

시계열 기반 국내 기록관리학 토픽 트렌드 분석*

Timeline-Based Topic Trend Analysis of Archives Management in Korea

박 준 형 (JunHyeong Park)**

류 법 모 (Pum-Mo Ryu)***

오 효 정 (Hyo-Jung Oh)****

목 차

- | | |
|----------------|------------------------|
| 1. 서론 | 3.2 연구대상 |
| 1.1 연구배경 및 필요성 | 3.3 데이터 전처리 |
| 1.2 선행연구 | 3.4 시계열 기반 토픽모델링 |
| 2. 이론적 배경 | 4. 연구결과 |
| 2.1 트렌드 분석 | 4.1 전·후기 10년 토픽 분석 결과 |
| 2.2 토픽모델링 | 4.2 5년 주기 영역별 토픽 분석 결과 |
| 3. 연구방법 | 4.3 시사점 |
| 3.1 연구모델 | 5. 결론 |

<초 록>

본 연구는 이전보다 더욱 빠르게 변화하고 있는 시대의 변화 속에서 국내 기록관리학 연구동향에 대한 심층적인 분석을 통해 향후 국내 기록관리학 연구의 발전 방향에 대해 모색하는데 목적이 있다. 이에 본 연구는 국내 기록관리학 관련 학술지 2종과 문헌정보학 관련 학술지 4종을 대상으로 1997년부터 2016년까지 기록관리학 관련 논문을 자동 수집하였다. 수집된 데이터는 전·후기 10년, 5년 주기로 나누고 각각 토픽모델링을 수행하였다. 5년 주기로 나눈 토픽모델링은 8가지 기록관리학 주제영역과 매핑하여 각 시기별로 특정 영역과 어느 정도 상관관계가 있는지를 파악하였다. 그 결과, 전자기록, 기록정보서비스, 다양한 기록물의 기록화 방안에 관한 연구가 점차 활성화되고 있는 것으로 나타났다.

주제어: 기록관리학, 토픽모델링, 트렌드분석, 시계열분석, HDP

<ABSTRACT>

The size and speed of archives and records information have recently increased. The purpose of this study is to find the direction of archives management research in Korea through an in-depth analysis of research trends in records management in the country. We collected articles related to archives management from 1997 to 2016 from two journals related to archives management and four journals related to library and information science in Korea. The collected articles from the first and second halves of the decade were subjected to five-year cycle analysis. As a result, research on electronic records, record information services, and archiving methods of various records has been gradually improved.

Keywords: archives management, topic modeling, trends analysis, time series analysis, HDP

* 이 논문은 2017년 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2017M3C4A7068186).

** 전북대학교 일반대학원 기록관리학과 박사과정(gulseori@naver.com) (제1저자)

*** 부산외국어대학교 동남아창의융합학부 부교수(pmryu@bufs.ac.kr) (공동저자)

**** 전북대학교 기록관리학과 조교수, 문화융복합아카이빙 연구소 연구원(ohj@jbnu.ac.kr) (교신저자)

■ 접수일: 2018년 1월 15일 ■ 최초심사일: 2018년 1월 31일 ■ 게재확정일: 2018년 2월 21일

■ 한국기록관리학회지 18(1), 29-47, 2018. <<http://dx.doi.org/10.14404/JKSARM.2018.18.1.029>>

1. 서론

1.1 연구배경 및 필요성

최근 사회는 4차 산업혁명 시대를 맞이하면서 이전보다 더욱 빠르게 변화하고 있다. 클라우드 컴퓨팅, 빅데이터 분석, 인공지능, 사물인터넷 등과 같은 컴퓨터 기술이 급성장하면서 이에 따라 정치, 경제, 사회를 포함한 모든 영역에 영향을 주고 있다. 특히 하드웨어와 소프트웨어의 발전으로 인해 기하급수적으로 증가하는 빅데이터에 대한 분석과 처리가 가능해지면서 재난 발생 예측, 주식 동향 파악, 이용자 맞춤형 서비스 등 다양한 분야에서 이를 활용하고 있다. 한편, 국민의 알권리가 존중되기 시작하면서 폐쇄적인 사회분위기에서 개방적이고 공개적인 사회분위기로 변화하기 시작하였다. 이러한 정보통신기술의 발달과 사회의 변화에 따라 기록관리학 연구 또한 시대의 흐름에 맞춰 변화하고 있다. 컴퓨터의 발전으로 인해 종이 기록에서 전자기록으로 기록관리의 패러다임이 변화하고 있으며, 정부의 투명성이 강조되는 사회가 되면서 기록의 폐쇄적인 관리 및 보존에서 공개, 열람, 활용 등 서비스 중심의 기록관리체제로 변화하기 시작했다(한국기록관리학회, 2013).

트렌드 분석(Trend Analysis)이란 시간에 따라 점차 변화하는 사회의 분위기와 환경에서, 특정 영역에 대해 당시 사람들에게 주목받거나 이슈가 되었던 트렌드를 분석하고 파악하는 것이다(김난도, 2008). 이를 통해 과거와 현재의 현황을 파악하고 미래의 트렌드에 대해 예측 및 전망하는 것이 목적이다. 기존의 트렌드 분석은

문헌조사, 설문조사, 전문가평가 등과 같은 방법을 사용하기 때문에 시간과 비용이 많이 소모된다는 한계점이 있다. 그러나 최근 데이터 마이닝과 같이 방대한 양의 데이터를 처리하는 빅데이터 기술의 발전으로 인해 빠르게 변화하는 시대에서 다양한 분야, 영역에 대한 트렌드를 분석하는데 자주 활용되고 있다.

이처럼 사회분위기의 변화와 정보통신기술의 발달로 인해 빠르게 변하고 있는 시대의 흐름 속에서 국내 기록관리학에 대한 트렌드 분석을 통해 시간에 따라 변화하는 국내 기록관리학 연구의 동향을 파악하고, 이를 기반으로 계속해서 변화하는 사회에서 국내 기록관리학 연구의 발전 방향에 대해 모색하는 것이 필요하다. 따라서 본 연구는 시계열에 따라 분석된 토픽모델링의 변화를 파악함으로써 국내 기록관리학 연구의 트렌드를 분석하고 이를 통해 기록관리학 연구 동향과 흐름에 대해 살펴보고자 한다.

1.2 선행연구

기존의 트렌드 분석 관련 선행연구는 문헌조사나 전문가평가와 같은 정성적인 분석방법을 활용한 연구와 텍스트마이닝 기법과 같은 정량적인 분석방법을 활용한 연구로 구분할 수 있다. 먼저 정성적인 분석방법을 활용한 선행연구를 살펴보면, 김종근(2014)은 2010년 초반의 사회문화를 대표하는 트렌드를 도출하기 위해서 57건의 미래전망자료를 수집하고, 키워드를 도출하고 공동집단으로 분류하였다. 그 결과 SNS 문화, 디지털 디스토피아, 웨저, 우먼파워, 웰빙, 매크로바이오틱, 미디어의 노예, 고령화, 싱글 문화, 스테이크이션, 휴식, 창조지식사회, 다문

화, 위버럭서리 등 14가지 대표적인 트렌드를 도출하고 이에 대해 세부적인 분석을 진행하였다. 이진백, 이충권, 차경진(2012)은 IT 트렌드를 분석하기 위해서 정보통신산업진흥원과 한국정보화진흥원에서 발표한 IT 트렌드 토픽 키워드를 기반으로 2013년 5월 1일부터 2013년 12월 31일까지 관련된 트윗 21,589건을 수집하여 분석을 실시하였다. 연구결과, 빅데이터, 클라우드, HTML5, 스마트홈, 태블릿PC, UI/UX와 같은 IT토픽이 시간이 지날수록 매우 빈번하게 언급되어지는 것을 확인하였다. 안소미 외(2009)는 점차 고령화되어가는 사회에서 고령친화 환경에 대한 트렌드를 분석하기 위해 『Environment For Aging 2009』에서 발표된 내용을 분석하였다. 그 결과, 주요 트렌드는 그린디자인, 새로운 기술, 소셜믹스, 증거기반 디자인, 자족적 커뮤니티 5가지가 도출되었다. 박지애, 박병철, 최안섭(2007)은 공동주택 경관조명의 트렌드를 분석하기 위해서 42개소의 공동주택을 직접방문하고, 조명설계회사의 자료를 분석하였다. 그 결과 건축물 형태의 변화에 의한 디자인의 다양화, LED의 사용증가에 의한 색과 패턴의 다양화 등이 도출되었다.

현재까지 정량적인 분석방법을 활용한 트렌드 분석 연구는 다양한 텍스트마이닝 기법을 활용하고 있으며 개수 또한 상당히 많기 때문에 모든 선행연구를 살펴보기에는 한계가 있다. 따라서 본 연구는 2016년부터 2017년 동안 진행된 토픽모델링 방법을 활용한 트렌드 분석 선행연구를 중점적으로 살펴보았다. 유리나, 황수철(2017)은 1984년부터 2015년까지 국내 농식품 유통분야에 대해서 빈도분석, 토픽분석, 연관성분석을 활용하여 트렌드 분석을 진행하

였다. 그 결과, 2000년대 이전에는 도매와 산지 연구가 활발하였으며, 2000년대 이후에는 소비, 식품, 수출입 연구가 활발한 것으로 나타났다. 윤소연, 윤동근(2017)은 재난과 안전사고에 대해 1990년부터 2016년간 배포된 뉴스 기사를 수집하고 토픽모델링을 수행하였다. 그 결과, 재난안전 분야는 특정 사건에 따라 크게 증감하는 패턴을 보이며, 2000년대 후반부터 재난안전교육, 체험, 훈련 등과 같은 토픽들이 관심이 높아지는 것을 확인하였다. 박주섭, 홍순구, 김종원(2017)은 AI(Artificial Intelligence) 분야에 대해 토픽모델링을 활용하여 20개의 세부 토픽을 추출하고 트렌드 분석을 진행하였다. 분석 결과, 최근 AI 분야에서 활발하게 연구되는 기술들은 텍스트 탐색, 컴퓨터 관리, 프로그래밍 구문, 네트워크 관리, 멀티미디어, 무선 네트워크 기술 등이 있었다. 나상태 외(2017)는 동적 토픽분석 기법을 활용하여 2009년부터 2016년까지 스마트그리드(Smart Grid) 기술에 대한 트렌드 분석을 진행하였다. 그 결과, 전반기에는 ESS, Service, ICT 순서로 연구가 활발하였으며, 하반기에는 Grid Operation, ESS, Grid Plan 순서로 연구가 활발하게 진행되었다. 김태경, 최희련, 이홍철(2016)은 핀테크(Fintech) 기술에 대해 토픽모델링을 활용하여 트렌드 분석을 진행하였다. 그 결과, 핀테크 시장이 성장할수록 모바일결제를 중심으로 금융소프트웨어와 데이터분석, 자산관리 부분이 증가하는 것을 확인할 수 있었다. 나상태 외(2016)는 토픽모델링을 활용하여 시뮬레이션 분야의 트렌드 분석을 진행하였다. 그 결과, 전기전자 분야에서 시뮬레이션이 활발하게 활용되고 있으며 사회과학 분야에서는 교육 및 오락 목적으로 활용

되고 있음을 파악하였다.

기록관리학 분야의 트렌드 분석 선행연구를 살펴보면, 이재운, 문주영, 김희정(2007)은 문헌정보학 관련 학술지 5종을 대상으로 기록관리학 관련 논문 145건을 수집하고 문헌 클러스터링과 문헌 유사도 네트워크 분석을 활용하여 트렌드 분석을 진행하였다. 이후 2009년부터 김규환, 장보성, 이현정(2009)은 3종의 기록관리학 관련 학술지를 대상으로 1999년부터 2008년까지 344편의 논문을 수집, 논문 제목의 구문과 의미구조를 파악하였으며, 남태우, 이진영(2009)은 4종의 관련 학술지를 대상으로 1997년부터 2007년까지 399편의 논문을 수집하여 주제별 분포, 간행시기별 분포, 학회지별 분포, 연구자별 분포 등을 분석하였다. 또한 김규환, 남영준(2009)은 3종의 기록관리학 관련 학술지에서 1999년부터 2009년 9월까지 374편의 논문을 대상으로 연구자 특성에 따른 트렌드 분석을 진행하였다. 장로사, 김유승(2009)은 여러 기록관리학 세부영역 중에서 전자기록물 및 정보학 영역에 관한 트렌드 분석을 진행하였다. 4종의 관련 학술지를 대상으로 2000년부터 2008년까지 99편의 논문을 대상으로 연구논문의 추이, 연도별, 연구자별, 학회지별 트렌드 분석을 진행하였다. 이소연(2011)은 2종의 기록관리학 관련 학회지와 4종의 문헌정보학 관련 학회지에서 2000년부터 2010년까지 전자기록을 다룬 논문 57편을 대상으로 전자기록에 관한 연구동향 분석을 진행하였다.

최근 연구로는 최이랑(2015)이 2004년부터 2013년까지 한국기록관리학회지, 기록학연구, 한국문헌정보학회지, 한국비블리아학회지, 한국도서관·정보학회지를 대상으로 기록관리학분

야 논문 479편을 대상으로 트렌드 분석을 진행하였다. 손혜인, 남영준(2016)은 한국기록관리학회지와 기록학연구 2종의 기록관리학 관련 학회지를 대상으로 2000년부터 2015년까지 게재된 681편의 논문으로 기록관리학 관련 연구자 배경과 주제 변화에 관한 연구를 수행하였다.

이상에서 정리한 선행연구를 살펴보면, 텍스트마이닝 기법을 활용한 트렌드 분석 연구가 점차 활발하게 진행되고 있는 추세이며, 최근 특정 연구분야의 트렌드 분석에 있어 토픽모델링 기법이 주요 연구방법으로 활용되고 있는 것을 확인할 수 있다. 그러나 지금까지 진행된 기록관리학 관련 트렌드 분석 선행연구는 대부분 기록관리학 관련 학술지 2종을 중심으로 진행되었으며, 수집된 논문 편수는 손혜인(2016)의 연구를 제외하면 500건 미만인 것을 확인할 수 있다. 또한 현재까지 국내 기록관리학 분야의 트렌드 분석 연구에서 토픽모델링을 활용한 연구는 진행되어지지 않았다. 이에 본 연구에서는 분석 대상과 기간을 확대하여 기록관리 연구간에 대한 거시적인 연구 동향과 각 기간별 미시적인 연구 동향 흐름을 파악하고자 한다.

2. 이론적 배경

2.1 트렌드 분석

트렌드는 사회, 경제, 문화 등 특정 영역에서 어떤 흐름이 진행되고 있거나, 어떤 방향으로 가고 있는 동향, 추세, 유행 등을 의미하며, 트렌드 분석이란 시간의 변화에 따라 생산된 과거와 현재의 데이터들을 다양한 방법을 활용하

여 분석하고 이를 통해 미래에 대한 가설을 만드는 것이라고 할 수 있다(ulf Pillkahn, 2007: 안병기, 2003). 트렌드 분석은 당시 사회의 공통적인 관심이나 견해 등을 의미하는 '시대정신'을 나타내며, 어떤 특정 영역이나 분야에서 시간에 따라 변화하는 공통적인 주요 관심과 이슈를 파악할 수 있다(김종근, 2014).

트렌드는 지속되는 시간적인 길이에 따라 마이크로트렌드(Microtrend), 패드(Fad), 트렌드(Trend), 메가트렌드(Megatrend)로 구분할 수 있다. 마이크로트렌드는 가장 기간이 짧고, 소수의 사람들만이 동조하는 작은 변화이다. 패드(Fad)는 '유행'을 뜻하며, 단기간에 확산되는 유행으로 수명이 짧다(이순중, 2010). 마이크로트렌드보다 지속기간이 길고, 사람들의 동조범위도 넓지만 대체로 지속기간이 1년 이내일 경우 패드라고 불린다(김종근, 2014). 패드와 같이 단순한 유행을 지나 짧게는 1년, 길게는 5년 정도 지속되며 많은 사람들이 동조하는 것을 일반적인 트렌드라고 할 수 있다. 반면, 메가트렌드(Megatrend)는 미국의 미래학자 Jhon Naisbitt(1983)가 언급한 용어로, 현대 사회에서 일어나고 있는 거대한 시대적 조류를 뜻한다. 보통 10년 이상 지속되는 유행이나 흐름을 메가트렌드로 보며, 이러한 메가트렌드는 세계적인 사회변동을 일으키거나 삶의 방향성을 나타내는 중요한 역할을 한다(Aburdene, 2005/2006). 이처럼 여러 가지 트렌드의 종류가 존재하지만 실제 분석하는데 있어 마이크로트렌드, 패드, 트렌드, 메가트렌드와 같이 세분하여 분석하기에는 한계가 있어 본 연구에서는 종합적으로 살펴보았다.

기존의 트렌드 분석은 주로 문헌조사, 전문

가평가, 델파이기법 등과 같은 정성적 연구방법을 사용하였다. 이러한 방법은 분석하여 결과를 도출하는데 시간과 비용이 많이 소요되고 연구자의 주관적인 견해가 반영될 수 있다는 한계점이 있다(김현정, 조남옥, 신경식, 2015). 최근 텍스트 마이닝(Text Mining)과 같은 정량적 연구방법이 많이 발전하게 되면서, 트렌드 분석 방법에서도 정성적 분석방법 뿐만 아니라 정량적 분석방법을 활용한 연구들이 점차 증가하게 되었다. 정량적 분석방법을 활용한 트렌드 분석은 대량의 데이터를 기반으로 보다 객관적이고 효율성이 높은 분석 결과를 확보할 수 있다(김현정, 조남옥, 신경식, 2015). 본 연구에서는 정량적 분석방법인 텍스트 마이닝 기법 중 하나인 토픽모델링을 활용하여 국내 기록관리학 토픽 트렌드를 분석하였다.

이처럼 트렌드 분석은 특정 영역에 관한 많은 양의 데이터를 시계열적으로 분석하여 시기별로 이슈가 되었던 트렌드를 파악하고 과거의 트렌드를 통해 현재의 트렌드가 무엇인지 분석하는 현상파악과 이로 인해 미래 트렌드를 전망하고 예측하는 것이 목적이다. 이에 따라 본 연구는 국내 기록관리학의 시간에 따른 토픽 트렌드 분석을 위해 토픽모델링 기법을 활용하여 과거와 현재의 토픽 트렌드를 파악하고 향후 국내 기록관리학 토픽 트렌드에 대해 제언을 하고자 한다.

2.2 토픽모델링

토픽모델링이란 텍스트마이닝 기법 중 하나로 특정 텍스트 문서집합을 대상으로 잠재된 토픽들을 추출하여 문서집합의 전반적인 주제들을

파악하는 기법이다. 최근 트렌드 분석, 연구동향 분석 등 다양한 연구에서 활용되면서 관심이 높아지고 있다. 기존의 텍스트 군집 분석은 각각의 문서가 하나의 주제에만 해당될 수 있기 때문에 개별의 문서와 다양한 주제와의 연관성을 파악하기가 어려웠다. 반면에 토픽모델링은 개별 문서가 여러 토픽과 연관성을 가질 수 있기 때문에 특정 문서의 복합적인 주제를 파악할 수 있다는 장점이 있다. 이러한 이유 때문에 토픽모델링은 기존의 다른 모델링보다 현실 세계에 더 적합한 기법으로 평가받고 있다(김남규, 이동훈, 최호창, 2017). 주요 토픽모델링 기법으로는 LSA(Latent Semantic), pLSA(Probabilistic Latent Semantic Analysis), LDA(Latent Dirichlet Allocation), HDP(Hierarchical Dirichlet Process) 등이 있다. 본 연구에서는 HDP(Hierarchical Dirichlet Process) 토픽모델링 기법을 사용하였다. HDP는 Dirichlet Process를 계층적으로 적용하여 주요 토픽을 분석하는 방법으로 연구자가 주제 개수인 K값을 임의로 설정하지 않아도 분석과정에서 적절한 K값을 설정하여 결과를 도출하는 방법이다. 기존의 토픽모델링 기법을 사용한 연구에서 주로 활용되었던 LDA 토픽모델링은 분석하는 데이터 크기에 따라 분석결과가 많은 영향을 받는다. 또한 특정 주제에서 공통적으로 다루는 키워드 분석에 유리해 주로 거시적 측면의 분석에 적합한 방법이다. 반면, HDP 토픽모델링은 데이터 크기에 따른 분석결과 변화가 LDA보다 적으며, 분석결과에 있어 LDA보다 일반적인 키워드, 중복된 키워드에 영향을 덜 받으며 각 토픽의 특징을 분명히 파악될 수 있는 변별력 있는 키워드가 뚜렷하게 나타나기 때문에 세부적

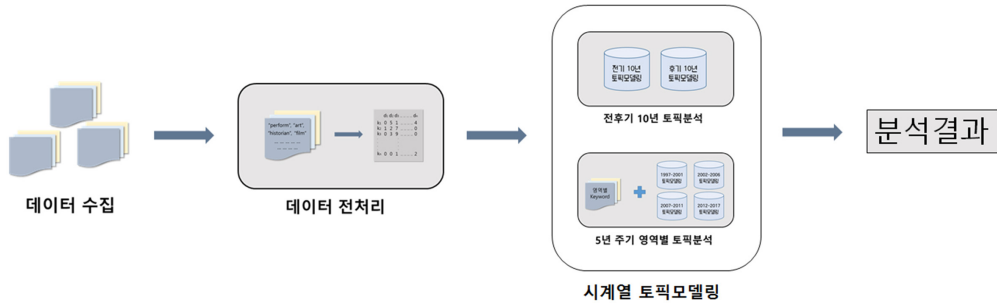
인 트렌드 분석에 LDA 보다 효율적이다(박준형, 오효정, 2017).

3. 연구방법

3.1 연구모델

본 연구에서 제안하는 시계열 기반 국내 기록관리학 토픽 트렌드 분석의 연구방법은 <그림 1>과 같다. 먼저 분석대상인 국내 기록관리학 관련 학술 논문지를 선정, 관련 논문을 자동 수집하였다. 수집한 데이터는 전처리과정을 통해 분석에 적합하도록 정제하였다. 정제된 데이터를 통해 시계열 토픽모델링을 수행한 후, 결과를 분석하여 시계열에 따른 기록관리학 주제 변화를 분석하였다. 구체적인 분석 내용은 다음과 같다.

첫째, 국내 기록관리학 관련 학술지를 대상으로 국내학술지 원문 DB사이트인 DBPIA에서 발간일부터 2016년까지 등재된 학술논문을 자동으로 수집하였다. 둘째, 수집된 데이터는 형태소분석, 명사추출 등과 같은 전처리 과정을 거쳐 토픽모델링을 수행하는데 적합한 포맷으로 정제하였다. 셋째, 정제된 데이터를 기반으로 시계열 토픽모델링을 수행하였는데 먼저, 1997년부터 2006년까지 전기 10년, 2007년부터 2016년까지 후기 10년으로 데이터를 나누고 각각 HDP 토픽모델링(박준형, 오효정, 2017)을 수행하여 전·후기 10년 기록관리학 주제 변화를 분석하였다. 다음으로 1997년부터 2016년까지 데이터를 5년 주기로 나누어 각각 토픽모델링을 수행한 후, 기록관리학 주제영역별 키워드를 매핑



〈그림 1〉 연구모델

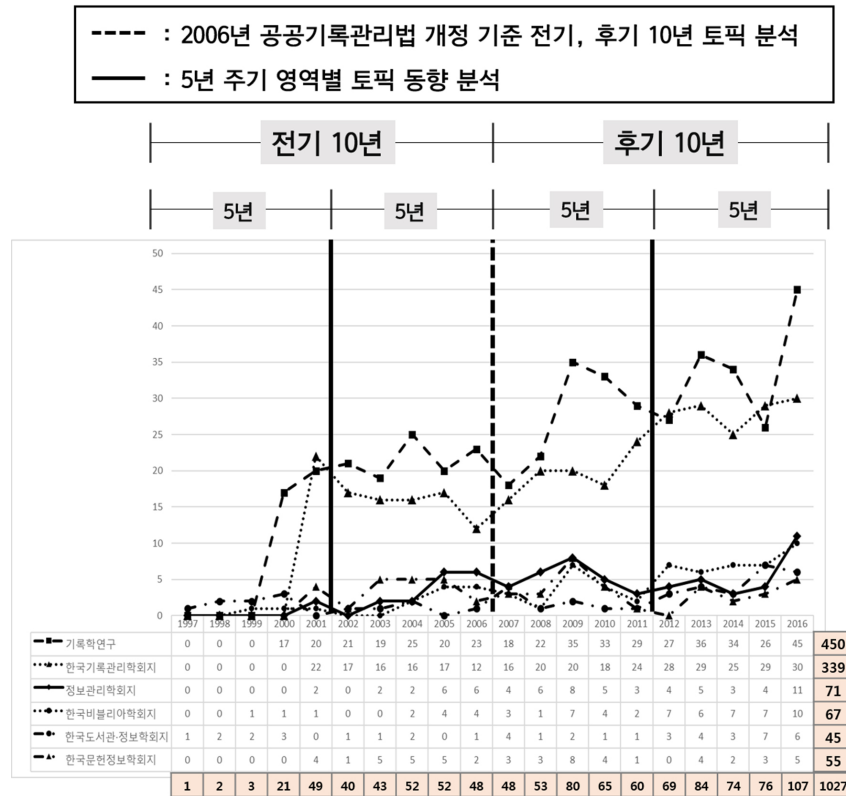
하여 시계열에 따른 기록관리학의 영역별 변화를 분석하였다.

3.2 연구대상

본 연구는 시계열 기반 국내 기록관리학 토픽의 트렌드를 파악하기 위해서 기록관리학 관련 학술지에서 등재한 학술논문을 수집대상으로 선정하였다. 기록관리학 관련 학술지인 ‘한국기록관리학회지’와 ‘기록학연구’ 2종의 학술지와 ‘한국문헌정보학회지’, ‘한국도서관·정보학회지’, ‘한국비블리아학회지’, ‘정보관리학회지’ 등 기록관리학 관련 학술논문이 빈번히 게재되는 문헌정보학 관련 학술지 4종을 포함하여 총 6종의 국내학술지에서 등재한 학술논문을 수집하였다. 각 학술지 발간일부터 2016년까지를 수집기간으로 선정하고, C# 기반으로 직접 구현한 웹크롤러를 활용하여 국내 학술DB 사이트인 DBPIA에서 제공하는 학술논문 정보들을 자동 수집하였다. 수집된 문헌정보학 관련 학술지 4종의 학술논문 중에서 기록관리학 관련 논문만을 추출하기 위해서 논문명과 키워드에 “기록”, “아카이브”, “아카이빙”, “아키비스트” 키워드가 포함된 논문만을 추출하였다.

본 연구에서 수집한 데이터의 종합 통계를 살펴보면 〈그림 2〉와 같다. 기록관리학 관련 학술지를 살펴보면 ‘기록학연구’는 발간연도인 2000년부터 2016년까지 450건의 논문을 수집하였고, ‘한국기록관리학회지’는 발간연도인 2001년부터 2016년까지 339건의 논문을 수집하였다. 문헌정보학 관련 학술지는 처음 기록관리학 관련 논문이 등재된 시기부터 2016년까지 관련 논문을 수집하였는데 구체적으로 살펴보면, 먼저 ‘정보관리학회지’는 2001년부터 2016년까지 71건의 논문을 수집하였으며, ‘한국비블리아학회지’는 1999년부터 2016년까지 67건의 논문을 수집하였다. ‘한국도서관·정보학회지’는 1997년부터 2016년까지 45건의 논문을 수집하였고 ‘한국문헌정보학회지’는 2001년부터 2016년까지 55건의 논문을 수집하였다. 종합적으로 수집된 데이터의 기간은 1997년부터 2016년까지 20년이며 총 1,027건의 기록관리학 관련 학술논문이 수집되었다.

〈그림 2〉의 상단 그래프는 연도별 기록관리학 관련 논문 수를 나타낸 것이다. 기록관리학 관련 학술논문은 1997년 ‘한국도서관·정보학회지’에서 처음 등장하였으며 2001년 기록관리학 관련 학술지가 등장하면서 발간되는 논문 수가



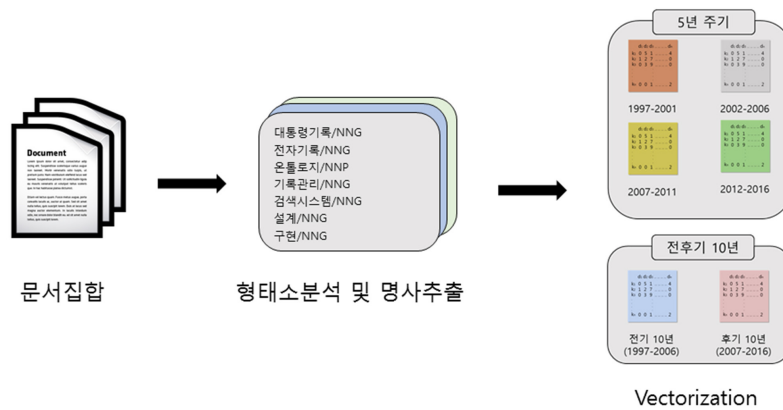
〈그림 2〉 수집 데이터 종합 통계 그래프

급격하게 증가하였으며 매년 논문 수가 지속적으로 증가하고 있는 것을 확인할 수 있다. 또한 토픽 분석은 그래프 위에 표시된 점선, 실선과 같이 2006년 공공기록관리법 개정 기준으로 전기(1997년~2006년), 후기(2007년~2016년) 10년 토픽 분석과 5년 주기 영역별 토픽 동향 분석으로 진행되어진다.

3.3 데이터 전처리

토픽모델링을 수행하기 전에 적절한 전처리 과정을 통해 수집된 데이터를 정제해야 한다. 토픽모델링 분석을 수행할 때, 데이터 전처리

는 분석의 정확도와 성능에 많은 영향을 주기 때문이다. 상세한 데이터 전처리 과정은 〈그림 3〉과 같다. 첫째, 1,027건의 기록관리학 관련 논문의 문서집합을 생성하고 각 문서에 대해 형태소 분석을 실시한다. 형태소분석은 파이썬 자연어 처리 라이브러리인 Konlpy의 Hannanum 형태소 분석기를 활용하였다. 둘째, 형태소분석이 완료된 문서집합에서 명사추출을 한다. 셋째, 1997년부터 2006년까지 전기 10년, 2007년부터 2016년까지 후기 10년, 1997~2001, 2002~2006, 2007~2011, 2012~2016 5년 주기로 데이터를 나누고 각각 토픽모델링 수행에 적합한 포맷으로 벡터화(Vectorization)한다.



〈그림 3〉 데이터 전처리 과정

3.4 시계열 기반 토픽모델링

본 연구는 토픽모델링을 활용한 국내 기록관리학의 시계열에 따른 주제 변화를 분석하기 위해서 먼저 1997년부터 2016년까지 20년의 수집데이터를 2007년 ‘공공기관의 기록물 관리에 관한 법률’에서 ‘공공기록물 관리에 관한 법률’로 개정된 시점을 기준으로, 1997년부터 2006년까지 전기 10년과 2007년부터 2016년까지 후기 10년의 시계열 토픽모델링을 수행하였다. 전·후기 10년에 대한 시계열 주제 변화 분석을 통해 거시적인 측면에서 기록관리학 연구의 흐름을 파악할 수 있다. 이후 1997~2001, 2002~2006, 2007~2011, 2012~2016 5년 주기로 각각 토픽모델링을 수행한 후 기록관리학 주제영역에 대한 각 키워드와 매핑을 하여 보다 상세한 시계열 주제 변화 분석을 진행하였다.

본 연구에서는 손혜인(2016)이 제안한 기록관리학 14개 주제분류 및 관련키워드와 「공공기록물 관리에 관한 법률」, 「대통령기록물 관리에 관한 법률」을 바탕으로 <표 1>과 같이 8가지 기록관리학 주제영역과 관련키워드를 선정하였

다. 기록관리학 주제영역은 크게 기록관리주기 측면과 기록물 측면으로 나눌 수 있다. 기록관리주기 측면은 생산, 관리 및 보존, 평가, 공개·열람·활용 등 4가지 영역으로 구성되어 있다. 기록물 측면은 공공기록물, 민간기록물, 대통령기록물, 전자기록물 4가지 영역으로 구성된다.

기록관리주기 측면은 「공공기록물 관리에 관한 법률」 제3조 3항(“기록물관리”란 기록물의 생산·분류·정리·이관(移管)·수집·평가·폐기·보존·공개·활용 및 이에 부수되는 모든 업무를 말한다.)에 근거하여 만들었으며, 기록물 측면은 가장 일반적인 구분이라고 할 수 있는 공공기록물과 민간기록물, 공공기록물과 동일하게 법령으로 규정하고 있는 대통령기록물, 「공공기록물 관리에 관한 법률 시행령」 제2조, 제5조, 제34조, 제36조 등에서 그 중요성이 강조되고 있는 전자기록물을 기록물 측면의 4가지 영역으로 선정하였다.

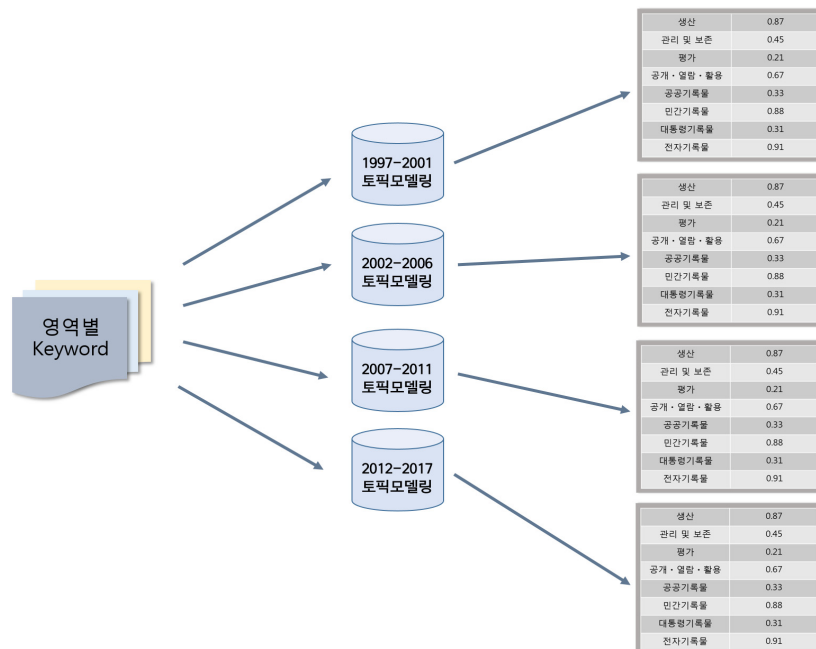
설정된 기록관리학 주제영역과 관련키워드를 각 5년 주기로 수행된 토픽모델링에 매핑하면 해당 키워드가 각 토픽모델링의 여러 토픽에 어느 정도 연관성이 있는지 수치로 확인할 수 있다. <그림 4>와 같이 각 영역이 해당 토픽모델링과

얼마나 연관성이 있는지 파악할 수 있다. 이후 다른 기록관리학 연구의 흐름과 이슈가 되는
 각 5년 주기 영역별 수치 비교를 통해 시계열에 주제 변화를 파악할 수 있다.

〈표 1〉 기록관리학 주제영역

	영역	키워드
기록관리주기 측면	생산	기록물의 등록 분류 편철, 출처주의, 전자레코드, 장서관리, 기록물처리일정표, 기록물 분류기준표, 업무모형, 수집정책, 아카이브 구축 방안, 기록화 현황, BRM
	관리 및 보존	기록물 관리의 원칙, 이관, 이중보존, 기록물의 회수, 폐기, 기록물평가심의회, 기록매체 및 용품 등의 표준 및 규격, 기록관리시스템, 보존포맷, 장기보존
	평가	평가제도, 평가모델, 젠킨슨, 웰렌버그, 평가, 평가선별론, 도큐멘테이션, 다큐멘테이션, 기능적평가, 베어만, 거시적평가
	공개·열람·활용	기록정보서비스, 탐색도구, 시소러스, 기록물전시, 견학, 홍보, 책자, 정보공개, 기록정보 콘텐츠
기록물 측면	공공기록물	중앙기록물관리기관, 기록물관리기관, 공공기관, 공공기록, 기록관, 특수기록관
	민간기록물	국가지정기록물, 메뉴스크립트, 로컬리티 기록화, 집단지역, 사회적지역, 개인기록물, 공동체 아카이브
	대통령기록물	대통령기록물, 대통령기록관, 대통령기록
	전자기록물	전자기록생산시스템, 디지털 아카이빙, 전자우편문서, OAS, ERMS, DIRKS, 디지털, 메타데이터, 오픈소스, 공개소프트웨어, 디지털아카이브

* 손혜인(2016) 참조.



〈그림 4〉 5년 주기 영역별 토픽 분석

4. 연구결과

4.1 전·후기 10년 토픽 분석 결과

공공기록관리법 개정 기준으로 전기 10년, 후기 10년에 대한 HDP 토픽모델링 결과는 <표 2>, <표 3>와 같다. 전·후기 10년에 대한 10개의 토픽과 각 토픽에서 가중치가 높은 상위 8개의 키워드를 나타낸 것이다. 해당 토픽은 임의의 순서로 정해진 것이 아니라 비율이 가장 높은 토픽부터 낮은 토픽까지 정렬된 결과다. 따라서 topic1이 가장 높은 비율을 가지며 topic2, topic3이 다음 순서로 비율이 높고 topic10이 가장 낮은 비율을 가지고 있다.

전기 10년(1997년~2006년) 토픽 분석 결과를 살펴보면, 먼저 “기록학”, “기록관리학”, “기록관리”, “기록물”, “아키비스트” 등 거시적 의미의 키워드가 전반적으로 나타나고 있는 것을 확인할 수 있다. “개념”, “고찰”은 topic1에서 가장 상위 키워드로 나타나고 있으며 topic9에서도 “고찰” 키워드가 높은 순위에 있다. 또한

topic9에서 “조선총독부”, “현대사”, “기록보존소” 등 역사에 관련된 키워드도 살펴볼 수 있다. 이러한 특징을 볼 때, 전기 10년의 연구는 주로 기록관리학의 전반적인 개념이나 고찰에 대한 연구, 기록관리학과 관련된 역사에 관한 연구가 진행되었음을 추론할 수 있다.

topic1의 “양성교육”, topic2의 “교육과정”, topic4와 topic10의 “양성제도”, topic5의 “교육프로그램”, topic8의 “교과과정” 등 전기 10년의 대부분 토픽에서 교육과 관련된 키워드가 높은 순위에 위치한 것을 확인할 수 있다. 이는 기록관리학 관련 교육, 아키비스트 양성교육 등 전기 10년의 기록관리학 연구가 교육 부문에 관심을 가지고 있었다는 것을 파악할 수 있다.

topic1의 “사례연구”, topic3의 “호주”, topic6의 “러시아”, “일본”, “사례”, “독일”과 같은 키워드가 높은 순위에 위치하고 있다. 이는 전기 10년의 연구가 국외 기록관리학에 대한 사례 관련 연구에도 관심을 보이고 있었다는 것을 파악할 수 있다.

후기 10년(2007년~2016년) 토픽 분석 결과

<표 2> 전기 10년(1997년~2006년) 토픽 분석 결과

topic1	topic2	topic3	topic4	topic5	topic6	topic7	topic8	topic9	topic10
개념	검색분류	전자정부	기록관리 제도	관리방안	기록관리 체제	기록관리 체제	교과과정	공공기록물	기록관리학
고찰	공공기관	공공기관	보존 기록물	일체화	기록물 분류기준표	메타데이터	대학기록물	전자정부	목록기술
기록학	보존	변천	구축	기록관리 제도	기록관리	디지털	지식경영	거버넌스	검색분류
설립	윤리	호주	접근권	검색분류	러시아	전자기록 관리	아키비스트	기록보존소	베트남
양성교육	교육과정	검색분류	기록관리	디지털화	일본	공공기록물	활성화	조선총독부	기록보존소
기준	표준화	표준화	양성제도	참조모형	사례	국가기록물	분류체계	현대사	진리탐구
표준	기록	정보화	기록물	목록기술	독일	일체화	메타데이터	보존	양성제도
사례연구	기술	기록관리 체제	전자기록	교육 프로그램	변천	전자기록	목록기술	고찰	일체화

〈표 3〉 후기 10년(2007년~2016년) 토픽 분석 결과

topic1	topic2	topic3	topic4	topic5	topic6	topic7	topic8	topic9	topic10
개인기록	기록관리 제도	대학기록관	큐레이션	운영	단위과제	교육 프로그램	정보공개 심의회	의무기록 관리	학교 기록물
교육 서비스	주립기록관	공공기록	기록전시	기업	정책	아키비스트	기록정보 서비스	서비스	평가론
기록물 분류기준표	기록분류	대통령기록	메타 데이터	전자기록 관리시스템	지역	기록사랑 마을	민간기록물	전자 레코드	개념
기록전시	시소러스	업무관리 시스템	재평가	오픈소스	에몰레이션	전자기록물	비밀기록	기록관	전통
기록화	클라우드	연구기록물 관리기관	연구 기록물	클라우드	기록정보 서비스	학습자료용	메타데이터	기록정보 콘텐츠	서양
과학	수집전략	디지털화	건설 기록물	평가제도	서비스	위험관리 체계	영상기록물	기록경영 시스템	미국
민주주의	표준기록 관리시스템	중앙 행정기관	오픈 액세스	지방기록 관리	활용	도구	국가기록 관리체제	연구기록	검색 도구
대학	레포지터리	법제화	소송기록	참여형	네트워크	컬렉션	정보공개 청구	평가기준	시대

를 살펴보면, 우선 전기 10년 토픽과 비교할 때 전반적인 키워드가 거시적인 의미의 키워드보다 지역적인 의미의 키워드가 많은 것을 확인할 수 있다. topic1의 “개인기록”, “기록화”, “과학”, “대학”, topic4의 “연구기록물”, “건설 기록물”, “소송기록”, topic8의 “민간기록물”, “영상기록물”, topic9의 “의무기록관리”, “연구기록”, topic10의 “학교기록물” 등 후기 10년의 기록관리학 연구는 개인기록, 소송기록, 의무기록 등 다양한 기록물에 대한 기록화, 기록관리 방안에 대한 연구가 활발하게 진행되고 있는 것을 파악할 수 있다.

topic1의 “교육서비스”, “기록전시”, topic4의 “큐레이션”, “기록전시”, “오픈엑세스”, topic5의 “참여형”, topic6의 “기록정보서비스”, “서비스”, “활용”, topic7의 “교육프로그램”, topic8의 “기록정보서비스”, “정보공개청구”, topic9의 “서비스”, “기록정보콘텐츠” 등 후기 10년에 기록정보서비스에 관한 연구가 많은 관심을 받

고 있는 것을 확인할 수 있다.

topic2의 “시소러스”, “표준기록관리시스템”, “레포지터리”, topic3의 “업무관리시스템”, “디지털화”, topic5의 “전자기록관리시스템”, “오픈소스”, “클라우드” 등 후기 10년 기록관리학은 기록과 기록관리시스템에 관한 디지털화에 대한 연구가 활성화 되고 있는 것을 파악할 수 있다.

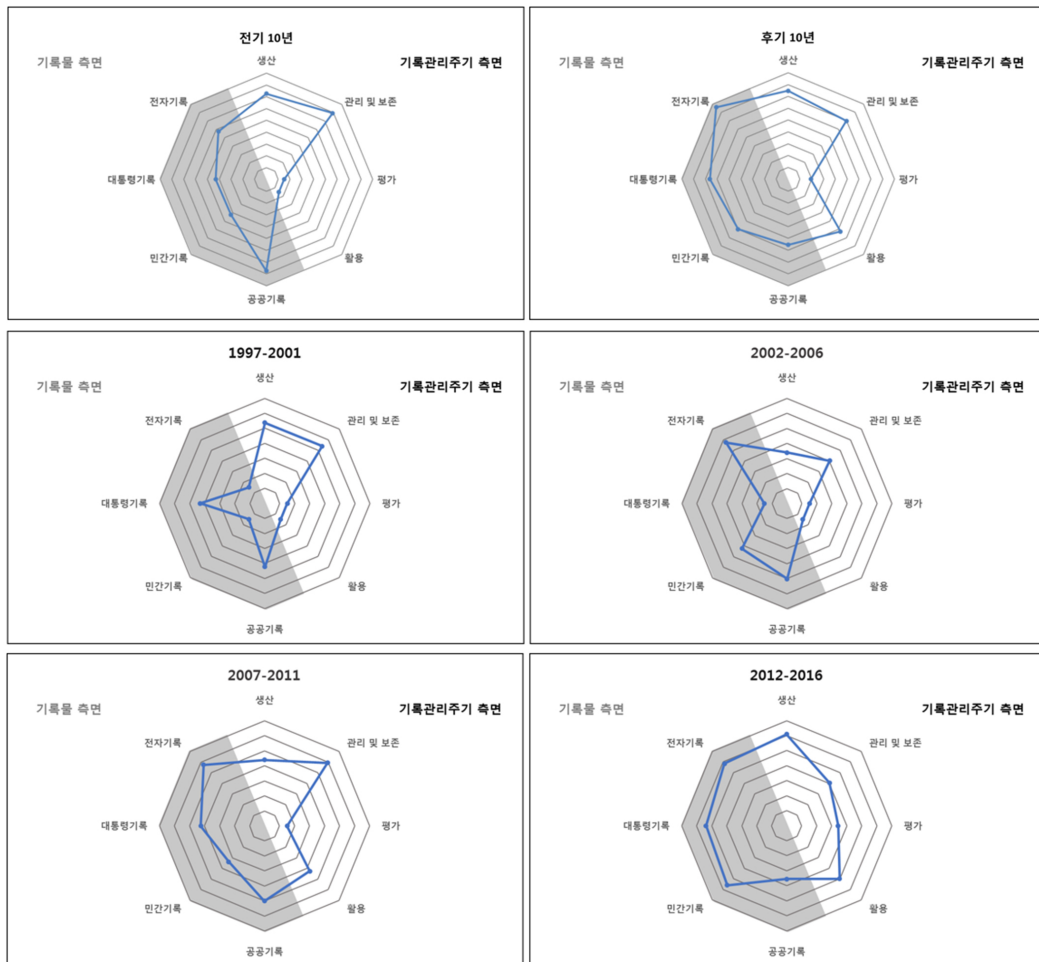
4.2 5년 주기 영역별 토픽 분석 결과

전·후기 10년 토픽 분석을 통해 거시적인 측면에서 국내 기록관리학의 흐름을 살펴봤다면, 5년 주기 영역별 토픽 분석은 본 연구에서 선정한 8가지 기록관리학 주제영역별 동향을 5년 주기로 파악하여 보다 세부적인 시계열 주제 변화 분석을 할 수 있다. 각 영역별로 수행된 토픽모델링을 〈표 1〉에서 나타난 기록관리학 주제영역별 키워드와 매핑하여 각 영역이 해당

토픽모델링과 얼마나 연관성이 있는지 수치화 하였다. 이를 <그림 5>와 같이 방사형 그래프로 표현하였으며 수치의 최대값은 1.0, 최소값은 0.0이다. 해당 토픽모델링이 특정 영역에 연관성이 높을수록 수치는 높아지며 반면, 연관성이 낮을수록 낮은 수치를 보인다. 우측 흰색 부분은 기록관리주기 측면에 관한 영역이고 좌측 회색 부분은 기록물 측면에 관한 영역을 나타내고 있다. 각 영역에 대해 높은 연관성을 보일수

록 바깥쪽 부분에 근접하며 낮은 연관성을 보일수록 안쪽 부분에 가깝게 된다.

먼저 기록관리주기 측면의 영역을 살펴보면, 1997~2001에는 생산, 관리 및 보존에 관한 연구가 활발하게 진행되었으며 반면에 평가, 활용에 관한 연구는 미비하게 진행되고 있음을 확인할 수 있다. 이러한 흐름은 2002~2006까지 계속되었다가 2007~2011부터 활용에 관한 연구가 점차 증가하기 시작하였다. 생산에 관한



<그림 5> 전후기 10년, 5년 주기 영역별 토픽 동향 방사형 그래프

연구는 2002~2006에 잠시 주춤하였지만 이후 다시 활발하게 진행되었다. 관리 및 보존에 관한 연구는 20년동안 꾸준히 관심을 받아왔으며, 평가에 관한 연구는 2011년까지 미비하게 진행되다가 점차 증가하게 되었다.

기록물 측면 영역의 흐름을 살펴보면, 1997~2001에는 대통령기록과 공공기록에 관한 연구가 활발하게 진행되었으나 전자기록, 민간기록에 관한 연구는 미비하였다. 2002년 이후 전자기록과 민간기록에 관한 연구가 지속적으로 증가하게 되었으며 2011년까지 공공기록이 민간기록보다 활발하게 진행되었으나 2012년 이후 공공기록에 대한 연구는 조금 감소하고 민간기록에 관한 연구는 크게 증가하였다.

특히 전자기록 영역은 1997~2001을 제외하면 항상 높은 관심을 받고 있는 것을 확인할 수 있다. 이에 대해 전자기록 영역이 처음 많은 관심을 받게 된 2002~2006과 가장 최근 시기인 2012~2016의 전자기록 영역에 대한 주요 키워드를 살펴보고 전자기록 영역에 대한 이슈와 흐름이 어떻게 변화했는지 분석하였다. 2002~2006, 2012~2016 전자기록 영역의 주요 키워드는 <표 4>와 같다. 2002~2006 전자기록 영역

의 주요 키워드를 살펴보면, “기록관리시스템”, “기록물분류기준표”, “디지털화”, “전자정부”, “공문서” 등 공공기록에 대한 디지털화, 공공기관의 기록관리시스템, 전자정부에 관한 연구가 주로 진행되었으며, 또한 “메타데이터”, “표준화”와 같이 전자기록에 대한 표준화에 대한 연구가 관심을 받고 있었다는 것을 알 수 있다. 반면 2012~2016 전자기록 영역의 주요 키워드를 살펴보면, “웹사이트”, “SNS”, “클라우드” 등 네트워크 환경의 발달로 인해 새로운 형태의 전자기록에 대한 관심이 높아지고 있으며, “ATOM”, “오픈소스”, “부산항”, “송전탑” 등 공개소프트웨어를 활용하여 디지털 아카이브를 구축하는 연구, 디지털 기록정보서비스에 대한 연구 등이 주로 진행되어지는 것을 파악할 수 있다.

4.3 시사점

분석결과를 종합적으로 살펴보면 다음과 같다. 전·후기 10년 토픽 분석 결과를 살펴보면, 1997년부터 2006년까지 전기 10년은 기록관리학이 탄생한 지 얼마 되지 않은 시점이기 때문에 기록관리학 전반적인 부분에 대한 개념적인 연

<표 4> 2002~2006, 2012~2016 전자기록 영역 주요 키워드

2002~2006 전자기록 영역	2012~2016 전자기록 영역
기록관리시스템	웹사이트
메타데이터	로컬리티
디지털화	기록정보서비스
전자정부	SNS
기록물분류기준표	ATOM
공문서	오픈소스
행정기관	클라우드
기록관리법	부산항
표준화	송전탑

구, 고찰 등이 활발하게 진행되었으며 기록관리학과 관련된 역사에 대한 연구, 아키비스트 양성과 교육에 대한 연구가 관심이 높았다. 또한 국외 기록관리학에 대한 사례연구를 통해 국내 기록관리학의 방향과 목적을 찾으려고 노력하였다. 이후 후기 10년 토픽은 2007년 공공기록관리법이 개정되어지는 시점부터 전자기록과 기록정보서비스에 대한 관심이 점차 집중되기 시작하였으며 학교기록, 소송기록, 연구기록 등 다양한 민간 기록물에 대한 기록화 방안에 대한 연구가 활발하게 진행되었다.

5년 주기 영역별 토픽 분석을 살펴보면, 기록관리주기 측면에서는 초기 생산, 관리 및 보존에 관한 연구는 활발하였으나 평가, 활용에 관한 연구는 미비하게 진행되었으나 2007년부터 활용에 관한 연구가 지속적으로 증가하였으며 2012년부터는 평가에 관한 연구도 점차 증가하였다. 생산, 관리 및 보존에 관한 연구는 2002~2006에 잠깐 줄어들었지만 점차 증가하였다. 기록물 측면에서는 초기 전자기록, 민간기록에 관한 연구는 많이 진행되지 않았고 대통령기록과 공공기록에 관한 연구가 많은 관심을 받았다. 2002년 이후 전자기록과 민간기록 영역이 지속적으로 증가하게 되었다. 2011년까지 공공기록 영역이 민간기록 영역보다 항상 앞서있었지만 2012년 이후 민간기록 영역이 공공기록 영역보다 앞서면서 더 많은 관심을 받게 되었다. 이러한 변화는 사회가 발전하고 민주주의에 대한 의식이 향상되어짐에 따라 민간기록에 대한 관리의 중요성이 높아졌기 때문이다. 종합적으로 초기에는 공공기록, 대통령기록, 생산, 관리 및 보존에 관한 연구가 주로 활발하게 진행되었지만 최근에는 전자기록, 민간기록, 활용에 관한

연구가 활발하게 진행되어지고 있는 것을 확인할 수 있다.

5. 결 론

본 연구는 최근 4차 산업혁명 시대로 인해 정보의 규모가 거대해지고 처리속도가 더욱 빨라지는 시대의 흐름 속에서, 트렌드 분석을 통해 시간의 변화에 따른 기존의 국내 기록관리학 연구동향에 대해 심층적으로 분석하고 이를 기반으로 향후 국내 기록관리학 연구의 발전 방향과 흐름에 대해 살펴보려고 하였다. 본 연구의 결과를 정리하면 다음과 같다.

첫째, 기록관리학 전·후기 10년 토픽분석 결과를 살펴보면, 전기 10년 기록관리학 연구는 주로 개념적인 연구, 학문적 고찰 등과 같은 연구가 진행되어졌다. 또한 기록관리학과 관련된 역사에 대한 연구와 국내 아키비스트 양성과 교육에 대한 연구도 관심이 높았으며, 국외 기록관리학에 대한 사례 분석을 통해 국내 기록관리학의 방향과 목적을 탐색하는 연구도 많이 수행되었다. 이후 기록관리학 연구는 2007년 공공기록관리법이 개정되는 시점부터 점차 전자기록과 기록정보서비스와 관련된 연구가 증가하기 시작하였으며 학교기록, 연구기록, 소송기록 등 특정 기록물에 대한 아카이빙 방안에 대한 연구가 활발하게 진행되었다.

둘째, 5년 주기 영역별 토픽 분석 결과를 살펴보면, 초기 기록관리학 연구는 생산 영역, 관리 및 보존 영역, 공공기록 영역, 대통령기록 영역에서 활발한 연구가 진행되었음을 확인할 수 있었다. 2002년 이후 전자기록 영역과 민간기

록 영역이 점차 상승세를 보이기 시작하였으며, 2007년 이후 활용에 관한 연구도 점차 증가하기 시작하였다. 2012년 이후에는 전자기록, 민간기록, 활용에 관한 연구가 가장 많은 관심을 받게 되었으며, 공공기록은 초기에 비해 상대적으로 하락세를 보였다.

본 연구는 토픽모델링 기반 트렌드 분석을

통해 국내 기록관리학 연구동향에 대해 살펴보았다. 토픽모델링은 텍스트마이닝 기반의 정량적 분석 방법이기 때문에 내용분석을 통한 세부적인 연구동향을 파악하는데 다소 한계점을 보인다. 따라서 향후 연구에서는 정성적인 분석 방법을 병행하여 국내 기록관리학 연구동향에 대해 보다 세부적인 분석이 진행되어야 할 것이다.

참 고 문 헌

- 김규환, 남영준 (2009). 국내 기록관리학 분야 학회지 논문 분석을 통한 연구동향 연구. 한국문헌정보학회지, 43(4), 217-239.
- 김규환, 장보성, 이현정 (2009). 우리나라 기록관리학 분야의 연구영역 분석 - 논문제목의 구문 및 의미 구조를 중심으로. 한국문헌정보학회지, 43(3), 417-439.
- 김난도, 권혜진, 김희정 (2008). 트렌드 코리아 2009. 서울: 미래의 창.
- 김남규, 이동훈, 최호창, William Xiu Shun Wong (2017). 텍스트 분석 기술 및 활용 동향. 한국통신학회논문지, 42(2), 471-492.
- 김종근 (2014). 대한민국 사회문화 트렌드 분석 - 2010년 초반에 대한 정성적 접근. 문화산업연구, 14(1), 45-56.
- 김태경, 최희련, 이홍철 (2016). 토픽 모델링을 이용한 핀테크 기술 동향 분석. 한국산학기술학회 논문지, 17(11), 670-681.
- 김현정, 조남욱, 신경식 (2015). 항공산업 미래유망분야 선정을 위한 텍스트 마이닝 기반의 트렌드 분석. 지능정보연구, 21(1), 65-82.
- 나상태, 김자희, 정민호, 안주언 (2016). 토픽 모델링을 이용한 시뮬레이션 연구 동향 분석. 한국시물레이션학회논문지, 25(3), 107-116.
- 나상태, 안주언, 정민호, 김자희 (2017). 동적 토픽분석을 활용한 스마트그리드 연구동향 분석. 전기학회 논문지, 66(4), 613-620.
- 남태우, 이진영 (2009). 우리나라 기록관리학 연구 동향 분석. 한국도서관·정보학회지, 40(2), 451-472.
- 박주섭, 홍순구, 김종원 (2017). 토픽모델링을 활용한 과학기술동향 및 예측에 관한 연구. 한국산업정보학회논문지, 22(4), 19-28.
- 박준형, 오효정 (2017). 국내 기록관리학 연구동향 분석을 위한 토픽모델링 기법 비교 - LDA와 HDP

- 를 중심으로. 한국도서관·정보학회지, 48(4), 235-258.
- 박지애, 박병철, 최안섭 (2007). 사례분석을 통한 공동주택 경관조명의 트렌드 연구. 대한건축학회 논문집 - 계획계, 23(5), 203-210.
- 손혜인, 남영준 (2016). 기록관리학 분야 국내 학술지의 연구동향에 관한 연구 - 『한국기록관리학회지』와 『기록학연구』를 중심으로. 정보관리학회지, 33(1), 85-110.
- 안병기 (2003). 패션전문용어의 이해. 서울: 경춘사.
- 안소미, 정화, 이지연, 이연숙 (2009). 최근 고령친화환경의 주요 트렌드 분석 연구. 한국살내디자인학회 학술대회논문집, 11(1), 192-197.
- 유리나, 황수철 (2017). 텍스트마이닝 기법을 이용한 국내 농식품유통 연구동향 분석. 한국산학기술학회 논문지, 18(10), 215-226.
- 윤소연, 윤동근 (2017). 토픽모델링을 이용한 재난 및 안전관리 동향 분석. 한국지형공간정보학회지, 25(3), 75-85.
- 이소연 (2011). 국내 전자기록 연구의 동향 분석 - 회고와 전망. 한국기록관리학회지, 11(2), 7-31.
- 이순중 (2010). 디자인의 시대 트렌드의 시대. 서울: 미래의창.
- 이재운, 문주영, 김희정 (2007). 텍스트 마이닝을 이용한 국내 기록관리학 분야 지적구조 분석. 한국문헌정보학회지, 41(1), 345-372.
- 이진백, 지충권, 차경진 (2015). 트윗 데이터를 활용한 IT 트렌드 분석. 지능정보연구, 21(1), 143-159.
- 장로사, 김유승 (2009). 한국의 기록관리학 연구동향 분석에 관한 연구 - 정보학 영역을 중심으로. 한국비블리아학회지, 20(1), 87-100.
- 최이량 (2015). 국내 기록관리학 연구동향에 관한 연구 - 최근 10년간(2004-2013) 학술논문을 중심으로. 기록학연구, 43, 147-177.
- 한국기록관리학회 (2013). 기록관리론: 증거와 기억의 과학. 성남: 아세아문화사.
- P. Aburdene (2006). 메가트렌드 2010. (윤여중 역). 서울: 청림출판. (원전발행년 2005).

[웹사이트]

- Gensim (2009). HDP Topic modeling API. 검색일자: 2017. 12. 15. <https://radimrehurek.com/gensim/>
- Konlpy (2009). Konlpy 파이썬 한국어 NLP. 검색일자: 2017. 12. 15. <http://konlpy.org/en/v0.4.4/>

• 국문 참고자료의 영어 표기

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

- Ahn, Byung-Gi (2003). Understanding Terminology of Passion. Seoul: Kyungchun.
- An, So-Mi, Zheng, Hua, Lee, Ji-Yeun, & Lee, Yeon, Sook (2009). Analysis of Major Trends

- appeared in recent Aging Friendly Environment. Korean Institute of Interior Design Proceeding, 11(1), 192-197.
- Chang, Ro-Sa, & Kim, You-Seung (2009). An analytical Study on Research Trends of Records and Archives Management Science in Korea: Focused on the Information Science. Journal of the Korean Biblia Society for Library and Information Science, 20(1), 87-100.
- Choi, Yilang (2015). A Study on the Research Trends of Archival Studies in Korea: Focused on Research Papers between 2004 and 2013. The Korean Journal of Archival Studies, 43, 147-177.
- Kim, Gyu-Hwan, Jang, Bo-Seong, & Yi Hyun-Jung (2009). A Study on Intellectual Structure of Records Management and Archives in Korea: Based on Syntactic and Semantic Structure of Article Titles. Journal of The Korean Society for Library and Information Science, 43(3), 417-439.
- Kim, Gyu-Hwan, & Nam Young-Joon (2009). A Study on the Research Trends of Records and Archives Management in Korea through an Analysis of Journal Articles. Journal of The Korean Society for Library and Information Science, 43(4), 217-239.
- Kim, Hyun-jung, Jo, Nam-ok, & Shin, Kyung-shik (2015). Text Mining-Based Emerging Trend Analysis for the Aviation Industry. Journal of Intelligence and Information Systems, 21(1), 65-82.
- Kim, Jong-keun (2014). Analysis of Korea Social Culture Trend: A Qualitative Research in the Early 2000. Journal of Korea Culture Industry, 14(1), 45-56.
- Kim, Namgyu, Lee, Donghoon, Choi, Hochang, & Willam Xiu Shun Wong (2017). Investigations on Techniques and Applications of Text Analytics. The Journal of Communications and Information Sciences, 42(2), 471-492.
- Kim, Nando, Gwon, Hyejin, & Kim, Hee-Jung (2008). Trend Korea 2009. Seoul: Mi Raeeuichang.
- Kim, TaeKyung, Choi, HoeRyeon, & Lee, HongChul (2016). A Study on the Research Trends in Fintech using Topic Modeling. The Journal of Korea Academy Industrial Cooperation Society, 17(11), 670-681.
- Lee, Jae-Yun, Moon, Ju-Young, & Kim, Hee-Jung (2007). Examining the Intellectual Structure of Records Management & Archival Science in Korea with Text Mining. Journal of The Korean Society for Library and Information Science, 41(1), 345-372.
- Lee, Soon-Jong (2010). Era of Design, Era of Trend. Seoul: Mi Raeeuichang.
- Lee, So-Yeon (2011). Trends Analysis of Electronic Records and Archives Research in Korea: Retrospect and Prospect. Journal of Korean Society of Archives and Records Management, 11(2), 7-31.

- Na, Sang-Tae, Ahn, Joo-Eon, Jung, Min-Ho, & Kim, Ja-Hee (2017). Research Trend Analysis for Smart Grids Using Dynamic Topic Modeling. *The transactions of The Korean Institute of Electrical Engineers*, 66(4), 613-620.
- Na, Sang-Tae, Kim, Ja-Hee, Jung, Min-Ho, & Ahn, Joo-Eon (2016). trend Analysis using Topic Modeling for Simulation Studies. *Journal of the Korea Society for Simulation*, 25(3), 107-116.
- Nam, Tea-Woo, & Lee, Jin-Young (2009). A Study on the Research Trends of Records and Archives Management in Korea. *Journal of Korean Library and Information Science Society*, 40(2), 451-472.
- Park, Ji-Ae, Park, Byoung-Chul, & Choi, An-Seop (2007). Case Studies on the Trend of the Outdoor Lighting Design in Apartment Complex. *Journal of the Architectural Institute of Korea Planning & Design*, 23(5), 203-210.
- Park, Ju Seop, Hong, Soon-Goo, & Kim, Jong-Weon (2017). A Study on Science Technology Trend and Prediction Using Topic Modeling. *Journal of the Korea Industrial Information Systems Research*, 22(4), 19-28.
- Park, JunHyeong, & Oh, Hyo-Jung (2017). Comparison of Thpic Modeling Methods for Analyzing Research Trends of Archives Management in Korea: focused on LDA and HDP. *Journal of Korean Library and Information Science Society*, 48(4), 235-258.
- Records Management & Archives Society of Korea (2013). *Records & Archives Management-The Science of Evidence & Memory*. Seongnam: The Asian Culture Press.
- Sohn, Hye In, & Nam, Young Joon (2016). A Study on the Research Trends of Archives Management in Korea: Focused on the Journal of Records - Management & Archives Society of Korea and The Korean Journal of Archival Studies. *Journal of Korea Society for Information Management*, 33(1), 85-110.
- Yi, Jin Baek, Lee, Choong Kwon, & Cha, Kyung Jin (2015). An Analysis of IT Trends Using Tweet Data. *Journal of Intelligence and Information Systems*, 21(1), 143-159.
- Yoo, Li-Na, & Hwang, Su-Chul (2017). A Trend Analysis of Agricultural and Food Marketing Studies Using Text-mining Technique. *Korea Academy Industrial Cooperation Society*, 18(10), 215-226.
- Yoon, So Yeon, & Yoon, Dong Keun (2017). A Trends Analysis on Disaster and Safety Management using Topic Modeling. *Journal of Korean Society for Geo-spatial Information Science*, 25(3), 75-85.