

제1차 기록관리 현안과제세미나 개최 결과 보고

□ 세미나 개요

- (주제) 구글 애널리틱스를 활용한 국가기록원 포털 웹로그 분석
 - 사회 : 김주연 연구관(연구협력과)
 - 발표 : 서경란 주무관(기록정보기반과), 김영희 주무관(기록서비스과)

발표 제목	발표자
구글 애널리틱스를 활용한 국가기록포털 웹로그 분석	서경란 주무관
구글 애널리틱스를 활용한 대통령기록포털 웹로그 분석	김영희 주무관

- 토론 : 오효정 교수(전북대학교), 원종관 과장(서울기록원)
- (일시/장소) '19. 5. 7.(화) 14:00~16:00 / 국가기록원 제1회의실(403-1호)
- (참석) 총 27명
 - (내부) 국가기록원장, 부서장, 관련 담당자 및 관심있는 직원 25명
 - (외부) 해당 분야 전문가 2명

□ 주요 논의 내용

- (국가기록포털) GA를 활용한 사용자 유입경로, 방문자·페이지 뷰 수, 접속 페이지 순위, 국가별 분포, 접속 기기 등 현황 및 향후 서비스 방안
- (대통령기록포털) 사용자 유입현황, 포털 내 이용자 이동 경로, 주요 키워드 및 이용 메뉴현황 및 접속환경 등 현황 및 향후 서비스 방안
- 브라우저/기기 분석의 활용 목적, 검색 활용도 개선 방향, 기록의 특화된 분석을 위한 툴 제작의 필요성 및 가능 여부, 분석전문가 양성 등

□ 향후 계획

- 제4차 기록관리 연구세미나 개최(정부대전청사 2동 207호) : '19. 5. 22.
 - ※ 행정정보 데이터세트 기록관리 방안(전자기록관리과 이주광 연구관)

○ (오효정 교수) ① 국가기록원이 웹로그 쌓기 시작한 것은 의미 있음. 다만 국가기록원과 대통령기록관의 포털 운영을 위해 목적성 설정이 보다 시급함. 현실적으로 일반인과 기록 관련 전문가의 유저를 동시에 만족시키기 어려우므로 target user를 설정한 후 UI를 설계할 필요가 있음. 사용자의 세션 시간이 짧다는 것은 원하는 정보를 못 찾았거나 없을 가능성이 높음. 일선 교사 80% 이상이 국가기록원 포털의 존재를 모르고, 기록포털 교육용 콘텐츠 활용도 5~10% 내외에 불과함.

② GA는 공공포털에 적합하지는 않지만 올해 했으니 내년에 다시 분석하는 시즈널 분석이 필요하며, 모바일이나 웹을 통한 질의를 축적해서 분석하면 좋을 듯함. 국가기록원은 검색 결과를 어떻게 잘 할 것인지를 포커싱해야 하며, 대통령기록관은 기록관의 소개 및 자판을 편리하게 볼 수 있게 하는 방향을 고려해야 할 것임. 아울러 주기적이고 장기적으로 로그 결과를 축적만 해도 향후 분석하는 데 용이할 것임.

○ (원종관 과장) ① 구글 애널리틱스는 웹사이트 트래픽 분석하는 도구로 출발했다가 마케팅 톨로 영역과 정체성을 확장하고 있음. 애널리틱스는 다른 통계 분석하는 엔진이나 기술보다는 유용하거나 콘솔을 보기 쉬운 경우는 있지만, 기관이 일정하게 마케팅 하지 않으면 애널리틱스를 사용하는 것은 분명히 한계에 부딪힐 것임. 따라서 기관이 무엇을 할 것인지 조직적으로 결정 내리는 것이 훨씬 중요한 문제임.

② 대통령기록관이 새로운 시도나 친절하게 제공하려고 노력하는 것을 많이 느끼고 있음. 최고 권력자의 사소한 일상을 담은 ‘대통령의 식탁’ 같이 일반 국민의 흥미를 유발할 수 있는 주요 기록을 메인 화면에 전진 배치하는 것은 좋은 일례이며, 사용자가 탐색에 집중할 수 있도록 앞으로 빼주는 것은 앞으로도 필요할 것임.

⇒ (이소연 국가기록원장) 이용자가 포털 메뉴를 많이 보는 것을 전진배치하는 것이 맞긴 하지만 수요가 많은데 뒤에 있을 수 있기 때문에 이용률 대비 전진배치라는 공식을 그렇게만 생각할 필요는 없다고 봄.

③ 외국 접속률이 10% 이하면 개인적으로 외국어 홈페이지 운영은 불필요하다고 생각함. 외국인 접속이 아니라 한국인이 외국에서 접속했을 가능성을 감안할 때 모바일의 운영체계보다는 PDF 뷰어 다운의 편의성 및 이미지의 텍스트 기반 전환 다운 방식을 잘 활용하면 더욱 효율적일 것임.

④ 모바일이 유행이지만 모바일 운영의 실효성을 생각할 때 많은 고민이 필요함. 5년간 모바일을 운영해 본 입장에서 반응형 웹 운영 효과가 적어서 데스크탑 웹만 운영할까 고민한 바도 있음. 연구자들은 웹사이트에서 자체 필터링이나 정렬이 잘 되어 있는 것을 원치 않고 엑셀로 다운 받는 것을 원하고 있음.

=> (오효정 교수) 사용자마다 니즈가 다른데 법령 사이트의 경우 이동 시 모바일을 수시로 이용하는 반면, 기록을 보고자 하는 입장에서는 모바일 접속은 당연히 안할 것임. 사진아카이브나 콘텐츠는 동영상으로 모바일 이용이 가능하지만 기록물은 잘 안 볼 것이므로 교사용 콘텐츠 등 몇 가지 케이스에 집중해서 모바일을 운영하면 좋을 듯하며, 모든 페이지에 반응형을 적용하는 것은 ROI(Return On Investment, 투자수익률)가 상당히 떨어질 것임.

=> (원종관 과장) 정보로 찾아오는 길, 주차장 안내, 전시 안내의 경우는 모바일이 상당히 효율적일 수 있음.

⑤ 대통령기록관은 메뉴가 많은데 촘촘히 조직되어 있지 않거나 일반적으로 콘텐츠를 소비하는 파이프가 설계되지 않아서 향후 GA 분석이 어려워질 수 있으므로 메뉴를 단순화 하는 것이 좋을 듯함. 아울러 검색 엔진 최적화를 제안하고 싶음. 개별 사이트에 검색엔진을 사서 채택하는 것은 ROI가 떨어질 수 있으므로 차라리 그런 정보들을 구글 검색 엔진에 잘 크롤링 될 수 있게 최적화된 검색엔진을 만들어 놓는 것이 더 중요하며, 그럴 경우 트래픽이 더 증가할 것으로 보임.

⑥ 국가기록포털의 경우, 오가닉 서치가 직접 접속보다 훨씬 높은 이유는 무엇인지? 66%의 세션이 첫 번째 화면 접속 후 이탈하는데 이것을 어떻게 받아들이고 있는지? 웹사이트를 기능분리하는 의도가 무엇인지?

=> (서경란 주무관) 기록원 사이트의 직접 접속은 용어가 어려워 적을 수 있음. 다만 오가닉 서치가 많은 이유는 기록원 자체 검색보다는 네이버 혹은 구글에서 본인이 원하는 특정 키워드를 넣고 검색했을 가능성이 있음. 66%의 세션이 첫 화면에서 나간 결과에 대해서도 네이버를 통해서 들어온 사용자들이 특정 페이지 한두 페이지만 검색하고 나간 유형이 많았다고 판단함.

=> (민선미 사무관) ① 네이버에서 포털에 들어온 첫 시작페이지는 메인 화면이 아니고, 포털에서 수집된 웹 페이지 형태로 링크가 나열되어 있는 기술계층별 검색이 첫 페이지였음. 기술계층별 검색이 높은 이유는 기술 계층별 검색 안에 담고 있는 텍스트 키워드 정보가 검색 시 걸리는 목록이 많아서 거기를 통해 유입되는 경향이 컸다고 봄. 그런 점에서 기술서의 의미를 잘 모르는 일반인의 경우 첫 페이지만 보고 이탈하는 경우가 많았을 것으로 추측되며, 66%는 첫 페이지 유입경로가 그렇다 보니 거기에 원인이 있었던 듯함.

② '11년 모바일 서비스를 실시했지만 이용이 저조해서 '16년도에 폐지함. 현재 모바일 이용률이 40% 이상 높아지고 있기 때문에 주요 페이지에 대해서 반응형 웹 형태로 포털에서 제공하는 기관 소개 등은 반응형으로 새롭게 구현할 계획임. 용량이 큰 기록정보들은 모바일 제공 대상은 아니며, 콘텐츠 등을 선별·제공하는 것이 좋을 듯함.

③ 첫 화면이 66% 나오는 것에 대해 GA가 어느 정도 디테일 하게 보고서를 줄 수 있을지 의문임. 공공기관의 유료 서비스는 우리에게 유용하다고 보기 어렵고, 이것보다 더 나은 통계를 잡을 수 있을지 모르겠음. 우리의 관심은 어떤 행위에 대한 절차나 진행 정도가 아니라 어떤 주제에 대해서 어느 시기에 어떤 것들을 먼저 활용 가능한 상태로 줄 수 있는지에 대한 것으로 그런 내용을 분석하는 틀은 아직 찾지 못했음.

④ 사용자의 기호를 고려한 포털 화면 구성은 가능하지만, 무엇을 서비스 할 것인가를 정하다 보면 검색 로그를 분석해야 하는데 검색 자체가 메타데이터가 약하다 보니 검색 결과도 제대로 안 나오고 사용자들이 원하는 서비스를 받았는지의 여부도 파악이 어려운 상황임.

=> (오효정 교수) GA의 용도는 매년 서비스하는 것을 받는 것이 아니라 리뉴얼, 신상품 출시, 매출 저하 시에 단기 컨설팅을 받기 위한 것임. 이번에는 처음 했으니까 여기까지 하고 1년 혹은 1년 6개월 후에 다시 하면 되며, 이것을 5~6년 계속할 필요는 없음.

○ (이소연 국가기록원장) ① 기록 이용에 관련된 데이터를 쌓고 분석하는 틀을 만들 필요가 있는지? 국가기록포털, 대통령기록포털 두 곳에서 수요가 있고 서울기록원, 경남기록원 등 기록을 가지고 포털 서비스를 제공하는 기관들이 범용으로 사용할 수 있는 GA와 차별되게 효용을 얻을 수 있는 틀을 만들 수 있을지?

=> (오효정 교수) 로그는 쌓기만 하는 것이고 분석은 연구진에 의해 이루어지는 것으로, 로그 쌓기와 분석은 별개로 운영해야 함.

=> (이소연 국가기록원장) 로그를 상시적으로 분석할 수 있는 업무가 대통령기록관은 기록서비스과에서 포털까지 담당하고 있고, 본원의 포털 관리는 기록정보기반과가, 서비스는 공개서비스과에서 담당하고 있어 오늘 발표는 공개서비스과도 함께 공유해야 할 내용으로, 특히 기술 계층별 검색 관련은 현재 상태에서도 분석할 수 있을만큼 공개서비스과와 공유해야 할 사항임. 그런 점에서 구글 애널리틱스로 기록의 특화된 분석이 불가능하다면 서비스 담당부서에서 정기적으로 데이터를 분석, 관리할 수 있는 틀은 있어야 된다고 생각함.

=> (원종관 과장) 분석 결과에서 여러 가지 그래프를 제시하고 있지만 6개월짜리 그래프는 큰 의미가 없고, 평소에 로그들을 쌓아놓을 수 있는데 매주, 매달 들여다 볼 필요는 없음. 우리가 어떤 행위를 했을 때 이 행위가 실효가 있는지를 판단하기 위해서 GA를 운영하는 것임. 따라서 GA는 기관이 바깥으로 콘텐츠를 내보내거나 검색엔진 추적을 통해서 트래픽을 끌어올릴 수 있다는 것을 검증하는데 활용했으면 좋겠음. 서울기록원도 GA를 크게 생각하고 있지 않음. 디지털 컴포넌트, 디지털 자원들을 얼마나 효율적으로 보여줄 것인지에 대해서는 기록이건 상품이건 큰 변동이 있다고는 보지 않음.

=> (오효정 교수) 툴은 쌓는 것은 일반적이므로 도구를 만들 수 있으며, 분석은 니즈에 맞추어야 하기 때문에 국가기록원과 대통령기록관의 전용 분석으로 범용을 사서 끼워서 되는 것이 아님.

=> (민선미 사무관) 분석을 하려면 어떤 통계가 나왔을 때 수치를 보고 이 수치가 나온 원인을 예측하기 위해서는 추가적인 데이터가 있어야 분석이 가능할 것임.

=> (이소연 국가기록원장) 기록의 특화된 툴이 필요한 것이 아니라 이 데이터를 좀더 분석하는 것이 중요함. 로그 데이터 갖고는 완벽한 툴이 있어도 얻어낼 수 있는 정보가 제한적임. 로그로 전반적인 내용을 본 다음에 집중된 user 스터디를 통해서 찾아낼 수 있는 것임. 포커스 뷰 인터뷰나 특정 세션 패턴을 갖고 있는 그룹들을 찾아 들어가서 할 수 있는 방법들이 있을 것임. 로그 데이터는 다른 방식의 유저 스터디와 병행 되었을 때 입체 적인 결과를 내는 것이 맞지만 구글 애널리 틱스만 가지고 할 수 있는 것보다 좀더 기록의 특화된 로그를 쌓고 분석을 지원할 수 있는 툴이 필요하다고 생각함.

② 월간운영 보고와 이 분석결과의 서식을 맞추어 필요가 있는지 검토 바람.

=> (민선미 사무관) 월간운영 보고는 유저 같은 경우 IP 로그를 이용해서 유지보수업체에서 로그를 쌓아서 그것을 분석한 것이고, 본 분석결과는 작년 8월부터 GA를 통해 쌓고 있는데 유저 식별방법부터 모두 다르기 때문에 둘 중 하나를 선택해야 함. 방향을 정해서 통계를 하나로 유지해주는 것이 맞을 듯함.

③ 기술계층별 검색이 12만건이라는 수치는 작은 것이 아니기 때문에 좀 나누어서 봐야 할 듯함. 이 내용은 정책부뿐만 아니라 서비스부에서도 같이 고민했으면 좋겠음.

④ 서울시 정보소통광장의 경우 반 이상의 사용자가 서울시 직원이었고, 대통령기록관도 성남에서 많이 검색했음. 이용자 분석을 시도할 때 내부 고객과 외부 고객을 나누어 볼 필요가 있으며, 외부 고객보다 내부 고객 수가 많다고 해서 절망할 것도 아니고 내부 고객이 결국은 더 직접적인

서비스를 제공하기 위해서 포털을 잘 사용한다는 것도 중요함.

⑤ 기록 포털에서 마지막 목표가 무엇인지 셋팅해 볼 수 있다고 봄. 찾아오는 길이나 전시 안내 같은 것은 거기가 목표 페이지라고 상정해도 좋을 듯하고, 질의어를 입력해서 3개의 결과 중 1~2개를 보고 나갔다면 이정도면 목표가 달성되었다고 볼 수 있을 것임. 목표도 반드시 100% 일치한다고 볼 수 없지만 로그 데이터를 분석하는 과정에서 나온 것을 가지고 목표를 가상으로 상정하면서 그 내용을 유저 스토리에서 입체적인 맥락이 닿는 액티브를 모아보면서 목표를 재상정하여 정렬해 나갈 수 있을 것임.

○ (김현진 연구사) 시작 페이지에서 첫 번째 페이지 그 사이, 그 이후까지 단계별로 되어 있는데 이것은 연속성이 있다고 봐야 하는지? 예를 들어 시작 페이지가 기술계층별 검색에서 12만여 건, 그다음이 2만6천여 건인데 12만건 중에 2만6천여 건이 페이지에 남아있다고 볼 수 있는지?

=> (서경란 주무관) 그렇지는 않고 이는 이 페이지에서 다음 페이지로 연결된 트래픽 수를 나타낸 것으로 정확히 1:1이 되지는 않음.

○ (이소연 국가기록원장) 오늘 놀랄만한 사실은 이용자의 기술계층별 검색 수치가 상당히 많이 나왔다는 것임.

=> (오효정 교수) 기술계층별 검색에서 검색 결과가 많이 나옴. 다만 이용자가 만족하는지는 모름. 검색기는 텍스트가 있는 것에 결과를 보이기 때문에 풀텍스트 서치를 하지 않는한 텍스트가 가장 많이 있는 곳이 계층별 기술서 쪽이어서 앞으로도 계속 검색하면 계층별 검색이 검색 결과로 가장 많이 나올 것임.

=> (민선미 사무관) 기록물 검색할 때 사용하는 메타데이터도 기술서 만큼의 텍스트 정보를 준다면 포털 검색 시 많이 걸릴 수 있음.

=> (김주연 연구관) 기술계층별 검색은 웹페이지 첫 화면이 아니라 두세번 째 화면에서 볼 수 있게 숨겨놨을 때는 검색율이 그리 높지 않았음. 홈페이지를 개편하면서 기술계층별 화면을 앞단으로 옮기고

나서부터 검색율이 높아졌고, 전거레코드의 경우도 마찬가지임.

○ (이소연 국가기록원장) 분석전문가가 양성될 필요성은 있다고 봄.

=> (원종관 과장) 분석의 결과가 실용적이려면 사이트맵이나 사이트 구조도 논리적이고 합리적으로 설계되어야 하는데 공공기관 사이트는 그렇지 않은 경우가 매우 많아서 그 부분은 잘 모르겠음.

=> (오효정 교수) 가장 중요한 것은 포털이 하고자 하는 바가 무엇이나임. 소장 기록물을 잘 보여주겠다, 검색을 많이 해주겠다가 목표인지, 사용자들이 원하고 궁금해 하는 몇 가지만 강조하는 것이 목적인지에 따라서 내용이 분명히 다를 수 있음.

○ (이소연 국가기록원장) 서울기록원은 어디를 목표로 하고 있는지?

=> (원종관 과장) 서울기록원은 웹사이트에서 제공하는 정보를 확인해야 되는 사람은 연구자라고 생각하고, 찾아오는 사람들을 대상으로 해서는 아주 쉽고 일부는 유치할 수도 있는 프로그램을 하고자 함.

=> (이소연 국가기록원장) 대통령기록관은 아직 연구자들에게 엑셀로 제공할 만한 공개된 데이터가 많지 않아서 일반인, 학생들 대상의 콘텐츠를 만들어 전면에 내세우고 있는 반면, 국가기록원은 1억1천 건에 해당하는 기록물을 어떻게 보여줘야 될지 고민하는 접근방식의 차이가 있음. 콘텐츠를 푸쉬하는 방식보다 연구자가 이 많은 기록의 범위 안에서 내용을 마이닝할 수 있게 돕는 방식으로 해야 할 것임.

=> (원종관 과장) 연구자들은 아카이브가 정리해낸 정보들에 많이 신뢰하지 않음. 대통령기록관 웹사이트를 보면서 의아한 것은 모 대학 모 교수님이 특정 사건에 대해 해제해 놓은 부분에 대해 소장 연구자들은 모 교수님의 의견을 전적으로 신뢰하지 않았음. 그래서 서울기록원에 바란 첫 번째는 원데이터를 달라는 것이었고, 두 번째는 원데이터가 어떻게 정리되어있는지(전거연대기, 행정기관 변천사 등), 자료들의 구조가 어떤 것이지를 원했었음.

=> (이강수 연구관) 대통령기록관의 올해 포털 운영 기본방침은 개정 시 모 교수를 중심으로 한 설명문은 전체 내리고 원데이터 중심으로

제시할 계획이며, 생산 이력, 조직변천 등에 대한 기본적인 기록에 대한 정보를 가급적 원데이터를 가공하지 않은 방향으로 나갈 것임. 대통령 기록관포털에서 변천사 내용이 빈약한데 이용 빈도수가 높아 적절히 활용할 수 있는 영역이라고 봄.

=> (이진영 지원정책과장) ① 연구자들이 원하는 방식의 기록정보를 제공하려면 포털을 어떻게 구축해서 정보를 줄 것이냐 하는 문제 이전에 기록원 내에서 우리가 가지고 있는 기록물을 어떻게 분류하고 정리, 기술할 것인가의 체계가 먼저 정돈될 필요가 있음.

② 기록물을 건부터 bottom-up해서 올라가서 정보를 기술하고 분류된 방식에 따라 검색해서 찾을 수 있도록 가공하고 그룹핑 하면 연구자들에게 정보가 제공될 텐데 지금은 제목이 들어있어야 기록물을 찾아볼 수 있는 형태임. 연구자들이 기록물관에서 제공한 정보를 신뢰하지 못한 것도 이러한 이유에서 비롯된 것으로 보임. 기관 변천연혁 등은 연구자나 개인이 접근할 수 있는 정보가 아님. 그런 것은 기관에서 만들어준 정보에 신뢰를 두고 믿는 것이고, 다만 기록관리 기관에서 포털을 통해서 제공하는 정보는 제공 방식의 문제이기 이전에 기록원에서 지적통제 하는 방식 자체가 그렇게 되어 있기 때문에 연구자가 전적으로 신뢰하기 어렵게 공백이 많은 방식으로 되어 있음. 그동안 소장 기록물 전반에 대해서 지적으로 컨트롤 하는 방식에 대한 심도있는 고민을 덜한 측면이 있음.

=> (이소연 국가기록원장) 국가기록원 본원에 있는 데이터들은 여러 가지 방식으로 지적 조직화를 시도한 흔적이 있지만 어떤 한 방향으로 모든 데이터가 정리되지는 않았음. 현단계에서 어느 나를 선택해서 전량을 조직화한 후 서비스 하면 너무 늦고, 이미 되어있는 작업들을 기본 단서로 해서 시도하는 것이 좋을 듯함. 정리기술팀에게 현재까지 나와있는 기술서 정보를 포털에서 어떻게 보여줄 것인지에 대한 방식을 고민해달라고 요구한 것도 그와 관련된 맥락임.

붙임 2

2019년 제1차 기록관리 현안과제세미나 모습

