

국내 기록분야 연구주제 분석: 2002~2023년간 기록관리학, 문헌정보학, 역사학 학술논문을 중심으로

Analysis of Research Topics in Archival Studies: Focusing on
Academic Papers in Archival Science, Library and
Information Science, and History from 2002 to 2023

김선욱(SeonWook Kim)

E-mail: sewokim@gmail.com

대구가톨릭대학교 문헌정보학과 강사



논문접수 2023-10-16
최초심사 2023-10-24
게재확정 2023-11-08

ORCID

SeonWook Kim
https://orcid.org/0009-0009-1630-6876

초 록

본 연구의 목적은 국내 기록관리학, 문헌정보학, 역사학 학술논문의 서지정보를 분석함으로써 기록분야 연구주제를 분석하는 데 있다. 이를 위해 1,173편의 학술논문을 수집한 뒤, 저자키워드 데이터로부터 네트워크 분석을 시행하고 초록 데이터로부터 토픽모델링을 진행하고 분석 결과를 시간의 흐름에 따라 정리하였다. 저자키워드로부터의 네트워크 분석 결과에 따르면, 주요 법령과 정책의 변화에 따라 연구주제 네트워크가 적극적으로 변화하는 것이 확인되었다. 초록으로부터의 토픽모델링 결과에 따르면, 전체 학술논문의 주제는 '레코드매니지먼트', '아카이빙', '국가기록정책'으로 구분된다. 2002~2009년 동안은 '레코드매니지먼트'와 '국가기록정책'이 상대적으로 우세하였으나 2009년부터 균형적인 양적 성장을 이루며 2019년에 정점을 이루었다.

ABSTRACT

This study aims to analyze research topics within the domain of archival studies by examining bibliographic information from academic papers in archival science, library and information science, and history. After collecting 1,173 academic papers, network analysis was performed based on author keyword data, topic modeling was conducted from abstract data, and the analysis results were organized over time. The network analysis results based on author keywords confirmed that the research topic network actively changed according to variations in major laws and policies. Moreover, topic modeling from the abstract showed that the subjects of the entire academic paper were divided into "Records Management," "Archiving," and "National Records Policy." Notably, from 2002 to 2009, "Records Management" and "National Records Policy" were relatively dominant, but it has achieved balanced quantitative growth since 2009, peaking in 2019.

Keywords: 기록관리학, 연구주제, 연구동향, 네트워크분석, 토픽모델링
archival science, research topic, research trend, network analysis, topic modeling

© 한국기록관리학회

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

1. 서론

지난 수십 년 동안 기록 업무 현장에서 축적된 기록관리 전문가의 경험과 학문적 발전에 이바지하고자 노력해 온 연구자의 업적에 힘입어, 국내 기록관리학은 그 학문 영역을 끊임없이 확장해 왔다. 미국의 경우, Cox(2009)는 30년 전까지만 하더라도 기록관리학 대학원에서 기록보존 기술 업무와 참조 서비스만을 가르쳐도 충분했으나, 이제는 교육 범위가 평가, 지식재산권, 법령, 변호에까지 확장되었다고 주장하였다. 그러나 이로 인한 부작용도 확인되었는데, Benoit와 Force(2019)는 현재와 같은 상황에서 대학원 교육 커리큘럼에 너무 많은 내용이 얹고 넓게 포함되어 있다고 지적하였다. 정영경(2005)은 이 같은 괴리를 선제적으로 타개하고자 미국과 한국의 기록관리학 대학원 교과과정을 검토하였다.

국내도 기록관리학 영역의 급격한 확장에 부합하지 못한 결과, 실무자가 기록관리 전문가를 양성하는 대학원의 교육은 여전히 과거 이론에 치중하고 있다고 인식할 수밖에 없는 상황을 초래하였다(김순희, 2009). 그 결과 계속교육을 통한 기록관리 전문가의 지식과 경험을 갱신해야 한다는 필요성이 나날이 커지고 있다(장혜리, 최재황, 2008). 그러나 실무 경험이 풍부한 기록관리 전문가에게 무엇을 더 가르쳐야 하는가에 대한 질문에 대학원이 답하기 위해서는, 기록관리학을 전공하는 대학원생에게 무엇을 가르쳐야 하는가 또는 현대 기록관리학의 영역과 연구주제는 무엇인가라는 질문으로 귀결될 수밖에 없다.

이를 탐구하기 위해서는 포용적이고 통찰력 있는 시각이 수반되어야 할 것이다. 우리나라 기록관리학의 역사는 관련 법령과 국가 제도의 도입과 발전 시기와 깊은 관계를 맺으므로 시기에 따라 연구주제를 구분하는 것이 필수적일뿐더러, 학문적으로 문헌정보학과 역사학과도 깊은 관련을 맺고 있다. 오향녕(2017)은 기록학은 역사학에 뿌리를 두고 있으며 역사학은 기록학에 토대를 두고 있는 유기적인 관계라고 강조하였으며 한미경, 노영희(2015)는 문헌정보학 교수의 기록관리학 관련 논문 게재 수가 늘고 있으며 문헌정보학과 기록관리학의 공동연구도 늘어나고 있음을 밝혀냈다.

따라서 본 연구는 기록관리학 학술논문이 게재된 기록관리학, 문헌정보학, 역사학 학술지에 2002년부터 2023년 8월까지 게재된 논문의 저자키워드와 초록으로부터 각각 네트워크 분석과 토픽모델링을 시행하고 그 결과를 시계열로 해석함으로써, 지난 약 22년간의 기록분야 학술논문의 연구 주제를 분석하고자 한다. 이를 통해 기록관리학의 연구 주제와 시기별 추이를 밝혀내어, 예비 기록관리 전문가에게는 연구 주제를 탐색하는데 이정표가 되고 기록관리 전문가와 연구자에게는 식견을 넓히는 지적 촉매 역할이 되길 기대한다.

2. 선행 연구

기록관리학 초기 연구주제를 분석하고자 시도한 연구는 내용분석을 기본으로 한 정량적 분석 방법이 주를 이루었다. 이재윤, 문주영, 김희정(2007)은 2001년부터 2006년까지 기록관리학 및 문헌정보학 학술지에 발표된 기록관리학 관련 145편의 논문 제목과 초록을 수집하여 군집화한 뒤 각 군집 간 관계를 밝혀냄으로써 기록관리학 내 지적구조를 구축하고자 하였다. 김규환, 남영준(2009)은 1999년부터 2009년까지 기록관리학 분야 374편의 학술논문을 대상으로 '연구배경정보'와 '연구자특성정보'라는 2개의 독립변인에 따른 '주제영역'의 종속변인을 확인하고자 빈도 분석과 대응 분석을 시행하여 시간에 흐름에 따른 연구주제를 분석하였다. 남태우, 이진영(2009)은 기록관리학 학술지 논문 399편을 내용 분석함으로써 주제를 분류하고, 통계 분석을 통해 각종 현황을 파악하고자 하였다.

네트워크 분석과 토픽모델링이 학계에서 적극적으로 사용되기 시작한 것은 이를 이용할 수 있는 분석 환경이 보급된 후부터라고 할 수 있다. 기록관리학 분야도 R과 Python과 같은 개발 환경에서의 관련 API가 등장하거나, Netminer와 NodeXL과 같은 분석 도구가 보급된 이후 연구 활동에 적극 채택되는 경향을 보인다.

박희진(2014)은 2009년부터 2014년까지 영국의 기록관리학 전문학술지 『Archivaria』에 실린 75편의 논문을

대상으로 연구자 분포 및 저자키워드 네트워크 분석을 실행함으로써 영국의 기록관리학 연구 동향을 국내 학계에 전달하고자 하였다. 최이랑(2015)은 2004년부터 2013년까지 기록관리학 및 문헌정보학 학술지에 발표된 기록분야 논문 479편으로부터 저자키워드와 저자소속기관을 추출하여 네트워크 분석과 내용분석을 토대로 기록관리학 연구 동향을 살피고자 하였다. 손혜인, 남영준(2016)은 저자 관계와 기관, 그리고 기관과 연구주제 영역 사이의 관계를 확인하기 위해 2000년부터 2015년까지 게재된 681편의 기록관리학 학술지 논문을 조사하여 빈도 분석과 네트워크 분석을 시행하였다. 김희섭, 강보라(2018)는 2001년부터 2017년까지 한국기록관리학회지에 게재된 390편의 논문의 저자키워드를 네트워크 분석하여 연결 중심성과 매개 중심성에 기반한 연구주제를 분석하였다. 토픽모델링을 이용한 분석 연구의 경우, 박준형, 오효정(2017)은 1997년부터 2016년까지 기록관리학 및 문헌정보학 학술지에 발표된 기록분야 논문 1,027건의 제목을 토픽모델링 분석하는데 LDA와 HDP, 두 가지 방법론을 동시에 사용하여 차이점을 도출하고자 하였다.

다만, 기록관리학 분야에서 진행된 대부분의 네트워크 분석 및 토픽모델링 연구는 전체 데이터를 하나의 집합으로써 분석했다는 한계를 지니며 시기별 또는 시계열 기준으로 분석하려는 시도는 많지 않았다. 박준형, 류범모, 오효정(2018)은 1997년부터 2016년까지 약 20년 동안의 기록관리학과 문헌정보학 학술논문 1,027편을 5년 단위의 중기 4구간 그리고 10년 단위의 장기 2구간으로 나눈 뒤, 논문 제목에서 추출한 명사 기반 토픽모델링 결과를 근거로 중장기 연구주제 영역 관련성을 파악하고자 하였다. 나연희, 박희진(2022)은 2000년부터 2021년까지 약 22년 동안 문헌정보학 및 기록관리학 학술지에 실린 기록보존 관련 논문 463편의 저자키워드를 7년 또는 8년 단위로 3개 시기로 구분하고 NetMiner를 사용해 각각 네트워크 분석함으로써 시기별 중요도 높은 주제어를 도출하고자 하였다.

3. 이론적 배경

3.1 네트워크 분석

네트워크 분석은 분석 대상의 계량적 특성을 추출한 뒤 기하학적 네트워크로 구성하여 의미적 관계를 해석하고 영향력이 높거나 중요한 지위를 갖는 노드를 찾는 방법론을 말하며, 여러 학문 분야에서 확대된 내용분석 방법론으로써 채택하고 있다(김성희, 장로사, 2010; 이수상, 2014; 이용구, 우윤희, 2015) 네트워크 내 분석 대상을 노드(node)라고 부르며, 노드 간 연결을 에지(edge)로 칭한다.

3.1.1 연결 중심성

일반적으로 네트워크 안에서 연결이 많은 노드는 중요하다고 인식될 수 있는데, 예를 들어 감염병 상황에서 감염자 집단 내에서 타인과 접촉회수가 높은 사람은 중요한 감염원으로 간주할 수 있다. 연결 중심성(Degree Centrality)은 수식 (1)의 정의와 같이 한 노드가 다른 노드와 연결된 정도를 측정하는 것으로, 규모가 다른 네트워크 사이에서 연결 중심성을 비교할 수 있도록 산술평균으로 정규화한 것이 수식 (2)의 표준화 연결 중심성이다(Freeman, 1978). n 은 전체 노드 개수이며, $\alpha(p_i, p_k)$ 는 노드 p_i 와 노드 p_k 가 연결되었을 때 1이 된다.

$$C_D(p_k) = \sum_{i=1}^n \alpha(p_i, p_k) \quad (1)$$

$$C'_D(p_k) = \frac{\sum_{i=1}^n \alpha(p_i, p_k)}{n-1} \quad (2)$$

3.1.2 근접 중심성

어떤 노드의 연결 중심성이 높더라도 실제로는 네트워크의 변방에 있을 수 있는데, 이 경우 거시적 관점에서 해당 노드의 중심성이 높다고 보기는 어렵다. 이 경우에는 한 노드가 네트워크 내 나머지 노드까지 얼마나 가까운지 측정하는 것이 필요하다. 무방향 네트워크에서 근접 중심성(Closeness Centrality)은 수식 (3)의 정의와 같이 한 노드가 네트워크의 다른 노드들에 도달하는 최단 경로 합의 역수이고, 규모가 다른 네트워크 사이에서 근접 중심성을 비교할 수 있도록 이론상 최댓값으로 나누어 정규화한 것이 수식 (4)의 표준화 근접 중심성이다(Beauchamp, 1965; Sabidussi, 1966). n 은 전체 노드 개수이며, $d(p_i, p_k)$ 는 노드 p_i 와 노드 p_k 의 최단 경로 내 에지의 합이다.

$$C_C(p_k) = \frac{1}{\sum_{i=1}^n d(p_i, p_k)} \quad (3)$$

$$C'_C(p_k) = (n-1)(C_C(p_k)) \quad (4)$$

3.1.3 매개 중심성

네트워크 안에서 노드를 연결할 때 반드시 거쳐야 하는 노드가 있다면 중요한 중개자 역할을 하므로 영향력이 높은 요소로 인식될 수 있다(이수상, 2013). 예를 들어 통신망에서 매개 중심성이 높은 노드는 통제력을 갖춘 중요한 통신 중계 장비로 간주한다. 매개 중심성(Betweenness Centrality)은 수식 (5)의 정의와 같이 한 노드가 다른 두 노드 사이에서의 중개자 역할을 수행한 정도를 측정하는 것으로, 규모가 다른 네트워크 사이에서 매개 중심성을 비교할 수 있도록 이론상 최댓값으로 나누어 정규화한 것이 수식 (6)의 표준화 매개 중심성이다(Brandes, 2001). n 은 전체 노드 개수이며, σ_{st} 는 노드 s 에서 노드 t 까지 최단 경로의 개수이고, $\sigma_{st}(v)$ 는 노드 s 에서 노드 t 까지 최단 경로 중에서 v 를 포함하는 경로의 개수이다.

$$C_B(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}} \quad (5)$$

$$C'_B(v) = \frac{C_B(v) \times 2}{(n-1)(n-2)} \quad (6)$$

3.1.4 아이겐벡터 중심성

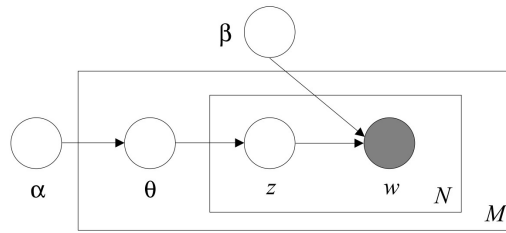
이상의 중심성 이론은 모두 네트워크 그래프 관점에서 노드의 연결 또는 비연결이라는 이진값(binary status)만을 중요시하는 것에 반해, 아이겐벡터 중심성(Eigenvector Centrality)은 어떤 노드에 연결된 노드의 개수뿐만 아니라 그 노드의 중요성도 함께 고려하는 방법으로써 수식 (7)처럼 연결된 전체 노드의 중심성 값의 합에 비례한다(Bonacich, 2007). n 은 전체 노드 개수이며, λ 는 최대 아이겐벡터 값이고, a_{ij} 는 노드 i 에서 노드 j 간 이진값이다.

$$\lambda C_E(i) = \sum_j a_{ij} C_E(j) \quad (7)$$

3.2 토픽모델링

토픽모델링(topic modeling)은 문서 집합 내 주제를 찾는 방법론을 통칭한다. 여러 분야의 연구에서 널리 쓰이는 통계 모델인 LDA(Latent Dirichlet Allocation)를 비롯해, 최근에는 비지도 학습 알고리즘에 기반한 차원축소법 응용하기도 한다(Grootendorst, 2022). 그중에서 LDA는 <그림 1>처럼 실제 인간이 글을 쓸 때 가장 먼저 주제(α)를 결정하고, 그 주제를 담을 어휘(β)를 고르는 과정을 모사한 것이다. 즉, 문서 안에 어떤 주제가 존재할지에 대한 확률적 모델로서 현재 관찰할 수 있는 단어를 기반으로 문서별 토픽 분포와 토픽별 단어 분포를 추정한다(Blei,

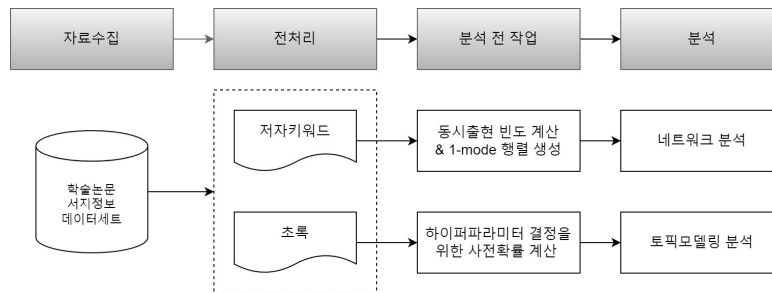
Ng & Jordan, 2003).



<그림 1> LDA 개념 (Blei, Ng & Jordan, 2003)

4. 연구 방법

본 논문의 연구설계는 <그림 2>과 같다.



<그림 2> 연구설계 개요

1단계 자료수집 과정에서는, 권위 있는 학술지의 논문을 정형화된 양식으로 수집하고자 KCI(한국학술지인용색인)를 이용하여 데이터를 수집하였다. 먼저 기록관리학 3대 학술지인 『한국기록관리학회지』, 『기록학연구』, 『기록과 정보·문화 연구』의 학술논문 서지정보를 모두 다운로드¹⁾하였다. 계속해서 KCI 검색엔진에서 기록관리학 학술지 내 최근 데이터 기준 시점인 2023년 8월까지 기간을 설정하고 주제 분류를 문헌정보학, 역사학을 한정하고 ‘기록’, ‘아카이브’, ‘아카이브즈’, ‘아카이빙’ 중 하나가 초록과 저자키워드에 모두 등장하는 조건으로 하여 검색하였다. 검색 결과인 1,540건(기록학 1,217건, 문헌정보학 264건, 역사학 172건)의 서지정보를 다운로드한 뒤 전수조사를 통해 제목과 초록의 내용이 기록분야에 해당하지 않은 논문과 초록 및 저자키워드가 빈(null) 논문을 제외한 결과, 최종적으로 1,173건의 서지정보를 확보하였다. 이 중 기록학 학술논문은 876편, 문헌정보학 학술논문은 247편, 역사학 학술논문은 50편이었고 기간은 2002년부터 2023년 8월까지였다. 최종 분석 데이터의 연도별 분포를 <그림 3>으로 시각화하였으며, 논문 편수 기준 상위 10위 규모의 학술지 현황을 <표 1>로 정리하였다.

2단계 전처리 과정에서는, 네트워크 분석에 입력될 저자키워드 데이터와 토픽모델링 분석에 입력될 초록 데이터에 서로 다른 기준과 방법을 적용하였다. 저자키워드는 쉼표나 세미콜론과 같은 구분자(delimiter)로 구분할 수 있는 명사 문자열을 의미하므로 자연어처리 측면에서 부하는 높지 않지만, 저자의 학풍에 따라 같은 대상이더라도 언어적 표현이 다를 수 있으므로 올바른 전처리 규칙을 만들어 적용하기 위한 노력이 필요할 수 있다. 연구 방향에 따라 연구자의 저자키워드 작성 의도를 존중할 목적으로 전처리를 최소화하는 예도 있지만(이혜경, 양기덕, 김선옥, 2019), 본 연구에서는 저자키워드에 대해 다음과 같은 전처리 규칙을 적용하였다.

1) 2023년 9월 20일 기준

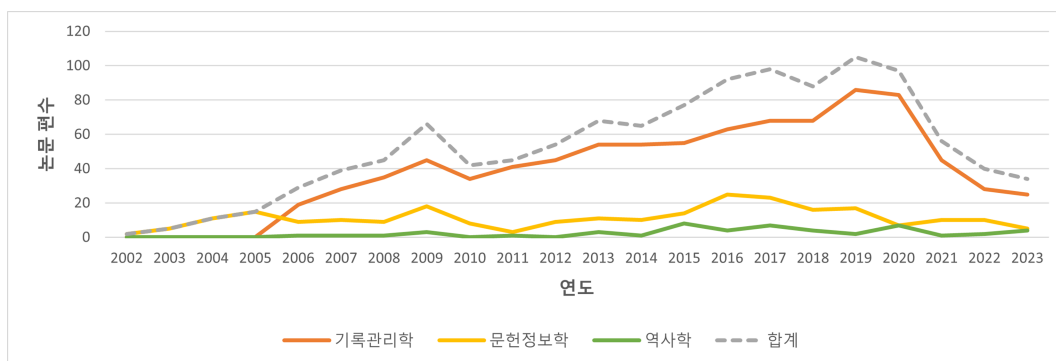
- 데이터셋 문제로 저자키워드에 외국어키워드가 들어있거나 한자로만 표기된 경우, 한글로 전환함.
- 영문은 모두 소문자 표기로 통일함.
- 가운데점 특수문자(유니코드를 기준으로 U+0387 Greek Ano Teleia, U+2024 One Dot Leader, U+22C5 Dot Operator, U+318D Hangul Letter Araea)를 올바른 문자(U+00B7 Middle Dot)로 통일함.
- 띄어쓰기 삭제함.
- 용어통일.

예 1. [‘기록학’, ‘기록관리학’]²⁾은 ‘기록학’으로 통일함. 단, 보존기록학과 민속기록학과 같은 세부 분야 명칭과 학술지명은 그대로 둠.

예 2. [‘archives’, ‘archive’, ‘아카이브즈’, ‘아카이브’]는 ‘아카이브’로 통일함.

예 3. [‘기록연구사’, ‘기록물관리전문요원’]은 ‘기록물관리전문요원³⁾’으로 통일함.

예 4. [‘기록관리사’, ‘기록관리자’, ‘기록전문직’, ‘기록관리전문직’, ‘기록관리전문인력’, ‘기록관리전문요원’, ‘기록전문요원’]⁴⁾은 ‘레코드매니저’로 통일함.



〈그림 3〉 최종 분석 데이터의 연도별 분포

〈표 1〉 최종 분석 데이터 내 상위 10위 규모 학술지

학술지명	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	합계
한국기록관리학회지	0	0	0	0	7	16	20	19	11	21	25	26	24	29	30	36	36	45	41	27	14	11	438
기록학연구	0	0	0	0	12	12	15	26	23	20	20	28	30	26	33	32	32	30	34	14	11	10	408
한국비블리아학회지	0	0	2	4	2	2	1	4	4	1	6	4	5	7	10	7	8	6	3	2	2	0	80
정보관리학회지	0	2	2	6	4	3	4	5	2	1	2	3	1	1	8	5	6	5	1	6	5	2	74
한국문헌정보학회지	1	3	5	5	2	3	3	7	2	1	0	3	1	2	3	3	2	2	3	2	1	2	56
한국도서관·정보학회지	1	0	2	0	1	2	1	2	0	0	1	1	3	4	4	8	0	4	0	0	2	1	37
기록과 정보·문화 연구	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	8	4	3	4	30
무형유산	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	1	0	1	0	0	0	6
역사문화연구	0	0	0	0	0	1	0	2	0	0	0	2	0	0	1	0	0	0	0	0	0	0	6
한국독립운동사연구	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	3
역사와경계	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
민속연구	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
한국독립운동사연구	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	3

2) 대괄호([])는 프로그래밍 언어에서 배열을 나열할 때 사용하므로, 같은 맥락에서 문자열 집합을 구분하기 위해 사용함.

3) 기록물관리전문요원은 공공기록물에 관한 법률에 따른 자격이며, 기록연구사는 공공기관에서 해당 자격자에게 부여되는 직위임.

4) 이 저자키워드가 언급되는 학술논문의 경우, 영문 논문명에 Record Manager로 주로 표기됨.

- 괄호가 사용된 경우, 단순 부연 설명하거나 어원을 쓰는 등 중의적인 성격이면 괄호의 내용을 생략함. 단, 괄호 안의 내용이 더 대표적이면 괄호의 내용만을 남김.

예 1. '가상현실(VR)'은 'VR', '국제기록물기술규칙(ISAD(G))'은 'ISAD(G)'으로 통일함.

- 한글 표현과 영문 약어가 혼용되는 경우, 영문 약어로 통일함. 예. erms, rms, cams.

영문 표현을 한글 음차로 표현하는 것이 표준인 경우는 이를 따름. 예. 코로나19⁵⁾

본 연구에서는 저자키워드 기준 데이터세트를 크게 4개의 시기로 분류하였다. 1기는 2002~2006년 동안 발간된 학술논문을 구분한 것으로, 공공기록물법 제정 이후이며 1999년 '공공기록물 관리에 관한 법률'이 제정되고 2000년 1월에 시행되면서 2005년에 최초로 중앙부처 45개 부서에 기록물관리전문요원이 배치된 시기이다. 2기는 2007~2014년 동안 발간된 학술논문을 구분한 것으로, 공공기록물법 개정 이후이다. 2006년 개정된 '공공기록물 관리에 관한 법률'이 2007년 4월 시행되고, 같은 해 12월에는 나라기록관과 대통령기록관이 개관한 시기이다. 더불어 국가기록원의 국가기록관리혁신 로드맵이 제안된 직후이다(국가기록원, 2006). 3기는 2015~2019년 동안 발간된 학술논문을 구분한 것으로, 2015년 1월에 국가기록관리 혁신의 일환으로서 권역별 통합기록관리 체계로 전환되었고(송정숙, 2017), 2017년부터는 국가기록관리 TF가 진행되었다(국가기록관리 혁신 T/F, 2017; 심성보, 2020). 4기는 2020~2023년 동안 발간된 학술논문을 구분한 것으로, 국가기록관리 중장기 발전계획 이후이다(국가기록원, 2020).

초록은 학술논문의 표지에 등장하는 저자의 논문 요약문에 해당하며, 다수의 문장이 단락 단위로 구분된 자연어 집합이므로 연구 목적에 부합하는 전처리 방안을 결정하는 것이 상당히 중요하다. 본 연구에서는 토픽모델링을 위해 초록에서 명사를 추출한 뒤, 등장빈도와 맥락적 의미에 따라 불용어 처리함으로써 연산 과정을 최적화하였다. 일반적으로 품사(Part of Speech)를 판별하기 위한 형태소분석기로서 은전한닢, 꼬꼬마, KoNLPy이 사용되곤 한다(김학래, 2022). 그러나 상기 언급된 형태소 분석기는 2007년 완료된 『21세기 세종 계획』에서 유래된 세종 말뭉치를 사용하고 있으므로, 다양한 영역에서 사용되는 현대 한국어를 정확히 처리한다고 기대하기에는 무리가 있다(이동주 외, 2010; 최민우, 강범모, 2008). 본 연구에서는 2020년 공개된 최신 대용량 엑소브레인 언어분석 말뭉치에 기반하여 동작하는 ETRI OPEN API⁶⁾를 사용하여 초록 전체로부터 UTF-8 캐릭터셋 기준 길이가 3 (한글 1글자, 영어 3글자) 이상인 한글 형태소 품사 태그표 기준 명사(NNG: 일반명사, NNP: 고유명사, SL: 외국어)만을 추출하였다. 이후 추출된 외국어는 표시형식을 소문자로 통일하였다.

3단계 분석 전 작업 과정은, 전처리 과정과 마찬가지로, 저자키워드 데이터와 초록 데이터에 각각 다른 작업이 필요하다. 먼저 시기별로 저자키워드 집합을 구분하고, 이재윤⁷⁾의 semi와 cooc 프로그램 v.0.4를 사용하여 동시출현 빈도를 계산한 뒤, 등장 횟수가 적은 저자키워드를 배제하고, 1-mode 행렬을 생성하였다. 초록의 경우, 최적의 토픽모델링에 필요한 하이퍼파라미터를 결정하기 위한 다중 연산을 시행하였다.

4단계 분석 과정에서는 본격적인 저자키워드의 네트워크 분석 및 초록의 토픽모델링 분석을 시행한다. 저자키워드 네트워크 분석은 1-mode 행렬로부터 Ucinet 프로그램으로 중심성을 계산하고, Gephi 프로그램을 사용하여 시각화하였다. 전체 데이터세트는 물론 4개의 시기별 데이터세트에 네트워크 분석을 적용함으로써, 시대적 흐름에 따른 맥락적 주요 연구 주제의 변화를 투영할 수 있다. 초록의 토픽모델링은 Python과 Gensim API를 이용해 분석용 프로그램을 작성하였다. 토픽모델링 방법론으로 많이 사용되는 LDA는 다양한 연구를 통해 접근방법이 개선되고 응용되고 있지만, 한편 최적의 결과를 얻기 위한 최적의 하이퍼파라미터를 설정하기가 어렵다고 알려져 있다(김선욱, 양기덕, 2022; 박자현, 송민, 2013; 박준형, 오효정, 2017). 하이퍼파라미터(α , β)를 결정하기 위해 참조하는 coherence score의 종류로는 u_mass, c_v, c_uci, c_npml 등이 있지만, 최적의 하이퍼파라미터를 결정하

5) <https://ncov.kdca.go.kr/>

6) <https://aiopen.etri.re.kr/>

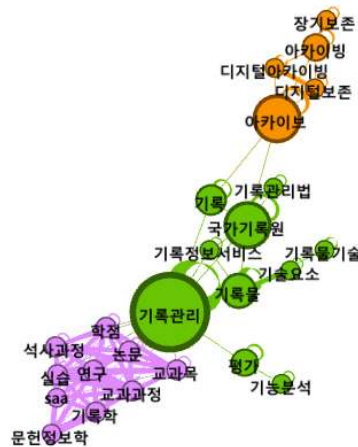
7) 명지대학교 문헌정보학과 교수

군집화하고 각 커뮤니티에 서로 다른 색상을 부여한 뒤, 이 중에서 연결도(degree)가 11인 이상만을 표현한 것이 <그림 4>이다. 각 군집 내 가장 높은 매개중심성을 갖는 노드를 군집의 규모 순으로 각각 나열하면 ‘기록관리’, ‘공동체아카이브’, ‘아카이브’, ‘기록학’, ‘국가기록원’, ‘메타데이터’, ‘기록평가’, ‘정보공개’, ‘기술’이다. 표준화 매개 중심성, 표준화 근접 중심성, 표준화 연결 중심성, 아이겐벡터 중심성을 기준으로 15위 순위를 <표 2>에 정리하였다. 4개 중심성 기준 공통으로 ‘기록관리’의 중심성이 가장 높았으며, 중심성 측정 기준에 따라 다소 차이가 있으나 ‘국가기록원’, ‘아카이브’, ‘기록정보서비스’가 뒤를 이었다. 이어서 ‘기록’, ‘기록관’, ‘평가’, ‘rms’, ‘아키비스트’, ‘기록물관리전문요원’, ‘대통령기록’, ‘대통령기록관’, ‘전자기록물’ 등이 중심도 점수가 높게 확인되었다.

요컨대, 국내 기록분야 학술논문의 저자키워드를 네트워크 분석한 결과, 그동안의 연구는 ‘기록관리’와 ‘국가기록원’과 ‘기록정보서비스’를 중심으로 ‘빅데이터’와 ‘일상아카이브’에 이르기까지 광범위한 영역을 다루고 있음을 확인할 수 있다.

5.1.2 제1기 (2002~2006) 공공기록물법 제정 이후

등장빈도 2회 이상인 저자키워드를 대상으로 시각화한 결과, 33개의 노드와 115개의 에지로 구성된 네트워크가 생성되었다. 네트워크의 모든 노드를 Louvain method에 기반한 Modularity Class로 7개로 군집화하고 각 커뮤니티에 서로 다른 색상을 부여한 것이 <그림 5>이다. 먼 거리에 있는 변방의 군집은 그림에서 생략하였다. 각 군집



<그림 5> 주요 저자키워드 네트워크 - 제1기 (2002~2006)

<표 3> 중심성 기준 주요 저자키워드 - 제1기 (2002~2006)

순위	표준화 매개 중심성	수치	표준화 근접 중심성	수치	표준화 연결 중심성	수치	아이겐벡터 중심성	수치
1	기록관리	33.468	기록관리	0.333	교과과정	0.219	교과과정	0.364
2	아카이브	16.129	국가기록원	0.314	연구	0.219	연구	0.364
3	국가기록원	15.625	기록물	0.299	실습	0.219	실습	0.364
4	기록물	9.073	기록	0.299	기록학	0.167	기록학	0.338
5	기록	6.552	기록학	0.296	석사과정	0.167	문헌정보학	0.275
6	평가	4.637	교과과정	0.296	교과목	0.167	석사과정	0.268
7	아카이빙	4.637	연구	0.296	학점	0.167	교과목	0.268
8	석사과정	0.756	실습	0.296	논문	0.167	학점	0.268
9	교과목	0.756	석사과정	0.296	문헌정보학	0.156	논문	0.268
10	실습	0.756	교과목	0.296	saa	0.146	기록관리	0.251
11	기록학	0.756	학점	0.296	기록관리	0.135	saa	0.245
12	교과과정	0.756	논문	0.296	기록물	0.063	기록물	0.02
13	학점	0.756	기록정보서비스	0.294	국가기록원	0.063	국가기록원	0.018
14	연구	0.756	아카이브	0.276	erms	0.052	기록	0.016
15	논문	0.756	평가	0.274	아카이브	0.052	기록정보서비스	0.016

내 가장 높은 매개중심성을 갖는 노드를 군집의 규모 순으로 각각 나열하면, 먼저 가장 큰 군집에서 8개 저자키워드가 ['기록학', '교과과정', '연구', '실습', '석사과정', '교과목', '학점', '논문'] 공동 1위였고 그다음 '기록관리', 'erms', '아카이브', 'oais참조모형', '기록관', '기능분류'이다. 다른 시기에 비해 네트워크 내 군집 간 연결은 느슨한 모양을 나타낸다.

표준화 매개 중심성, 표준화 근접 중심성, 표준화 연결 중심성, 아이겐벡터 중심성을 기준으로 15위 순위를 정리한 것이 <표 3>이다. 중심성 측정 기준과 상관없이 15위 내 위치한 저자키워드는 그 내용에 큰 차이가 없었다. 표준화 매개 중심성과 표준화 근접 중심성 기준으로는 '기록관리', '아카이브', '국가기록원', '기록', '기록물'이 상위 5위를 차지하였으나, 표준화 연결 중심성과 아이겐벡터 중심성을 기준으로는 '교과과정', '연구', '실습', '기록학'이 상위 4위였다. 교과과정과 관련된 키워드인 '석사과정', '교과목', '실습', '기록학', '교과과정', '학점', '연구', '논문'가 표준화 연결 중심성과 아이겐벡터 중심성에서는 상위권을 장악하고 있는 데 반해, 표준화 매개 중심성과 표준화 근접 중심성에서는 상대적으로 낮은 순위를 차지하고 있었다.

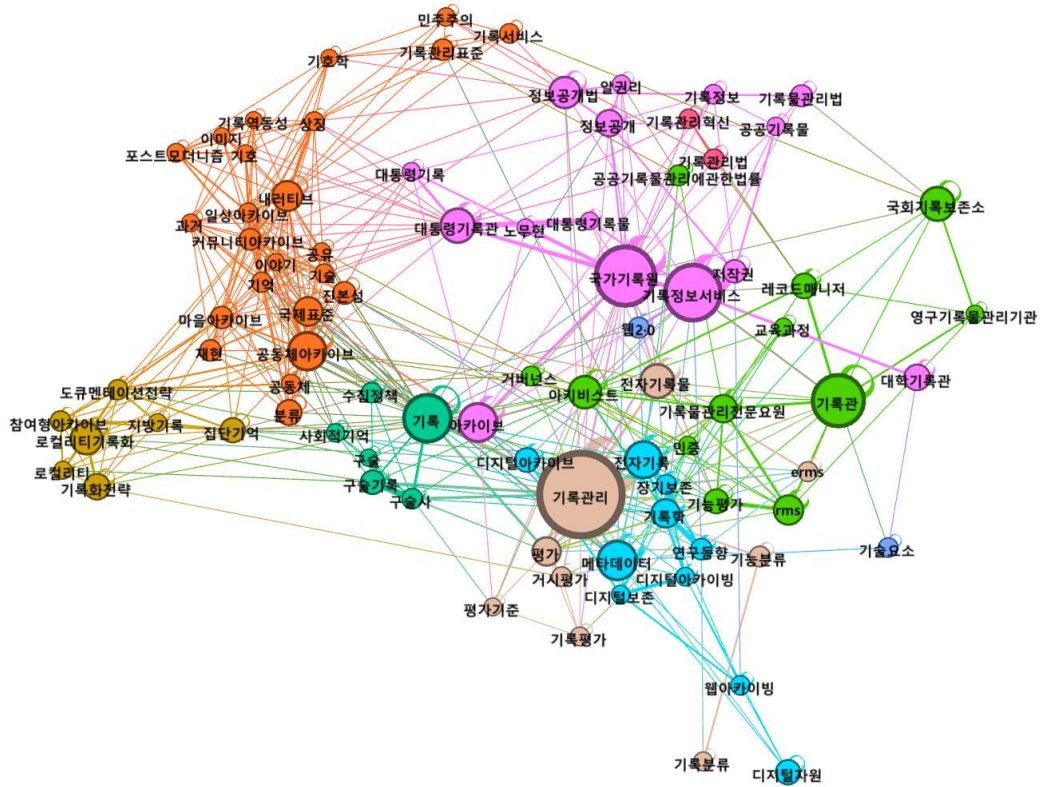
제1기의 저자키워드를 분석한 결과를 정리하면, 당대 '기록관리', '아카이브', '국가기록원', '기록', '기록물'이 저자키워드의 중요한 구심점과 통제점 역할을 하면서 '교과과정', '연구', '실습', '기록학'과 같이 기록분야 교육에 관련된 저자키워드가 동시 등장하는 학술논문이 활발하게 게재되었음을 의미한다. 이는 자격을 갖춘 기록물관리전문요원을 시급히 양성해야 하는 당시 시대적 상황을 투영하며, 더불어 레코드매니지먼트와 아카이빙에 관한 초기 연구 주제의 스펙트럼이 제한적이었던 초기 상황을 나타낸다고 할 수 있다.

5.1.3 제2기 (2007~2014) 공공기록물법 개정 이후

등장빈도 2회 이상인 저자키워드를 대상으로 시각화한 결과, 285개의 노드와 1,488개의 에지로 구성된 네트워크가 생성되었다. 네트워크의 모든 노드를 Louvain method에 기반한 Modularity Class로 12개로 군집화하고 각 커뮤니티에 서로 다른 색상을 부여한 뒤, 이 중에서 연결도(degree)가 11 이상인 관계만을 표현한 것이 <그림 6>이다. 각 군집 내 가장 높은 매개중심성을 갖는 노드를 군집의 규모 순으로 각각 나열하면 '국가기록원', '기록관', '메타데이터', '공동체아카이브', '기록관리', '기록', '로컬리티기록화', '기록관리법', '웹2·0', '대한제국', '전자저널', ['디지털전시', '웹전시']이다.

표준화 매개 중심성, 표준화 근접 중심성, 표준화 연결 중심성, 아이겐벡터 중심성을 기준으로 15위 순위를 정리한 것이 <표 4>이다. 중심성 측정 기준과 상관없이 15위 내 위치한 저자키워드는 그 내용에 큰 차이가 없었다. 표준화 매개 중심성, 표준화 근접 중심성, 표준화 연결 중심성과 아이겐벡터 중심성 네 가지 모두 '기록관리', '국가기록원', '기록정보서비스', '기록'이 상위 5위 내에 등장하였다. 표준화 매개 중심성에만 등장한 '국회기록보존소'와 '정보공개법'은 기록분야 내 새로운 통제점 역할을 하는 연구주제의 등장을 의미한다. 표준화 근접 중심성에서만 당대 '국제표준'이 국내 기록분야에 영향력을 미친 주제 중 하나라는 점에서 주목할 필요가 있다. '기록관'과 '대통령기록관'은 표준화 연결 중심성에서 4위와 6위를 차지하는 등 다른 중심성 순위에서도 새롭게 등장한 저자키워드였다. 이외에도 '평가', 'rms', '기록물관리전문요원'이 저자키워드로 처음 등장하였다.

제2기의 저자키워드를 분석한 결과를 정리하면, '기록관리'를 중심으로 전자기록과 표준, 국가기록원과 대통령기록관을 중심으로 한 기록정보서비스에 관한 연구, 기록관 업무와 기록물관리전문요원의 역할에 관한 연구부터 공동체 아카이브에 이르기까지 학술 스펙트럼이 전체적으로 확장되고 세분되기 시작한 시기로 볼 수 있다.



<그림 6> 주요 저자키워드 네트워크 - 제2기 (2007~2014)

<표 4> 중심성 기준 주요 저자키워드 - 제2기 (2007~2014)

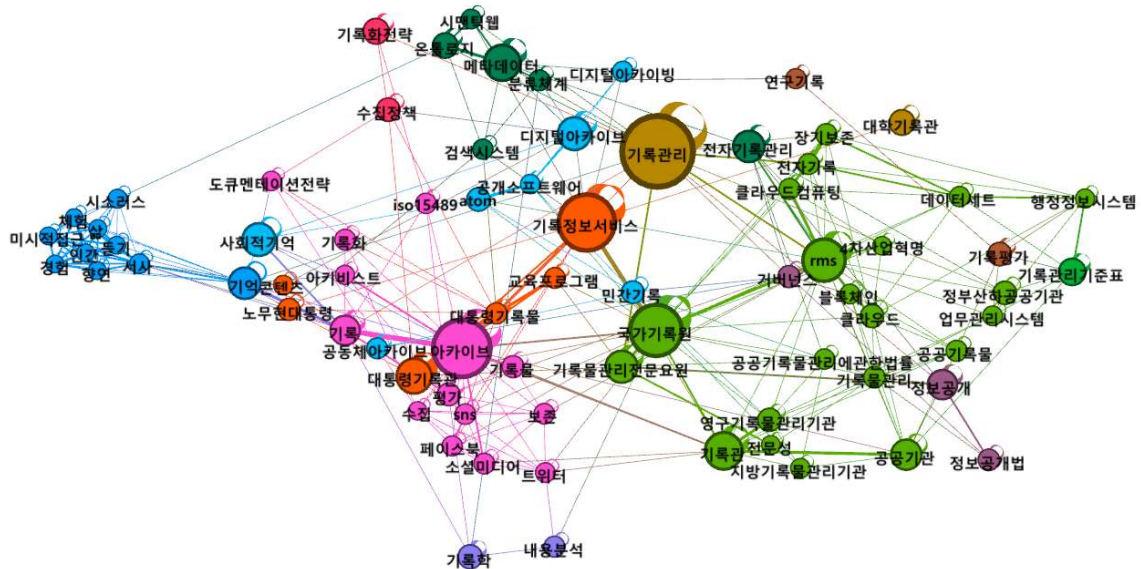
순위	표준화 매개 중심성	수치	표준화 근접 중심성	수치	표준화 연결 중심성	수치	아이겐벡터 중심성	수치
1	기록관리	17.572	기록관리	0.481	기록관리	0.034	기록관리	0.75
2	국가기록원	11.184	국가기록원	0.46	국가기록원	0.032	국가기록원	0.277
3	기록정보서비스	10.347	기록	0.46	기록정보서비스	0.026	기록정보서비스	0.236
4	기록관	9.237	기록정보서비스	0.444	대통령기록관	0.025	기록	0.18
5	기록	8.085	아카이브	0.441	기록	0.024	평가	0.164
6	메타데이터	5.885	평가	0.435	기록관	0.021	전자기록물	0.153
7	아카이브	5.731	전자기록물	0.428	공동체아카이브	0.021	구술사	0.144
8	공동체아카이브	5.625	공동체아카이브	0.427	전자기록	0.02	메타데이터	0.142
9	전자기록	5.311	기록관	0.423	커뮤니티아카이브	0.019	구술기록	0.13
10	대통령기록관	4.663	아카이브	0.417	내러티브	0.019	전자기록	0.116
11	국회기록보존소	4.579	메타데이터	0.414	전자기록물	0.018	대통령기록관	0.107
12	전자기록물	4.57	기능평가	0.414	로컬리티기록화	0.017	기록물관리전문요원	0.1
13	아카이브	4.206	전자기록	0.409	아카이브	0.016	기록관	0.097
14	로컬리티기록화	3.95	대통령기록관	0.404	rms	0.016	아카이브	0.087
15	정보공개법	3.914	국제표준	0.403	평가	0.016	rms	0.08

5.1.4 제3기 (2015~2019) 국가기록관리 혁신

등장빈도 2회 이상인 저자키워드를 대상으로 시각화한 결과, 312개의 노드와 1,406개의 에지로 구성된 네트워크가 생성되었다. 네트워크의 모든 노드를 Louvain method에 기반한 Modularity Class로 15개로 군집화하고 각 커뮤니티에 서로 다른 색상을 부여한 뒤, 이 중에서 연결도(degree)가 11 이상인 관계만을 표현한 것이 <그림 7>이다. 각 군집 내 가장 높은 매개중심성을 갖는 노드를 군집의 규모 순으로 각각 나열하면 '국가기록원', '아카이브', '디지털아카이브', '기록정보서비스', '메타데이터', '기록관리', '기록화전략', '기록학', '정보공개', '기억', '기록평가', '기록관리기준표', '증거적가치', '문화재청', '국립공문서관'이다.

표준화 매개 중심성, 표준화 근접 중심성, 표준화 연결 중심성, 아이겐벡터 중심성을 기준으로 15위 순위를 정리한 것이 <표 5>이다. 중심성 측정 기준과 상관없이 10위 내 위치한 저자키워드는 그 내용에 큰 차이가 없었다. 표준화 매개 중심성, 표준화 근접 중심성, 표준화 연결 중심성과 아이겐벡터 중심성 네 가지 모두 ‘기록관리’, ‘아카이브’, ‘기록정보서비스’, ‘국가기록원’, ‘rms’, ‘기록관’, ‘대통령기록관’이 상위 10위 내에 등장하였다. 표준화 매개 중심성에만 등장한 ‘사회적기억’, ‘기억’, ‘정보공개’는 당대 새로운 통제점 역할을 하는 연구주제의 등장을 의미한다. 표준화 근접 중심성에서만 당대 ‘장기보존’이 국내 기록분야에 영향력을 미친 주제 중 하나라는 점에서 주목할 필요가 있다. 특히 ‘rms’는 이전 시기에 비해 표준화 연결 중심성을 비롯한 다른 중심성 기준에서도 높은 순위를 보였다. 이외에도 아이겐벡터 중심성을 제외한 3개 중심성 순위에 ‘디지털아카이브’가 처음 등장하였다. 아이겐벡터 중심성에만 등장한 저자키워드로는 ‘sns’, ‘클라우드컴퓨팅’, ‘교육프로그램’이 있다.

제3기의 저자키워드를 분석한 결과를 정리하면, 1기와 2기에서 연구의 중심에 위치하던 ‘기록관리’가 더 이상 독점적인 위상을 지니지 못하고 ‘기록정보서비스’와 ‘아카이브’의 중요도가 함께 상승하였으며 ‘국가기록원’의 중요도는 감소한 것을 확인할 수 있다. 이는 국내 국가기록원의 위상이 낮아지거나 영향력이 낮아졌다고 보기보다는,



<그림 7> 주요 저자키워드 네트워크 - 제3기 (2015~2019)

<표 5> 중심성 기준 주요 저자키워드 - 제3기 (2015~2019)

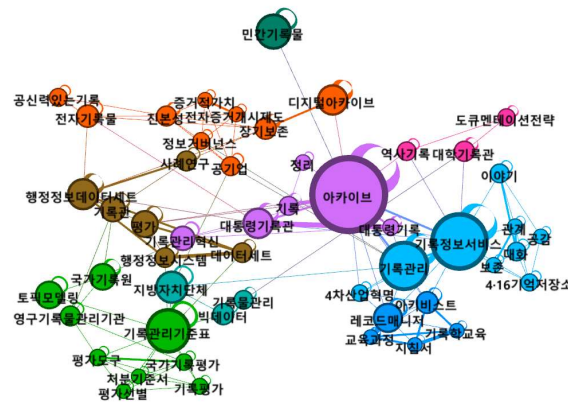
순위	표준화 매개 중심성	수치	표준화 근접 중심성	수치	표준화 연결 중심성	수치	아이겐벡터 중심성	수치
1	기록관리	18.241	기록관리	0.437	국가기록원	0.025	국가기록원	0.628
2	아카이브	13.529	아카이브	0.426	기록정보서비스	0.023	기록정보서비스	0.467
3	기록정보서비스	13.15	기록정보서비스	0.415	아카이브	0.023	아카이브	0.284
4	국가기록원	11.573	국가기록원	0.411	rms	0.022	대통령기록관	0.227
5	rms	8.446	rms	0.404	기록관리	0.02	rms	0.208
6	기록관	6.65	기록관	0.385	기록관	0.017	기록관리	0.177
7	메타데이터	6.226	대통령기록관	0.383	대통령기록관	0.016	기록관	0.172
8	대통령기록관	5.891	기록물관리전문요원	0.378	메타데이터	0.016	기록물관리전문요원	0.149
9	디지털아카이브	5.203	디지털아카이브	0.373	기록	0.016	대통령기록물	0.143
10	사회적기억	5.158	기록	0.372	기록물관리전문요원	0.014	기록	0.086
11	기록물관리전문요원	4.923	전자기록관리	0.365	기록관리기준표	0.013	레코드매니저	0.062
12	기억	4.724	메타데이터	0.362	디지털아카이브	0.012	sns	0.061
13	전자기록관리	4.678	기억	0.362	대통령기록물	0.011	클라우드컴퓨팅	0.057
14	공공기관	4.317	공공기관	0.36	기록학	0.011	전자기록	0.057
15	정보공개	4.033	장기보존	0.357	전자기록관리	0.011	교육프로그램	0.057

저자키워드의 영역이 방대해지고 연구 중심성이 국가기록원을 벗어나기 시작했다는 데 의의가 있다. 예를 들어 ‘아카이브’의 경우 전통적 기록물 이외에도 소셜 미디어 영역까지 기록관리 대상으로 확장하기에 이르렀다.

5.1.5 제4기 (2020~2023) 국가기록관리 중장기 발전계획

등장빈도 2회 이상인 저자키워드를 대상으로 시각화한 결과, 145개의 노드와 467개의 엣지로 구성된 네트워크가 생성되었다. 네트워크의 모든 노드를 Louvain method에 기반한 Modularity Class로 18개로 군집화하고 각 커뮤니티에 서로 다른 색상을 부여한 뒤, 이 중에서 연결도(degree)가 7 이상인 관계만을 표현한 것이 <그림 8>이다. 각 군집 내 가장 높은 매개중심성을 갖는 노드를 군집의 규모 순으로 각각 나열하면 ‘아카이브’, ‘기록관리기준표’, ‘디지털아카이브’, ‘기록정보서비스’, ‘민간기록물’, ‘행정정보데이터세트’, ‘레코드매니저’, ‘대학기록관’, ‘지방자치단체’, ‘[디지털기록관’, ‘가상박물관’, ‘만월대], [‘영상기록물’, ‘5·18], [‘광주학생독립운동’, ‘디지털인물아카이브’, ‘기록관리현황’, ‘기록전시’, ‘지방아카이브’, ‘기록보존’, ‘유네스코세계기록유산’, ‘연구데이터’이다.

표준화 매개 중심성, 표준화 근접 중심성, 표준화 연결 중심성, 아이겐벡터 중심성을 기준으로 15위 순위를 정리한 것이 <표 6>이다. 모든 중심성에서 1위부터 3위는 ‘기록관리’, ‘아카이브’, ‘기록정보서비스’였으나 그 외 순위에서는 저자키워드 분포 형태가 이전에 비해 다양한 형태를 보인다. 표준화 매개 중심성 기준으로는 ‘민간기록물’이 5위, ‘지방자치단체’가 6위, ‘기록관리혁신’이 15위로 기록분야 내 새로운 통제점 역할을 하는 연구주제가 등장한



<그림 8> 주요 저자키워드 네트워크 - 제4기 (2020~2023)

<표 6> 중심성 기준 주요 저자키워드 - 제4기 (2020~2023)

순위	표준화 매개 중심성	수치	표준화 근접 중심성	수치	표준화 연결 중심성	수치	아이겐벡터 중심성	수치
1	아카이브	28.157	아카이브	0.314	아카이브	0.061	아카이브	0.787
2	기록정보서비스	19.093	기록관리	0.297	기록정보서비스	0.045	기록정보서비스	0.319
3	기록관리	15.481	기록정보서비스	0.292	행정정보데이터세트	0.033	기록관리	0.235
4	기록관리기준표	13.328	대통령기록관	0.283	레코드매니저	0.031	대통령기록관	0.201
5	민간기록물	9.228	대통령기록	0.277	기록관리	0.03	대통령기록	0.196
6	지방자치단체	7.91	평가	0.277	교육과정	0.028	기록	0.161
7	행정정보데이터세트	7.451	레코드매니저	0.274	대통령기록관	0.024	참여정부	0.13
8	대통령기록관	7.182	아키비스트	0.273	디지털아카이브	0.023	라이프사이클	0.102
9	디지털아카이브	7.111	행정정보데이터세트	0.271	대통령기록	0.021	레코드컨텐츠업	0.102
10	레코드매니저	6.563	4차산업혁명	0.271	기록관리기준표	0.021	레코드매니저	0.097
11	토픽모델링	5.879	디지털아카이브	0.27	진본성	0.021	아키비스트	0.085
12	평가	5.544	기록관리혁신	0.27	기록	0.021	디지털아카이브	0.075
13	국가기록원	5.438	지방자치단체	0.264	데이터세트	0.021	4차산업혁명	0.072
14	기록관	4.651	민간기록물	0.26	장기보존	0.019	역사기록	0.068
15	기록관리혁신	4.452	디지털화	0.26	평가	0.019	평가	0.067

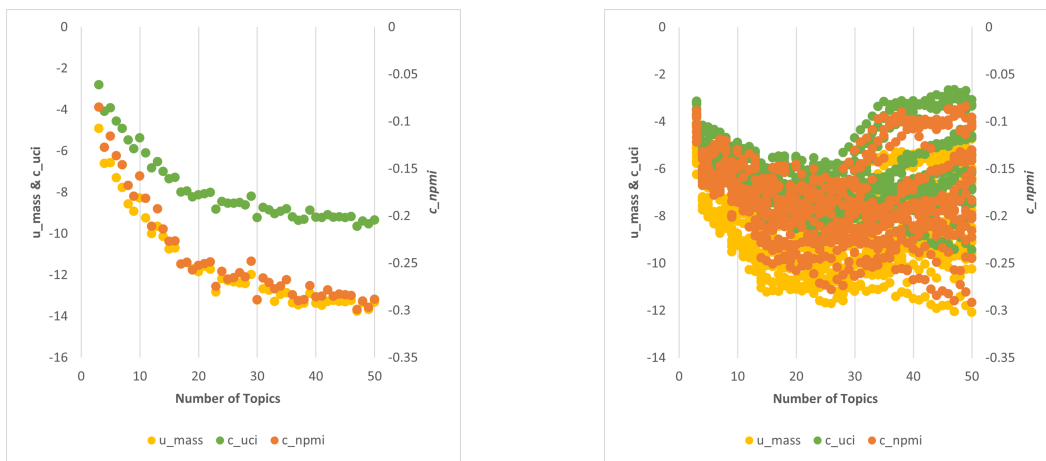
것을 시사한다. 표준화 연결 중심성 기준으로는 ‘행정정보데이터세트’가 3위, ‘레코드매니저’가 4위, ‘교육과정’이 6위 등 레코드매니지먼트와 기록물관리전문요원 양성 및 재교육에 관련된 저자키워드가 동시 등장하는 학술논문이 활발하게 게재되었음을 의미한다.

제4기의 저자키워드를 분석한 결과를 정리하면, 과거와 달리 ‘아카이브’가 독점적인 중심성 지위를 획득하였으며 ‘기록관리’는 ‘기록정보서비스’를 중심으로 한 레코드매니지먼트 영역으로 통합 확장되었고, 나머지는 ‘평가’와 ‘행정정보데이터세트’ 및 ‘디지털아카이브’ 등으로 연구주제가 재편되었다. 또한 중심 네트워크와 연결되지 않은 위성 네트워크가 다수 확인⁸⁾됨에 따라 기존 기록분야 연구주제와 접점을 갖지 않는 새로운 연구주제가 등장한 것을 짐작할 수 있다. 저자키워드 네트워크의 크기는 3기에 비해 줄어든 상태이며, 1~3기와는 다른 저자키워드 네트워크 형태를 나타내고 있다. 요컨대, 디지털 환경에서의 아카이빙과 레코드매니지먼트를 전제로 이전보다 다양한 이해당사자가 반영된 연구 주제로 재편되고 있다고 판단된다.

5.2 토픽모델링

5.2.1 토픽 개수(k) 결정

최적의 토픽모델링 결과를 얻기 위해서는 Coherence score를 기준으로 토픽 개수(k) 및 하이퍼파라미터(α, β)를 결정해야 한다. <그림 9>는 토픽 개수(k) 3~50개 범위에서 각각의 토픽 개수(k)에 대한 하이퍼파라미터(α, β) 변위에 따른 u_mass , c_uci , c_npmi 값을 시각화한 것이다. 이 중에서 (가) 실험 1은 각 토픽 개수(k)에 대해 Gensim API가 예측한 corpus 사전확률분포를 하이퍼파라미터(α, β)로 반영한 계산 결과이며, (나) 실험 2는 토픽 개수(k)에 대해 하이퍼파라미터(α, β)가 각각 [0.1, 0.5, 0.9, 1]일 때의 16개 조합을 계산한 결과이다. (가) 실험 1과 (나) 실험 2 모두 토픽 개수(k)가 3일 때 최댓값을 보였다. (가) 실험 1은 지속해서 값이 감소하는 추이를 보이거나, (나) 실험 2는 토픽 개수(k)가 28일 때 종속변수 상승부는 증가 추세로 전환된 뒤 35 이후 작은 범위에서 진동하는 추이를 보이지만 전체적으로는 산개하는 형태를 나타내었다. 이를 종합하면, 기록분야 학술논문의 초록에 대한 최적의 토픽 개수(k)는 3으로 결정하는 것이 타당하며 이후 토픽모델링 분석은 (가) 실험 1의 결과를 기원을 두었다.



(가) 실험 1

(나) 실험 2

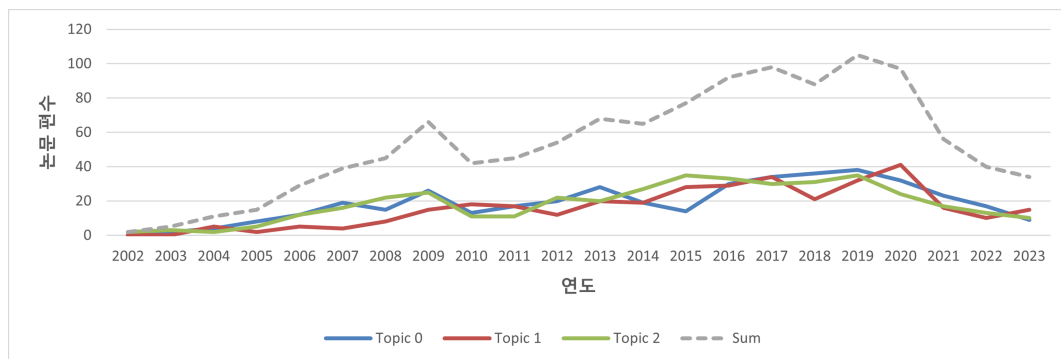
<그림 9> 토픽 개수(k) 결정을 위한 Coherence Score 계산 결과

8) <그림 8>의 외곽에 존재하는 위성 네트워크. 인쇄 지면상의 한계로 그림에는 포함하지 못함.

5.2.2 토픽모델링 분석 결과

LDA에 의해 추출된 각 토픽은, 그 토픽을 구성하는 토픽 용어(topic words)의 군집으로 대표된다. <표 7>은 본 연구에서 추출된 3개의 토픽과 이를 구성하는 토픽 용어를 확률순위로 정리한 것이다. 첫 번째 토픽인 Topic 0의 10대 주요 토픽 용어는 '기록관리', '기록물', '메타데이터', '전자기록', '생산', '문서', '기록물관리', '데이터', '정보', '법률'이다. 이어서 두 번째 토픽인 Topic 1의 10대 주요 토픽 용어는 '아카이브', '수집', '기억', '역사', '아카이빙', '공동체', '정보', '기록학', '기록물', '전시'이다. 마지막으로 세 번째 토픽인 Topic 2의 10대 주요 토픽 용어는 '기록물', '기록관', '기록관리', '평가', '이용자', '국가기록원', '대통령기록관', '기록정보서비스', '미국', '기록관리학'이다.

각 토픽에 대한 명칭은, 수집된 토픽 용어를 근거로 Topic 0은 '레코드매니지먼트', Topic 1은 '아카이빙', Topic 2는 '국가기록정책'으로 결정하였다. LDA는 확률모델이므로 분석 대상이 어떤 토픽에 속할지 지배적인 확률수치로 판정할 수 있는데, 이를 우세 토픽(dominant topic)이라 한다. 학술논문을 우세 토픽 현황을 연도별로 정리하면 <그림 10>과 같이 시계열 토픽 동향을 판별할 수 있다. 2009년까지는 Topic 0 '레코드매니지먼트'와 Topic 2 '국가기록정책'이 다소 우세인 상황으로 증가추세를 보이다가 2010년 이후부터는 3개의 토픽이 서로 앞서거나 뒤처지거나 하는 모습을 보인다. 2015년에는 Topic 1 '아카이빙'과 Topic 2 '국가기록정책'이 다소 우세를 보이다가, 2018년에는 Topic 0 '레코드매니지먼트'와 Topic 2 '국가기록정책'이 강세를 보이고, 2020년에는 Topic 1 '아카이빙'만이 강세를 보였다. 거시적 관점에서 3개의 토픽은 큰 치우침 없이 균형 잡힌 양적 성장을 꾸준히 이어 나가고 있었으나, 2020년 이후 감소세로 접어든 양상을 나타낸다.



<그림 10> 연도별 학술논문 우세토픽 동향

<표 7> 토픽별 주제 분석

토픽 순위	Topic 0 레코드매니지먼트		Topic 1 아카이빙		Topic 2 국가기록정책	
	토픽 용어	확률	토픽 용어	확률	토픽 용어	확률
1	기록관리	0.092	아카이브	0.144	기록물	0.065
2	기록물	0.043	수집	0.035	기록관	0.061
3	메타데이터	0.027	기억	0.032	기록관리	0.049
4	전자기록	0.025	역사	0.029	평가	0.040
5	생산	0.024	아카이빙	0.028	이용자	0.033
6	문서	0.020	공동체	0.027	국가기록원	0.030
7	기록물관리	0.019	정보	0.019	대통령기록관	0.018
8	데이터	0.016	기록학	0.017	기록정보서비스	0.017
9	정보	0.015	기록물	0.017	미국	0.015
10	법률	0.014	전시	0.016	기록관리학	0.015

6. 결론

본 연구는 현대 기록관리학 연구주제 현황을 파악하기 위해 2002년부터 2023년 8월까지 약 22년간 KCI에 등재된 기록분야 1,173편의 논문(기록학 876편, 문헌정보학 247편, 역사학 50편)의 서지정보를 수집하고, 저자키워드와 초록 데이터로부터 각각 네트워크 분석과 토픽모델링을 시행한 뒤, 최종 결과를 시계열로 해석하였다. 첫째, 약 22년간 전체 저자키워드 네트워크는 '기록관리', '공동체아카이브', '아카이브', '기록학', '국가기록원', '메타데이터', '기록평가', '정보공개', '기술'로 각각 대표되는 9개 군집으로 구분되었으며 중심성 측정법에 상관없이 가장 높은 중심성을 갖는 것은 '기록관리'였다.

둘째, 시기별 네트워크의 특징을 정리하면 다음과 같다. 1기(2002~2006년 공공기록물법 제정 이후)는 '기록관리', '아카이브', '국가기록원', '기록', '기록물'을 중심으로 '교과과정', '연구', '실습', '기록학'과 같이 기록관리 전문가 양성 교육에 관련된 연구가 활발했음을 보여준다. 2기(2007~2014년 공공기록물법 개정 이후)는 국제표준에 부합하는 기록관리 연구와 그 적용 방안을 본격적으로 심사숙고하기 본격화된 시기로 사료되었다. 3기(2015~2019년 국가기록관리 혁신)는 국가기록원을 필두로 한 기록관리, 정보공개에 대한 논의, 디지털 기록 보존, 국민의 기록 접근권에 대한 논의가 확대된 시기로 인식된다. 4기(2020~2023년 국가기록관리 중장기 발전계획)는 지방자치단체 수준에서의 기록관리와 이를 위한 기록관리기준표 개선, 행정정보데이터세트를 고려한 기록관리, 4차산업혁명을 고려한 전자기록관리 및 기록정보서비스, 디지털아카이브 환경에서의 레코드컨티뉴엄 연구처럼 기록관리학 영역이 한층 더 확장되고 있는 중요한 시기로서 빅데이터에 기반한 철저한 기록화, 이용자 편의를 고려한 기록정보서비스 확대, 디지털 기반 영구기록관리체계 확립에 대한 논의가 심화되고 있다.

셋째, 토픽모델링에서의 토픽 개수(k)는 u_mass , c_uci , c_npmi 수치에 따라 3으로 결정하였다. 각 토픽을 대표하는 토픽 용어 집합을 검토한 결과, 3개의 토픽은 각각 '레코드매니지먼트', '아카이빙', '국가기록정책'을 상징하는 것으로 확인되었다. 1,173편의 논문의 우세 토픽을 결정한 결과를 시계열 분석한 결과, 2002~2009년 동안은 '레코드매니지먼트'와 '국가기록정책'이 '아카이빙'에 비해 다소 우세인 상황이었으나 2009년 이후에는 3개 토픽이 함께 어우러지며 양적 증가하는 추이를 나타내었다. 그러나 2019년을 정점으로, 2020년 이후부터 최근까지 모든 토픽이 급격한 하강 추이를 보인다.

이상의 연구 결과를 종합하면 다음과 같다. 첫째, 국내 기록분야 연구 주제는 크게 공공기록물법과 국가기록정책의 영향을 받는 경향을 보인다. 대표적으로 1999년 법률제정 이후 기록물관리전문요원 양성과 관련된 연구주제가 갑자기 주를 이루었던 사례나, 2020년 국가기록원의 행정정보데이터세트 사업추진 일정과 맞물려 이에 관한 연구가 급증한 것이 한 예이다. 2000년대 중반 이후, 공동체 아카이브와 예술 아카이브처럼 공공기관 외의 영역인 민간 영역에서 자발적으로 싹이 튼 사례는 매우 고무적이라 할 수 있으나, 여전히 국가기록정책의 영향을 받는 공공기관의 기록물관리가 연구의 중심이라는 사실을 투영한 것으로 볼 수 있다.

둘째, 2020년 이후 기록분야 연구는 과거와는 다른 추이를 보인다. 기록분야 연구논문의 게재 건수는 2019년을 정점으로 정점을 기록한 뒤 다소 감소 추이를 보이는 한편, 저자키워드 네트워크가 크게 재편되는 경향을 보이는 점에 주목할 필요가 있다. 다만 토픽모델링을 통해 아카이빙, 레코드매니지먼트, 국가기록정책으로 구분된 3대 연구 주제(토픽) 내 변화는 큰 차이가 없었다. 이는 기록분야 지식 실천공동체(Community of Practice)의 세대교체, 기록관리학 세부 영역의 전문화와 확장, 디지털 환경의 정착 등이 반영된 결과로 볼 수 있다.

본 연구의 한계를 바탕으로 한 후속 연구 제언은 다음과 같다. 첫째, 연구자마다 서로 다른 경험과 지식이 다른 만큼 논문의 연구 범위나 연구 주제에 따라 같은 용어가 사용되더라도 실제로 다른 의미인 경우가 있었다. 따라서 연구자의 의도를 고려한 분석을 위해서는, 기록관리학 학술용어 전자통제를 보완할 필요가 있다. 둘째, 본 연구의 기록분야 논문 출처를 기록관리학, 문헌정보학, 역사학으로 한정하였으나 행정학도 기록관리학과 관련된 학제이며(이윤정, 정연경, 2021), 법학과 의약학을 비롯한 다양한 분야에서 기록관리의 중요성이 강조되고 있다(이인영,

2011; 전영주, 2013). 따라서 추가 연구를 통해 한층 더 광범위한 데이터로부터 학계 내 기록분야 연구 추이를 분석해야 할 것이다. 셋째, 본 논문에서 확인한 기록관리학 영역의 확장과 재편과 관련하여 최신 저자 네트워크를 분석함으로써 연구 기여도와 연구자 영향력을 분석해야 할 필요성을 확인하였다.

본 연구는 기록관리학을 중심으로 문헌정보학과 역사학 학술논문을 망라하였으며, 저자키워드뿐만 아니라 초록에 대해서 네트워크 분석 및 토픽모델링을 시행함으로써, 지난 22년간 기록분야 전문가와 연구자가 쌓아 올린 연구 업적을 분석하고 숨겨진 연구주제를 밝혀냈다는 데 의의가 있다. 더불어 시대적으로 요구되고 있는, 기록관리 전문가의 계속교육과 대학원에서의 교과과정 개선에 반영될 수 있다는 점에서도 그 의미를 더할 수 있을 것이다.

참고문헌

- 국가기록관리 혁신 T/F (2017). 국가기록관리 혁신 방안.
- 국가기록원 (2006). 「국가기록관리혁신 로드맵」추진과 각급기관의 역할.
- 국가기록원 (2020). 신뢰받는 기록관리를 위한 국가기록관리 중장기 발전계획(안) - 2020~2024 -.
- 김규환, 남영준 (2009). 국내 기록관리학 분야 학회지 논문 분석을 통한 연구동향 연구. 한국문헌정보학회지, 43(4), 217-239. <https://doi.org/10.4275/KSLIS.2009.43.4.217>
- 김선옥, 양기덕 (2022). LDA와 BERTopic을 이용한 토픽모델링의 증강과 확장 기법 연구. 정보관리학회지, 39(3), 99-132. <https://doi.org/10.3743/KOSIM.2022.39.3.099>
- 김성희, 장로사 (2010). 사회 연결망 분석 연구동향 및 정보학 분야에서의 활용가능성에 관한 연구. 정보관리학회지, 27(4), 71-87. <https://doi.org/10.3743/KOSIM.2010.27.4.071>
- 김순희 (2009). 기록관리 전문가의 실무능력 제고를 위한 교육 프로그램 연구. 한국도서관 정보학회지, 40(3), 271-293. <https://doi.org/10.16981/kliss.40.3.200909.271>
- 김학래 (2022). 기록관리 분야에서 한국어 자연어 처리 기술을 적용하기 위한 고려사항. 한국기록관리학회지, 22(4), 129-149. <https://doi.org/10.14404/JKSARM.2022.22.4.129>
- 김희섭, 강보라 (2018). 『한국기록관리학회지』 논문의 연구 주제 분석 - 2001년부터 2017년까지 -. 한국기록관리학회지, 18(4), 183-204. <https://doi.org/10.14404/JKSARM.2018.18.4.183>
- 나연휘, 박희진 (2022). 국내 기록보존 연구동향 분석: 2000~2021년 학술논문을 중심으로. 한국기록관리학회지, 22(4), 175-196. <https://doi.org/10.14404/JKSARM.2022.22.4.175>
- 남태우, 이진영 (2009). 우리나라 기록관리학 연구 동향 분석. 한국도서관 정보학회지, 40(2), 451-472. <https://doi.org/10.16981/kliss.40.2.200906.451>
- 박자현, 송민 (2013). 토픽모델링을 활용한 국내 문헌정보학 연구동향 분석. 정보관리학회지, 30(1), 7-32. <https://doi.org/10.3743/KOSIM.2013.30.1.007>
- 박준형, 류법모, 오효정 (2018). 시계열 기반 국내 기록관리학 토픽 트렌드 분석. 한국기록관리학회지, 18(1), 29-47. <https://doi.org/10.14404/JKSARM.2018.18.1.029>
- 박준형, 오효정 (2017). 국내 기록관리학 연구동향 분석을 위한 토픽모델링 기법 비교 - LDA와 HDP를 중심으로 -. 한국도서관 정보학회지, 48(4), 235-258. <https://doi.org/10.16981/kliss.48.4.201712.235>
- 박희진 (2014). Archivaria 리뷰를 통한 국외 기록관리학 연구의 동향 분석. 한국기록관리학회지, 14(4), 99-125. <https://doi.org/10.14404/JKSARM.2014.14.4.099>
- 손혜인, 남영준 (2016). 기록관리학 분야 국내 학술지의 연구동향에 관한 연구: 「한국기록관리학회지」와 「기록학연구」를 중심으로. 정보관리학회지, 33(1), 85-110. <https://doi.org/10.3743/KOSIM.2016.33.1.085>

- 송정숙 (2017). 국가기록원과 권역별 기록관의 역할 재정립. 한국기록관리학회지, 17(4), 155-175.
<https://doi.org/10.14404/JKSARM.2017.17.4.155>
- 심성보 (2020). 2017~2020년 기록관리 혁신 평가: 국가기록원을 중심으로. 기록학연구, 65, 7-46.
<https://doi.org/10.20923/kjas.2020.65.007>
- 오향녕 (2017). 역사학과 기록학: 학문의 인연, 학제의 괴리. 기록학연구, 54, 179-210.
<https://doi.org/10.20923/kjas.2017.54.179>
- 이동주, 연종흠, 황인범, 이상구 (2010). 꼬꼬마: 관계형 데이터베이스를 활용한 세종 말뭉치 활용 도구. 정보과학회논문지: 컴퓨팅의 실제 및 레터, 16(11), 1046-1050.
- 이수상 (2013). 연구자 네트워크의 중심성과 연구성과의 연관성 분석 : 국내 기록관리학 분야 학술논문을 중심으로. 한국도서관정보학회지, 44(3), 405-428. <https://doi.org/10.16981/kliss.44.3.201309.405>
- 이수상 (2014). 언어 네트워크 분석 방법을 활용한 학술논문의 내용분석. 정보관리학회지, 31(4), 49-68.
<https://doi.org/10.3743/KOSIM.2014.31.4.049>
- 이용구, 우윤희 (2015). 대만 문헌정보학 분야의 공저 네트워크 분석 연구. 정보관리학회지, 32(2), 167-192.
<https://doi.org/10.3743/KOSIM.2015.32.2.167>
- 이윤정, 정연경 (2021). 국내의 기록관리학 대학원 교육과정에 관한 비교 연구. 한국문헌정보학회지, 55(1), 567-591.
<https://doi.org/10.4275/KSLIS.2021.55.1.567>
- 이인영 (2011). 전자의무기록에 관한 법적 문제. 법학논총, 28(1), 75-98.
- 이재윤, 문주영, 김희정 (2007). 텍스트 마이닝을 이용한 국내 기록관리학 분야 지적구조 분석. 한국문헌정보학회지, 41(1), 345-372. <https://doi.org/10.4275/KSLIS.2007.41.1.345>
- 이혜경, 양기덕, 김선옥 (2019). 국내 문헌정보학 분야의 공동연구 동향 분석. 한국도서관·정보학회지, 50(2), 191-214.
<https://doi.org/10.16981/kliss.50.2.201906.191>
- 장혜리, 최재황 (2008). 레코드매니저의 계속교육 프로그램 개발. 한국기록관리학회지, 8(1), 123-142.
<https://doi.org/10.14404/JKSARM.2008.8.1.123>
- 전영주 (2013). 전자의무기록(EMR)의 활용과 환자정보보호. 보건의료산업학회지, 7(3), 213-224.
<https://doi.org/10.12811/kshsm.2013.7.3.213>
- 정연경 (2005). 한국과 미국의 대학원 기록관리학 교과과정에 관한 비교 연구. 한국기록관리학회지, 5(2), 5-25.
<https://doi.org/10.14404/JKSARM.2005.5.2.005>
- 최민우, 강범모 (2008). 세종 형태 (의미) 분석 말뭉치 구축의 실제. 민족문화연구, 48, 337-372.
- 최이량 (2015). 국내 기록관리학 연구동향에 관한 연구 -최근 10년간(2004-2013) 학술논문을 중심으로. 기록학연구, 43, 147-177.
- 한미경, 노영희 (2015). 문헌정보학과의 기록관리 교과목과 문헌정보학 교수의 기록관리학 논문 출판 동향 분석. 한국기록관리학회지, 15(4), 125-149. <http://doi.org/10.14404/JKSARM.2015.15.4.125>
- Aouichaty, S., Maleh, Y., Hajami, A. & Allali, H. (2023). Automated decision classification model for tax appeals commission in Morocco using latent dirichlet allocation. Indonesian Journal of Electrical Engineering and Computer Science, 31(3), 1811-1820.
<https://doi.org/10.11591/ijeecs.v31.i3.pp1811-1820>
- Beauchamp, M. A. (1965). An improved index of centrality. Behavioral Science, 10(2), 161-163.
<https://doi.org/10.1002/bs.3830100205>
- Benoit, E. & Force, D. C. (2019). One Size Does Not Fit All: Graduate Archival Education in the Twenty-First Century. The American Archivist, 82(1), 24-52. <https://doi.org/10.17723/0360-9081-82.1.24>
- Blei, D. M., Ng, A. Y. & Jordan, M. I. (2003). Latent dirichlet allocation. Journal of machine Learning research, 3(Jan), 993-1022.
- Bonacich, P. (2007). Some unique properties of eigenvector centrality. Social networks, 29(4), 555-564.

<https://doi.org/10.1016/j.socnet.2007.04.002>

- Brandes, U. (2001). A faster algorithm for betweenness centrality. *The Journal of Mathematical Sociology*, 25(2), 163-177. <https://doi.org/10.1080/0022250X.2001.9990249>
- Cao, L. & Ren, J. (2022). Machine learning shows that the Covid-19 pandemic is impacting US public companies unequally by changing risk structures. *Plos one*, 17(6), <https://doi.org/10.1371/journal.pone.0269582>
- Cox, R. J. (2009). Unpleasant things: teaching advocacy in archival education programs. *InterActions: UCLA Journal of Education and Information Studies*, 5(1), <https://doi.org/10.5070/D451000643>
- Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, 1(3), 215-239 [https://doi.org/10.1016/0378-8733\(78\)90021-7](https://doi.org/10.1016/0378-8733(78)90021-7)
- Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv. Available: <https://doi.org/10.48550/arXiv.2203.05794>
- Sabidussi, G. (1966). The centrality index of a graph. *Psychometrika*, 31(4), 581-603. <https://doi.org/10.1007/BF02289527>

• 국문 참고자료의 영어 표기

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

- Choi, Minwoo & Kang, Beom-mo (2008). Practice in Constructing Sejong Morph (Sense) Analysis Corpora. *Korean Cultural Studies*, 48, 337-372.
- 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.
- Chung, Yeon Kyoung (2005). A Comparative Study of Graduate Level Archival Curriculum in Korea and the United States of America. *Journal of Korean Society of Archives and Records Management*, 5(2), 5-25. <https://doi.org/10.14404/JKSARM.2005.5.2.005>
- Han, Mi-kyung & Noh, Younghee (2015). An Analysis of Records Management Courses offered at the Department of Library and Information Science and Publication Trend of Records Management Articles by Library and Information Science Professors. *Journal of Korean Society of Archives and Records Management*, 15(4), 125-149. <https://doi.org/10.14404/JKSARM.2015.15.4.125>
- Jang, Hye-Li & Choi, Jae Hwang (2008). A Study on the Continuing Education Program for Records Managers. *Journal of Korean Society of Archives and Records Management*, 8(1), 123-142. <https://doi.org/10.14404/JKSARM.2008.8.1.123>
- Jeun, Young-Ju (2013). EMR System and Patient Medical Information Protection. *The Korean Journal of Health Service Management*, 7(3), 213-224. <https://doi.org/10.12811/kshsm.2013.7.3.213>
- Kim, Gyuhan & 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. <https://doi.org/10.4275/KSLIS.2009.43.4.217>
- Kim, Haklae (2022). Considerations for Applying Korean Natural Language Processing Technology in Records Management. *Journal of Korean Society of Archives and Records Management*, 22(4), 129-149. <https://doi.org/10.14404/JKSARM.2022.22.4.129>
- Kim, Heesop & Kang, Bora (2018). An Analysis of the Research Topics of the Academic Papers Published in the Journal of Korean Society of Archives and Records Management: From 2001 to 2017. *Journal of Korean Society of Archives and Records Management*, 18(4), 183-204.

- <https://doi.org/10.14404/JKSARM.2018.18.4.183>
- Kim, Seonghee & Chang, Ro-Sa (2010). The Study on the Research Trend of Social Network Analysis and the its Applicability to Information Science. *Journal of the Korean Society for Information Management*, 27(4), 71-87. <https://doi.org/10.3743/KOSIM.2010.27.4.071>
- Kim, SeonWook & Yang, Kiduk (2022). Topic Model Augmentation and Extension Method using LDA and BERTopic. *Journal of the Korean Society for Information Management*, 39(3), 99-132. <https://doi.org/10.3743/KOSIM.2022.39.3.099>
- Kim, Soon-Hee (2009). A Study on a Teaching Program to Improve the Working Ability for Records Manager. *Journal of Korean Library and Information Science Society*, 40(3), 271-293. <https://doi.org/10.16981/kliss.40.3.200909.271>
- Lee, Dongjoo, Yeon, Jongheum, Hwang, Inbeom & Lee, Sang-goo (2010). KKMA : A Tool for Utilizing Sejong Corpus based on Relational Database. *KIISE Transactions on Computing Practices*, 16(11), 1046-1050.
- Lee, Hyekyung, Yang, Kiduk & Kim, SeonWook (2019). Analysis of Collaborative Research Trends in Library and Information Science in Korea. *Journal of Korean Library and Information Science Society*, 50(2), 191-214. <https://doi.org/10.16981/kliss.50.2.201906.191>
- Lee, In-Young (2011). An Analysis of the Major Issues on Electronic Medical Records. *Hanyang Law Review*, 28(1), 75-98.
- 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. <https://doi.org/10.4275/KSLIS.2007.41.1.345>
- Lee, Soo-Ssang (2013). Analytical Study on the Relationship between Centralities of Research Networks and Research Performances. *Journal of Korean Library and Information Science Society*, 44(3), 405-428. <https://doi.org/10.16981/kliss.44.3.201309.405>
- Lee, Soo-Ssang (2014). A Content Analysis of Journal Articles Using the Language Network Analysis Methods. *Journal of the Korean Society for Information Management*, 31(4), 49-68. <https://doi.org/10.3743/KOSIM.2014.31.4.049>
- Lee, Yong-Gu & Woo, Yunhee (2015). A Study on Analyzing Co-authorship Networks of Library and Information Science in Taiwan. *Journal of the Korean Society for Information Management*, 32(2), 167-192. <https://doi.org/10.3743/KOSIM.2015.32.2.167>
- Lee, Yun-Jung & Chung, Yeon Kyoung (2021). A Comparative Study on the Curriculum of Graduate Schools of Archival Sciences in Korea and the Foreign Countries. *Journal of the Korean Society for Library and Information Science*, 55(1), 567-591. <https://doi.org/10.4275/KSLIS.2021.55.1.567>
- Na, Yonwhae & Park, Hee Jin (2022). A Study on the Research Trends of Archival Preservation Papers in Korea from 2000 to 2021. *Journal of Korean Society of Archives and Records Management*, 22(4), 175-196. <https://doi.org/10.14404/JKSARM.2022.22.4.175>
- Nam, Tae-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. <https://doi.org/10.16981/kliss.40.2.200906.451>
- National Archives of Korea (2006). Implementation of the “National Records Management Innovation Roadmap” and the roles of institutions at each level.
- National Archives of Korea (2020). National records management mid- to long-term development plan (draft) for trustworthy record management - 2020~2024 -.
- National Records Management Innovation T/F (2017). Innovation plan for national records management.

- Oh, Hang-Nyeong (2017). History and Archives: Colleagues or Strangers?. *The Korean Journal of Archival Studies*, 54, 179-210. <https://doi.org/10.20923/kjas.2017.54.179>
- Park, Hee Jin (2014). A Study on the Research Trends of Records and Archive Management in North America through the Review of Archivaria. *Journal of Korean Society of Archives and Records Management*, 14(4), 99-125. <https://doi.org/10.14404/JKSARM.2014.14.4.099>
- Park, Jahyun & Song, Min (2013). A Study on the Research Trends in Library & Information Science in Korea using Topic Modeling. *Journal of the Korean Society for Information Management*, 30(1), 7-32. <https://doi.org/10.3743/KOSIM.2013.30.1.007>
- Park, JunHyeong & Oh, Hyo-Jung (2017). Comparison of Topic 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. <https://doi.org/10.16981/kliss.48.4.201712.235>
- Park, JunHyeong, Ryu, Pum-Mo & Oh, Hyo-Jung (2018). Timeline-Based Topic Trend Analysis of Archives Management in Korea. *Journal of Korean Society of Archives and Records Management*, 18(1), 29-47. <https://doi.org/10.14404/JKSARM.2018.18.1.029>
- Shim, SungBo (2020). Evaluation of Records and Archives Management Innovation in 2017~2020. *The Korean Journal of Archival Studies*, 65, 7-46. <https://doi.org/10.20923/kjas.2020.65.007>
- 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 the Korean Society for Information Management*, 33(1), 85-110. <https://doi.org/10.3743/KOSIM.2016.33.1.085>
- Song, Jung-Sook (2017). Reestablishing the Roles of the National Archives of Korea and the Branch Archives. *Journal of Korean Society of Archives and Records Management*, 17(4), 155-175. <https://doi.org/10.14404/JKSARM.2017.17.4.155>