

텍스트마이닝을 활용한 공공기록물 관련 뉴스 기사 분석*

An Analysis of News Articles Related to Public Records Using Text Mining*

한승희(Seunghee Han)¹, 이해원(Hyewon Lee)²

E-mail: hanshee@swu.ac.kr, hwlee@swu.ac.kr



1 제1저자 서울여자대학교 사회과학대학 문헌정보학과 교수
2 교신저자 서울여자대학교 사회과학대학 문헌정보학과 교수

논문접수 2025.01.31
최초심사 2025.02.03
게재확정 2025.02.18

ORCID

Seunghee Han
https://orcid.org/0000-0002-4748-4968

Hyewon Lee
https://orcid.org/0000-0002-3571-9491

© 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.

■ 이 논문은 서울여자대학교 학술연구비의 지원에 의한 것임(2024-0166)

초 록

본 연구는 2000~2024년 사이에 발행된 공공기록물 관련 뉴스 기사 5,050건을 대상으로 텍스트마이닝 기법을 활용하여 공공기록물과 사회·정치적 이슈 간 연관성을 밝히고자 하였다. 빈도분석 결과, 뉴스 기사는 공공기록물 자체보다 사건·사고 및 정치적 이슈를 중심으로 보도하며, 가장 많이 언급된 기록물 유형은 회의록이었다. 바이그램 네트워크 분석에서는 ‘대화록’, ‘회의록’, ‘기록물’, ‘공공기록물법’이 핵심 키워드로 나타났으며, 네트워크 군집분석 결과 ‘공공기록물 관리’, ‘남북정상회담 회의록’, ‘공공기록물법 관련 논란’의 세 가지 주요 주제가 도출되었다. 감성분석 결과, 긍정 기사에서는 기록관리 우수기관 사례를, 부정 기사에서는 공공기록물법 위반 및 기록관리 미흡 문제를 주로 다루고 있었음을 확인하였다. 마지막으로 공공기록물 관련 연구 동향과 뉴스 기사 분석결과를 비교하였고 이를 바탕으로 공공기록물 관련 사회적 문제를 해결하고 개선하기 위한 정책적 제언을 도출하였다.

ABSTRACT

This study examines the relationship between public records and sociopolitical issues by applying text mining techniques to 5,050 news articles related to public records published from 2000 to 2024. The frequency analysis results indicate that news articles primarily focus on incidents, accidents, and political issues rather than public records, with meeting records being the most frequently mentioned record type. In the Bigram network analysis, the main keywords identified were “dialogue records,” “meeting minutes,” “records,” and “Public Records Management Act.” The network clustering analysis revealed three major themes: (1) public records management, (2) inter-Korean summit meeting records, and (3) controversies surrounding the Public Records Management Act. The sentiment analysis results indicate that positive news articles mainly highlight cases of exemplary record management institutions. Conversely, negative articles focused on Public Records Management Act violations and records management deficiencies. Finally, the study compared research trends with findings from the news article analysis and provided policy recommendations to address and improve social issues related to public records.

Keywords: 공공기록물, 뉴스 빅데이터 분석, 텍스트마이닝, 네트워크 군집분석, 감성분석
Public Records, News Big Data Analysis, Text Mining, Network Cluster
Analysis, Sentiment Analysis

1. 서론

『공공기록물 관리에 관한 법률』(이하, 공공기록물법)에서 공공기록물은 공공기관이 업무 수행 과정에서 생산하거나 접수한 모든 기록물을 의미한다. 이는 행정의 투명성을 확보하고, 기관의 설명책임성을 강화하며, 국민의 알권리를 보장하기 위한 중요한 법적 근거를 제공한다. 또한, 공공기록은 국가의 역사적 기록으로서 미래 세대를 위한 지식 자원으로 활용되며, 민주주의 사회에서 사회혁신의 강력한 수단이자 동시에 사회문화 발전에 빼놓을 수 없는 중요한 도구라고 할 수 있다(한국기록관리학회, 2024).

그러나 공공기록의 미생산, 무단 공개, 무단 유출, 무단 파괴 등 공공기록물과 관련된 사회적 논란이 주기적으로 뉴스에 등장하면서, 국민들은 국가 기록관리 체계의 완전성과 신뢰성에 대한 의구심을 품을 수밖에 없게 되었다. BigKinds에서 최근 10년간 생산된 공공기록물 관리에 대한 뉴스 기사를 검색해 보면 공공기록물 관리 부실을 지적하는 기사가 큰 비중을 차지하고 있음을 알 수 있다. 그런데 이러한 기사가 특정 시기와 특정 사건에만 집중해서 출현한 것이 아니라 해마다 여러 사건과 관련되어 지속적으로 출현하고 있다는 것은 생각해 볼 문제이다. 이는 단순히 특정 사건에 대한 논란을 넘어, 공공기록물이 법과 제도를 통해 일관되게 관리되고 있는지에 대한 근본적인 문제를 제기하는 계기가 된다. 특히, 정치적 이해관계에 따라 기록의 생산·보존·공개 과정이 달라지는 사례가 반복되면서, 국민들은 공공기록물이 투명하게 관리되고 있는지에 대해 신뢰하기 어려운 상황이 되었다(김태경, 2025; 전진한, 2024). 『공공기록물법』이 제정된 지 25년이 다 되어감에도 불구하고 여전히 법령 위반을 주요 소재로 한 뉴스가 정치적 상황에 따라 반복적으로 보도되는 일은 매우 안타까운 일이 아닐 수 없다.

이러한 문제의 근본 원인 중 하나는 공공기록물의 법적 정의가 모호하다는 점에 있다(이경남, 2023). 앞에서 언급한 바와 같이 『공공기록물법』에서는 공공기록물을 공공기관이 업무 수행 과정에서 생산하거나 접수한 모든 기록물로 정의하고 있으나, 특정한 기록물의 공개 여부나 관리 기준에 대한 해석이 일관되지 않거나 명확하지 않은 경우가 많다. 이러한 모호성은 법 적용의 차이를 초래하고, 결과적으로 특정 기록물이 정치적·행정적 이유로 은폐되거나 임의로 관리되는 사례를 발생시킬 가능성을 높인다. 따라서 기록관리 체계의 완전성을 확보하기 위해서는 공공기록물의 정의와 관리 기준을 명확히 하고, 법적·제도적 개선을 지속적으로 추진할 필요가 있다.

이러한 배경 속에서 공공기록물과 관련된 사회·정치적 논의가 어떻게 형성되고 있는지를 분석하는 연구가 필요하다. 공공기록물 관리와 관련된 사회적 논란은 단편적인 사건이 아니라, 시간이 지남에 따라 반복적으로 등장하는 구조적인 문제일 가능성이 크다. 따라서 특정 사건만을 개별적으로 분석하기보다는, 관련 뉴스 기사를 종합적으로 분석함으로써 거시적 관점에서 공공기록물 관련 이슈를 이해할 필요가 있다. 특히, 공공기록물에 관련 연구 동향을 살펴보면, 공공기록물 관련 기존 연구들은 주로 공공기록물의 법적·제도적 측면이나 기록관리 체계 개선 방안에 집중되어 있었으며, 공공기록물이 사회적 논의에서 어떻게 다뤄지고 있는지를 포괄적으로 다룬 연구는 부족하다.

뉴스 기사는 단순한 정보 전달의 역할을 넘어, 공공기록물의 중요성에 대한 대중의 인식을 형성하는 데 영향을 미칠 수 있다. 따라서 뉴스 매체가 공공기록물 관련 사건을 어떤 시각에서 보도하고 있는지를 분석하면, 공공기록물에 대한 사회적 이해도를 높이는 데 중요한 시사점을 도출할 수 있다.

이러한 관점에서 본 연구의 목적은 2000년부터 2024년까지의 공공기록물 관련 뉴스 기사를 분석하여 공공기록물이 사회·정치적 이슈와 어떤 연관성을 가지는지를 규명하는 것이다. 이를 위해 본 연구는 텍스트마이닝 기법을 활용하여 뉴스 기사에서 다루지는 공공기록물 관련 주요 주제를 분석하고, 공공기록물이 사회·정치적 이슈 속에서 어떻게 활용되거나 논쟁의 대상이 되고 있는지를 살펴보고자 한다. 그리고 이러한 분석을 통해 기록관리학과 실무 현장에서 고려해야 할 시사점을 도출하고자 한다.

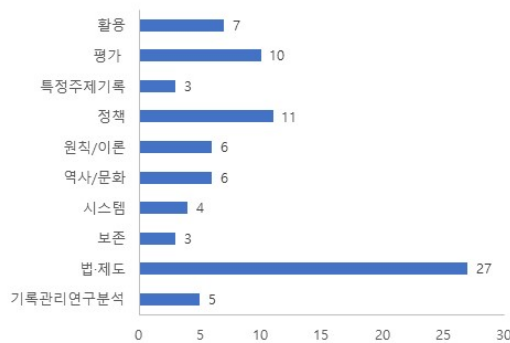
2. 이론적 배경

2.1 공공기록물 관련 연구 동향

본 장에서는 공공기록물 관련 선행 연구를 분석하였는데, 이를 위해 KISTI ScienceOn 상세검색 기능을 이용(2025년 1월 10일 기준)하였다. 국내 논문으로 범위를 지정하고 논문명에 ‘공공기록물’이 포함되어 있는 논문을 검색한 결과 46건의 논문이 제시되었다. 그러나 본 연구에서는 기록관리 분야 외의 다른 연구 분야에서 발견되는 논문을 포함하고 공공기록물에 대한 검색 범위를 확장하기 위해 ‘공공기록’이라는 키워드를 사용하였다. 제목에 ‘공공기록’이 포함된 국내 논문의 검색 결과는 108건이었고, 그 중 중복을 제거하고 본 연구에서는 82건의 논문을 분석 대상으로 삼았다([부록] 참고).

대상 논문들은 범주, 출처의 연구 분야, 출판연도 등으로 구분하여 분석하였다. 범주 분석은 공공기록이라는 주제와 관련된 세부 영역을 확인하고 공공기록의 연구 범위를 개괄적으로 정리한 것이었다. 범주 분석은 기록관리 분야 전문가 2인이 대상 논문의 제목 및 초록을 확인한 후 범주를 지정하고 상호 확인 절차를 통해 개념 범주를 설정하였다.

<그림 1>에서 제시된 바와 같이 ‘법·제도’ 범주가 27건으로 가장 많았고, 다음은 ‘정책(11건)’이었다. 그리고 ‘평가(10건)’, ‘활용(7건)’, ‘원칙/이론(6건)’, ‘역사/문화(6건)’, ‘기록관리연구분석(5건)’, ‘시스템(4건)’, ‘특정주제기록(3건)’, ‘보존(3건)’ 순이었다. [부록]을 살펴보면 ‘법·제도’는 논문명에 법이나 제도라는 용어가 포함되어 있는 경우가 대부분이어서 법률 제정 및 개정, 제도화 등의 내용이 주 연구 주제였다. ‘정책’ 범주에 있는 논문은 법이나 제도를 수행하거나 실행하는 부분에 초점이 맞추어져 있었다. ‘평가’에 속한 논문들은 ‘법·제도’ 및 ‘정책’ 범주에도 포함될 수 있지만 ‘평가’라는 주제만으로도 논문의 수가 상당하여 범주를 따로 설정하였다. 기록관리에 대한 개념, 원칙, 원리, 범위에 해당하는 논문들은 ‘원칙/이론’에 분류하였으며, 역사, 문화, 정치적인 관점에서 접근하는 것은 ‘역사/문화’로 포함하였다. 기록관리 분야의 선행 연구들을 분석하는 논문들은 ‘기록관리연구분석’의 범주에 할당하였으며, 소프트웨어나 시스템과 관련된 논문은 ‘시스템’으로 추가하였다. 공공극장, 유아교육, 탄핵 등과 같은 특정 주제를 다루는 논문은 ‘특정주제기록’으로 분류하였으며, 보존 방법 및 장비 등에 관한 연구는 ‘보존’에 연결하였다.



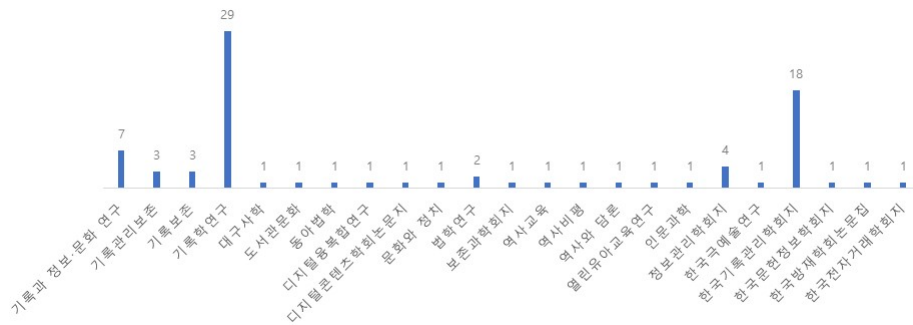
<그림 1> 공공기록물 관련 연구: 범주

10개의 범주 중 ‘법·제도’ 및 ‘정책’에 속한 논문의 비중이 높은 것은 이경남(2023)이 제시한 공공기록의 범위에도 부합하였다. 이경남(2023, 114)은 이론적으로 공공기록물의 생산과 등록의 의미를 검토하여 공공기록물에 대한 개념 범위를 명확하게 정의해야 하고 공공기록물로서 효력을 갖는 시점을 법제화할 필요가 있음을 강조하였다.

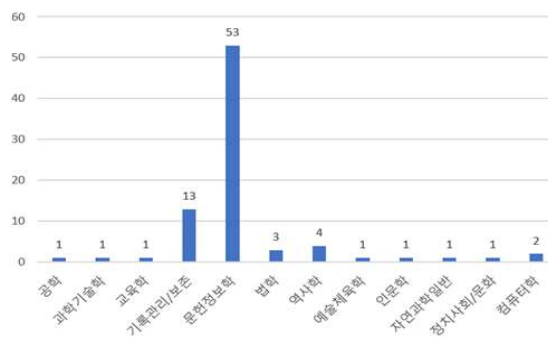
출처의 연구 분야 분석에서는 공공기록을 다루는 학문 분야를 파악하고자 하였다. 먼저 <표 1>과 <그림 2> 등과 같이 공공기록 관련 연구의 출처를 살펴보았다. 학술지 단위로 출처의 순위를 매기면 ‘기록학연구(29건)’, ‘한국기록관리학회지(18건)’, ‘기록과 정보·문화 연구(7건)’ 순이었다. 논문 출처인 저널/프로시딩명을 확인한 후 KCI 학술지 검색을 통해 연구 분야를 조사하였다. 분석 대상 82건 중 한국연구재단 KCI 학술지 검색을 통해 확인할 수 없었던 저널 및 프로시딩은 ‘記錄管理保存(기록관리보존; 한국기록관리협회)’, ‘記錄保存(기록보존; 국가기록원)’, ‘圖書館文化(도서관문화; 한국도서관협회)’ 등으로 총 3건이었다. 3건에 대해서는 저널 및 프로시딩의 표제에 포함된 주제어를 기준으로 학문 분야를 할당하였다. ‘記錄管理保存’과 ‘記錄保存’은 기록관리/보존으로 ‘圖書館文化’는 문헌정보학으로 분류하였다(<표 1> 참고). KCI 연구 분야 활용 시, 계층 구조를 고려하여 제일 하위 계층의 분류명을 활용하였으며 만약 하위 계층이 ‘기타’인 경우에는 바로 위 단계 분류명을 사용하였다.

<표 1> 공공기록 관련 연구: 출처의 연구 분야

저널/프로시딩명	한국연구재단 KCI 연구 분야	출처의 연구 분야	건수
기록과 정보·문화 연구	복합학 > 문헌정보학 > 기록관리/보존	기록관리/보존	7
記錄管理保存*		기록관리/보존	3
記錄保存*		기록관리/보존	3
기록학연구	복합학 > 문헌정보학	문헌정보학	29
대구사학	인문학 > 역사학	역사학	1
圖書館文化*		문헌정보학	1
東亞法學	사회과학 > 법학	법학	1
디지털융복합연구	복합학 > 과학기술학 > 기타과학기술학	과학기술학	1
디지털콘텐츠학회 논문지	공학 > 컴퓨터학	컴퓨터학	1
문화와 정치	사회과학 > 정치외교학 > 비교정치 > 정치사회/문화	정치사회/문화	1
법학연구	사회과학 > 법학 > 기타법학	법학	2
보존과학회지	자연과학 > 자연과학일반	자연과학일반	1
역사교육	인문학 > 역사학	역사학	1
역사비평	인문학 > 역사학	역사학	1
역사와 담론	인문학 > 역사학	역사학	1
열린유아교육연구	사회과학 > 교육학	교육학	1
인문과학	인문학 > 기타인문학	인문학	1
정보관리학회지	복합학 > 문헌정보학	문헌정보학	3
한국극예술연구	예술체육학 > 기타예술체육	기타예술체육학	1
한국기록관리학회지	복합학 > 문헌정보학 > 기록관리/보존	문헌정보학	18
한국문헌정보학회지	복합학 > 문헌정보학	문헌정보학	1
한국방재학회논문집	공학 > 기타공학	공학	1
한국전자거래학회지	공학 > 컴퓨터학	컴퓨터학	1
한국정보관리학회 학술대회 논문집	복합학 > 문헌정보학	문헌정보학	1
	합계		82



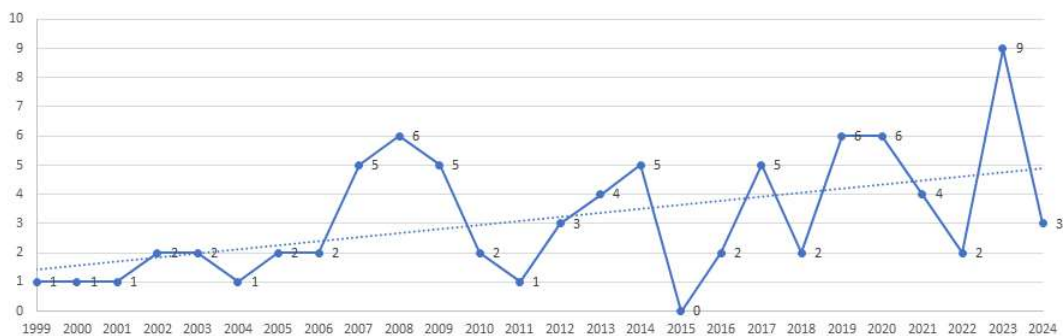
<그림 2> 공공기록 관련 연구: 출처



<그림 3> 공공기록 관련 연구: 출처의 연구 분야

<그림 3>을 살펴보면 ‘문헌정보학’이 53건으로 가장 높았고, ‘기록관리/보존(13건)’이 그다음이었다. ‘문헌정보학’이 가장 높은 이유는 ‘기록학연구’의 영향이었다. <표 1>에 제시된 바와 같이 ‘기록학연구’의 KCI 연구 분야가 ‘복합학 > 문헌정보학’로 지정되어 있어 <그림 3>의 ‘문헌정보학’ 연구 분야가 가장 높은 것으로 나타났다.

출판연도 분석은 공공기록에 대한 연구를 시계열로 정리한 것으로 <그림 4>와 같다. 건수가 많은 순으로 2023년에는 9건이, 2008년, 2019년, 2020년에는 6건이 출판되었다. 점선으로 표시된 추세선을 확인하면 지속적으로 공공기록에 대한 출판이 증가하고 있음을 알 수 있다.



<그림 4> 공공기록 관련 연구: 출판연도

본 장에서 제시한 공공기록 관련 선행 연구 분석은 3장의 공공기록물 관련 뉴스 데이터 분석 결과와 비교해서

살펴볼 필요가 있겠다. 즉 공공기록에 대한 기록관리학계의 연구 방향과 사회 전반에서 관심을 받고 있는 주제 영역에는 어떤 차이가 있는지를 확인하고자 하였다.

2.2 텍스트마이닝을 활용한 뉴스 기사 분석 선행 연구

정보기술의 발전으로 텍스트 빅데이터의 수집과 분석이 수월해지면서 텍스트마이닝 기법을 활용한 뉴스 빅데이터 분석 연구가 국내외 다양한 분야에서 이루어지고 있다. 특히 특정 사회현상이나 사건에 대한 뉴스 기사를 대상으로 텍스트마이닝 기법을 적용하여 언론에서 이들을 어떤 시각으로 바라보고 있는지를 분석하는 연구들이 증가하고 있다.

Salloum et al.(2017)은 중동 걸프 지역 24개 신문사에서 포스팅한 62,327건의 페이스북 신문 기사를 대상으로 빈도분석과 클러스터링 기법을 적용하여 뉴스 기사에서 다룬 주요 이슈와 언론사 군집을 분석하는 연구를 실시하였다. 그 결과 언론사 그룹 별로 주로 다루는 기사의 주제가 다르다는 점을 밝혀냈다. Krawczyk et al.(2021)은 코로나-19 관련 온라인 뉴스 빅데이터를 대상으로 양적 분석 및 주요 하위 주제 식별, 감성분석 등을 진행하였다. Febro-Naga와 Tinam-isan(2022)은 필리핀에서 2015~2020년 사이에 발행된 3,506건의 여성 및 여아폭력 관련 뉴스 기사를 대상으로 빈도분석과 토픽모델링을 실시하여 뉴스를 구성하는 주요 개념과 하위 주제를 식별하였다.

이러한 연구들은 뉴스 기사에 텍스트마이닝 기법을 적용함으로써 특정 주제 뉴스에 담긴 사회적 현상이나 담론의 구조를 분석할 수 있다는 것을 전제로 한다(윤호영, 2020). 텍스트마이닝 기법을 뉴스 기사에 적용하는 연구들은 일반적으로 단어의 빈도분석을 통해 전체 기사에 출현한 주요 단어를 식별함으로써 뉴스 기사의 대략적인 주제를 파악하고, 이후 텍스트마이닝 기법을 적용하여 전체 기사를 구성하는 하위 주제를 식별하는 형식을 취한다. 이러한 관점에서 볼 때 텍스트마이닝 기법은 내용분석 방법론의 하나로 간주할 수 있다. 이 연구에서는 최근 5년간 발표된 뉴스 기사 대상 텍스트마이닝 연구 중 특정 사회적 문제나 이슈를 다룬 연구를 검토하여 본 연구의 방법론적 타당성을 검증하고자 한다.

텍스트마이닝 기법을 활용하여 특정 주제 뉴스 빅데이터를 분석한 최근 5년간(2020~2024)의 국내 논문을 대상으로 분석 주제, 뉴스 수집 기간, 분석 뉴스 건수, 분석 방법 등을 조사하였다. <표 2>에서 보는 바와 같이 다양한 사회 문제(기후변화, 아동학대, 혐오, 미세먼지, 정보 프라이버시, 노인자살 등)와 이슈(남북관계, 인공지능, 국내 이주노동자, 외국인 유학생, 메타버스, 5·18민주화운동 등)에 대해 연구가 이루어진 것을 알 수 있다. 절반 이상의 연구가 10년 이상의 긴 기간을 대상으로 뉴스 기사를 수집하였으나, 뉴스 기사 수는 차이가 큰 것으로 나타났다. 텍스트마이닝에는 다양한 기법이 존재하나, 뉴스 빅데이터 분석에서는 주로 빈도분석, 언어 네트워크 분석, 토픽 모델링 등이 주로 많이 적용되고 있으며, 그 밖에 네트워크 군집분석과 N-gram 분석, 감성분석을 적용한 연구도 있는 것을 확인할 수 있다.

<표 2> 텍스트마이닝을 활용한 특정 주제 뉴스 빅데이터 분석 국내 선행 연구(2020~2024)

저자(출판년도)	분석 주제	뉴스 수집기간	분석된 뉴스 기사 수	분석방법
윤호영(2020)	남북관계	2000~2019	435,649	빈도분석 토픽모델링 언어네트워크분석
조현국, 이봉우(2021)	기후변화	2001~2020	7,516	빈도분석 토픽모델링
오세현, 강현아(2021)	아동학대	2014~2019	18,058	빈도분석 토픽모델링
김혜진(2022)	혐오	1990~2020	17,867	빈도분석 네트워크 군집분석
박종화 외(2022)	인공지능	2016~2021	11,116(한국) 6,179(미국)	빈도분석 토픽모델링 언어네트워크분석
양지연(2022)	미세먼지	1995~2021	63,185	빈도분석 감성분석
최한별 외(2022)	정보 프라이버시	1990~2021	8,865	빈도분석 토픽모델링 언어네트워크분석
엄연용 외(2023)	성역할	1990~2021	8,049	빈도분석 언어네트워크분석 토픽모델링
박주연, 정도현(2022)	메타버스	????~2022	41,366	빈도분석 토픽모델링
박주현 외(2024)	5·18민주화운동	1990~2022	48,138	빈도분석 언어네트워크분석
이정은(2024)	노인 자살	2011~2020	7,079	빈도분석 N-gram 분석 언어네트워크분석

이 외에 기록관리학 분야에서 뉴스 기사를 대상으로 텍스트마이닝 기법을 활용한 연구는 이소현, 구본진(2022), 정상준 외(2019), 한승희(2018)의 연구가 있다. 이 연구들은 기록관리학 분야의 관심 주제인 ‘잊힐 권리’, ‘대통령기록물’, ‘기록관리’에 대한 뉴스 기사들을 대상으로 빈도분석, 언어 네트워크 분석, 토픽 모델링을 적용하여 언론에 나타난 기록관리 관련 사회적 현상과 담론의 구조적 특성을 분석하고자 하였다.

이 연구에서는 내용분석의 접근법의 한 방법으로 활용될 수 있는 텍스트마이닝 기법을 활용하여 공공기록물 관련 뉴스 기사의 내용적 특성을 분석하고자 하였다. 아직 기록관리학 분야에서 뉴스 기사 분석을 통해 공공기록물에 대한 내용분석을 시도한 연구는 없다.

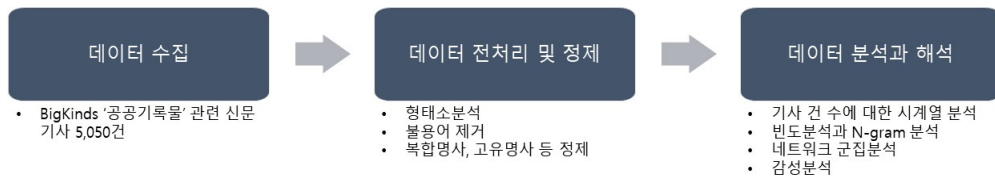
3. 연구 방법

3.1 데이터 수집 및 전처리

이 연구에서 진행된 전체적인 데이터 분석 절차는 <그림 5>와 같다. 먼저 한국언론진흥재단의 BigKinds(<https://www.bigkinds.or.kr/>)에서 기사 제목이나 본문에 ‘공공기록물’이라는 단어가 들어간 뉴스 기사

전체를 수집하였다. 수집 기간은 2000년 1월 1일부터 2024년 12월 31일까지로 하였는데, 『공공기록물법』이 2000년 1월 1일부터 시행되면서 이때부터 본격적으로 공공기록물과 관련된 기사가 생산되기 시작했다. 해당 기간 기사 제목이나 본문에 ‘공공기록물’이라는 단어가 들어간 기사의 총수는 5,465건이었다. 이 중 중복 수집된 기사를 제외한 최종 5,050건 기사를 대상으로 데이터 전처리와 정제 후 텍스트마이닝 기법을 활용하여 데이터를 분석하였다.

이 연구에서는 빅데이터 분석 플랫폼 텍스트톰(<https://www.textom.co.kr/>)을 이용하여 데이터 전처리와 분석을 실행하였다. 먼저 수집된 데이터를 대상으로 형태소 분석을 하였는데, 이때 텍스트 자체 형태소 분석기인 MeCab-IMC를 활용하여 명사만 추출하였다. 형태소 분석 후, 불용어(한 글자 단어(‘일’, ‘월’, ‘년’, ‘차’ 등), 비주제어(‘이번’, ‘관련’ 등))를 제거하고 마지막으로 색인 오류와 띄어쓰기 오류, 단어의 이형 표기(‘공정위’, ‘공정거래위’ 등) 등으로 인해 발생한 복합명사와 고유명사 등을 정제하는 작업을 진행하였다. 단어 정제 작업의 효율성을 고려하는 동시에 기사 자체에서 선호하는 단어나 표현 방식을 살려서 분석하기 위해 뉴스 기사에 표현된 단어 그대로 활용하는 것을 단어 정제의 기본적인 방침으로 삼았다. 예를 들어, 학계에서는 ‘기록물’이라는 표현은 공공기록물법 외에서는 잘 쓰지 않는다. 그러나 뉴스 기사에서는 ‘기록’이라는 표현보다 ‘기록물’이라는 표현을 훨씬 더 많이 쓴다. 통제어회집을 구축하고 이형표기를 하나로 통일하여 단어를 정제하면 주제적 분석의 유의미성을 더 크게 얻을 수 있겠으나, 이 연구에서는 뉴스에서 표현된 어휘 자체를 그대로 활용하는 것이 더 적합하다고 판단하였고, 또한 데이터의 규모가 적은 편이 아니기 때문에 굳이 하나의 대표 어휘로 통일하여 분석하는 것은 지양하였다. 다만 뉴스기사 특성상 동일 기관명이나 고유명사를 다양한 약어로 표현하는 경우가 많아서 기관명과 고유명사에 대해서는 정식명칭으로 정제하였다. 또한 고빈도어 중 기록관리와 밀접한 관련이 없는 단어(예: 기사)는 제거하였다.



<그림 5> 데이터 수집 및 분석 과정과 절차

3.2 데이터 분석 방법

이 연구에서는 전처리된 데이터를 대상으로 시계열 분석, 빈도분석, N-gram 분석과 함께 뉴스 기사 빅데이터 분석에서 많이 활용되고 있는 네트워크 군집분석과 감성분석을 실시함으로써 공공기록물 관련 뉴스 기사의 주제적 특성을 발견하고자 하였다. 시계열 분석을 통해서서는 연도별 기사의 발행 건수를 살펴봄으로써 공공기록물 관련 주제에 대한 언론의 관심 추세를 확인할 수 있다.

(1) 빈도분석과 N-gram 분석

공공기록물 관련 뉴스 기사에서 어떤 내용들이 주로 다루어졌는지를 개괄적으로 분석하기 위해 빈도분석을 실시하였다. 또한 기사에서 공공기록물과 밀접하게 다룬 관련 이슈와 주제를 분석하기 위해 N-gram 분석을 실시하였다. N-gram 분석은 문자열에서 N개의 연속된 요소를 추출하는 통계 기반 언어 분석모델이다. 일반적으로 특정 단어의 앞뒤로 연속해서 출현한 단어들 사이에는 밀접한 관련이 있으며, N-gram 빈도가 높을수록 이 단어들

사이에는 관련성이 높다고 할 수 있다(Jurafsky & Martin, 2009). 특히 특정 단어의 앞 또는 뒤에 연속으로 출현하는 바이그램(bi-gram) 단어 간의 관계를 네트워크 형태로 시각화하여 분석하는 바이그램 네트워크 분석(Bi-gram Network Analysis)은 텍스트 네트워크 분석법 중 하나로, 관련성이 있는 단어 간의 관계 파악에 유용하게 활용될 수 있다(정우경, 신동희, 2021). 이 연구에서는 뉴스 기사에 출현한 공공기록물 관련 단어들의 앞 또는 뒤에 연속으로 출현한 바이그램 단어 간의 관계를 네트워크로 시각화하여 분석함으로써 공공기록물 관련 뉴스 기사에서 주로 다루어 온 주요 이슈와 주제를 확인하고자 하였다.

(2) 네트워크 군집분석

커뮤니티 탐지(community detection) 기법(Girvan & Newman, 2002)이라고도 하는 네트워크 군집분석(network cluster analysis)은 복잡한 네트워크 구조를 이해하는 데 사용되는 분석법 중 하나로, 네트워크 분석을 기반으로 커뮤니티 구조를 찾아 연관 노드들을 하위 군집으로 그룹화한다(이준석 외, 2018). 뉴스 기사에서 추출된 단어들을 대상으로 키워드 네트워크를 생성하고 모듈성(modularity)이 높은 키워드들을 연관 키워드 군집으로 그룹화하여 분석함으로써 분석 텍스트를 구성하는 여러 하위 주제를 탐지할 수 있다. 여기서 모듈성이 높다는 것은 커뮤니티에 속한 내부 노드 간에는 연결 강도가 높고 커뮤니티 외부에는 연결 강도가 낮은 것을 의미한다.

이 연구에서는 Leiden 알고리즘을 이용하여 뉴스 기사의 키워드 네트워크를 기반으로 한 네트워크 군집분석을 실시하였다. Leiden 알고리즘은 커뮤니티 탐지 알고리즘 중 가장 많이 알려진 Louvain 알고리즘의 단점인 커뮤니티 내 노드 간 연결이 끊어지는 문제를 해결하는 동시에 네트워크 연결 정확성과 속도를 모두 개선한 알고리즘이다(Traag et al., 2019). 이 연구에서는 공공기록물 관련 뉴스 기사에서 추출한 키워드들을 대상으로 네트워크를 생성하고 Leiden 알고리즘을 적용하여 하위의 연관 용어 군집을 추출함으로써 공공기록물 관련 전체 뉴스 기사를 구성하는 하위 주요 주제 그룹을 확인하고자 하였다.

(3) 감성분석

기계학습 기법을 이용한 텍스트의 감성 자동분류는 특정 이슈에 대한 뉴스 기사의 논조나 감정을 확인하는 데에 활용될 수 있다(감미아, 송민, 2012). 이 연구에서는 공공기록물 관련 뉴스 기사의 감성을 자동분류하여 분석하기 위해 GPT-4o-mini(beta) 모델을 활용하였다. 이 모델은 Open AI사에서 개발한 LLM(Large Language Model)으로, 방대한 양의 텍스트 데이터를 학습하여 자연어를 이해하고 생성할 수 있는 인공지능 모델이다(Open AI Platform, 2023). 이 모델은 원래 텍스트를 대상으로 생성·요약·분석·질의응답·번역 등 다양한 작업을 수행할 수 있도록 훈련된 범용 모델이나, 텍스트 분류, 감성분석, 주제분류 등과 같은 자동분류도 가능하며, 그 성능도 우수한 것으로 나타났다(김형도, 2023). 기존의 자동분류 알고리즘을 기반으로 한 뉴스 기사의 감성분류 연구에서는 뉴스 기사의 감성을 긍정/중립/부정으로 사전 분류한 학습집단을 구축하고 이를 바탕으로 검증집단을 대상으로 테스트를 한 후 감성분류를 실행해야 하는 여러 단계의 과정이 필요했다. 그러나 GPT-4o-mini(beta)를 활용하면 대규모 사전 학습을 통해 텍스트의 문맥을 깊이 있게 이해하고 이를 바탕으로 텍스트의 감성을 분석할 수 있기 때문에 편리하고 정확하게 감성분석을 할 수 있다는 다는 장점이 있다.

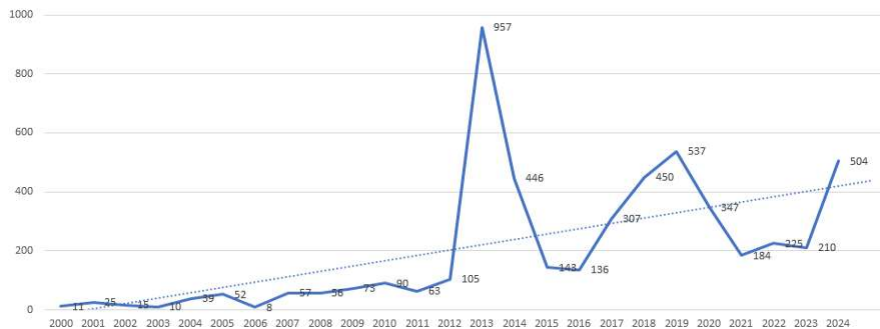
이 논문에서는 GPT-4o-mini(beta)를 활용하여 공공기록물 관련 뉴스 기사를 긍정/중립/부정으로 자동분류하고 그 결과를 분석함으로써 공공기록물 관련 기사들의 논조와 해당 이슈에 대한 감성 분포를 확인하고자 한다.

4. 연구 결과

4.1 빈도분석과 N-gram 분석

2000년부터 2024년까지의 공공기록물 관련 뉴스 기사 건수의 추이를 살펴보면 <그림 6>과 같다. 그림에서 점선으로 표시된 추세선으로 볼 때 공공기록물 관련 뉴스 기사는 시간이 지날수록 지속적으로 증가하는 추이를 보인다. 2013년(박근혜 정부, 957건), 2019년(문재인 정부, 537건), 2024년(윤석열 정부, 504건), 2018년(문재인 정부, 450건), 2014년(박근혜 정부, 446건) 순으로 많은 기사를 발행했는데, 시기별 특징을 자세히 살펴보면 다음과 같다.

먼저, 2013~2014년은 노무현 정부 때 있었던 남북정상회담 서해 북방한계선(NLL) 대화록 공개와 관련된 정치적 이슈가 전 사회에 관심을 불러일으켰던 시기이고, 2019년에는 문재인 전 대통령의 개별 대통령기록관 설립, 국가기록원 50주년, 서울기록원 개원, 신재민 전 기획재정부 사무관의 청와대 의혹에 대한 기밀 유출과 폭로사건 등 공공기록물과 관련된 다양한 이슈가 있었다. 2024년에는 월성원전 관련 문건 파기, 의대 증원 관련 회의록 논란과 비상계엄 관련 기록의 부실 관리 이슈가 있었으며, 2018년에는 박근혜 정부의 사법농단 재판과 관련하여 대법원 기밀문서의 외부 유출과 무단 파기, 이명박 정부 당시 4대강 문건의 무단 파기와 경성남도기록원 개원 등의 이슈가 있었다. 이를 통해, 뉴스 기사에서는 공공기록물을 정치·사회적 이슈와 매우 밀접한 관련이 있는 대상으로 다루고 있음을 알 수 있다. 이를 보다 구체적으로 확인하기 위해 뉴스 기사에 출현한 단어의 빈도분석을 실시하였다.



<그림 6> 공공기록물 관련 뉴스 기사의 연도별 건 수 추이

뉴스 기사에서 추출한 상위 50위 빈도의 단어와 빈도수를 나타내면 <표 3>과 같다. ‘검찰’, ‘수사’, ‘사건’, ‘고발’, ‘혐의’, ‘위반’ 등의 단어가 높은 빈도를 나타내고 있는 것을 볼 때 공공기록물 관련된 뉴스 기사들은 기록물 자체보다는 공공기록물과 관련된 사건이나 사고를 많이 다루어 왔음을 알 수 있다. 또한 정당명이나 ‘국회’, ‘야당’, ‘여당’ 등의 단어가 많이 출현한 것을 볼 때 공공기록물은 정치적 이슈와도 매우 밀접한 관련이 있음을 알 수 있다. 공공기록물 관련 기사에 가장 많이 언급된 대통령은 ‘노무현대통령’으로, 이는 남북정상회담 회의록 유출과 퇴임 후 봉하마을 이지원 시스템 반출 관련 대통령기록물법 위반 여부에 대한 이슈로 인한 것이라고 볼 수 있다.

이를 통해, 뉴스 기사에서는 공공기록물을 정치·사회적 이슈와 매우 밀접한 관련이 있는 대상으로 다루고 있음을 알 수 있다. 정치·사회적 이슈와 연관된 뉴스기사들의 유형과 내용을 분석해보면 다음과 같다. 먼저, 특정한 사회·정치적 이슈와 연관지어 공공기록물의 부실 관리를 지적하는 기사가 발행되거나(예: 4대강 국정감사에서 관련 기록물이 무단 폐기된 사건의 기사화 등), 특정 정치·사회적 이슈와 관련하여 공공기록물법이나 대통령기록물법의 위반 여부에 대한 논쟁이 기사화되는(예: 북방한계선 관련 남북정상회담 회의록 열람과 공개 등) 경우가 대부분인

것으로 확인되었다. 이러한 점은 공공기록물이나 관리의 문제를 직접 다루고 있다기보다는 특정 정치·사회적 이슈와 관련지어 공공기록물과 공공기록물법을 다루고 있다는 것으로 해석할 수 있다. 특정 이슈에 대한 관심을 불러일으키기 위해 공공기록물과 관련법이 지속적으로 언급되고 있다는 것은 특정 이슈에 관계없이 공공기록물 관리와 관련 법에 대한 문제가 지속적으로 존재하고 있음을 시사한다.

<표 3> 공공기록물 관련 전체 뉴스 기사에 출현한 상위 50위 빈도 단어 리스트

순위	단어	빈도	순위	단어	빈도	순위	단어	빈도	순위	단어	빈도	순위	단어	빈도
1	검찰	1,988	11	기록	1,206	21	협의	887	31	기록물 관리	714	41	발언	562
2	공개	1,660	12	서울	1,175	22	국회	874	32	공공 기록물	703	42	북방 한계선	528
3	의원	1,606	13	사건	1,140	23	공공 기록물법	820	33	논란	667	43	노무현 대통령	522
4	대화록	1,556	14	국가 기록원	1,054	24	행정	800	34	조사	632	44	회의	509
5	수사	1,498	15	정부	1,016	25	기관	789	35	사업	592	45	추진	501
6	기록물	1,492	16	자료	979	26	민주당	783	36	내용	586	46	사무관	495
7	국가 정보원	1,395	17	청와대	975	27	관리	756	37	기획재정부	580	47	결과	494
8	회의록	1,297	18	유출	957	28	새누리당	754	38	기록관리	571	48	위반	494
9	남북정상 회담	1,248	19	보존	904	29	폐기	744	39	국가	569	49	신재민	494
10	의혹	1,242	20	고발	895	30	문건	736	40	야당	567	50	기록관	493

또한 뉴스 기사에서 가장 많이 언급되고 있는 공공기록물의 유형은 ‘회의록’인 것으로 나타났다. 『공공기록물 관리에 관한 법률 시행령』 제18조에 따르면, 회의록은 공공기관의 모든 주요 회의에서 의무적으로 생산해야 하는 기록물이다. 그러므로 이에 대한 생산뿐만 아니라, 공개와 보존, 처분 등 여타의 기록관리 행위 역시 법령에 근거하여 이루어져야 한다. 그럼에도 불구하고 회의록이 제대로 생산되지 않거나 함부로 공개 또는 유출되는 등 회의록과 관련된 위법 행위를 다룬 기사들이 많다는 것은 현재 공공기관 회의록의 생산과 관리 체계 및 이를 규정한 법률의 구속력에 문제가 있다는 것을 의미한다. 기록관리 학계와 현장에서는 이 문제의 해결 방안을 함께 논의하고 제안함으로써 공공기록물로서의 회의록 생산과 관리 수준을 향상하고 더 나아가 정치·사회적 논란에 기록물이 악용되는 것을 막을 수 있어야 한다.

이어서, 공공기록물과 밀접하게 다루고 있는 관련 주제들을 확인하기 위해 뉴스 기사에 출현한 공공기록물 관련 단어들을 대상으로 바이그램 네트워크 분석을 실시하였다. 뉴스 기사 전체에서 가장 많은 바이그램 빈도를 나타낸 단어쌍은 ‘남북정상회담-대화록(563회)’, ‘남북정상회담-회의록(413회)’이었다. 단어쌍 분석 과정에서 공공기록물과 직접적인 관련이 없는 단어쌍(예: 중앙-지방검찰청(192회))을 제외한 후 바이그램 빈도 50 이상의 46개의 단어쌍을 네트워크 시각화하였다. 노드 간을 연결하는 링크가 굵을수록 두 노드의 바이그램 빈도가 높다는 것을 의미한다.

<그림 7>에서 보는 바와 같이, ‘대화록’, ‘회의록’, ‘기록물’, ‘공공기록물법’이 바이그램 네트워크에서 주요 노드의 역할을 하는 것을 알 수 있다. 주요 노드와 바이그램으로 연결된 연관 단어들을 분석한 결과(<표 4>), 공공기록물 관련 기사에서는 어떤 정치·사회적 이슈(예: 남북정상회담, 비상계엄, 의대 증원 등)와 관련된 기록물에 대한

제1군집은 공공기록물의 관리와 관련된 단어들로 주로 포함되어 있다. 공공기록물 관리와 관련된 기관(‘국가’, ‘감사원’, ‘기록관’, ‘국회’, ‘국가기록원’, ‘교육청’, ‘정부’, ‘청와대’)과 기록관리 행위(‘파기’, ‘폐기’, ‘선정’, ‘관리’, ‘보존’, ‘구축’, ‘생산’, ‘보관’, ‘교육’)와 관련된 단어들로 이 군집에 포함되어 있다.

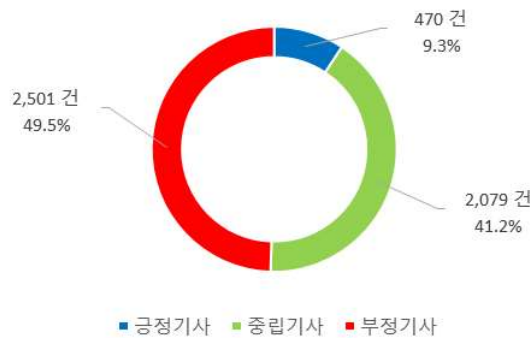
제2군집은 주로 남북정상회담 회의록과 관련된 단어들로 구성되어 있다. 남북정상회담 회의록 관련 주제(‘북방한계선’, ‘포기’, ‘발언’, ‘의혹’), 인물(‘김무성’, ‘노무현대통령’, ‘박근혜대통령’, ‘정문헌’), 기록 관련 행위(‘열람’, ‘유출’)와 정치적 행위(‘기소’, ‘개입’, ‘위반’, ‘고발’, ‘재판’) 관련 단어들로 이 군집에 포함되어 있다.

제3군집은 공공기록물법 관련 논란과 관련된 단어들로 구성되어 있다. 남북정상회담 회의록 외에 공공기록물법 위반 논란과 관련된 기관(‘기획재정부’, ‘검찰’, ‘민주당’, ‘의대’, ‘대법원’, ‘경찰’)과 행위(‘수사’, ‘작성’, ‘공개’) 관련 단어들로 이 군집에 속해 있다.

네트워크 군집분석 결과를 종합적으로 분석해 보면, 24년간 발행된 공공기록물 관련 뉴스는 크게 ‘공공기록물 관리’, ‘남북정상회담 회의록’, ‘공공기록물법 관련 논란’의 세 가지 하위 주제로 나누어 볼 수 있다. 이러한 결과는 앞에서 살펴본 빈도분석과 N-gram 분석의 결과를 뒷받침해 주는 것으로, 특정 사회정치적 이슈에 관련된 공공기록물에 대한 관리 행위와 관련된 문제와 이에 따른 공공기록물법 위반 여부에 대한 논란과 관련된 내용들이 뉴스 기사에서 주로 다루어져 왔음을 다시 한번 확인할 수 있다.

4.3 감성분석

공공기록물 관련 5,050개의 뉴스 기사를 대상으로 감성 자동 분류를 한 결과는 <그림 8>과 같다. 부정기사가 전체 기사의 절반에 가까운 비율을 차지하고 있는 반면, 긍정기사는 전체 기사의 10%에 못 미치는 것으로 나타났다. 중립기사는 긍정어휘와 부정어휘가 혼합되어 있거나 사실 전달을 목적으로 하는 기사를 의미하는데, 전체 기사의 약 40%를 차지하는 것으로 나타났다.



<그림 8> 공공기록물 관련 뉴스 기사의 감성 자동분류 결과

부정기사가 전체 기사의 절반에 가까운 비율을 차지하고, 중립기사의 일정 비율이 부정기사에서 다루고 있는 이슈에 대한 사실을 전달하는 기사라면 공공기록물 관련 뉴스 기사는 긍정적인 이슈보다는 부정적인 이슈를 다루고 있을 가능성이 높다고 할 수 있다. 고빈도어를 중심으로 분석이 이루어진 빈도분석과 N-gram 분석, 네트워크 군집분석에서 부정적 어조의 기사들이 주된 분석의 대상이었을 가능성이 높아 공공기록물 관련 기사에서 다른 긍정적인 이슈에는 어떠한 것이 있는지 확인하기 어렵다. 그러므로 각 기사 유형별로 빈도분석을 통해 어떤 주제를 중요하게 다루었는지를 살펴볼 필요가 있다.

각 기사 유형별 상위빈도어 50개를 워드 클라우드로 표현한 결과는 <그림 9>와 같다. 그림에서 보는 바와 같이 긍정기사에 ‘기록관리’, ‘평가’, ‘표창’, ‘기록관’, ‘우수’, ‘최우수’, ‘포상’, ‘시스템’, ‘구축’, ‘보존’, ‘추진’ 등의 단어가 많이 출현한 것으로 나타났다. 이를 통해 긍정기사에는 기록관리 기관 평가에 따른 기관 표창이나 포상, 기록관리 업무 우수 혁신 사례, 기록관리 시스템 구축 등과 관련된 이슈가 주로 다루어졌음을 알 수 있었다.



<그림 9> 공공기록물 관련 긍정/중립/부정기사의 워드 클라우드 비교(상위 50개 단어)

부정기사에서는 앞에서 살펴본 공공기록물 관련 여러 논란과 관련된 단어들 많이 출현하였음을 다시 한번 확인하였다. 중립기사에 출현한 단어들 긍정기사와 부정기사 중 어느 쪽과 더 많이 공통으로 출현했는지를 알아본 결과, <표 6>에서 보는 바와 같이 긍정기사보다 부정기사에 공통으로 출현한 단어가 더 많은 것을 알 수 있다. 이것은 중립기사의 많은 부분이 부정기사에서 다룬 주제에 대한 사실 보도 중심의 기사라는 것을 의미한다. 공공기록물 관련 뉴스 기사의 감성 자동분류 결과 예시는 <표 7>과 같다.

<표 6> 긍정/중립/부정기사 간 공통 출현 단어

기사 유형	공통 출현 단어
긍정-중립-부정	국가기록원, 기록, 기록물, 정부
긍정-중립	공공기록물, 관리, 교육, 교육청, 구축, 기관, 기록관, 기록물관리, 보존, 사업, 업무, 역사, 추진
부정-중립	검찰, 고발, 공개, 공공기록물법, 국가정보원, 국회, 남북정상회담, 내용, 노무현대통령, 대화록, 민주당, 발언, 북방한계선, 사건, 새누리당, 서울, 수사, 유출, 의원, 의혹, 자료, 조사, 청와대, 폐기, 회의록

<표 7> 공공기록물 관련 뉴스 기사 감성 자동분류 결과 예시

기사 유형	기사 제목
긍정기사	진주시 경상남도 2024년 기록관리 평가 우수기관 선정 창원시 전국 기초지방자치단체 첫 기록원 건립 추진 디지털 기록관리 국제연수 시스템 수출증대 국가 간 소통 기여할 것 공공기록물관리 솔루션 선도 부울경 대표 IT 기업 도약 발판 인천국제공항공사 기록경영시스템 ISO 30301 인증
중립기사	국가기록원 채수근상병 이태원참사 기록물 폐기 금지 한국외대 정보기록학연구소 공공기록물법 개정안 긴급토론회 개최 인천 서구 전 직원 기록물관리 역량강화 교육 국가기록관리위원회 8일 출범 위원장에 강규형 교수 광주 서부교육지원청 비전자기록물 6만 8000여권 폐기

부정기사	비상계엄 증거가 사라지고 있다 시민단체 여인형 방첩사령관 고발 민주 파기했다던 대통령 취임식 초청명단 있는 것으로 드러나 4대강 문건 파기한 한국수자원공사 기록물법 위반으로 수사 조치 공공기록물 민간 위탁 추진 논란 전문가 보안성 문제 청와대 3년 전 조선왕조실록 계승 큰소리치더니 속기록 없는 대한민국 국무회의
------	--

4.4 공공기록물 관련 연구 동향과의 비교분석과 시사점

이 절에서는 2.1의 공공기록 관련 국내 연구 동향과 공공기록물 관련 뉴스 기사의 분석 결과를 비교함으로써 기록관리학과와 현장, 그리고 사회의 공공기록물 관련 관심에 나타난 공통점과 차이점을 살펴보고자 한다. 그리고 이를 바탕으로 공공기록물 관련 사회 문제를 해결하기 위해 앞으로 기록관리학과와 현장이 고려해야 할 시사점에 대해 기술하고자 한다. 연구동향과 뉴스 기사를 비교 분석하는 이유는 뉴스 기사에서 관심 갖는 주제와 학술연구에서 관심 갖는 주제가 무엇이 다르고 유사한지를 확인함으로써 실천학문으로서의 기록관리학이 사회적 요구의 어떤 부분에 좀 더 관심을 가져야 하는지를 제안하기 위함이다.

먼저 발행 연도 추이의 측면에서 연구와 뉴스 기사 모두 지속적으로 증가추세를 보이는 공통점이 있으나, 연구 논문과 뉴스 기사의 발행 추이는 일치하지 않은 것으로 나타났다. 연구 논문이 가장 많이 생산된 2023년에 관련 뉴스 기사가 다른 해에 비해 더 많이 발행되지 않았다. 뉴스 기사가 많이 생산된 해 중 하나인 2019년에는 다른 해에 비해 공공기록물 관련 연구가 많이 이루어졌으나 연구 주제와 뉴스 기사 이슈 간에는 관련성이 없는 것으로 나타났다. 기록관리학은 현장 중심의 실천학문으로 기록과 관련된 사회적 요구를 수용하여 연구와 현장에 반영해야 할 의무가 있다. 공공기록물에 대한 사회적 관심이 높아지고 있는 만큼 관련 지속적으로 많은 연구가 이루어져야 할 것이다. 특히 연구자들은 공공기록물에 대해 사회가 갖고 있는 관심 주제에 대해 연구자의 시선으로 바라보고 그에 대한 문제해결 방안을 제시할 수 있어야 한다.

공공기록물 관련 연구의 주된 관심 분야는 ‘법·제도’, ‘정책’이다. 그리고 이는 공공기록물 관련 기사에서 주로 다루고 있는 이슈인 ‘공공기록물법’, ‘공공기록물관리’와 관련이 있다. 공공기록물법과 공공기록물관리는 공공기록물과 관련된 사회·정치적 문제를 야기하는 핵심 개념이다. 공공기록물법 위반 논란과 공공기록물 관리 미흡 등과 같은 사회·정치적 문제는 공공기록물법 제정 이후 지속적으로 반복되고 있고, 이에 따라 뉴스 기사에 자주 등장하는 주제가 되었다. 공공기록물법이 기록관리 현장에서 잘 지켜졌다면 여러 사회·정치적 이슈에 동반되어서 언급되는 빈도가 줄어들 것이다. 그런데 현재 공공기록물법이 기록관리 현장에서 잘 지켜지기 어려운 여러 이유 중에는 공공기록물법 자체의 한계와 문제 때문인 것도 있다(이경남, 2023). 그리고 이를 해결하기 위해서는 학계와 현장에서 요구하는 바람직한 방향으로 공공기록물법의 개정이 이루어져야 할 필요가 있다. 학계와 현장에서 지속적으로 공공기록물법 개정에 대한 목소리를 내고 있으나 결과론적으로 볼 때 이 의견들이 반영되지 않고 있는 현 상황은 지속적으로 기록관리에 대한 문제를 양산할 수밖에 없다. 공공기록물과 관련된 사회·정치적 문제가 지속적이고 반복적으로 발생하지 않도록 학계와 현장에서는 이에 대한 해결 방안을 제시해야 할 책임이 있다. 특히 이러한 문제를 해결할 수 있도록 공공기록물법의 개정에 학계와 현장의 의견이 반영되어야 한다.

공공기록물 관련 연구의 또 다른 관심 분야는 ‘평가’와 ‘활용’이다. 먼저 연구 분야에서는 기록의 평가선별 차원에서 공공기록의 평가에 대한 연구가 주를 이루고 있는 반면, 뉴스 기사에서는 기록관리 기관 평가 결과에 관심을 갖고 있는 것으로 나타났다. 기록관리 기관 평가의 문제점을 분석하고 개선 방안을 제시한 연구를 바탕으로 기록관리 현장에서는 기록관리 기관 평가의 과정과 결과를 지속적으로 개선함으로써 공공기관의 전반적인 기록관리 수준을 향상할 필요가 있다. 기록관리 기관 평가의 지속적 개선을 위해서는 공공기록물 관련 사회 문제를 해결할

방안들이 평가항목으로 고려되어야 할 것이다.

또한 공공기록의 활용에 대한 연구는 주로 정보공개, 기록정보서비스, 기록관리시스템 관련 연구가 주를 이루는데, 이 중 기록정보서비스와 관련된 뉴스 기사는 찾아보기가 어려웠다. 기록정보서비스의 활성화를 위해 기록정보서비스의 이론적 접근법을 현장에서 적극적으로 활용할 수 있도록 하는 방안이 필요하다. 더 나아가 공공기록물과 기록관리에 대한 인식 제고를 위해 기록정보서비스 우수 사례들을 발굴하여 뉴스 매체를 통해 알리고 소개하는 노력이 더욱 필요하다.

5. 결론 및 제언

이 연구는 2000년부터 2024년까지 발행된 공공기록물 관련 뉴스 기사 5,050건을 분석하여 공공기록물이 사회·정치적 이슈와 어떤 연관성을 가지는지를 밝히고자 하였다. 이를 위해 텍스트마이닝 기법을 적용하여 뉴스 기사에서 다루어진 주요 주제를 분석하였다.

이론적 배경에서는 최근 5년간 발표된 뉴스 기사 대상 텍스트마이닝 연구 중 특정 사회적 문제나 이슈를 다룬 연구를 검토하여 본 연구의 방법론적 타당성을 검증하였다. 또한, 빈도분석, N-gram 분석, 네트워크 군집분석, 감성분석을 활용하여 24년간 공공기록물 관련 뉴스 기사에서 다루어진 주요 주제와 특징을 분석하였다. 주요 연구 결과는 다음과 같다.

첫째, 공공기록물 관련 뉴스 기사는 지속적으로 증가하는 경향을 보였다.

둘째, 연도별 기사 양과 기사에 출현한 단어의 빈도분석 결과를 종합하면, 공공기록물 관련 기사는 공공기록물 자체보다는 공공기록물과 관련된 사건, 사고, 정치적 이슈를 주로 다루었으며, 기사에서 가장 많이 언급된 공공기록물 유형은 회의록이었다.

셋째, 바이그램 네트워크 분석 결과, ‘대화록’, ‘회의록’, ‘기록물’, ‘공공기록물법’이 주요 노드 역할을 하며, 공공기록물 관련 기사는 정치·사회적 이슈 및 기록관리 과정에서 발생하는 문제에 초점을 맞추고 있음을 확인하였다.

넷째, 네트워크 군집분석 결과, 공공기록물 관련 뉴스는 ‘공공기록물 관리’, ‘남북정상회담 회의록’, ‘공공기록물법 논란’의 세 가지 하위 주제로 구성됨을 확인하였다. 이러한 결과는 빈도분석 및 N-gram 분석 결과를 뒷받침한다.

다섯째, 감성분석 결과, 긍정 기사(약 10%)는 기록관리 평가 우수기관 표창 및 혁신 사례를, 부정 기사(약 50%)는 공공기록물법 위반 및 기록관리 미흡 문제를 주로 다루고 있었다. 중립 기사(약 40%)에서 많이 출현한 단어들이 부정 기사에서도 중복해서 출현하는 경향을 보였으며, 이는 중립 기사가 부정 기사에서 주로 다루는 공공기록물 관리 문제와 법적 이슈를 함께 다루면서 논조에서 중립적인 입장을 유지하는 경우가 많았음을 의미한다.

또한 본 연구에서는 공공기록물 관련 국내 연구 동향과 뉴스 기사 분석 결과를 비교하였다. 그 결과, 연구와 뉴스 기사 간 발생 추이가 일치하지 않으며, 연구에서 강조하는 주제와 뉴스에서 집중하는 이슈 간에도 차이가 있는 것으로 나타났다. 이를 바탕으로 기록관리학과와 현장이 고려해야 할 몇 가지 시사점을 제안하면 다음과 같다.

첫째, 공공기록물과 관련된 사회·정치적 문제가 지속적으로 발생하지 않도록 학계와 현장은 해결 방안을 제시해야 한다. 이를 위해 공공기록물법 개정 과정에서 학계와 현장의 의견이 반영될 필요가 있다.

둘째, 기록관리 기관 평가의 문제점을 분석하고 개선 방안을 제시한 연구들을 토대로, 기록관리 현장은 기록관리 기관 평가 과정과 결과를 지속적으로 개선해야 한다. 또한, 공공기록물 관련 반복되는 사회 문제에 대한 해결 방안이 기록관리 기관 평가 항목에 반영될 수 있다.

셋째, 기록정보서비스 활성화를 위해 이론적 접근법을 실무에서 적극적으로 활용할 수 있도록 하는 방안을 마련해야 한다. 나아가 공공기록물과 기록관리에 대한 인식을 높이기 위해, 기록정보서비스 우수 사례를 발굴하고 이를 뉴스 매체를 통해 알리는 노력이 필요하다.

기록관리학은 현장 중심의 실천학문으로, 학계와 실무 현장은 기록과 관련된 사회적 요구를 수용하여 연구와 실무에 반영해야 할 책임이 있다. 따라서 학계와 실무자들은 기록과 관련된 사회적 문제의 해결 방안을 제시하는 동시에, 이를 법·제도·정책 등의 변화에 반영하기 위해 지속적으로 목소리를 내야 한다. 이를 통해 공공기록물 관리에 대한 사회적 이해도를 높이고, 관련 정책을 지속적으로 개선하며, 나아가 기록관리에 대한 사회적 인식을 개선할 수 있을 것이다.

본 연구는 텍스트마이닝 기법을 이용하여 공공기록물 관련 뉴스 기사를 분석하고 주요 주제와 감성 분포를 도출하는 데 초점을 맞추었다. 그러나 다음과 같은 한계점이 있다.

첫째, 뉴스 기사 전처리 과정에서 활용된 형태소 분석기의 한계와 뉴스 텍스트의 오타 및 띄어쓰기 차이로 인해 복합명사나 고유명사의 처리가 완벽하게 이루어지지 못했다. 이로 인해 특정 개념이 과소 또는 과대 평가될 가능성이 있으며, 연구 해석 시 이를 보완적으로 고려해야 한다.

둘째, 본 연구는 특정 빅데이터 분석 플랫폼이 제공하는 텍스트마이닝 기법만을 활용하여 분석을 수행하였기 때문에, 각 기법의 정확성과 성능을 완벽하게 검토하지 못했다.

이러한 한계를 보완하기 위해 추후에는 보다 정교한 언어처리 기법을 적용하고, 텍스트마이닝 기법의 성능 검증을 위해 다양한 플랫폼을 활용하여 분석 결과를 비교·검증하는 방법을 고려할 필요가 있다. 또한, 향후 연구에서 뉴스와 미디어에서 공공기록물을 다루는 방식과 프레임링을 심층적으로 분석하고, 법 개정 전후의 보도 경향 변화를 시계열적으로 분석하는 것도 가능할 것이다. 아울러, 해외 사례와 비교하여 한국의 공공기록물 관리 및 보도 경향을 분석하는 것도 의미 있는 연구가 될 것이다.

본 연구는 24년간의 공공기록물 뉴스 기사를 분석하여 공공기록물과 관련된 사회적 이슈의 흐름을 조망하고, 이를 바탕으로 추상적이나마 기록관리 연구와 정책의 역할과 방향을 제안했다는 점에서 의의가 있다. 향후 연구를 통해 공공기록물 관리 및 정책에 대한 보다 실질적인 개선 방안이 마련되기를 기대한다.

참고문헌

감미아, 송민 (2012). 텍스트 마이닝을 활용한 신문사에 따른 내용 및 논조 차이점 분석. *지능정보연구*, 18(3), 53-77.

<https://doi.org/10.13088/JIIS.2012.18.3.053>

공공기록물 관리에 관한 법률. 법률 제20309호.

공공기록물 관리에 관한 법률 시행령. 대통령령 제34487호.

김태경 (2025. 1. 19.). [강남시선] 요원한 국가기록원 독립성. 파이낸셜뉴스.

<https://www.fnnews.com/news/202501191839554571>

김형도 (2023). GPT-4 기반 제로샷 영화 리뷰 감성 분석. *한국콘텐츠학회논문지*, 23(11), 185-196.

- <https://doi.org/10.5392/JKCA.2023.23.11.185>
- 김혜진 (2022). 텍스트마이닝 기법을 이용한 한국 사회의 혐오 양상 분석. 한국문헌정보학회지, 56(4), 431-453.
<https://doi.org/10.4275/KSLIS.2022.56.4.431>
- 더아이엠씨 (발행년불명). 텍스트롬. 출처: <https://www.textom.co.kr/>
- 박중화, 김민성, 김정환 (2022). 언론은 인공지능(AI)을 어떻게 다루는가?: 뉴스 빅데이터를 통한 한국과 미국의 보도 경향 분석. 정보시스템연구, 31(1), 175-195.
- 박주연, 정도현 (2022). 텍스트마이닝 기법을 활용한 교육관점에서의 메타버스 관련 이슈 탐색-뉴스 빅데이터를 중심으로. 산업융합연구, 20(6), 27-35. <https://doi.org/10.22678/JIC.2022.20.6.027>
- 박주현, 박현지, 김영범 (2024). 빅카인즈를 활용한 5.18 관련 국내 기사 분석 연구. 정보관리학회지, 41(1), 107-132.
<https://doi.org/10.3743/KOSIM.2024.41.1.107>
- 양지연 (2022). 텍스트마이닝을 이용한 미세먼지 관련 신문기사 분석. 디지털융복합연구, 20(1), 1-13.
<https://doi.org/10.14400/JDC.2022.20.1.001>
- 엄연용, 김경인, 송원영 (2023). 성역할 관련 연구동향과 여론 변화 분석: 텍스트마이닝 분석방법을 활용하여. 통계연구, 28(2), 45-70. <https://doi.org/10.22886/jkos.2023.28.2.45>
- 오세현, 강현아 (2021). 아동학대 정책 변화 시점 별 뉴스 분석-텍스트마이닝을 활용하여. 한국아동복지학, 70(3), 1-31.
<https://doi.org/10.24300/jkscw.2021.09.70.3.1>
- 윤호영 (2020). 남북 관계 뉴스 제목 텍스트 마이닝 분석: 227개월 44만 건 보도 관심 주제. Journal of The Korean Data Analysis Society, 22(6), 2517-2534. <https://doi.org/10.37727/jkdas.2020.22.6.2517>
- 이경남 (2023). 공공기록의 개념 및 범위에 관한 논의. 정보관리학회지, 40(1), 95-119.
<https://doi.org/10.3743/KOSIM.2023.40.1.095>
- 이소현, 구분진 (2022). 텍스트마이닝을 활용한 "잊힐 권리"의 토픽 분석. 정보관리학회지, 39(2), 275-298.
<https://doi.org/10.3743/KOSIM.2022.39.2.275>
- 이정은 (2024). 텍스트 마이닝을 활용한 한국의 노인자살 현상에 대한 뉴스 빅데이터 분석. 한국콘텐츠학회 논문지, 24(7), 70-81. <https://doi.org/10.5392/JKCA.2024.24.07.070>
- 이준석, 강윤석, 김상욱 (2018). 소셜 네트워크를 위한 커뮤니티 탐지 알고리즘 성능 비교. 2018년 한국소프트웨어종합학술대회 논문집, 219-220.
- 전진한 (2024. 8. 25.). [지금,여기] 회의록 부실 생산, 이대로 좋은가. 경향신문.
<https://www.khan.co.kr/article/202408252036005>
- 정상준, 윤보현, 오효정 (2019). 언어 네트워크 기반 대통령기록물 관련 이슈 및 매체별 특성 분석. 한국비블리아학회지, 30(1), 181-207. <https://doi.org/10.14699/kbiblia.2019.30.1.181>
- 정우경, 신동희 (2021). 토픽 모델링 및 바이그램 네트워크 분석 기법을 통한 여대생의 건강관리 및 웨어러블 디바이스 인식에 관한 연구. 정보관리학회지, 38(4), 129-152. <https://doi.org/10.3743/KOSIM.2021.38.4.129>
- 조현국, 이봉우 (2021). 텍스트 마이닝을 활용한 기후변화에 대한 20년간의 뉴스 기사 분석. 에너지기후변화교육, 11(2), 153-163. <https://doi.org/10.22368/ksecce.2021.11.2.153>
- 최한별, 정윤혁, 김성철 (2022). 텍스트 마이닝을 활용한 정보 프라이버시 의제 분석: 1990~2021년 뉴스 빅데이터를 중심으로. 정보사회와 미디어, 23(2), 69-113.
- 한국과학기술정보연구원 (발행년불명). ScienceOn. 출처: <https://scienceon.kisti.re.kr/>
- 한국기록관리학회 엮음 (2024). 기록관리의 세계: 이론에서 실천까지. 과주: 한올아카데미.
- 한국언론진흥재단 (발행년불명). BigKinds. 출처: <https://www.bigkinds.or.kr>
- 한국연구재단 (발행년불명). 학술지검색. KCI. 출처: <https://www.kci.go.kr/kciportal/po/search/poSereSear.kci>

- 한승희 (2018). 뉴스 빅데이터를 이용한 우리나라 언론의 기록관리 분야 보도 특성 분석: 1999~2018 뉴스를 중심으로. *정보관리학회지*, 35(3), 41-75. <https://doi.org/10.3743/KOSIM.2018.35.3.041>
- Febro-Naga, J. D. & Tinam-isan, M. A. C. (2022). Exploring cyber violence against women and girls in the Philippines through mining online news. *Comunicar: Media Education Research Journal*, 30(70), 121-133. <https://doi.org/10.3916/C70-2022-10>
- Girvan, M. & Newman, M. E. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12), 7821-7826. <https://doi.org/10.1073/pnas.122653799>
- Jurafsky, D. & Martin, J. H. (2009). *Speech and Language Processing* (2nd ed.). New Jersey: Pearson Prentice Hall.
- Krawczyk, K., Chelkowski, T., Laydon, D. J., Mishra, S., Xifara, D., Gibert, B., Flaxman, S., Mellan, T., Schwämmle, V., Röttger, R., Hadsund, J. T., & Bhatt, S. (2021). Quantifying online news media coverage of the Covid-19 pandemic: text mining study and resource. *Journal of Medical Internet research*, 23(6), e28253. <https://doi.org/10.2196/28253>
- Open AI Platform (2023). Models. Available: <https://platform.openai.com/docs/models>
- Salloum, S. A., Al-Emran, M., Abdallah, S., & Shaalan, K. (2017). Analyzing the Arab gulf newspapers using text mining techniques. *Proceedings of the International Conference on Advanced Intelligent Systems and Informatics 2017*, 396-405. http://doi.org/10.1007/978-3-319-64861-3_37
- Traag, V. A., Waltman, L., & van Eck, N. J. (2019). From Louvain to Leiden: guaranteeing well-connected communities. *Scientific Reports*, 9, 5233. <https://doi.org/10.1038/s41598-019-41695-z>

• 국문 참고자료의 영어 표기

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

- Choi, Hanbyul, Jung, Yoonhyuk, & Kim, Seongcheol (2022). A text mining analysis of information privacy agenda: focusing on news big data from 1990 to 2021. *Information Society & Media*, 23(2), 69-113.
- Enforcement Decree of the Public Records Management Act. Presidential Decree No.34487.
- Eom, Yunyong, Kim, Gyoungin, & Song, Wonyoung (2023). Analysis of trends research on gender-role: using text mining analysis. *Journal of the Korean Official Statistics*, 28(2), 45-70. <https://doi.org/10.22886/jkos.2023.28.2.45>
- Han, Seunghee (2018). An Analysis of news report characteristics on archives & records management for the press in Korea: based on 1999~2018 news big data. *Journal of the Korean Society for Information Management*, 35(3), 41-75. <https://doi.org/10.3743/KOSIM.2018.35.3.041>
- Jeon, Jinhan (2024, August 25). [Now, here] Is it okay to produce defective minutes?. *The Kyunghyang Shimmun*. <https://www.khan.co.kr/article/202408252036005>
- Jeong, Woo-kyoung & Shin, Donghee (2021). Analyzing female college student's recognition of health monitoring and wearable device using topic modeling and bi-gram network analysis. *Journal of the Korean Society for Information Management*, 38(4), 129-152. <https://doi.org/10.3743/KOSIM.2021.38.4.129>
- Jho, Hunkoog & Lee, Bongwoo (2021). Analysis of news articles about climate change based on text mining for the last two decades. *Journal of Energy and Climate Change Education*, 11(2), 153-163. <https://doi.org/10.22368/ksecce.2021.11.2.153>

- Jung, Sang jun, Yun, Bo-hyun, & Oh, Hyo-jung (2019). Analysis of social issues and media-specific characteristics related to presidential records based on semantic network. *Journal of the Korean Biblia Society for Library and Information Science*, 30(1), 181-207. <https://doi.org/10.14699/kbiblia.2019.30.1.181>
- Kam, Miah & Song, Min (2012). A study on differences of contents and tones of arguments among newspapers using text mining analysis. *Journal of Intelligence and Information Systems*, 18(3), 53-77. <https://doi.org/10.13088/jiis.2012.18.3.053>
- Kim, Hea-jin (2022). Analyzing the phenomena of hate in korea by text mining techniques. *Journal of the Korean Society for Library and Information Science*, 56(4), 431-453. <https://doi.org/10.4275/KSLIS.2022.56.4.431>
- Kim, Hyoungdo (2023). Zero-shot sentiment analysis of movie reviews based on GPT-4. *The Journal of the Korea Contents Association*, 23(11), 185-196. <https://doi.org/10.5392/JKCA.2023.23.11.185>
- Kim, Taegyong (2025, January 19). [Gangnam View] The independence of the National Archives that seems distant. *Financial News*. <http://www.fnnews.com/news/202501191839554571>
- Korea Institute of Science and Technology Information (n.d.). ScienceOn. Available: <https://scienceon.kisti.re.kr/>
- Korea Press Foundation (n.d.). BigKinds. Available: <https://www.bigkinds.or.kr>
- Korean Society of Archives & Records Management (2024). *Archives and Records Management: Theories and Practices*. Paju: Hanwool Academy.
- Lee, Jung-eun (2024). News big data analysis of elderly suicide in korea using text mining. *The Journal of the Korea Contents Association*, 24(7), 70-81. <https://doi.org/10.5392/JKCA.2024.24.07.070>
- Lee, Junseok, Kang, Yoonsuk, & Kim, Sang-wook (2018). Performance comparison of community detection algorithms in social networks. *Proceedings of KSC 2018*, 219-220.
- Lee, Kyungnam (2023). A study on the legal concept and the scope of public records. *Journal of the Korean Society for Information Management*, 40(1), 95-119. <https://doi.org/10.3743/KOSIM.2023.40.1.095>
- Lee, So-hyun & Koo, Bonjin (2022). Topic analysis of the “right to be forgotten” using text mining. *Journal of the Korean Society for Information Management*, 39(2), 275-298. <https://doi.org/10.3743/KOSIM.2022.39.2.275>
- National Research Foundation of Korea (n.d.). Journal Search. KCI. Available: <https://www.kci.go.kr/kciportal/po/search/poSereSear.kci>
- Oh, Sehyun & Kang, Hyunah (2021). Analysis of news on child abuse based on the major policy changes: using text mining. *Journal of the Korean Society of Child Welfare*, 70(3), 1-31. <https://doi.org/10.24300/jkscw.2021.09.70.3.1>
- Park, Jonghwa, Kim, Min sung, & Kim, Junghwan (2022). How does the media deal with artificial intelligence?: analyzing articles in korea and the us through big data analysis. *The Journal of Information Systems*, 31(1), 175-195.
- Park, Ju yeon & Jeong, Do-heon (2022). Exploring issues related to the metaverse from the educational perspective using text mining techniques – focusing on news big data. *Journal of Industrial Convergence*, 20(6), 27-35. <https://doi.org/10.22678/JIC.2022.20.6.027>
- Park, Ju-hyeon, Park, Hyunji, & Gim, Youngbum (2024). An analysis of domestic newspaper articles on 5.18 using the Bigkinds system. *Journal of the Korean Society for Information Management*, 41(1), 107-132. <https://doi.org/10.3743/KOSIM.2024.41.1.107>
- Public Records Management Act. Law No.20309.
- The IMC (n.d.). Textom. Available: <https://www.textom.co.kr/>

- Yang Ji-yeon (2022). An analysis of newspaper articles on fine particle matter using text mining techniques. *Journal of Digital Convergence*, 20(1), 1–13. <https://doi.org/10.14400/JDC.2022.20.1.001>
- Yoon, Ho young (2020). Text mining news analysis on north–south korea relation: news coverages for 227 months. *Journal of The Korean Data Analysis Society*, 22(6), 2517–2534. <https://doi.org/10.37727/jkdas.2020.22.6.2517>

[부록] 공공기록 관련 논문 목록

번호	범주	논문명	저널/프로시딩명	저자	출판 연도
1	기록 관리 자체 분석 연구	공공기록관리분야의 빅데이터 활용 방법과 시사점 제안	한국기록관리학회지	홍덕용	2021
2		공공기록관리 제도의 형성과 변화에 관한 언론보도 분석	한국기록관리학회지	이주영	2013
3		공공기록관의 소셜미디어 이용 현황 및 이용자 관심도 분석: 국가기록원과 대통령기록관을 중심으로	정보관리학회지	최정원, 강주연, 박준형, 오효정	2016
4		'공공기록'으로서 방송·영상 콘텐츠의 '아카이브 효과' 연구	기록과 정보·문화 연구	최효진	2021
5		대중 매체에 비친 아카이브의 공공 인식: 아카이브와 관련된 직업군, 기관, 기록물에 대한 2003-2004년간의 국내보도	한국기록관리학회지	이주현	2005
6	법 · 제도	記録保存(기록보존)소에서의 영구기록관리 업무절차 : 이관·정리·기술·보존 : ICA 공공부문 기록관리를 중심으로	記録保存	이상민	2003
7		공공기록물관리예에 관한 법률시행령 전부개정령(안)	記録管理保存	편집부	2007
8		공공기록물관리법과 특수기록관리제도에 관한 고찰	기록학연구	윤은하	2024
9		한국 공공기록관리의 쟁점과 전망 2013년 기록관리체제를 위하여	기록학연구	안병우, 이상민, 심성보, 남경호, 김진성, 오동석, 정태영	2012
10		공공 기록의 처분 동결 제도 도입을 위한 연구	기록학연구	현문수	2017
11		우리나라 공공기록물 관리에 관한 법규의 변천	한국기록관리학회지	김세경	2007
12		공공업무의 체계적 기록화를 위한 보유일정표 설계 방안	한국문헌정보학회지	설문원	2006
13		공공기록물법 개정을 위한 방향과 과제	기록학연구	현문수	2017
14		공공기록물 관리에 관한 법령 개정을 위한 면담 연구	기록학연구	현문수, 정상희, 박민영, 황진현, 이소연	2017
15		전문성 강화를 위한 공공기록관리 조직 및 제도 혁신 방향	기록학연구	설문원, 김형국, 이원규	2018
16		기록물 보존장소 변경에 따른 제도개선 방안 공공기록물법령을 중심으로	기록학연구	임진수	2020
17		민간시설 활성화를 위한 공공기록법 개정(안) : 쟁점과 과제	기록학연구	이소연	2016
18		폐교(초·중·고)기록물 이관현황과 개선방안에 관한 연구: 공공기록물법 제정 이후	한국기록관리학회지	전다솜, 이지수	2023
19		기록관리 공공표준의 적정성 검토 체계 고찰을 통한 개선 방안 연구	기록학연구	김보람, 김순희	2023
20		기록으로의 공공데이터 관리를 위한 제도적 고찰 - 『공공데이터의 제공 및 이용 활성화에 관한 법률』 분석을 중심으로 -	한국기록관리학회지	김유승	2014
21		"공공기록물 관리에 관한 법률" 개선 방향에 관한 연구: KS X ISO 15489 표준에 입각하여	정보관리학회지	정기애, 김유승	2009

번호	범주	논문명	저널/프로시딩명	저자	출판 연도
22		기록정보서비스 관점에서 공공저작물 자유이용 법제화의 의미와 한계에 대한 고찰	한국기록관리학회지	정경희	2014
23		한국의 국가기록물관리와 역사연구 - 「공공기록물관리법」(2007)을 중심으로 -	대구사학	김정하	2008
24		행정기록물의 심사범위 확정에 따른 행정의 사법적 통제 - 미국의 공공데이터 기록·관리 법제를 중심으로-	법학연구	방정미	2020
25		페이퍼리스 전자정부를 위한 공공기록물 정보관리에 관한 법적 검토	법학연구	양영식, 송인방, 박상혁	2019
26		기록관리 국가 표준과 공공 표준 내용 불일치 문제에 대하여	기록과 정보·문화 연구	박춘원, 이선옥	2023
27		『공공기록물 관리에 관한 법률』의 제정 의의와 개선방안 - 국가기록원의 위상과 국민의 알권리를 중심으로 -	한국기록관리학회지	김유승	2008
28		교육자치의 이념에 토대한 지방아카이브 논의 공공기록물법 제11조의 개정 필요성을 중심으로	기록학연구	정상명	2022
29		공공기록물법의 전부 개정 필요에 대한 논의: 디지털플랫폼정부를 중심으로	기록과 정보·문화 연구	정상명	2022
30		공공기록물법 개정안 비평: 2024년도 정부입법안을 중심으로	기록과 정보·문화 연구	성민결	2024
31		미국 공공기록관리법(PRA)을 통한 기록관리 제도에 관한 연구: 연방과 주를 중심으로	한국기록관리학회지	조애란, 김민경, 윤은하	2019
32		한국 공공분야 '기록보유(Recordkeeping)'체제 전망 -기록물분류기준표의 제도적 의의와 특성	기록학연구	이승익	2001
33		기록물 보존용품 및 장비에 몇 가지 제언 - 공공기록물관리에 대한 법률 및 법률 시행령 개정에 즈음하여	記錄管理保存	박인신	2007
34	보존	대량 공공 기록물의 선별적 소독을 위한 미생물 활성화 분석 연구	보존과학회지	김대운, 박가영, 김지원, 강대일	2018
35		산중에 위치한 공공보존기록관의 산림화재 방호를 위한 드렌치노즐 수막방호시스템의 실험연구	한국방재학회논문집	이영재, 박형주	2013
36		공공기록물에서의 PDF의 활용	記錄管理保存	박정수	2005
37	시스템	블록체인 기반 대학 공공기록물 관리 서비스 설계 연구	한국전자거래학회지	홍기완, 장항배	2021
38		공공사이트 게시관 자료의 기록관리를 위한 자동 분류 시스템	한국정보관리학회 학술대회 논문집	남은경, 안혜림, 송민	2013
39		공공데이터 기록관리 활성화와 큐레이션 활용방안 연구 - 공공데이터포털, 해안시스템을 중심으로-	기록과 정보·문화 연구	이상현	2020
40		비밀문화와 기록문화 - 공공영역 기록화의 문화사적 의미	역사와 담론	이종흡	2008
41		구술 기록의 공공성	기록학연구	김택호	2024
42	역사 / 문화	역사를 위하여: 아키비스트와 역사가의 역할 -공공記錄保存(기록보존)소를 중심으로-	기록학연구	이상민	2002
43		인권 정치와 기록레짐화: 글로벌 공공재로서의 세계기록유산	문화와 정치	노선영, 최은봉	2020
44		기록영상물의 공공재화와 영상역사 쓰기- '제국-국민국가' 서사를 넘어서	역사비평	허은	2014

번호	범주	논문명	저널/프로시딩명	저자	출판연도
45		아테네의 공공 기록 속 문자와 형상 - 고전기 아테네의 기록 부조(document reliefs)의 사례를 중심으로 -	인문과학	김혜진	2017
46		ICA사례연구 -강명숙·김명훈·정은이·황정원 공역-, 『공공부문의 기록물 관리: 사례연구 3』, 진리탐구,2008 -전상숙 역-, 『공공부문 기록관리를 위한 기초자료』, 진리탐구,2008	기록학연구	이은영	2008
47		공공기록물의 분류원리 : 출처주의에 대한 이론적 검토	記録保存	김명훈	2003
48	원칙 / 이론	『공공부문의 기록관리 : 원칙과 맥락』 (이상민 역,진리탐구,2008)	기록학연구	남경호	2009
49		공공기록의 개념 및 범위에 관한 논의	정보관리학회지	이경남	2023
50		한나 아렌트 정치이론에 비취 본 記録保存(기록보존)소(Archives)의 공공성 검토	기록과 정보·문화 연구	박태선	2023
51		노근리사건 규명·노근리정신 정립과 기록의 힘 - 레코드 컨티뉴엄 및 공공역사 개념을 적용한 분석과 이해 -	기록과 정보·문화 연구	노명환	2023
52		지방 공공기록관리의 정책과제	기록학연구	이원규	2010
53		공공기록물관리'에 대한 오늘의 과제	圖書館文化	배현숙	1999
54		공공기록관의 경제적 효과	한국기록관리학회지	Yakel, Elizabeth	2012
55		디지털 전환 시대의 공공기록정책 기록자산으로서 정보의 관리	기록학연구	설문원	2020
56		공공기록물 정보서비스 발전방안에 관한 연구	한국기록관리학회지	이영숙	2007
57	정책	공공기록물 관리에 있어 이명박정부의 책임과 '업적'	기록학연구	이승휘	2008
58		한국 공공기록관리 정책의 연대기적 검토	한국기록관리학회지	서혜란	2009
59		SNS 게시물에 대한 공공기록화 수용에 관한 연구	디지털융복합연구	윤승욱, 장준갑, 김건	2019
60		공공기록물의 수집·이관과 아키비스트의 역할	기록학연구	이원규	2000
61		기록연구사의 직무특성과 공공봉사동기가 혁신행동에 미치는 영향에 관한 연구	한국기록관리학회지	김미용, 김정현	2023
62		국정과제로서의 디지털플랫폼정부 추진과 대응 방향 공공기록물 관리에 미치게 될 영향을 중심으로	기록학연구	정상명	2023
63	특정주제 기록	공공극장의 역할과 가치에 대한 탐색—김옥란, 『남산예술센터 관객의 기록 2009_2020』 (연극과인간, 2021)	한국극예술연구	박미란	2021
64		기록작업(documentation)의 의미에 대한 재조명 - 21세기 한국 유아교육의 공공성 구성을 위한 탐색	열린유아교육연구	김희연	2008
65		공공영역의 2004년 대통령 탄핵사건 기록	기록학연구	오명진	2012
66	평가	공공기록 평가제도의 구조와 쟁점 과제	한국기록관리학회지	설문원, 이승익	2020
67		공공기록 평가의 책임과 권한의 분배 '시민참여 평가'를 중심으로	기록학연구	이경래	2019

번호	범주	논문명	저널/프로시딩명	저자	출판연도
68		공공기록에 대한 '거시평가(Macroappraisal)' 적용 시론	기록학연구	이승억	2006
69		공공기록물 평가제도의 문제점과 개선방안	한국기록관리학회지	최재희	2014
70		공공기록물의 평가·폐기에 대한 감독제도 연구 미국의 기록관리 감독·보고 프로그램을 중심으로	기록학연구	설문원, 박인선	2019
71		디지털시대의 공공기록평가에 관한 정책적 고찰 영국 TNA 사례를 중심으로	기록학연구	김유승	2019
72		단위과제 기반 공공기록물 평가제도의 문제점과 개선방안	한국기록관리학회지	설문원	2013
73		공공기록물의 평가체제에 대한 이론적 검토 -선별 방식 및 가치 범주를 중심으로-	기록학연구	김명훈	2002
74		영국 공공기록물 선별체계의 역사적 흐름과 의미	역사교육	최재희	2017
75		공공기록물 재평가 제도 보완 방안	기록학연구	류신애, 이승휘	2010
76		외국의 공공기록정보 공개제도	記録保存	이상민	2004
77		공공기록물의 저작권과 이용증진 (시귀선 엮음, 한국국가기록연구원,2011)	기록학연구	허선	2011
78		공공기록물의 이용과 저작권보호에 관한 연구	한국기록관리학회지	시귀선	2009
79	활용	공공기록물 관리 블록체인에서 타임스탬프 기반의 효율적 탐색 방법	디지털콘텐츠학회 논문지	손민영, 손기봉, 김영학	2023
80		공공기록에 대한 접근성 제고 방안에 관한 연구 - 대통령기록관 주제시소러스 개발 사례를 중심으로 -	한국기록관리학회지	이혜영, 권용찬, 성효주,유병후	2014
81		영기록물관리기관 공공프로그램의 역할과 운영방안 연구	기록학연구	정은진	2007
82		정보공개법상 법령에 의한 비공개정보에 관한 고찰 - 공공기록물관리에 관한 법률과의 관계를 중심으로	동아법학	박진우	2009