수를 측정하는 수준에 그치고 있다. 이는 웹이용자 분석에 대한 실무 현장에서 인식이 부족하고 인력 및 예산의 불충분이 그 이유로 추정된다. 따라서 기록관리 담당자의 업무 효율을 높이면서 예산이 적은 기관에서도 활용될 수 있는 분석 도구의 도입이 필요하다.

특정 웹사이트에 방문한 이용자가 어떤 행태를 보였는지를 아는 방법 중 기본적인 접근법은 웹사이트의 로그를 분석하는 것이다. 웹사이트의 로그는 이용자가 해당 사이트에 접속하여 남긴 모든 흔적을 축적한 데이터세트이다. 로그정보에는 이용자의 접속시간, IP주소, 브라우저

식별자 등 다양한 정보가 쌓이게 되는데 이를 이용하여 이용자별로 웹 사이트 이용행태를 추적해볼 수 있다. 웹로그를 이용하여 이용자 분석을 도와주는 도구에는 여러 종류가 있다.

그 중에서도 ‘구글애널리틱스(Google Analytics)’는 많은 정부·공공기관 프로젝트에서 사용하고 있다. 2014년 Fortune 500 기업의 로그분석툴 사용현황에 따르면 67%로 부동의 1위를 차지하였으며¹⁾, 2012년 미국 대선에서 오바마 대통령 캠프에서도 사용하였다고 한다.

구글애널리틱스에는 기본 기능을 제공하는 무료 버전과 고급 기능을 제공하는 유료 버전이 있다. 무료 버전으로도 기관 성격에 따라 행태 추적을 위한 항목의 맞춤설정이 가능하며 웹이용자에 대한 다양한 항목의 행태 분석이 가능하다. 따라서 예산의 제약이 있는 기록관리기관도 이러한 무료 버전의 도구를 이용하여 기초적인 웹이용자 분석을 할 수 있다. 다만, 구글애널리틱스와 같은 분석도구를 이용하여 이용자의 행태를 제대로 파악하기 위해서는 먼저 도구가 제공하는 데이터를 이해하고, 해당 웹사이트의 특성에 맞춰 해석할 수 있어야 한다. 또한, 이용자군을 판별해보고 이들이 원하는 기록정보 콘텐츠에 대해 예측해보는 시도를 해야 한다. 이용자에 대한 해석과 예측에 기반하여 새로운 콘텐츠 서비스를 기획하고 실행해볼 수 있어야 하며, 그 결과를 분석하여 이용자군에 대한 판별과 특성에 대한 예측을 정교화 해나가야 한다. 반복적 분석과정을 거쳐야만 이용자의 특성에 대해 가정을 하는 수준에서 점차 확신하는 수준으로 발전해갈 수 있다.

본 논문에서는 웹사이트를 통해 기록정보 콘텐츠를 제공하고 있는 기록관리기관이 웹애널리틱스를 이용하여 이용자분석을 하여 서비스를 개선해가는 절차와 방법을 제시하고자 한다. 이를 위한 먼저, 예산투자 없이 도입할 수 있는 웹애널리틱스를 중심으로 이 도구가 제공하는 데

1) <http://www.e-nor.com/blog/google-analytics/google-analytics-maintains-lead-with-fortune-500-in-2014>, [인용 날짜: 2015. 4. 17].

이터를 어떻게 해석하면 되는지 소개하고자 한다. 이러한 해석은 기록 관리기관의 웹사이트 목적, 기록정보 콘텐츠의 유형 등 여러 맥락에 따라 달라지게 된다. 이 논문에서는 웹애널리틱스를 사용하여 이용자 분석을 시도한 아카이브의 사례를 상세히 살펴봄으로써 맥락에 따른 해석의 방법을 제시하고자 한다. 또한, 향후 웹사이트를 운영하는 기록 관리기관들이 이용자분석 결과에 따라 서비스 전략을 어떻게 변화시켜갈 수 있는지를 제시하고자 한다. 이 논문에서는 기록물관리기관의 범위를 현용·준현용 기록물관리기관보다 일반인의 접근이 더 많은 비현용 기록물관리기관인 아카이브에 한정시켰다. 몇몇 현용·준현용 기록물관리기관에서도 웹애널리틱스를 이용하여 자관의 웹로그분석을 통해 웹사이트와 이용자 서비스 개선에 힘쓰고 있다. 향후 더 많은 기관에서 웹로그분석에 대한 인식이 확산되어 웹이용자에게 적극적인 서비스가 이루어지길 바란다.

2. 아카이브 이용자 분석과 세분화 필요성

1) 이용자 세분화 개념의 등장과 의미

근대의 산업 생산 시스템이 대량생산체제였다면 현대의 제품 생산방식은 다품종소량생산체제로 변함에 따라 모든 소비자를 만족시킬 수 없게 되었다. 그래서 소비시장 전체를 공략하는 대신 기업의 미션과 전략에 따라 세분화된 시장을 선택하게 되었다. 게다가 현대의 마케팅은 전통적인 기업의 생존이라는 목적을 달성하는 것이 아닌 소비자의 욕구를 충족시키는 의미로 변화하였다. 개인화라는 개념의 등장에 따라 소비자 개별의 욕구를 분석하기 위해서는 시장 전체를 세분화하여 즉,

타겟을 선정하여 공략하는 기법을 선택하게 되었다. 따라서 마케팅 분야에서 전략을 세울 때 중요시하는 부분이 시장 세분화이다. 시장 세분화(Market Segmentation)란 1950년대 등장한 개념으로 마케팅이나 광고 분야에 지대한 영향을 미치며 고객을 여러 segment로 나누는 과정 또는 유사한 요구를 가진 고객을 몇 개의 segment로 나누는 것을 의미한다.

종래의 마케팅개념이란 마케팅의 목표를 기업의 이윤추구에 두는 판매지향개념과 생산지향개념을 포괄하는 의미로서 생산자 또는 판매자가 생산한 상품이나 서비스를 강력한 촉진활동을 매개로 하여 소비자(또는 구매자)에게 전달하는 마케팅개념이다.²⁾ 그러나 대부분의 상품과 서비스의 공급이 수요에 비해 급속도로 확대되면서 시장은 완전한 공급과잉상태에 이르러 종전의 대량생산에 의한 공급관리체계에서 수요관리체제로 바뀌어져 기업경영의 변혁이 요구되는 것이다. 수요에 해당하는 소비자의 요구가 무엇인지 파악하는 것이 마케팅에서 중요하게 된 이유이다.

이처럼 소비자요구의 본질적 이해를 위해서는 특정 소비자 그룹을 선택하고 그들의 요구를 이해하여 상품과 서비스를 전달함으로써 소비자 만족을 실현하게 되는 것이다. 새로운 마케팅 개념이란 고객의 요구를 효과적인 방향으로 충족시키는 기업 전체의 활동이라는 인식을 바탕으로 하고 있다.³⁾

통상적으로 마케팅 분야에서 시장세분화의 핵심과정이라고 하면 다음 <표 1>과 같이 시장세분화의 기준변수를 설정하는 것이다. 연령, 성별, 직업 등의 인구통계적 변수와 지리적 변수, 심리적 변수, 행동적 변수로 나눈다. 마케팅관리자는 시장을 분류하는데 있어 이 중 하나 혹은 둘 이상을 이용하여 시장세분화 기준을 적용한다.

2) 박명호, 박종무, 윤만희, 「마케팅원론」, 2014, 경문사, 11쪽.

3) 위의 글, 14쪽.

〈표 1〉 시장세분화의 기준변수

기준 변수	구체적 변수
인구통계 변수	연령, 성별, 직업, 소득, 종교, 교육수준
지리적 변수	거주지역, 지역규모, 인구, 기후, 지리적 특성
심리적 변수	사회적 계층, 개성, 라이프스타일
행동적 변수	구매계기, 사용량, 사용경험여부, 재구매의사 등

웹이 확산되면서 온라인으로 제품 및 서비스를 검색하거나 구매하는 비율이 높아졌다. 통계청에서 배포하는 연간 온라인 쇼핑 동향만 보더라도 매년 꾸준히 그 비율이 증가하고 있다.⁴⁾ 이에 따라 기업에서도 자체 홈페이지를 제작하고 온라인으로 서비스를 제공하는 마케팅에 더욱 열을 올리고 있다. 이렇게 마련한 홈페이지에 어떤 고객이 들어왔는지를 알기 위한 웹사이트의 효율성을 평가할 필요가 있다. 웹방문자를 분석하기 위해서는 그들의 정보를 알아야하는데 회원제로 운영되지 않는 경우는 그조차 불가능하다. 어떤 사람이 우리 웹사이트를 방문했고 얼마나 머물렀으며 어떤 서비스를 이용했는지 전혀 알 수 없다. 예를 들어, 백화점에 고객이 물건을 사게 되면 물건을 파는 점원 입장에서 성별, 연령대 정도의 정보를 파악할 수 있다. 그러나 온라인 쇼핑물의 경우 회원제로 운영되지 않는 이상 웹방문자에 대한 어떠한 정보도 알 수 없다.

그래서 Sterne(2002)은 기업에서 그들의 웹사이트의 효율성을 평가하기 위하여 웹애널리틱스가 널리 사용된다고 하였다. 영리를 추구하는 기업이나 사이트에서는 이윤 추구하고 시장 점유율을 극대화하기 위하여 경영 성과의 예측과 평가를 시도하고 있고 그를 위해 고객의 정보를 분석하는 도구로 웹애널리틱스를 사용하고 있는 것이다.

4) 통계청, 온라인쇼핑동향, http://kostat.go.kr/portal/korea/kor_nw/2/11/3/index.board, [인용 날짜: 2015. 4. 17].

2) 아카이브 분야의 이용자 세분화 연구동향

기록학 분야도 종래의 기록물을 쌓아두고 이용자를 기다리는 보존·관리 중심의 체제에서 이용자에게 먼저 다가가는 적극적인 서비스가 확대되면서 이용자의 요구사항을 파악해야 한다는 개념이 확산되고 있다.

미국아키비스트협회(SAA)의 용어사전에서는 ‘정보서비스(reference service)’를 기록을 이용해서 이용자를 도와주는 활동 전반이라고 정의하고 있다. Pugh(2005)는 기록정보서비스를 현재 이용자와 잠재적 이용자가 소장기록을 활용하여 원하는 정보를 찾도록 돕는 것, 즉 이용자 요구를 충족시킬 수 있도록 기록과 이용자를 연결시켜 주는 아키비스트의 활동이라고 정의하였다.⁵⁾ 이는 기록정보서비스가 잠재적 이용자에게도 기관의 존재를 알리고 기록물을 활용할 수 있도록 안내하는 넓은 의미의 활동을 의미한다고 볼 수 있다.

이처럼 다양한 요구를 가진 이용자를 파악하기 위해서는 요구가 발생하게 된 원인과 배경에 대한 정보가 필요하다. 이를 위해 이용자의 직업으로 유형을 구분하여 이용자 요구를 파악하기 위한 연구와 마케팅 분야의 시장 세분화 기법을 적용하여 이용자 유형을 구분하는 연구들이 진행되어 왔다.

Hallem Smith(2003)⁶⁾는 PRO(Public Record Office)의 고객초점서비스의 흐름을 연구함으로써 이용자 세분화 과정을 거쳐야만 이용자의 요구를 더 잘 이해할 수 있고 효과적인 타겟팅이 가능하다고 주장하면서 마케팅 분야의 이론을 아카이브에 적용할 것을 제안하였다. Pugh(2005)는

5) Pugh, Mary Jo, *Providing Reference Services for Archives and Manuscripts*, Chicago : Society of American Archivists.

6) Elizabeth Hallem Smith, “Customer Focus and Marketing in Archive Service Delivery: theory and practice”, *Journal of the Society of Archivists* Vol.24(1), 2003, pp.35-53.

아카이브의 이용자 유형을 크게 어떤 이득을 얻기 위해 아카이브를 이용하는 직업적 이용자와 취미로 관심을 가지고 자료를 수집하는 비직업적 이용자로 구분하였다. 이후 직업적 이용자인 모기관의 직원, 전문직 이용자, 학자, 학생, 교육자들에 관한 연구와 비직업적 이용자에 대한 연구가 잇따라 발표되었다.

아카이브 이용자에 대한 연구가 대부분 직업적 이용자 유형을 중심으로 전개된 반면 Yeo(2005)는 런던대학에서 수행한 LEADERS 프로젝트⁷⁾에 관해 발표한 논문⁸⁾에서 아카이브와 기록관리의 이용자 그룹을 이해하기 위한 접근방법으로 시장 세분화에 대한 설명을 제공하고자 하였다. 마케팅 영역에서 고객의 요구를 분석하여 제품을 제공하는 것과 마찬가지로 아카이브도 이용자 세분화를 통해 이용자 유형에 맞게 서비스를 제공해야 한다고 하였다. 이용자를 세분화하는 기준으로 연령, 성별, 직업 등과 같은 인구통계학적 변수를 제안하였다.

국내에서도 이용자 세분화와 관련된 많은 연구물이 발표되고 있다. 이소연(2004)은 문화시설의 경영에 마케팅 기법 적용을 제안하며 시장 조사와 함께 시장분할(market segmentation), 즉 수요자 집단을 세분화하여 그들의 구체적이고 특수한 요구를 파악해야 한다고 하였다.⁹⁾ 정경희(2007)¹⁰⁾는 현대 마케팅 기법인 표적 마케팅을 기록정보서비스에 적용 가능성에 대하여 논의하였다. Yeo, LEADERS 프로젝트, Pugh, Rowley

-
- 7) 시장 세분화 기술을 채용한 사례로 보존 문서의 이미지와 전문을 EAD(Encoded archival Description)와 함께 관리하는 데모 시스템과 툴킷을 개발. LEADERS (Linking EAD to Electronically Retrievable Sources) 프로젝트는 맥락 속에서 보존 자료를 표현하는 방법을 제공함으로써 아카이브에 대한 원격 이용자의 접근성 향상을 목적으로 함.
 - 8) Geoffrey Yeo, "Understanding User and Use: A Market Segmentation Approach", *Journal of the Society of Archivists*, Vol.26(1), 2005, pp.25-53.
 - 9) 이소연, 「지역사회 주민을 위한 공공도서관의 문화 서비스 활성화 방안」, 『한국문헌정보학회지』 38(3), 2004, 23-43쪽.
 - 10) 정경희, 「시장세분화 기반의 기록정보서비스에 관한 고찰」, 『한국도서관·정보학회지』 38(3), 2007, 277-296쪽.

의 세분화 변수를 기반으로 우리나라 기록관 이용자 세분화를 위한 변수를 인구통계 변수, 지리적 변수, 심리적 변수, 이용 변수 네 가지로 구분하여 아래 표와 같이 제안하였다. 이용자의 요구를 파악하지 않은 세분화는 가능하지 않다고 보며 아카이브 이용자에게 대한 정보를 지속적으로 수집하고 이를 분석해야 이용자 세분화가 가능하다고 역설하였다.

〈표 2〉 기록관 이용자 세분화 변수

변수명	변수	설명
인구통계 변수	연령, 직업, 교육수준, 종교	기록정보의 이용목적이나 관심사와 밀접한 관련이 있는 변수임
지리적 변수	행정구역, 기록관과의 물리적 거리	이용자와 관련된 행정구역은 관심사에, 기록관과의 물리적 거리는 직접이용인가 간접이용인가에 영향을 미침
심리적 변수	관심사, 정보이용 스타일	동일한 인구통계집단이나 행정구역에 속한 이용자들을 세분할 수 있는 변수
이용 변수	이용목적 및 동기, 이용경험 및 빈도, 이용능력(시스템, 정보해석), 질문유형	기록 이용과 관련한 태도, 지식, 경험

* 출처 : 정경희, 「시장세분화 기반의 기록정보서비스에 관한 고찰」, 『한국도서관정보학회지』 38(3), 2007, 295쪽.

위의 〈표 2〉와 같이 제시된 이용자 세분화 변수가 중요하다는 것은 쉽게 이해할 수 있다. 그러나 변수의 종류를 아는 것은 아카이브 이용자 분석의 이론적 접근이다. 웹사이트에 접속하는 이용자가 누구인지를 파악하는 것과 이용자별 변수 값에 따라 이용행태나 요구사항이 어떻게 달라지는지 알아낼 수 있어야 실무적인 이용자분석이라 할 수 있다. 이론적 접근 혹은 열람이용자를 기반으로 도출된 이용자세분화 변수가 웹이용자에게도 그대로 적용될 수 있을지 확인해야 하며 이를 위해 웹애널리틱스와 같은 도구를 이용할 필요가 있다.

기록학 분야에서 이용자 세분화는 내·외부 설문조사나 보고서를 위해 통계를 수집하는 방식으로 많이 이용되었으나 최근 효과적인 서비스 정책을 위한 전제조건으로 인지되고 있다. 기록관리법의 제정과 알 권리에 대한 인식이 확산됨에 따라 이용자의 기록정보 요구도 급증하고 있는데 점점 더 다양해지는 이용자의 요구에 따라 개인화된 서비스에 대한 필요성이 증대되고 있다. 그만큼 이용자 세분화와 정보요구에 대한 세분화는 전략적이고 분석적인 기록정보서비스의 기반이라 할 수 있다.

3) 아카이브 웹이용자 세분화의 필요성과 요건

웹이 대중에게 확산·보급되어 아카이브도 각 기관별로 홈페이지를 구축하고 웹을 통한 서비스에 적극 나서고 있다. 앞서 살펴본 마케팅 분야처럼 이용자 요구의 본질적 이해를 위해서는 특정 이용자 그룹을 선택하고 그들의 요구를 이해하여 서비스를 전달함으로써 이용자 만족을 실현할 수 있는 것이다.

그러나 온라인 쇼핑물은 최소한 상품을 구매한 소비자의 특성을 알 수 있지만 비영리집단인 아카이브는 어떤 사람이 홈페이지를 다녀갔으며 어떤 서비스를 이용하였는지 전혀 알 수 없다. 아키비스트 입장에서는 눈에 보이지 않는 그림자와 같은 존재라 할 수 있다. 눈에 보이지 않는 이 그림자가 어떤 사람의 것인지를 추적하기 위해서는 웹로그를 분석해야 한다.

‘웹로그(Web Log)’란 이용자가 사이트에 접속하였을 때 서버에 남는 흔적이다. 웹이용자들의 성향을 파악하여 그들의 이용행태를 알 수 있게 한다. 이는 일종의 ‘개인화’ 개념에서 비롯되었고 이로 인해 개인화된 솔루션이 필요하게 되었다.

단순히 웹로그만으로는 이용자에 대한 기초적인 데이터만 제공하여

이용자가 웹 상에서 ‘무엇’을 하였는지만 알려준다. 실행 가능하고 통찰력 있는 데이터를 얻기 위해서는 무엇을 하였는지도 중요하지만 이용자가 ‘왜’ 그랬는지를 알아야 한다. 즉 웹사이트 개선이나 웹이용자의 이용행태를 파악하기 위한 용도로 웹이용 행태를 분석하는 것이 ‘웹로그분석’이다.

웹로그분석을 실행하기 위해서는 각 기관이 소장하고 있는 기록물의 성격에 따라 기관의 특성이 달라지고 이용자의 행태도 달라진다. 아카이브 웹이용자를 세분화하기 위해서는 기관의 현황을 파악하는 것이 중요하다.

우선, 기관의 운영형태를 파악해야한다. 기록관리기관에는 비현용 기록물을 관리하는 기관인 아카이브와 현용·준현용 기록물관리기관이 있는데 기관이 어떤 형태인지부터 파악해야 한다. 비현용 기록물관리기관인 아카이브는 현용·준현용 기록물관리기관에 비하여 일반인의 접근이 용이하고 이용률이 높다는 특징을 가지고 있다. 반면 현용·준현용 기록물관리기관은 기관에서 생산되는 공문서 위주의 관리가 이루어지고 있어 일반인의 접근이나 이용률이 다소 낮은 편이다. 그러므로 해당 기관의 웹이용자를 분석하는 변수도 달라지게 된다.

둘째, 해당 기관의 웹사이트가 회원제로 운영되는지에 대한 여부가 중요하다. 회원제로 운영되는 경우 회원가입 시 기관 측에서 요구하는 이용자의 연령, 성별, 관심분야를 입력하게 되어 이용자의 개개인에 대한 정보가 파악되지만 비회원제로 운영되는 기관의 경우 이용자에 대한 상세한 정보까지는 파악하기 힘들다. 따라서 웹애널리틱스에서 웹로그로 분석되는 정보로 웹이용자를 세분화하여 그룹화한 뒤 그들의 정보요구를 파악해야 한다.

위와 같이 기관의 형태, 웹사이트의 회원제 여부를 파악한 뒤 해당 기관의 웹이용자를 세분화하기 위한 기준을 세우고 웹로그로 얻어진 정보를 분석하여 적극적인 서비스 전략을 준비할 수 있다.

3. 이용자분석을 위한 웹애널리틱스 기능 검토

1) 웹애널리틱스의 종류와 특성

웹로그분석협회 WAA(web analytics association)에서는 최초로 ‘Web Analytics’의 표준화된 정의를 다음과 같이 결정하였다.

“Web analytics is the measurement, collection, analysis and reporting of web data for purposes of understanding and optimizing web usage.

웹로그분석은 웹 사용을 이해하고 최적화하려는 목적을 위해 인터넷 데이터를 측정하고 수집하고 분석하며 보고하는 것이다.”

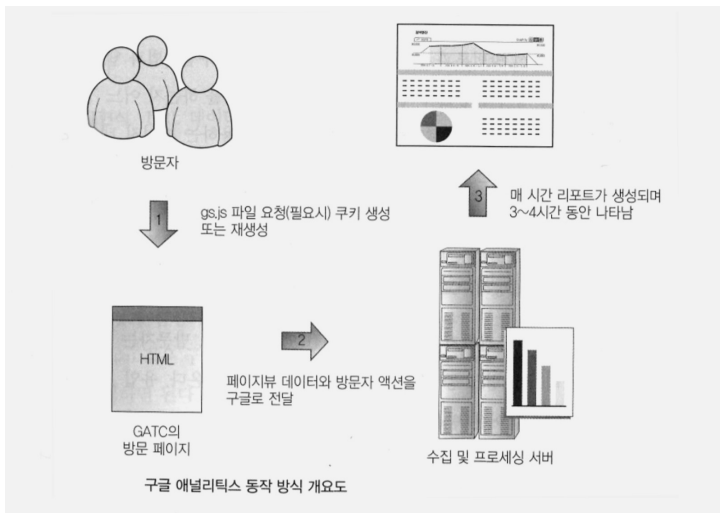
앞서 언급했듯이 웹애널리틱스는 크게 해외업체와 국내업체로 구분해 볼 수 있다. 해외업체에서 제공되는 것은 구글애널리틱스, 어도비 디지털 마케팅 수트(Adobe Digital Marketing Suite), 알렉사(Alexa), IBM 코어메트릭스(Coremetrics), 웹트렌드(Webtrends), 야후(Yahoo) 등이 있고 국내 업체에서 제공되는 것은 네이버 애널리틱스, 에이스카운터 등이 있다. 이렇게 다양한 웹애널리틱스 중에서도 구글애널리틱스는 많은 정부·공공기관 프로젝트에서 사용하고 있다. 2014년 Fortune 500 기업의 로그분석 툴 사용현황에서 사용률 67%로 1위를 차지하였으며¹¹⁾, 2012년 미국 대선에서 오바마 대통령 캠프에서도 사용해 더욱 유명해졌다.

2005년 4월, 구글이 ‘얼친’이라는 소프트웨어 기업을 인수하여 얼친의 기술을 토대로 2006년부터 본격적으로 웹사이트 분석 툴(Tool)을 서비스하게 되는데, 이것이 현재의 구글에서 제공하고 있는 웹로그분석 툴

11) <https://www.e-nor.com/blog/google-analytics/google-analytics-maintains-lead-with-for-tune-500-in-2014>, [인용 날짜: 2015. 4. 17].

인 구글애널리틱스(Google Analytics)이다. 구글은 웹 검색뿐 아니라 클라우드 컴퓨팅과 광고 등 다른 영역까지 세력을 확장하며 세계적인 기업이 되었다. 다른 업체에서 제공되는 웹애널리틱스는 대부분 유료로 제공되지만 구글은 웹로그분석에 관한 방대한 정보를 무료로 제공하고 있다. 구글애널리틱스의 작동방식을 설명하면 다음 <그림 1>과 같다.

<그림 1> 구글애널리틱스 동작 방식 개요도



* 출처 : 브라이언 클리프튼, 박종채·이세현 역, 구글애널리틱스, 개정판, 2014, 102쪽.

1. 웹이용자가 사이트를 방문하면 구글 서버의 자바 스크립트(JAVA Script)¹²⁾를 호출하게 되는데 이때 쿠키(Cookie)¹³⁾가 생성된다.

- 12) 브라우저 자체에 내장된 해석기능을 이용한 클라이언트(Client) 기반의 스크립트 언어. 작고 빠르기 때문에 웹문서를 동적으로 꾸밀 때 널리 쓰인다.(NEW 경제용어사전, 2006, 미래와경영)
- 13) 웹이용자와 사이트 간 매개 정보를 담은 4KB 이하의 임시 파일. 웹이용자가 인터넷에서 어떤 내용을 봤는지 등의 정보가 기록되어 온라인 광고업체가 쿠키를 이용한 웹이용자의 선호도를 분석하여 광고 전략을 세우기도 한다.

2. 구글애널리틱스 설정 시 사이트 내 HTML¹⁴⁾ 사이에 삽입된 GAT C¹⁵⁾가 구글 서버로 페이지뷰 데이터와 웹이용자의 행동 등의 정보를 전달한다.

3. 수집 및 프로세싱 서버가 사이트로부터 전달받은 데이터를 처리하여 매시간 리포트가 생성되고 3~4시간 동안 나타나게 된다.

2) 웹애널리틱스의 주요 기능

애널리틱스는 웹이용자의 행태를 추적하기 위해 웹이용자가 접속한 국가나 지역부터 언어, 성별, 연령대, 유입경로, 접속시간 등 여러 종류의 분석항목을 제시하고 있다. 구글애널리틱스의 경우 기본적으로 제공하는 분석항목을 표로 정리하면 다음과 같다.

〈표 3〉 구글애널리틱스 웹이용자 분석 항목

잠재고객(Audience)	인구통계 및 지역(Demographics & Geo)
	관심분야(Interests)
	방문형태(Behavior)
	기술환경(Technology)
	모바일(Mobile)
	사용자 흐름(Users Flow)

14) Hypertext Markup Language. 인터넷 서비스의 하나인 월드 와이드 웹을 통해 볼 수 있는 문서를 만들 때 사용하는 프로그래밍 언어의 한 종류이다. 문서의 글자크기, 글자색, 글자모양, 그래픽, 문서이동(하이퍼링크) 등을 정의하는 명령어로서 홈페이지를 작성하는 데 쓰인다(네이버 지식사전, 두산백과).

15) Google Analytics Tracking Code, 자바 스크립트 코드로 정보를 수집하여 데이터 수집서버로 전송하는 역할을 한다.

획득(Acquisition)	채널(Channel)
	키워드(Keywords)
	소셜(Social)
방문형태(Behavior)	행동흐름(Behavior Flow)
	사이트 콘텐츠(Site Contents)
	사이트 속도(Site Speed)
	사이트 검색(Site Search)
전환(Conversion)	목표(Goals)
	전자상거래(Ecommerce)
	다채널 유입경로(Multi-Channel Funnels)
	기여(Attribution)

‘잠재고객’ 항목에서는 웹사이트를 방문한 이용자의 성별, 연령, 언어, 접속지역 등 웹이용자의 인구통계 변수에 해당하는 정보와 방문빈도, 접속 환경에 대한 정보까지 보여준다. 웹사이트 분석에 대한 많은 연구에서도 웹이용자의 성별, 연령, 지역, 언어는 인구학적 세분화(Demography Segmentation)를 가능하게 하여 웹사이트 이용도 분석에 유의미한 영향을 미치는 것으로 나타난다. 따라서 아카이브의 서비스 전략을 세우는데 중요한 변수라 할 수 있다. 예를 들어 기관의 홈페이지에 학부생의 방문자가 많다면 학부생 수준에 맞는 서비스를 제공할 필요가 있을 것이다. 구글은 비회원으로 접속한 웹이용자의 연령과 성별도 분석한다. 웹이용자가 ‘구글 디스플레이 네트워크’라는 파트너 사이트(구글 디스플레이 네트워크 광고가 노출되는 사이트)를 탐색할 때 쿠키를 통해 이용자의 웹 브라우저를 구분해주는 고유번호를 저장하여 방문 정보를 기록한다. 해당 웹이용자가 이전에 방문했던 사이트를 통해 성별과 연령을 유추하는 것이다. 그러므로 비회원으로 접속한 이용자의 정보는 ‘추정된 값’으로 100% 정확하지 않다고 볼 수 있다. 따라서 가능한 한 회원 가입제로 웹사이트를 운영하여 정확한 인구통계 정보를 수집하는 것이 좋다.

‘획득’ 메뉴에서는 방문자가 웹사이트에 접속하게 된 경로를 추천채널, 검색키워드, 방문페이지 등으로 분석하여 방문자가 어떤 경로를 통하여 사이트에 접속하게 되었는지에 대한 데이터를 보여주는데, 채널(Channel), 키워드(Keywords), 소셜(Social) 세 가지 방법으로 구분한다. ‘기본채널그룹(Default Channel Grouping)’은 획득한 채널의 유형으로 기본채널그룹은 자연어검색(Organic Search)¹⁶⁾, 추천방문(Referral)¹⁷⁾, 직접방문(Direct)¹⁸⁾ 등으로 나뉘며, 각 유형을 클릭하면 분석 기준이 소스와 방문페이지로 나누어진다. 우리 사이트는 어떤 키워드로 상징되며, 방문자들은 어떤 인식을 가지고 있는지는 ‘키워드’로 분석할 수 있다. 많이 검색되는 키워드를 분석하여 이와 관련된 콘텐츠를 사이트의 메인 화면이나 인기 있는 페이지에 제공하는 서비스를 고려해 볼 수 있다. 최근 SNS(Social Network Service, 소셜 네트워크 서비스) 사용자가 늘어나면서 SNS를 통하여 사이트로 유입되는 웹이용자도 늘어나고 있다. 이에 따라 구글애널리틱스도 어떤 SNS에서 유입되어 사이트 내에서 어느 페이지를 많이 방문하였는지를 한 눈에 파악할 수 있는 표와 그래프로 보여준다.

‘방문형태’ 항목에서는 획득한 웹이용자가 어떤 경로로 이동하였는지, 얼마나 웹에 머물렀으며 어떤 서비스를 이용하였는지에 관해 보여준다. ‘행동흐름(Behavior Flow)’은 웹이용자가 어떻게 사이트에 들어오게 되었는지부터 시작하여 어떤 페이지를 경로로 이동하였는지를 보여준다. ‘사이트 콘텐츠(Site Contents)’는 방문자가 사이트 내에서 어떤 페이지를 방문하였는지를 URI¹⁹⁾로 표시하고, 페이지 제목도 측정기준으로 따로

-
- 16) 사이트로 유입되기 직전에 포털 검색창에서 검색엔진 결과를 통하여 유입된 유형이다.
 - 17) 애널리틱스는 사이트로 유입되기 직전에 방문한 사이트의 도메인 이름을 추천 트래픽 소스로 표시한다.
 - 18) 웹이용자가 사이트의 URL을 직접 입력하여 들어온 유형이다.

제시하며, 이를 다시 콘텐츠 드릴다운, 방문 페이지, 이탈 페이지의 자세한 메뉴로 보여준다. 방문 페이지(Landing Page)에서는 방문자가 해당 사이트에 들어오기 전 방문했던 페이지를 URI 형식으로 표시하며, 이탈 페이지(Exit Page) 또한 URI 형식으로 방문자가 해당 사이트에서 마지막으로 방문했던 페이지를 보여준다. ‘사이트 속도(Site Speed)’ 기능에서는 페이지 로드(Load) 시간을 측정하여 보여주는데 평균 리디렉션 시간, 평균 도메인 조회 시간, 평균 서버 연결 시간, 평균 서버 응답 시간, 평균 페이지 다운로드 시간까지 초 단위로 보여준다. 만약 페이지 로드 시간이 긴 브라우저라면 이를 이용하는 웹이용자의 불편이 발생하므로 해당 브라우저에서의 페이지 로드 시간을 줄이기 위한 개선 작업이 필요하다라는 판단을 내릴 수 있다.

애널리틱스에서 ‘전환’이란 사이트 관리자가 설정한 목표치의 성공을 위해 필요한 웹사이트에서의 활동을 완료하는 것이다. 주로 전자상거래 분야에 초점을 맞춘 분석 항목이지만 이를 아카이브에 맞게 활용하는 법을 생각해볼 수 있다. 예를 들어, 해당 아카이브에서 목표 1을 ‘회원가입하기’로 설정하고 해당 목표의 금전적 가치를 500원으로 책정한다. 목표2는 회원가입을 완료한 웹이용자가 ‘마이 아카이브(My Archives)’에 하나의 콘텐츠를 담을 때마다 금전적 가치를 300원으로 책정한다. 목표1의 달성률을 금전적 가치로 10,000원을 예상하였으나 5,000원에 그쳤다면 목표달성률은 50%로 계산되며, 목표2의 달성률을 3,000원으로 예상하였으나 6,000원에 도달하였다면 200% 달성한 것으로 볼 수 있다. 이는 아키비스트가 해당 아카이브와 이용자의 특성을 잘 파악하여 활용할 수 있는 부분으로 목표치를 금전적 가치로 환원하여 목표 달성률을 분석할 수 있는 항목이라 할 수 있다.

19) URI는 ‘Uniform resource identifier’의 약자로 인터넷 서비스를 전제로 한 인터넷 상의 통일된 정보 자원의 식별 체계로 쉽게 말해, 페이지의 도메인 이름 다음에 표시되는 부분을 말한다.

구글애널리틱스는 이와 같은 이용자분석 항목 외에도 관리자의 업무 편의를 위한 기능을 제공하고 있다. ‘실시간’ 보고서는 개요, 위치, 트래픽 소스, 콘텐츠, 이벤트, 전환 6개의 세부항목을 통하여 현재 사이트를 방문 중인 사용자 수, 최근 30분간 분당 페이지뷰수, 최근 60초 초당 페이지뷰수가 6개 세부항목에 공통으로 제공된다. 예를 들어 ‘트래픽 소스’ 항목에서는 현재 사이트에 접속 중인 웹이용자가 사이트에 들어오기 직전 어떤 웹사이트에 접속했었는지 보여준다.

또, 구글애널리틱스가 제공하는 분석항목들을 한 눈에 파악할 수 있도록 한 화면에 배치하는 ‘대시보드’ 기능이 있다. 분석항목들을 이용하여 다각적인 분석을 가능하게 하는 ‘맞춤설정(Customization)’ 기능도 있는데 이를 활용하여 맞춤형 보고서를 쉽게 작성할 수도 있다. 예를 들면, 어떤 검색어를 통하여 우리 사이트에 들어왔는지 알아보기 위해 ‘측정항목 그룹’을 ‘자연 검색 수’로, ‘측정기준 드릴다운’을 ‘키워드’와 ‘도착 페이지’로 설정하여 특정 검색어로 접속한 웹이용자의 수를 파악할 수 있다.

지금까지 살펴본 구글애널리틱스의 웹이용자 분석 기능은 앞서 살펴본 것처럼 연구자들이 제시한 아카이브 이용자 세분화 변수와 일치하거나 유사한 항목이 많다. 별도의 예산없이 추적코드 몇 줄만으로 웹이용자의 이용행태를 전반적으로 분석할 수 있다는 것이 구글애널리틱스의 최대 강점일 것이다. 이를 아카이브에 도입한다면 업무 효율성을 높이고 서비스 개선 방안을 제시하는데 좋은 수단이 될 것이다.

3) 기록관 이용자 변수 적용 가능성 검토

앞에서 살펴본 구글애널리틱스 웹이용자 분석 항목은 정경희(2007)의 기록관 이용자 세분화 변수와 유사한 점이 많다. 기록관 이용자 세분화 변수와 구글애널리틱스 웹이용자 분석 항목의 일치하는 부분을 표로

비교하여 보면 다음 <표 4>와 같다. 이용자의 연령, 성별 등 인구통계 변수, 지리적 변수, 관심분야 정보, 방문빈도 및 방문행태 등 일치하거나 유사한 항목이 많다.

<표 4> 기록관 이용자 세분화 변수와 구글애널리틱스 웹이용자 분석 항목 비교

기록관 이용자 세분화 변수		구글애널리틱스 웹이용자 분석 항목
변수명	변수	
인구통계 변수	연령, 직업, 교육수준, 종교	연령, 성별, 언어
지리적 변수	행정구역 기록관과의 물리적 거리	위치
심리적 변수	관심사 정보이용 스타일	관심분야 (사용자)행동흐름
이용 변수	이용목적 및 동기 이용경험 및 빈도 이용능력(시스템, 정보해석) 질문유형	관심분야 획득 채널 계재빈도 및 방문빈도 신규방문과 재방문 참여도 기술환경(브라우저, 운영체제) 키워드

선행연구에서도 언급했듯이 아카이브에 방문하는 이용자를 세분화하여 그들의 요구를 파악하는 것은 양질의 서비스를 제공할 수 있는 밑거름이 된다. 더욱이 형체를 파악할 수 없는 아카이브의 웹이용자에 대한 정보와 그들의 요구를 추정하기 위해서는 웹애널리틱스를 구동하여 웹로그 분석 정보를 이용하는 것이 효율적이다. 웹애널리틱스로 얻은 웹로그 정보로 이용자별 요구를 파악하는 것은 중요한 작업이지만 고난이도의 기술(technic)을 요구하지도 않기 때문이다.

그렇다면 구글애널리틱스가 가진 많은 기능은 과연 아카이브에 유용한 것인가. 아카이브 웹사이트를 드나드는 형체를 알 수 없는 웹이용자들을 고난이도의 능력이 없어도 분석이 가능할까. 구글애널리틱스를 아카이브에 적용 시 비용적 측면, 이용편리성, 효율성 측면에서 유용성

을 검토하였다.

첫째, 비용적 측면은 구글애널리틱스가 가진 최대 강점으로 무료로 사용이 가능하다는 것이다. 무료서비스는 누구나 사용할 수 있다는 장점을 가진다. 다른 웹애널리틱스 서비스 기업들도 다양한 어플리케이션과 분석 서비스를 제공하지만 예산이 부족한 기관이 사용하기에는 비용의 장벽이 높다. 그런 면에서 비용은 중요한 선택 기준이다. 물론 구글애널리틱스는 유료서비스인 프리미엄 기능도 제공하고 있지만 무료로 제공되는 기능만으로도 웹이용자를 분석하는데 충분하다. 다른 웹애널리틱스가 주요 분석항목을 유료로 제공하는 것과 대비되는 점이기도 하다.

둘째, 구글애널리틱스를 사용하여 웹이용자를 분석하는데 있어 고난이도의 디지털마케팅 기술이나 IT지식이 없어도 손쉽게 이용가능하다. 웹분석이나 디지털마케팅을 전공하지 않은 일반인도 웹이용자를 분석하기 쉽도록 생소한 개념이나 낯선 단어에 대해 첫 사용시점부터 이용자가 쉽게 이해할 수 있도록 도움말을 제공하고 있다.²⁰⁾

셋째, 웹이용자를 분석하기란 여간 어려운 일이 아니다. 웹이라는 가상공간의 특성상 웹이용자들이 사이트에 방문하여 어떤 흐름으로 이용하였는지, 얼마나 머무르다 나갔는지는 눈으로 볼 수 없기 때문이다. 이렇게 눈에 보이지 않는 웹이용자들의 행태를 단 몇 줄의 HTML 코드 삽입하면 가시적인 수치로 보여주니 분석하는데 소요되는 시간을 줄일 수 있어 업무 효율성을 향상시켜주는 도구라 할 수 있다.

넷째, 웹이용자들은 자신의 정보를 누출시키지 않고도 맞춤형 서비스를 제공받을 수 있다. 구글애널리틱스는 방문자의 주소, 이름, 신용카드정보 같은 개인정보는 수집하지 않고, 로그 데이터와 방문자의 웹이용행태에 대한 정보만을 익명으로 수집한다. 로그 정보는 IP주소, 접속

20) 웹로그 분석 도움말, <http://support.google.com/analytics>

시간, 브라우저 및 운영체제 같은 인터넷 표준 로그정보이고, 방문자의 웹이용행태에 대한 정보는 페이지뷰 수, 언어 설정, 스크린 해상도 설정에 대한 정보부터 목표 전환 수치와 실패 관련 정보까지 포함한다.

구글애널리틱스는 특정 사용자를 추적하는 것이 아니라 특정 웹사이트에 대한 방문 정보를 추적하므로 사이트별로 쿠키를 사용한다. 즉, 구글애널리틱스가 제공하는 리포트에는 IP에 기반한 접속지역에 대한 정보(코드)와 접속시간 등 다양한 데이터를 포함하지만 익명의 통계 리포트이므로 개별 방문자를 식별할 수 있는 신원정보는 제공하지 않는다.²¹⁾

구글애널리틱스는 많은 장점을 가지고 있는 반면 몇 가지 한계점도 가지고 있다.

우선 웹이용자의 연령대 항목에서 만 18세 미만은 측정이 불가능하다는 것이다. 아카이브는 그 특성상 교육용 콘텐츠로 제공되는 것이 많아 유아부터 고등학생의 연령대의 분석이 꼭 필요하다. 그러나 전 세계적으로 18세 이하 아동 및 청소년에 대한 데이터를 수집·보유하는 것은 법적으로 위험한 일일 수 있기 때문이다.²²⁾

따라서 회원제로 운영되는 아카이브에서 웹이용자가 회원가입을 하도록 하고 연령정보를 입력하도록 유도하는 것이 좋다. 앞 장에서도 언급하였듯이 비회원제 아카이브의 경우에도 추가 추적코드를 해당 페이지에 삽입하면 연령대 분석은 가능하나 추정값이므로 100% 신뢰할 수는 없다는 점도 고려해야 한다.

둘째, 많은 분석항목을 제시하여 관리자에게 유용한 분석서비스를 제공하는 반면, 오히려 너무 많은 항목으로 인터페이스가 복잡하여 사

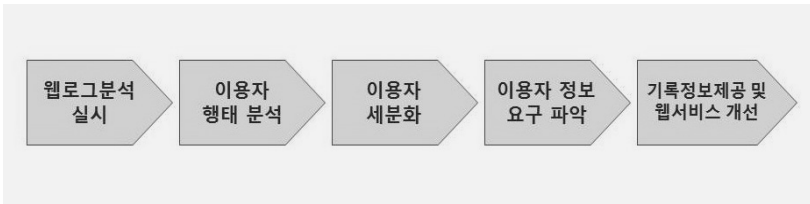
21) 브라이언 클리프튼, 이세현·박종채 역, 『구글애널리틱스』, 2010, 71-72쪽.

22) Children's Online Privacy Protection Rule("COPPA"), <https://www.ftc.gov/enforcement/rules/rulemaking-regulatory-reform-proceedings/childrens-online-privacy-protection-rule>, [인용 날짜: 2015. 4. 17].

용자 편의가 부족하다는 지적도 있다. 위에서 언급하였던 ‘대시보드’ 기능을 이용하여 필요한 분석항목만을 보고서로 제공받는 방법을 활용하거나 ‘맞춤설정’ 기능으로 업무에 적합한 항목을 구분하여 사용하면 편리하고 단점을 극복할 수 있을 것이다.

셋째, 구글애널리틱스는 상업용에 초점이 맞춰져있는 웹로그분석 도구이다. 아카ιβ는 특성 상 비영리기관이 대다수인데 구글애널리틱스는 분석항목 중 ‘제품실적’, ‘구매까지의 소요일수’ 등과 같이 상업적 용어가 많이 등장하고 구글의 ‘애드센스’나 ‘애드워즈’와 같은 유료 광고 캠페인과의 연동 서비스를 제공하는데 필요하지 않은 항목을 제외하거나 유료 서비스임을 확인하고 활용해야 혼동이 생기지 않는다. 몇 가지 한계점이 있음에도 불구하고 구글애널리틱스가 가진 장점은 효율적이며 예산과 인력이 부족한 아카ιβ의 관리자 입장에서는 매력적인 도구가 아닐 수 없다. 비용절감과 업무에 소용되는 시간을 단축시켜주는 것이 강점이라 할 수 있다.

〈그림 2〉 웹로그분석을 통한 아카ιβ 웹이용자 분석 프로세스



웹애널리틱스를 이용하여 아카ιβ 웹이용자를 분석하고 서비스를 개선해나가는 프로세스를 도식화하면 위의 〈그림 2〉와 같다. 구글애널리틱스가 가진 주요기능을 아카ιβ 웹사이트에 적용할 때는 웹애널리틱스의 분석기능을 바탕으로 이용자들의 웹행태를 파악하고 행태 특성별로 세분화하여 그들의 정보요구를 파악해야 한다. 이용자 그룹별로

정보요구가 파악이 되었다면 해당 아카이브의 소장 기록물의 특성을 고려하여 이용자가 만족할만한 기록정보를 제공하고 웹서비스를 개선해 나가야 한다.

4. 아카이브 웹이용자 분석 사례 : 민주화운동기념사업회를 중심으로

디지털마케팅이나 IT를 전공한 사람이 아니면 다소 어렵게 느껴질 수 있는 구글애널리틱스이지만 실제 구동하고 있는 사례를 보면 쉽게 이해할 수 있다. 따라서 다양한 웹콘텐츠 서비스를 실시하면서 구글애널리틱스를 구동하고 있는 아카이브 기관인 민주화운동기념사업회 사료관의 사례를 소개하고자 한다.

1) 민주화운동기념사업회 오픈아카이브 웹로그분석 현황

민주화운동기념사업회 사료관의 ‘오픈아카이브’는 2010년 ‘오픈아카이브 구축사업 대비 데이터베이스 정리’를 시작으로 90만 건에 이르는 사료를 수집하였고²³⁾ 누구나 사료를 이용할 수 있도록 체계적으로 기록을 수집하고 정리하였다. 현재 우리나라 메타스크립트 기관 중 소장 사료의 조직과 활용 측면에서 서비스가 잘 이루어지고 있는 곳이기도 하다.

2011년 4월 22일부터 민주화운동기념사업회 사료관 사이트에 구글애널리틱스를 구동하여 웹이용자 분석을 하고 있다. 사이트가 제공하고

23) 민주화운동기념사업회 오픈아카이브 홈페이지(<http://archives.kdemo.or.kr/Archives>), 사료관 소개.

있는 메뉴 중 ‘자료로 배우는 민주화운동’에 접속하는 연령대가 궁금하여 해당 정보를 찾아보았지만 연령대의 분석은 하고 있지 않았다. 구글 애널리틱스에서 제공하는 기본코드 외에 추가코드를 부여받아 삽입하면 인구통계에 대한 정보가 분석되는데 추가 코드를 삽입하지 않아 연령대의 분석이 불가능하였다.

이러한 문제점들로 보아 현재 오픈아카이브는 구글애널리틱스를 통해 ‘보고서’ 기능이 제공하는 통계적인 수치만 측정하고 이를 업무에 반영시키지는 않고 있는 것으로 파악된다. 물론 웹사이트의 방문자수가 몇 명인지, 어떤 검색어를 통하여 웹사이트에 방문하는지 등의 수치만으로도 업무에 도움을 줄 수 있다. 하지만 좋은 콘텐츠 서비스를 제공하고 있는데 반해 애널리틱스가 가진 본질적인 기능을 활용하지 못하고 있는 점은 아쉬운 점이다. 오픈아카이브가 구글애널리틱스를 구동한 2011년 4월 이후로 꾸준히 증가하는 세션 수만 보더라도 웹이용자의 참여도를 증가시키고 이들의 정보요구를 파악하여 더 나은 웹서비스를 위해 웹이용자에 대한 다각적인 분석이 이루어져야 한다. 따라서 본 논문에서는 오픈아카이브의 2014년 8월 한 달간의 구글애널리틱스 기본 지표 데이터를 분석하고 소셜네트워크를 통한 서비스 시점을 전후하여 웹이용자를 다각적으로 분석하여 그들의 정보요구를 추측한 서비스 개선방안을 제시하고자 한다.

2) 민주화운동기념사업회 오픈아카이브 기본지표 분석

앞 장의 ‘〈표 4〉 구글애널리틱스를 이용하여 민주화운동기념사업회 오픈아카이브 기본지표’를 2014년 8월 한 달간 분석하였다. 위젯을 이용한 대시보드를 활용한 기본지표 화면은 다음 〈그림 3〉과 같다.

〈그림 3〉 구글애널리틱스를 이용한 민주화운동기념사업회
오픈아카이브 기본지표 분석(2014. 8. 1~2014. 8. 31)



위의 그림을 살펴보면 가장 많은 접속국가는 한국(South Korea)이 약 92%였다. 웹이용자들은 오픈아카이브 접속 시 Internet Explorer와 Chrome을 많이 이용하였고, 운영체제는 Windows를 선호하는 것으로 나타났다.

민주화운동기념사업회 웹사이트에 어떤 ‘키워드’로 검색하여 유입되었는지 분석해보면 ‘(net set)’이나 ‘(not provided)’²⁴⁾로 나타나는 경우가 많다. 이때 ‘도착페이지’를 두 번째 측정기준으로 설정해보면 (그림 4)와 같이 주로 기관명과 관련된 단어(메인페이지로 나타나는 경우 주로 기관명을 검색했을 가능성이 높다)이거나, ‘지강헌’, ‘동일방직사건’, ‘김대중’ 등 ‘민주화운동’과 관련된 단어가 많았다.

〈그림 4〉 도착페이지 URL을 통해 유추한 오픈아카이브 키워드

기본 측정기준: 소스/매체 소스 매체 키워드 기관		정렬 유형: 기본		
키워드	도착 페이지	획득		
		세션	새로운 세션 %	신규 방문자
		4,711 전체 대비 비율(%) 100.00% (4,711)	76.42% 평균 조회: 76.37% (0.09%)	3,600 전체 대비 비율(%) 100.06% (3,598)
<input type="checkbox"/> 1. (not set) 메인페이지	/	757 (16.07%)	55.75%	422 (11.72%)
<input type="checkbox"/> 2. (not set) "탈주범 지강헌 사건" 관련 노트/View?pRegNo=00730736	/	77 (1.63%)	98.70%	76 (2.11%)
<input type="checkbox"/> 3. (not set) 사료로 배우는 민주화운동	/Contents	59 (1.25%)	67.80%	40 (1.11%)
<input type="checkbox"/> 4. (not provided) 메인페이지	/	57 (1.21%)	38.60%	22 (0.81%)
<input type="checkbox"/> 5. (not set) 동일방직사건-"똥을 먹고 살 수 없다"	/RecordContentsView?pid=33	57 (1.21%)	94.74%	54 (1.50%)
<input type="checkbox"/> 6. (not set) 검색결과값	/Search	56 (1.19%)	3.57%	2 (0.06%)
<input type="checkbox"/> 7. (not set) 김대중 대통령 후보 보라매공원 유세현장	/PhotoView?pPhotoId=00756574	38 (0.81%)	100.00%	38 (1.06%)
<input type="checkbox"/> 8. 민주화운동기념사업회	/	36 (0.76%)	13.89%	5 (0.14%)
<input type="checkbox"/> 9. (not provided) 부천시 상고문 사건	/Collection?evtNo=10000072	31 (0.66%)	90.32%	28 (0.78%)
<input type="checkbox"/> 10. (not set) 1987년 노동자대투쟁 - 울산 현대중공업	/contents/1987	31 (0.66%)	70.97%	22 (0.61%)

24) ‘(not provided)’는 구글에서 이루어진 ‘자연어 검색’에 대한 키워드를 구글이 제공하지 않도록 감춰놓았다는 의미이다. 구글이 이렇게 한 이유는 구글애널리틱스를 이용해서 데이터 분석 후 SEO(검색엔진최적화, Search Engine Optimization) 상에서 키워드 최적화라는 방식으로 인위적으로 유입을 증가시키거나 어뷰징(인터넷 포털 사이트에서 언론사가 의도적으로 검색을 통한 클릭수를 늘리기 위해 동일한 제목의 기사를 지속적으로 전송하거나 인기검색어를 올리기 위해 클릭수를 조작하는 것 등이 이에 해당된다, 한경 경제용어사전, <http://s.hankyung.com/dic/searchList.php>)을 방지하기 위한 목적이다.

오픈아카이브의 기본지표를 분석하던 중 특이한 점이 몇 가지 있었다. 국내 검색엔진 점유율은 ‘네이버’가 80%정도인데 반해 오픈아카이브에 들어오게 된 계기 중 구글 검색엔진을 경유한 웹이용자가 전체 절반 이상이라는 것이다. 이는 기관의 담당자와의 인터뷰에서 이유를 알아낼 수 있었는데 민주화운동기념사업회측에서 2년 전, 구글 사이트에 기관 사이트 내 소스를 크롤링(crawling)²⁵할 수 있도록 허가했던 이력이 있다는 점으로 추정할 수 있었다. 당시 구글이 크롤링한 소스를 검색결과 상위에 링크함으로써 민주화운동기념사업회 오픈아카이브 페이지로 유입되는 원인 중 하나로 추측하고 있다.

다른 한 가지는 8월 한 달간 한 번만 방문한 이용자가 76%, 처음 방문이거나 1년 이상 방문을 하지 않은 이용자가 88%, 재방문자의 비율이 24%인 점은 전체 웹사이트에서 이용자의 충성도가 대체적으로 낮은 것으로 평가되었다. 재방문자들은 사이트에 머무르는 시간도 신규방문자보다 길었고 사이트 충성도가 높은 것으로 추정된다. 따라서 웹이용자의 눈길을 끌만한 콘텐츠를 메인 페이지 또는 외부 사이트로부터 유입되었을 때 처음 보이는 도착 페이지에 게재하여 세션시간을 늘릴 수 있다. 또 민주화운동에 관심 있는 이용자층을 겨냥한 적극적인 서비스 정책을 세워 문제점을 해결해나갈 수 있을 것이다. 구글애널리틱스를 통하여 살펴본 민주화운동기념사업회 사료관 오픈아카이브의 웹이용자들의 기본지표를 ‘기록관 이용자 세분화 변수’에 따라 분석하여 그 결과를 표로 정리하면 다음 <표 5>와 같다.

25) 무수히 많은 컴퓨터에 분산 저장되어 있는 문서를 수집하여 검색 대상의 색인으로 포함시키는 기술(IT용어사전, 한국정보통신기술협회).

〈표 5〉 구글애널리틱스를 이용한 오픈아카이브 웹이용자 세분화

기록관 이용자 세분화 변수명	구글애널리틱스 웹이용자 분석 항목	오픈아카이브 웹이용자 분석결과 (가장 많은 비율을 차지한 결과임)	
인구통계 변수	연령, 성별	추가코드 삽입하지 않아 분석 불가	
	언어	한국어 90% (※영어 8% / 기타 2%)	
지리적 변수	위치	한국 92%	
심리적 변수	(사용자)행동흐름	메인페이지 → 검색결과페이지 → 메인페이지 → 검색결과페이지	
이용 변수	관심분야	추가코드 삽입하지 않아 분석 불가	
	획득 채널	자연어 검색 57% (※ 구글 검색엔진 51%)	
	계재빈도 및 방문빈도	세션수 1: 76%	
		마지막 세션 이후 경과 일수 0: 89%	
	신규방문 vs 재방문	76% vs 24%	
	참여도	세션 시간 0-10초: 69%	
	기술환경	브라우저	Internet Explorer 37% / Chrome 35%
		운영체제	Windows 59.97%
키워드	기관명 관련 단어, 민주화운동 관련 단어		

* 비율의 소수점 이하는 반올림함.

민주화운동기념사업회 오픈아카이브는 온라인으로 다양한 콘텐츠 서비스를 제공하고 있다. 이러한 콘텐츠 서비스를 누가, 어떻게 활용하고 있는지 알기 위해서는 애널리틱스를 활용한 정밀한 분석이 필요하다. 이를 위해서는 구글애널리틱스의 ‘보고서’가 제공하는 통계적 수치와 이를 응용한 ‘맞춤설정’ 항목을 함께 이용한 다각적인 분석을 하는 것이 필요하다.

최근 모바일을 통한 웹이용자가 늘어남에 따라 민주화운동기념사업회 자료관도 대표적인 소셜네트워크서비스인 페이스북과 트위터를 통하여 자료관이 소장 중인 자료를 활용한 콘텐츠 서비스를 실시하고 있다. 이는 기관의 홍보 효과와 함께 일반인들도 쉽게 민주화운동 관련 자료에 접근할 수 있는 기회를 제공하고 있는 것으로 평가된다.

민주화운동기념사업회 사료관은 지난 2014년 9월, 페이스북에 사료관에 소장중인 '문익환의 삶'과 '청춘 문익환' 두 가지 콘텐츠에 대한 소개와 '구글오픈갤러리채널'로의 링크를 게재하였다. 페이스북 페이지 메인 화면에는 민주화운동기념사업회 사료관 홈페이지의 URL이 링크되어 있고 민주화운동기념사업회 사료관의 구글오픈갤러리채널에도 민주화운동기념사업회 사료관 홈페이지로 연결되는 URL이 게재되어 있다. 민주화운동기념사업회 사료관 홈페이지와 구글오픈갤러리채널, 페이스북 페이지에 대한 연결이 되어 있었다.

우선 페이스북 페이지에 '문익환의 삶'을 게시한 2014년 9월 3일 전후 두 달간 즉, 2014년 7월 3일~9월 2일과 2014년 9월 3일~11월 2일의 잠재고객 개요를 비교해보았다.

페이지뷰 수는 39,897에서 47,529로 19.13%나 증가하였다. 세션 수는 8,992에서 10,267로 14.18% 증가하였으며 세션 당 페이지 수는 4.34% 증가하여 사이트에 방문하여 보는 페이지수가 다소 증가하였음을 알 수 있다. 웹이용자가 머무르는 시간인 '평균세션시간'도 2분 54초에서 3분 17초로 미미하지만 평균 23초 늘어났다. 다른 항목은 모두 수치가 증가한데 반해 이탈률은 3.14% 감소하였다. 수치가 감소하였다고 하여 부정적인 의미로 착각할 수 있지만 이탈률이 줄어들어 들었다는 것은 웹이용자가 처음 접속한 페이지만 보고 사이트 접속을 종료한 수치가 줄어들었으므로 긍정적인 의미로 평가해야 한다. 즉 '문익환의 삶' 콘텐츠 게시 이후 모든 측정 항목에서 긍정적인 효과를 거둔 것으로 평가할 수 있다.

페이스북에 '문익환의 삶' 콘텐츠 게시 일자로 살펴본 결과 웹이용자 수치가 늘어났다. 더불어 페이스북의 콘텐츠를 보고 오픈아카이브에 관련 자료를 이용하기 위해 페이스북을 통해 오픈아카이브로 방문한 웹이용자를 측정하기 위해 '획득' 항목의 '네트워크 추천'으로 전후 기간의 수치를 비교하여 보았다. 페이스북으로 유입된 세션은 2배 정도 증

가하였으나 페이지뷰 수는 약 7%, 평균세션시간은 두 배 이상 감소하였고, 세션 당 페이지 수는 약 22% 감소한 것으로 측정된다. 이는 페이스북으로 유입된 방문자만 증가하였을 뿐 사이트에 접속하여 시청한 페이지 숫자나 머무른 시간 등이 감소하였다. 그에 반해 눈에 띄는 점은 트위터(Twitter)²⁶⁾로 유입된 세션이 5배나 증가하였고 페이지뷰 수, 평균 세션시간에서 큰 폭의 비율로 증가한 것이다. 민주화운동기념사업회가 페이스북에 구글오픈갤러리채널로의 연결을 시도하였음에도 불구하고 트위터로의 유입이 큰 폭으로 증가한 것으로 미루어 앞으로 페이스북 보다 트위터에 구글오픈갤러리채널 또는 오픈아카이브 웹사이트로의 연결을 위한 콘텐츠 게시에 노력하는 것이 더 효과적일 것으로 추측된다.

민주화운동기념사업회 사료관은 웹이용자 분석을 위해 국내 다른 기록물관리기관보다 앞서 구글애널리틱스를 도입하였다. 구글애널리틱스를 처음 도입한 2011년 4월 이후 약 4년이 넘는 시간동안 쌓인 웹로그 데이터를 잘 분석하여 기관의 웹서비스 상태를 점검하고 웹이용자들의 정보요구와 문제점을 진단하여 서비스의 개선점을 처방해야 할 것이다. 한 달여간의 시간동안 구글애널리틱스를 통하여 민주화운동기념사업회 오픈아카이브의 웹이용자를 분석한 사례를 통해 기록정보 서비스 측면에서의 개선점을 정리해보면 다음과 같다.

첫째, 인구통계 추적코드를 추가 삽입하여 웹이용자의 연령대와 관심분야에 대한 세밀한 관찰이 필요하다. 웹사이트 분석에 대한 많은 연구에서 웹사이트 방문자의 인구통계변수는 인구학적 세분화를 가능하게 하며 웹사이트 이용 분석에 긍정적인 영향을 미치는 것으로 나타났기 때문이다. 민주화운동기념사업회의 사이트는 이미 회원제로 운영하고 있으므로 로그인을 하여 방문하는 웹이용자들에 대한 인구통계 변수에 대한 정보를 분석할 필요가 있다. 그리고 그들의 연령대와 성별

26) 트위터, <http://twitter.com/>

뿐만 아니라 직업적 유형과 관심분야, 접근의도 등을 파악하고 그들이 어떤 정보요구를 가지고 사이트에 방문하는지 추측해내는 시도가 필요하다.

둘째, 민주화운동기념사업회 오픈아카이브는 '민주화운동'과 관련한 사료를 웹으로 서비스하는데 있어 선구적인 역할을 해온 기관이다. 기관이 보유한 사료를 지금처럼 꾸준히 웹으로 서비스하고 더불어 관련 있는 콘텐츠의 연결, 다른 관련된 사이트와의 연계 서비스 등 웹이용자가 사이트에 오래 머무르고 좀 더 많은 콘텐츠 이용을 유도하기 위한 서비스 개선을 통하여 신규방문자를 재방문자로 전환시키고 방문자의 사이트 충성도를 증가시킬 필요가 있다.

셋째, 민주화운동기념사업회 오픈아카이브 사이트는 페이스북이나 트위터 등 소셜네트워크서비스에 콘텐츠를 게재한 이후 유입된 웹이용자 수가 증가한 것으로 평가되었다. 웹이용자수가 증가하였을 뿐만 아니라 처음 접속한 페이지만 보고 사이트를 떠난 웹이용자 비율이 전체적으로 줄어든 것으로 나타났다. 소셜네트워크서비스를 통하여 지속적으로 콘텐츠 서비스를 제공하여 기관의 홍보뿐만 아니라 사이트로의 유입을 유도하여 사이트 충성도가 높은 웹이용자를 많이 끌어들이는 전략을 세울 수 있다.

넷째, 구글애널리틱스는 유료로 제공하는 프리미엄 기능도 제공하고 있다. 무료버전과 다른 점은 365일 24시간 내내 서비스를 지원받을 수 있고, 하루가 지나야 보고되는 데이터를 4시간에 한 번씩 업데이트되어 보고받을 수 있다. 또, 구글 전문가로부터 교육을 받거나 전문가의 업무지원을 받을 수 있다는 점이다. 무료버전은 데이터를 처리하는 양이 한 달에 1천만개이고 보고서에 이용가능한 데이터도 5만개에 한정되지만, 프리미엄 버전은 데이터 처리량이 10억개이며 보고서에 이용가능한 데이터도 샘플링되지 않은 보고서에 2M(메가바이트)의 용량을 할당해 준다. 이런 기능들 중에서도 주목할 점은 샘플링되지 않은 보고서를 제

공한다는 점이다. 구글애널리틱스는 그 특징상 수집된 모든 데이터를 보여주는 것이 아니라 보이지 않는 웹이용자들이 사이트에 남기고 한 웹로그를 분석한 데이터를 압축하여 비율적으로 환산한 수치이다. 그러므로 100% 정확한 수치가 아니라 예상수치들을 보여주는 것이므로 이를 관리자가 기관의 특성에 맞게 분석해야 하는 것이다. 즉, 프리미엄 기능을 이용하면 분석전문가가 이런 관리자의 분석 역할까지 대신 처리해준다. 하지만 연간 사용료가 15만달러이므로 중견기업이나 예산이 풍족한 기관이 아니고서는 업무를 쉽게 내지 못한다. 민주화운동기념사업회 사료관에서도 예산의 확보가 가능하다면 프리미엄 버전을 도입하여 좀 더 정확한 웹이용자 데이터로 기록정보서비스를 제공할 것을 제안한다.

5. 결론

웹애널리틱스를 활용하여 웹이용자의 행태를 추적하면 바로 방문자 수, 페이지뷰 수 등의 기본지표를 얻을 수 있다. 여기에서 얻은 수치를 세분화하는 것은 웹이용자를 세분화한다는 것과 같은 의미이며 더 나아가 웹사이트가 단순히 페이지만 제공하는 것 이상의 성과를 거두고 있는지 확인하기 위함이다. 게다가 앞에서 언급하였던 정경희(2007)의 ‘기록관 이용자 세분화 변수’와 구글애널리틱스의 웹이용자 분석 항목은 유사한 점이 많은 것으로 보인다. 이는 구글애널리틱스를 통하여 아카이브 웹이용자를 세분화하고 이들의 정보요구를 유사한 그룹으로 나누어 그들의 이용행태를 파악하는데 유용한 도구라는 점을 확인할 수 있다. 이를 토대로 아카이브 이용자를 세분화하고 그들의 요구를 파악하여 서비스를 개선해나가는데 있어 몇 가지 전략을 제안하고자 한다.

첫째, 연령, 성별, 관심분야 등 인구통계 변수를 통한 주요 이용자층

을 확인해야 한다. 주로 교육용 콘텐츠를 많이 서비스하고 있는 아카이브 기관의 경우 웹이용자 연령대의 파악이 중요하다. 국가기록원의 경우 콘텐츠 유형, 이용자 그룹별로 콘텐츠 서비스를 하고 있다. 예를 들어, ‘조선왕의 수결’이라는 주제로 그림이 게재되어 있는데 초등학교의 경우 그림과 텍스트만으로는 이해하기 어려울 수 있다. 학생들은 ‘수결’이라는 주제에 대해 좀 더 쉽게 이해를 도울 수 있는 콘텐츠를 원할 것이다. 따라서 나레이션을 통하여 왕과 대통령의 수결이 만들어지는 과정을 동영상 형식의 콘텐츠로 제공하는 방안을 모색해보면 좋을 것이다.

둘째, 방문자수의 다각적인 응용을 통한 이용자 요구를 분석해야 한다. 방문자의 숫자만으로는 웹서비스를 개선하기 위한 유용한 정보가 될 수 없다. 예를 들어 방문자수라는 단순한 지표를 사이트에 머물렀던 시간으로 세분화하여 방문만 했던 웹이용자와 사이트에 오래 머무르며 콘텐츠를 이용한 웹이용자를 구분하여 분석하여 의미있는 활동을 한 웹이용자의 비율을 측정하는 것이다. 또 유입경로를 분석하여 유입이 많은 사이트와의 연계 서비스를 고려해 볼 수 있고, 키워드를 분석하여 어떤 콘텐츠가 인기많은 것인지 알아낸 다음 이를 활용한 서비스를 제공하는 것도 방법이 될 수 있다.

셋째, 최근 소셜네트워크서비스와 모바일로 접속하는 이용자가 증가함에 따라 소셜네트워크나 연관 사이트를 통한 기관 또는 기관에 소장 중인 기록물을 홍보하는 서비스 전략을 세울 수 있다. 소셜네트워크는 웹이용자가 직접 콘텐츠를 게시하고 다른 콘텐츠에 댓글도 작성할 수 있으며 콘텐츠의 공유 또한 쉽다. 이는 이용자의 적극적인 참여를 유도할 수 있고 아카이브 관리자와 이용자의 쌍방향 소통이 이루어진다는 점에서 효과적인 매체라 할 수 있다.

넷째, 웹사이트 내 콘텐츠 서비스를 개선하는 방안에 대한 전략이 필요하다. 우선 웹이용자들이 보는 웹인터페이스의 구조를 이용자 편의

에 초점을 맞추어 설계하여 웹이용자들의 사이트 충성도를 높일 수 있다. 또 웹이용자들이 사이트 접속 시, 어떤 브라우저와 운영체제를 활용하는지 분석하여 좀 더 쉽게 접근할 수 있는 웹 환경을 조성한다. 또한 콘텐츠를 지속적으로 갱신하여 웹이용자로 하여금 해당 사이트가 지루하지 않게 느껴지도록 한다. 연관 콘텐츠를 배너에 삽입하여 사이트에 머무르는 시간을 증가시키는 것도 방법이 될 수 있다. 그리고 콘텐츠를 제공하는 기관 입장뿐만 아니라 이를 이용하는 웹이용자의 의견을 듣기 위한 쌍방향 소통이 가능한 공간을 마련한다. 예를 들어 사이트 내에 자유게시판을 마련하여 이용자들이 자유롭게 의견을 피력할 수 있게 하고 콘텐츠에 댓글을 달 수 있는 공간을 삽입하여 콘텐츠에 대한 반응을 보일 수 있게 한다.

아카이브는 웹의 기술적 발전에 따라 기록물의 형태, 운영방식, 서비스 등 모든 면에서 전환점을 맞이하고 있다. 아카이브의 서비스가 웹 중심으로 변하고 이에 따라 기록정보의 웹서비스의 중요성도 부각되고 있다. 시대의 흐름에 따라 아카이브 웹서비스를 이용하는 웹이용자의 정보요구 또한 점점 더 다양해지고 전문화되고 있다. 이러한 웹이용자의 행태를 분석하고 다양한 정보요구를 파악하기 위한 수단으로 웹애널리틱스가 있다. 여러 종류의 웹애널리틱스가 있지만 그 중에서도 유료 분석 도구가 가진 고급기능을 제공하면서 무료로 사용할 수 있는 구글애널리틱스를 아카이브 웹이용자 분석에 도입할 것을 제안한다.

사례 조사를 위하여 웹로그분석을 하고 있는 기록관리기관을 찾아보았지만 이를 실시하고 있는 국내 아카이브 기관은 단 몇 군데 밖에 없었고, 실제 웹이용자 서비스 개선을 위한 전략을 수립하는데 적용하는 기관은 없었다. 이는 웹로그분석 서비스에 관한 예산, 전담 인력, 아키비스트의 인식 부족의 문제가 주요 원인으로 추측된다. 기록관리기관의 웹서비스의 중요성이 부각되고 있는 시점에서 아키비스트의 인식 제고가 시급하다. 본 연구는 웹애널리틱스가 어려운 기술(technic)이 아

니라 아카이브의 업무 효율을 증대시킬 수 있는 좋은 도구이라는 점을 강조하고자 한다. 이와 함께, 아키비스트의 인식 제고와 함께 충분한 예산과 인력의 확충도 뒷받침 되어야 할 것이다.

ABSTRACT

A Case study analysing the users of archives through web analytics

Lee, Hyoeun · Yim Jin Hee

Record Information Services is an aggressive action of connecting documentaries focusing on the information needs of user. However, recent studies on the parliament's written information service recognize the necessity that it should segment the user's information requests, and provide personalized service, but have not discussed for specific cases or measures.

While the importance of Web services written with the proliferation of information and popularization of the Web is emerging right to know but, it is not being performed properly by lack of sufficient manpower and budget along with lack of recognition in hands-on sites upon the user analysis. So, while increasing the efficiency of the hands-on workers of Record Information Services, the introduction of analytical tools that can be utilized in low budget agencies is needed.

Web analytics is to analyze the behavior by analyzing Web logs which web users have left you visit the site. To estimate the behavior they want to request information of the analyzed Web user aims to provide a Web service, the Web service further continued improvement. There are several types that include among them Google Analytics offering a variety of analysis items for free and all over the world, many people are already using.

This study introduces a Google Analytics web analytics focused and proposes a service improvement plan with specific web user segmentation analyzes the cases of Korea Democracy Foundation of Open Archives introduces them to the actual institutions.

Key words : Web Analytics, Web log analysis, Google Analytics, Archives, User analysis, User segments, Archival Information Service