

키워드 네트워크 분석을 활용한 국내 리포지터리 연구 동향 분석*

An Analysis of Research Trends in Institutional Repositories in South Korea

조 은 정 (Eun Jeong Jo)**

정 유 경 (Yoo Kyung Jeong)***

〈 목 차 〉

I. 서 론
II. 선행연구
III. 연구방법

IV. 연구 결과
V. 결론 및 제언

요 약: 본 연구의 목적은 국내 리포지터리 연구의 동향을 분석하는 것이다. 이를 위해 한국 학술지 인용색인에 등록된 리포지터리 관련 논문을 대상으로 키워드 빈도 분석과 키워드 네트워크 분석을 수행하였다. 연구결과, 키워드 빈도 분석을 통해 기관 리포지터리, 오픈엑세스, 공동보존서고, 데이터 리포지터리, xml 등 자료의 관리와 보존, 접근, 활용 측면의 키워드가 중점적임을 확인하였다. 키워드 네트워크에서는 클러스터 분석을 통해 주요 연구주제를 도출하고자 하였으며, 7개의 클러스터가 도출되었다. 주요 클러스터로는 공동보존서고가 가장 큰 영역을 차지하고 있었으며, 디지털 리포지터리, 데이터 리포지터리, 오픈엑세스 및 기관 리포지터리, 데이터 표준 기술, 교육학술정보 시스템, 메타데이터와 관련된 클러스터들이 주요 연구 영역들로 나타났다. 키워드의 시기별 분석 결과, 초기에는 xml 관련 연구가 주를 이루었으나 최근에는 연구데이터와 데이터 리포지터리 관련 연구가 증가하는 추세를 보였다. 본 연구는 국내 리포지터리 연구의 현황을 진단하고, 동향을 체계적으로 분석하여 향후 리포지터리 연구 방향을 제시하는 데 기여할 것으로 기대한다.

주제어: 리포지터리, 기관 리포지터리, 연구동향, 오픈엑세스, 네트워크 분석

ABSTRACT: The purpose of this study is to analyze the research trends of repositories in Korea. We conducted keyword analysis and network analysis on repository-related papers indexed in the Korea Citation Index. The keyword frequency analysis identified major keywords including institutional repository, open access, collaborative preservation repository, data repository, and XML. Through keyword network analysis, seven clusters were identified. The largest cluster was related to cooperative shared storage, followed by clusters focusing on digital repository, data repository, open access and institutional repository, data standard technology, educational information systems, and metadata. A time series analysis showed that while XML-related research had been dominant in the early stage, recent trends indicate an increase in research related to research data and data repositories. This study is expected to contribute to diagnosing the current state of repository research in Korea, systematically analyzing trends, and suggesting future research directions.

KEYWORDS: Repository, Institutional Repository, Research Trends, Open Access, Network Analysis

* 이 논문은 2017년 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2017S1A6A3A01079581).

** 한남대학교 문헌정보학과 석사과정(bts2001415@naver.com) (제1저자)

*** 명지대학교 문헌정보학과 조교수(yk.jeong@mju.ac.kr / ISNI 0000 0004 9285 4707) (교신저자)

• 논문접수: 2024년 11월 23일 • 최초심사: 2024년 12월 4일 • 게재확정: 2024년 12월 16일

• 한국도서관·정보학회지, 55(4), 169-192, 2024. <http://dx.doi.org/10.16981/kliss.55.4.202412.169>

* Copyright © 2024 Korean Library and Information Science Society

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.

I. 서론

리포지터리는 대학과 연구기관에서 생산된 학술정보를 보존하고 확산하기 위해 설계된 지식 저장소이다(정영미, 이상기, 2010). 이러한 리포지터리는 정보자원의 개방적 이용을 목적으로 하며, 오픈엑세스와 밀접하게 연관되어 있다. 특히, 전통적 학술 출판의 한계를 보완하고 연구 결과의 신속한 공유와 접근성을 높이는 역할로 주목받고 있다. 더불어 디지털 기술의 발전은 리포지터리의 기능을 단순 저장소에서 기관 연구 성과의 가시화와 영향력 확대를 위한 도구로 확장시켰다(김예현, 김지현, 2023).

그러나 2024년 현재, OAK 리포지터리를 도입한 국내 기관은 66개에 불과하며, 이는 해외에 비해 낮은 도입률을 보이고 있었다(OAK 국가리포지터리, 2024). 스페인 국립연구위원회(CSIC) 산하의 Cybermetrics Lab에서 제공하는 구글 스칼라 기반의 국제 리포지터리 순위¹⁾에서도 서울대학교 S-Space와 KAIST의 KOASAS를 제외하면 국내 기관의 상위권 진입 사례는 드물게 나타났다(Ranking Web of Repositories, 2024).

이러한 현상의 원인을 파악하고 해결 방안을 모색하기 위해서는 국내 리포지터리 연구의 현황과 특성을 체계적으로 분석할 필요가 있을 것이다. 특히 리포지터리는 공동보존서고, 기관 리포지터리, 데이터 리포지터리 등 다양한 유형이 존재하며, 각각의 목적과 기능이 상이하다. 따라서 국내 리포지터리의 전반적인 연구동향과 함께 연구에 드러나 각 리포지터리별 특성을 종합적으로 파악하는 것이 선행되어야 할 것이다.

본 연구는 국내 리포지터리 연구의 동향과 핵심 주제를 파악하기 위해, 한국 학술지 인용색인(Korea Citation Index, KCI)에 등록된 리포지터리 관련 논문을 대상으로 키워드 빈도 분석과 키워드 네트워크 분석을 수행하고자 한다. 키워드 네트워크 분석은 연구 주제들 간의 관계성을 시각적으로 표현할 수 있으며, 특히 클러스터 분석을 통해 리포지터리의 다양한 유형과 특성을 체계적으로 분류하고 파악할 수 있다는 장점이 있다. 이를 통해 국내 리포지터리 연구의 현황을 진단하고, 발전 방향을 제안하고자 한다. 구체적인 연구 문제는 다음과 같다.

Q1: 국내 리포지터리 관련 연구에서 나타난 주요 키워드와 시기별 키워드의 변화는 어떠한가?

Q2: 키워드 네트워크 분석을 통해 드러난 국내 리포지터리 연구의 핵심 주제는 무엇인가?

본 연구는 국내 리포지터리 연구의 발전 수준과 관심 주제를 파악함으로써, 리포지터리 도입과 운영에서의 주요 이슈와 해결 과제를 도출하고, 정책 및 기술적 지원 방향을 제안하기 위한 기초 도구로 활용될 수 있을 것이다.

1) <https://repositories.webometrics.info/>

II. 선행연구

국내에서 수행된 리포지터리에 관한 연구는 크게 리포지터리 구축 및 운영에 관한 연구, 리포지터리에 수록된 자료의 내용을 분석한 연구, 영향력 평가 연구, 오픈엑세스 및 리포지터리 동향 연구로 분류해 볼 수 있다. 초기 연구들은 주로 리포지터리 구축 사례와 운영 정책에 초점을 맞추었으며(김현희 외, 2005; 김현희, 정경희, 김용호, 2006; 이재민, 손청기, 이미영, 2008; 정영미, 이상기, 2010; 황혜경, 이지연, 2009), 이후 콘텐츠 품질 관리(백종명, 이은지, 2013)와 글로벌 영향력 평가(신은자, 2017) 등으로 연구 범위가 확장되고 있다. 최근에는 데이터 리포지터리 활성화 및 인증(김예현, 김지현, 2023; 김주섭, 양성준, 김선태, 2022; 박형주, 2024; 이해림 외, 2024)과 운영 정책 수립을 위한 내부 환경 분석(김규환, 김지현, 2024) 등 보다 구체적이고 전문화된 연구가 진행되고 있는 것을 알 수 있다.

우선, 리포지터리 구축 및 운영에 관한 선행연구들은 국내외 사례 분석을 통해 리포지터리의 필요성과 효과적인 운영 방안을 제시하고 있었다. 김현희 외(2005)는 SPARC 체크리스트 분석과 16개 국내외 대학 리포지터리 사례 연구를 통해 리포지터리 운영에 중요한 6가지 요인(학내 합의 형성과 운영 체제, 콘텐츠, 지식 재산권, 학내 동종 사업과의 제휴, 기술과 시스템, 도서관)을 도출하여 대학 리포지터리 구축 및 활성화를 위한 지침을 제시하였다.

이재민, 손청기, 이미영(2008)은 KAIST의 KOASAS(KAIST Open Access Self-Archiving System) 리포지터리 구축 및 운영 사례를 분석하였다. 이 연구에서는 KOASAS의 구축 배경, 시스템 구성, 운영 정책 등을 상세히 살펴보고, 이를 바탕으로 국내 대학의 리포지터리 구축 필요성과 활용방안을 제시하였다. 특히 연구자들의 자발적 참여를 유도하기 위한 인센티브 제도와 저작권 관리 방안 등 실질적인 운영 전략을 제안하였다.

황혜경과 이지연(2009)은 문헌조사와 사례연구, 전문가 심층면담을 통해 기관 리포지터리 성공을 위한 주요 요인들을 조직적, 정책적, 기술적 측면에서 분석하였다. 이외에도 정영미와 이상기(2010)는 성공적인 리포지터리 사례들의 운영정책을 비교 분석하였다. 미국, 영국, 호주 등 주요 국가의 대표적인 리포지터리 운영 정책을 콘텐츠 수집, 저작권 관리, 시스템 운영 등의 측면에서 체계적으로 분석하였다. 이외에도 개인기록의 전자출판이나 교육자료 관련 리포지터리 구축(이종덕, 신규용, 유진철, 2013; 추기숙, 남영준, 2017) 등 특정한 분야의 리포지터리 구축에 관련된 연구들도 꾸준히 이루어지고 있었다.

최근의 연구로, 김예현, 김지현(2023)은 과학기술 분야 기관 연구데이터 리포지터리의 운영 현황을 파악하고 활성화 방안을 제안하였다. 이 연구에서는 국내 주요 과학기술 분야 연구기관들의 리포지터리 운영 실태를 조사하고, 연구데이터의 수집, 관리, 공유 과정에서 발생하는 문제점들을 분석하였다. 특히 연구데이터의 특성을 고려한 메타데이터 표준화, 데이터 품질 관리, 인센티브 제도 등 구체적인 활성화 방안을 제시하였다. 김규환과 김지현(2024)은 OAK(Open Access Korea) 사업 참여 기관들의 기관 리포지터리(Institutional Repository) 운영 관련 규정과 현황을 분석하였다.

OAK 사업에 참여하고 있는 다양한 유형의 기관별 IR 운영 정책을 비교 분석하여 국내 실정에 맞는 IR 운영 정책 수립의 필요성을 강조하였다. 이외에도 데이터 리포지터리의 신뢰성 및 인증 관련 연구(김주섭, 강효숙, 김선태, 2023; 이해림 외, 2024)들도 수행되고 있었다.

이렇게 리포지터리의 구축과 운영에 관한 초기 연구들이 주로 해외 사례 분석과 성공요인 분석, 국내 도입 필요성 제시에 초점을 맞추었다면, 최근의 연구들은 연구데이터와 같이 보다 세분화된 영역과 구체적인 운영 방안에 주목하고 있다.

리포지터리 콘텐츠 분석에 관한 연구로는 김현희, 정경희, 김용호(2006)의 연구와 백종명과 이은지(2013)의 연구를 들 수 있다. 김현희, 정경희, 김용호(2006)는 대학 기관 리포지터리의 콘텐츠 분석 및 개선방안에 관한 연구를 수행하였다. 이들은 국내 8개 대학 리포지터리의 콘텐츠를 분석하고, 이용자와 관리자를 대상으로 설문조사를 실시하여 개선방안을 도출하였다. 연구결과, 콘텐츠의 다양성 확보와 품질 관리, 이용자 인터페이스 개선 등이 필요한 것으로 나타났다. 백종명과 이은지(2013)의 연구는 국내 11개 D-Space 기반 리포지터리의 콘텐츠 현황을 조사하고 분석하여 콘텐츠 품질 관리 개선방안을 제시하였으며, 국내 리포지터리에 구축된 콘텐츠의 유형과 특성을 체계적으로 분석하여 품질관리 방안을 제시하였다.

리포지터리의 영향력과 오픈엑세스 동향에 관한 선행연구로는 신은자(2017)의 연구가 있다. OpenDOAR(Directory of Open Access Repositories), ROAR(Registry of Open Access Repositories), RWR(Ranking Web of Repositories) 등의 사이트를 활용하여 국내 대학 리포지터리의 현황을 파악하고, 이를 아시아 주요국과 비교 분석하였다. 연구 결과, 일본이나 대만과 비교했을 때 국내 대학 리포지터리의 수가 현저히 적고, 국제적 영향력도 낮은 것으로 드러났다. 이외에도 정영미(2014)는 국내 14개 OAK(Open Access Korea) 리포지터리의 현황을 조사하고, 이를 일본의 사례와 비교 분석하였다. 연구결과, 콘텐츠 구축과 원문 제공 측면에서 국내 리포지터리의 발전 속도가 더딘 것으로 드러났다.

국내에서는 리포지터리의 연구 동향을 살펴본 연구는 없었으나, 해외에서는 리포지터리 관련 연구 동향과 관련하여 다수의 선행연구가 수행되어왔다. Pinfield 외(2014)는 2005년부터 2012년까지의 전 세계 오픈엑세스 리포지터리 성장을 OpenDOAR 데이터를 통해 분석하였다. 초기에는 미국, 영국, 독일, 호주, 일본 중심에서 2010년 이후 대만, 브라질, 폴란드 등으로 확대되었음을 밝혔으며, 대부분의 리포지터리가 기관 단위, 다학제적, 영어 기반이며 오픈소스 OAI 호환 소프트웨어를 사용하는 것으로 나타났다. Ammarukleart와 Kim(2017)은 2005년부터 2015년까지의 기관 리포지터리와 관련된 연구 603편을 분석하여 연구 동향을 파악하였다. 분석 결과, 연구 생산량이 꾸준히 증가하였으며, 2011-2015년에는 연구데이터 관리, 링크드 오픈 데이터 등 새로운 연구 주제가 등장한 것으로 나타났다. Jain(2011)은 기관 리포지터리 구축의 장단점, 사서와 저자의 역할, 오픈엑세스에 대한 문헌들을 검토하였으며, 기관 리포지터리가 학술 커뮤니케이션의 중요한 도구이자 기관의 가시성을 높이는 지식

관리 도구로 인식되고 있음을 확인하였다. Zibani, Rajkoomar, Naicker(2022)는 2015년부터 2021년까지의 문헌을 체계적으로 분석하여 고등교육기관의 연구 리포지터리 현황을 평가하였으며, 학과 단위 연구 리포지터리의 개발이 학술 연구의 공유, 분석, 평가 및 보존을 강화하는 데 기여하는 것으로 나타났다.

국내에서는 리포지터리와 연관된 오픈엑세스 분야의 연구 동향이나 지적구조를 살펴본 연구는 다수가 수행되어왔다. 초기에는 전반적인 오픈엑세스의 동향을 통계 분석하거나(윤희윤, 김신영, 2007), 동시출현단어 분석을 사용하여 연구 동향을 살펴보았으며(서선경, 정은경, 2013), 이후 프로파일링 기법(김판준, 2021)이나 토픽 모델링을 활용한 연구(정재민, 김원중, 2022), 해외 학술지의 오픈엑세스 전환과 관련된 동향 분석 연구(이재윤, 정경희, 2020) 등으로 그 범위가 확대되었다.

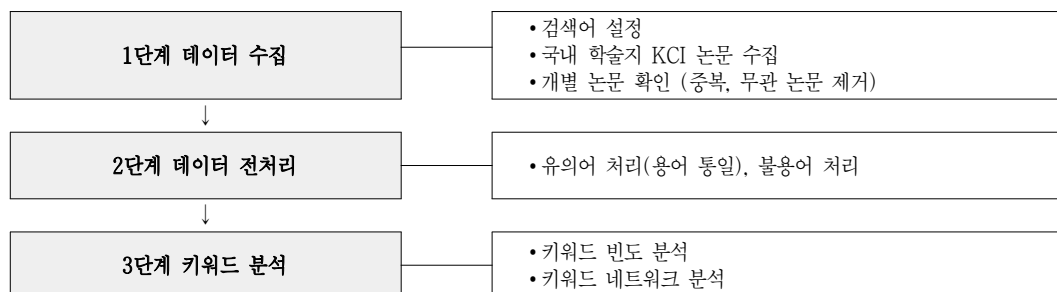
연관 분야인 연구데이터에 관한 연구로, 한상우(2023)는 연구데이터 관련 국내 연구의 동향을 파악하기 위해 키워드 네트워크 분석을 수행하였다. 분석 결과, 연구데이터 관련 연구는 주로 문헌 정보학을 중심으로 한 복합학 분야(81%)에서 수행되었으며, ‘연구데이터관리’, ‘연구데이터공유’, ‘데이터 리포지터리’, ‘오픈사이언스’ 등이 주요 키워드로 도출되었다.

연구 동향과 관련된 선행연구들은 전반적으로 최근 들어 오픈엑세스 출판이 빠르게 증가하고 있다는 점을 언급하고 있으며, 향후 리포지터리와 오픈엑세스의 중요성이 더욱 커질 것임을 언급하고 있었다.

이와 같은 선행연구들은 국내 리포지터리의 구축 현황과 운영 실태를 파악하고, 활성화를 위한 방안을 제시하는 데 초점을 맞추고 있었다. 그러나 최근의 리포지터리와 관련된 동향을 반영한 포괄적인 분석은 아직 부족한 실정이다. 따라서 본 연구에서는 최신 데이터를 바탕으로 국내 리포지터리의 전반적인 동향을 분석하고자 한다.

Ⅲ. 연구방법

국내 리포지터리 연구의 최신 동향을 분석하기 위해 논문의 키워드를 수집하여 키워드 빈도 분석과 키워드 네트워크 분석을 실시하고자 한다. 본 연구의 분석 과정은 <그림 1>과 같다.

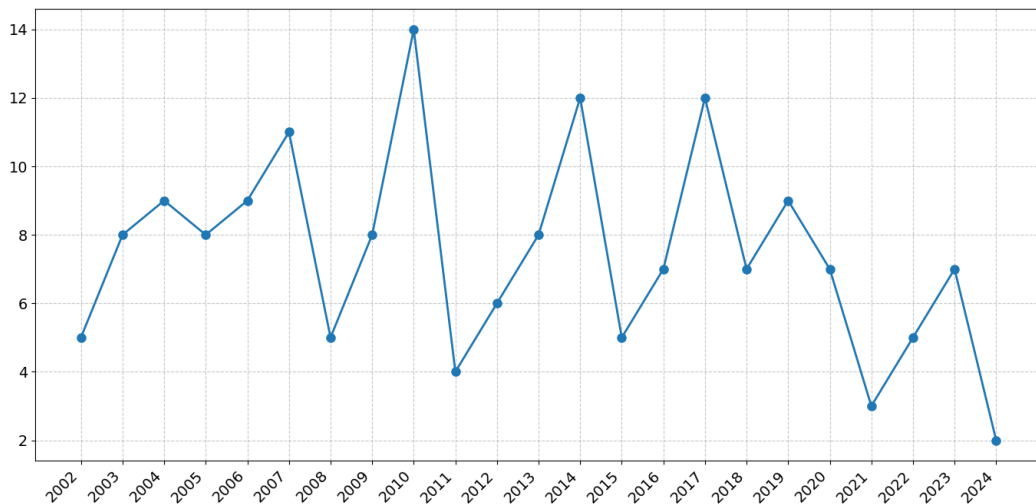


<그림 1> 연구방법의 개요

1. 데이터 수집

데이터 수집을 위해 한국 학술지 인용색인(KCI)을 이용하였으며, ‘리포지터리’와 관련된 논문을 수집하였다. KCI는 국내 학술지 논문에 대한 종합적인 색인 데이터베이스로, 다양한 학문 분야의 연구논문을 포함하고 있어 리포지터리 관련 연구를 광범위하게 수집하기에 적합하다고 판단하였다. ‘리포지터리’ 용어의 명칭은 연구에 따라 다양하게 사용하고 있었다. 리포지터리와 유사한 의미로 사용되는 용어를 논문에서 찾아 검색어로 7개를 선정하였다. 검색어는 ‘리포지터리’, ‘리포지토리’, ‘레포지터리’, ‘레포지토리’, ‘지식정보저장소’, ‘IR’, ‘repository’이다. 검색 필드의 논문제목과 키워드에 7개의 검색어를 설정하여 총 2,144건의 논문을 수집하였다. 수집된 2,144건의 논문에 대해 개별적인 검토 과정을 진행하였다. 중복된 논문과 연구주제와 무관한 논문을 제거한 결과, 최종적으로 244건의 리포지터리 관련 논문을 분석 대상으로 설정하였다.

244건의 논문에 대한 분야별 분포를 분석한 결과, 대부분의 분야에서 10건 이하의 연구가 수행되었다. 이에 본 연구에서는 상위 4개 분야(문헌정보학, 컴퓨터학, 전자/정보통신공학, 공학)에 집중하여 연구 동향을 파악하고자 하였다. 상위 4개 분야의 논문은 총 176건이었으며, 이 중 원문에 저자 키워드가 작성되지 않은 4건을 제외하였다. 결과적으로 172건의 논문이 본 연구의 최종 분석 대상으로 선정되었다. 연도별 논문 수의 추이를 살펴보면 다음 〈그림 2〉와 같다.



〈그림 2〉 연도별 논문 수의 추이

2000년대 초반(2002-2007)에는 연간 5-11편 사이의 논문이 꾸준히 발표되며 점진적인 증가 추세를 보였다. 특히 2007년에는 11편으로 이 시기 최고치를 기록하였다. 이는 당시 학술연구에서

관련 분야에 대한 관심이 점차 높아지고 있었음을 시사한다. 이후 2008-2010년에는 일시적인 정체를 겪었으며, 연간 4-8편 사이로 논문 발표 수가 감소하는 경향을 보였다. 2010년에 14편으로 일시적인 증가를 보였으나, 이후 다시 감소세로 돌아섰다. 2011-2020년 구간에서는 다시 연구 활동이 활발해지는 모습을 보였다. 특히 2014년과 2017년에는 각각 12편의 논문이 발표되어 연구 활동이 상대적으로 활발한 시기였다. 2021-2024년 전반부는 앞선 시기에 비해 상대적으로 논문 발표 수가 감소하는 경향을 보였으며, 연간 2-5편 수준으로 나타났다.

2. 데이터 전처리

분석 대상인 172건의 논문에서 총 643개의 저자 키워드를 추출하였고, 분석 시 키워드의 통일성을 위해 키워드 전처리 과정을 수행하였다. 첫째, 한글(영어) 또는 (영어)한글로 키워드가 작성되어 있는 경우, 괄호 안의 키워드를 제거하였다. 둘째, ‘리포지터리’는 검색어로 설정해 논문 수집 시 사용한 키워드이므로 모든 논문에 포함되어 있어 불용어로 처리하였다. 셋째, 동일한 의미로 사용된 한글, 영어 키워드를 모두 포함하여 하나의 키워드로 통일하였다. 예를 들어 ‘기관 리포지토리’, ‘institutional repository’는 ‘기관 리포지터리’로, ‘오픈 액세스’, ‘open access’는 ‘오픈엑세스’로, ‘데이터 저장소’, ‘idr’는 ‘데이터 리포지터리’로, ‘공동보존소’, ‘공동보존자료관’은 ‘공동보존서고’로, ‘셀프아카이빙’, ‘저자 셀프 아카이빙’ 등은 ‘셀프 아카이빙’으로 수정하였다. 총 80개의 키워드를 유의어 처리하였으며(〈표 1〉 참조), 데이터 전처리 과정을 거쳐 643개의 키워드 중 총 565개의 키워드를 추출하여 연구에 사용하였다.

〈표 1〉 저자 키워드 전처리 주요 예시

구분	사례	전처리 결과
유의어 처리	오픈 액세스, 오픈 액세스, open access	오픈엑세스
	오픈 액세스 잡지, 오픈 액세스 학술지	오픈엑세스 저널
	공동보존소, 공동보존자료관	공동보존서고
	연구 데이터, research data	연구데이터
	데이터리포지터리, 데이터 레포지토리, 데이터 리포지토리, 데이터 저장소, data repository, idr	데이터 리포지터리
	기관 리포지토리, 기관리포지터리, 기관리포지토리, 기관 레포지터리, 기관 레포지토리, 기관레포지토리, institutional repository, ir	기관 리포지터리
	연구데이터 레포지터리, 연구데이터 리포지터리, 연구데이터 리포지토리, rdr	연구데이터 리포지터리
	data management plan, dmp, 데이터관리계획	데이터 관리 계획
	셀프아카이빙, 셀프아카이브, 저자 셀프 아카이빙, self archiving	셀프 아카이빙
	oak 센트럴, oak 포털, 오픈엑세스코리아	oak
불용어 처리	리포지터리, 리포지토리, 리파지토리, 레포지터리, 레파지토리, 저장소, 정보저장소, 정보 저장소, 지식정보저장소, repository	검색키워드 삭제

3. 키워드 빈도 및 네트워크 분석

키워드 빈도 분석과 키워드 네트워크 분석을 이용하여 리포지터리 관련 키워드의 빈도와 키워드 간의 연결된 네트워크를 분석하였다. 키워드 빈도 분석은 키워드가 텍스트 내에서 얼마나 자주 사용되었는지를 측정하는 분석기법으로(이다현, 신동희, 2024), 논문에서 빈번하게 등장하는 키워드를 빈도수로 파악하여 주요 키워드를 확인하고자 하였다.

또한, 키워드 간의 관계와 키워드 클러스터를 파악하기 위해 키워드 네트워크 분석을 활용하였다. 이 분석기법은 단어와 단어 사이의 연결 관계를 네트워크로 시각화하여 분석 대상의 의미를 해석하는 방법으로(한상우, 2023), 특히 네트워크 클러스터 분석을 통해 연구 대상의 다양한 유형과 특성을 체계적으로 분류하고 파악할 수 있다는 장점이 있다. 본 연구에서는 논문의 저자 키워드를 대상으로 네트워크 분석을 통해 링크로 표시된 키워드 간의 연결 관계를 심층적으로 분석하고자 하였다. 저자 키워드를 노드로 하며 키워드 간의 공기빈도를 엣지로 하는 네트워크를 구축하였으며, 네트워크에 나타난 키워드의 클러스터를 파악하여 주요 연구주제를 식별하였다. 네트워크 내의 키워드의 중요도를 정량적으로 평가하기 위해 페이지랭크(PageRank) 알고리즘을 적용하여 분석을 수행하였다. 최종적으로 2회 이상 등장한 단어를 기준으로 키워드 네트워크를 구축하여 69개의 키워드 노드, 공기빈도에 기반한 185개의 엣지, 7개의 네트워크 클러스터를 식별하였다.

본 연구의 네트워크 분석을 위해 VOSviewer와 Gephi 두 가지 도구를 활용하였다. VOSviewer는 서지데이터를 기반으로 네트워크를 구성하고 시각화하는 도구로, 학술 문헌의 키워드 네트워크를 활용한 연구 동향 파악에 효과적이다. Gephi는 다양한 유형의 네트워크 분석과 시각화가 가능한 오픈소스 도구로, 중심성 분석(연결정도, 매개, 근접 중심성), 클러스터링 계수 등 다양한 네트워크 지표 계산이 가능하다. 이를 활용하여 키워드 네트워크의 구조적 특성을 분석하고자 하였다.

IV. 연구 결과

1. 키워드 빈도 분석

논문에 등장하는 총 565개 키워드의 빈도를 분석하였다. 빈도수 10회 이상인 키워드는 8개였으며, 이외에는 모두 빈도수가 10회 미만이었다. 총 565개의 키워드 중 빈도수 1회인 키워드가 가장 많았으며, <표 2>에서 빈도수 4회 이상의 키워드 18개를 살펴보았다.

<표 2>를 통해 리포지터리 관련 연구의 주요 키워드를 살펴보면, 기관 리포지터리가 39회로 가장 높은 빈도를 보였으며, 오픈엑세스(26회), 공동보존서고(16회), 데이터 리포지터리(13회),

xml(12회) 순으로 나타났다. 빈도수 4회 이상의 키워드를 전체적으로 살펴본 결과, 자료의 보존, 데이터 관리 및 공유와 관련된 키워드임을 유추할 수 있었다.

〈표 2〉 빈도수 4회 이상의 키워드 (총 18개)

순위	키워드	빈도수	순위	키워드	빈도수
1	기관 리포지터리	39	9	대학도서관	7
2	오픈엑세스	26	9	지역대표도서관	7
3	공동보존서고	16	12	자료보존	6
4	데이터 리포지터리	13	12	re3data	6
5	xml	12	14	공공도서관	5
5	메타데이터	12	14	보존서고	5
7	연구데이터	10	14	dcollection	5
7	셀프 아카이빙	10	17	coretrustseal	4
9	공동보존도서관	7	17	연구데이터 리포지터리	4

특히 자료의 보존 측면에서 리포지터리 연구가 활발히 진행되었음을 알 수 있었다. 공동보존서고(16회), 공동보존도서관(7회), 자료보존(6회), 보존서고(5회) 등의 키워드가 높은 빈도로 출현하였다. 이를 통해 도서관 자료의 효율적인 보존이 리포지터리 연구의 주요 관심사였음을 알 수 있다. 특히 ‘공동보존’이라는 용어가 연구에 반복적으로 등장하는 것은 개별 도서관 차원을 넘어선 협력적 보존 체계에 대한 연구가 중점적으로 이루어졌음을 추측해볼 수 있다.

데이터 관리 및 공유 측면의 연구도 중요한 축을 형성하고 있었다. 데이터 리포지터리(13회), 메타데이터(12회), 연구데이터(10회), re3data(6회) 등은 디지털 환경에서의 데이터 관리와 공유에 대한 관심을 반영한 키워드이다. 특히 연구데이터와 메타데이터에 대한 관심은 학술 커뮤니케이션의 변화와 오픈 사이언스의 움직임을 반영한 것이라고 볼 수 있다.

주목할 만한 또 다른 점으로는 도서관 유형별 키워드의 분포이다. 대학도서관(7회), 지역대표도서관(7회), 공공도서관(5회) 등 다양한 관종의 도서관이 연구 대상으로 다루어졌다. 이는 리포지터리 구축과 운영이 도서관의 유형에 관계없이 보편적인 관심사로 자리 잡았음을 보여준다.

이러한 키워드 분석 결과는 국내 리포지터리 연구가 물리적 자료의 보존이라는 전통적 관심사와 디지털 데이터의 관리라는 새로운 주제를 다양하게 다루고 있었음을 보여준다. 또한 다양한 관종의 도서관을 연구 대상으로 포함함으로써, 이론과 실무를 연계하는 실용적 연구가 이루어지고 있음을 알 수 있다.

다음으로는 시기별 주요 키워드와 해당 키워드의 빈도 분석을 진행하였다(〈표 3〉 참조). 5년 단위로 논문을 분류하여 해당 시기별 주요 키워드를 정리하였다.

〈표 3〉 시기별 주요 키워드 빈도 분석

시기	논문 수	주요 키워드
2002-2005년	31	xml(8회), 기관 리포지터리(4회), 메타데이터(3회), 셀프 아카이빙(3회), 관계정보(2회), 디자인패턴(2회)
2006-2010년	47	기관 리포지터리(15회), 오픈엑세스(9회), xml(4회), dcollection(3회), 메타데이터(3회), 셀프 아카이빙(3회)
2011-2015년	35	기관 리포지터리(9회), 공동보존서고(8회), 공동보존도서관(6회), 오픈엑세스(5회), 공공도서관(4회), 지역대표도서관(4회)
2016-2020년	42	기관 리포지터리(9회), 데이터 리포지터리(6회), 오픈엑세스(6회), 공동보존서고(5회), 메타데이터(5회), 연구데이터(5회)
2021-현재	17	데이터 리포지터리(5회), 연구데이터(4회), 오픈엑세스(4회), coretrustseal(3회), re3data(3회), 공동보존서고(3회)

시기별 논문 수를 살펴본 결과, 가장 많은 연구가 이루어진 시기는 2006년에서 2010년 사이인 것으로 나타났다. 해당 시기에 나타난 주요 키워드는 ‘기관 리포지터리’, ‘오픈엑세스’, ‘셀프 아카이빙’ 등으로 리포지터리와 관련된 주요 개념들이 소개되고 정착된 시기이다. 특히, ‘기관 리포지터리’는 시기 전반에 걸쳐 가장 많이 연구되는 핵심 키워드로 나타났으나, 2021년 이후에는 ‘기관 리포지터리’ 대신 ‘데이터 리포지터리’가 핵심 키워드가 되는 것을 확인할 수 있었다. 데이터 리포지터리는 연구데이터를 수집하고 관리하여 보존을 통해 재사용할 수 있는 플랫폼으로(김주섭, 강효숙, 김선태, 2023), 연구데이터의 중요성이 강조되면서 리포지터리 분야에도 이를 반영한 연구가 증가하고 있었음을 나타낸다.

2000년부터 2010년 시기 사이, xml를 키워드로 사용한 연구가 다수 진행되었다. 관련 연구에서 xml과 함께 해당 시기에 진행된 키워드를 살펴본 결과, 주로 xml 저장소, 객체저장소, ebxml 저장소와 같이 사용되었다. 그러나, xml를 키워드로 사용한 연구가 2010년 이후에는 수행되지 않았다. 또한 꾸준한 빈도로 시기별 주요 키워드에 포함되고 있는 키워드로는 메타데이터, 셀프 아카이빙 등이 있었다.

공동보존서고와 관련된 연구는 주로 2011년부터 현재까지 진행되고 있으며, 그중 2011년부터 2015년 사이에 공동보존서고의 건립과 관련한 연구(서혜란, 김선애, 강은영, 2015; 윤희윤, 2013; 2014)가 가장 활발하게 수행되었다. 2011년부터 2015년 사이의 시기에 공동보존서고와 함께 주요 키워드로 ‘공공도서관’, ‘지역대표도서관’이 등장하였으며, 이를 통해 이 2개의 키워드가 공동보존서고의 도입과 관련된 장소의 역할을 하는 것으로 추측해 볼 수 있다.

오픈엑세스는 2000년부터 현재까지 모든 시기에 걸쳐 주요 키워드로 유지되었다. 리포지터리는 오픈엑세스의 대표적인 실천 방안으로 제시되고 있으므로, 리포지터리 연구에서 오픈엑세스가 리포지터리와 함께 자주 사용되는 경향이 있었다.

2. 키워드 네트워크 분석 결과

가. 키워드 네트워크 클러스터 분석

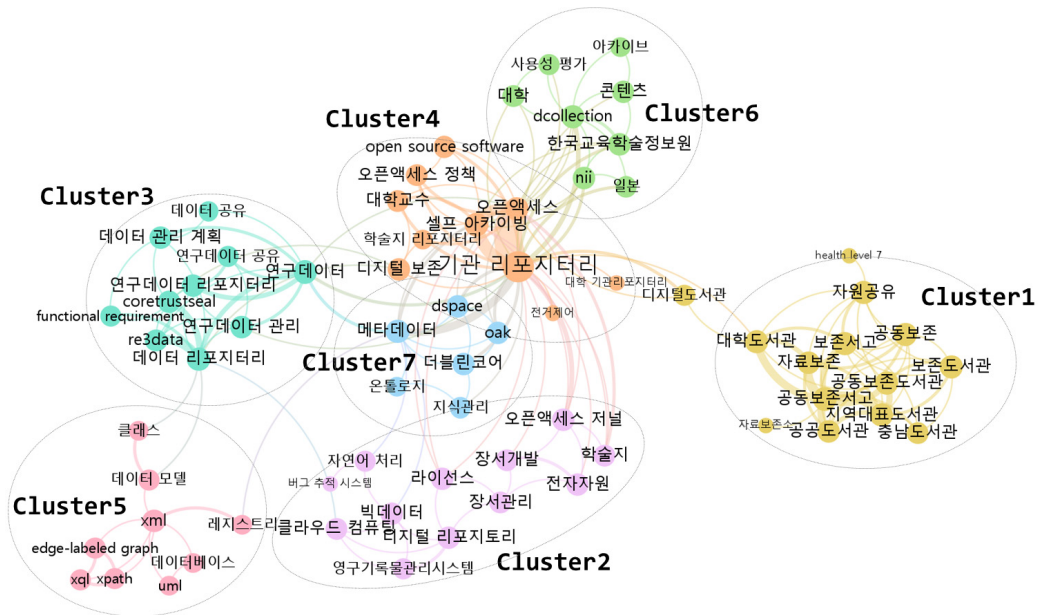
키워드 간의 연결 관계와 네트워크를 파악하기 위해 빈도수 2회 이상 출현한 단어를 기준으로 공기빈도를 측정하여 키워드 네트워크 분석을 진행하였다. 결과적으로 키워드 네트워크 분석을 통해 총 69개의 키워드 노드, 185개의 엣지와 총 7개의 네트워크 클러스터를 파악하였다. 클러스터는 키워드 간의 동시출현 빈도를 정규화한 유사성을 기반으로 추출되었다(Waltman, van Eck, & Noyons, 2010). 다음 <표 4>는 각 클러스터의 대표 키워드 및 클러스터에 속한 키워드의 수이다.

<표 4> 클러스터의 대표 키워드 및 키워드 수

클러스터	대표 키워드	키워드 수
1	공동보존서고, 대학도서관, 지역대표도서관, 공동보존도서관, 자료보존, 디지털 도서관 등	14개
2	디지털 리포지토리, 오픈엑세스 저널, 전자자원, 라이선스, 빅데이터, 장서개발 등	12개
3	데이터 리포지터리, 연구데이터, re3data, coretrustseal, 연구데이터 리포지터리 등	10개
4	기관 리포지터리, 오픈엑세스, 셀프 아카이빙, 대학교수, 오픈엑세스 정책 등	10개
5	xml, 데이터 모델, 레지스트리, xql, xpath 등	9개
6	dcollection, nii, 한국교육학술정보원, 대학, 일본 등	8개
7	메타데이터, dspace, 온톨로지, oak, 더블린코어, 지식관리	6개

<표 4>의 네트워크 클러스터 키워드를 살펴보면, 7개의 각 클러스터는 서로 다른 주제적 영역을 대표하고 있었다. 클러스터 1은 공동보존서고, 대학도서관, 지역대표도서관, 공동보존도서관 등 14개의 키워드로 구성되어 도서관 자료의 공동보존 및 관리가 국내 리포지터리 연구에서 가장 큰 영역을 차지하는 주제임을 알 수 있었다. 클러스터 2는 디지털 리포지토리, 오픈엑세스 저널, 전자자원, 라이선스 등 12개의 키워드를 포함하고 있었으며, 이를 통해 디지털 자원 관리와 관련된 내용이 도서관 자료의 공동보존 및 관리 영역과 더불어 두 번째로 많이 연구되는 영역임을 알 수 있었다. 다음으로 연구데이터 및 관리와 관련된 클러스터 3(데이터 리포지터리, 연구데이터, re3data, coretrustseal 등), 기관의 오픈엑세스 정책을 다루는 클러스터 4(기관 리포지터리, 오픈엑세스, 셀프 아카이빙 등)가 각각 10개의 키워드로 구성된 주요 클러스터로 나타났다. 클러스터 5는 xml, 데이터 모델, 레지스트리 등 9개의 키워드로 데이터 표준 기술을 다루었으며, 클러스터 6은 dcollection, nii, 한국교육학술정보원 등 8개의 키워드로 교육 학술정보 시스템이 주요 주제로 나타났다. 마지막으로 클러스터 7은 메타데이터, dspace, 온톨로지 등 6개의 키워드로 디지털 도서관의 지식 관리 시스템과 관련된 주요 연구 영역으로 나타났다.

다음으로 네트워크 시각화를 통해 클러스터 간의 관계를 살펴보았다. 아래 <그림 3>은 리포지터리 관련 연구의 키워드 네트워크이다.



<그림 3> 리포지터리 관련 연구의 키워드 네트워크

네트워크 시각화 결과, 네트워크 우측에는 공동보존서고 관련 클러스터가 독립적으로 위치하고 있었고, 좌측에 나머지 6개의 클러스터가 밀집해 있는 구조를 보였다. 즉, 클러스터 1의 구조적 특성이 다른 클러스터들과 뚜렷한 차이를 보이는 것으로 나타났다. 클러스터 1은 내부 키워드 간의 연결 밀도가 상대적으로 높은 반면, 다른 클러스터와의 외부 연결은 제한적이었다. 네트워크의 중심에는 ‘기관 리포지터리’가 위치하고 있었으며, 이는 전체 네트워크에서 가장 높은 연결중심성을 보이는 핵심 키워드였다. 또한 ‘기관 리포지터리’는 다수의 키워드와 링크로 연결되어 있어 높은 매개 역할을 수행하고 있었다. 이는 기관 리포지터리가 다른 연구 주제들과 결합하여 다양한 맥락에서 연구되고 있음을 의미한다. ‘대학도서관’과 ‘디지털도서관’ 두 키워드만이 다른 클러스터의 키워드들과 연결관계를 형성하고 있었다. 이는 클러스터 1에 포함된 연구들이 상대적으로 독립적인 연구 영역을 형성하고 있음을 의미한다. 이 두 키워드는 동시에 클러스터 4(기관 리포지터리)의 ‘기관 리포지터리’와 ‘오픈엑세스’ 키워드와도 연결관계를 형성하고 있었다. 이는 디지털도서관이 가진 온라인 정보의 접근성이라는 특성이 기관 리포지터리와 오픈엑세스의 개념과 밀접하게 연관되어 있다는 것을 의미한다. 즉, 공동보존서고 관련 연구가 독자적인 연구 영역을 형성하고 있으면

서도, 기관 리포지터리를 매개로 하여 디지털 환경에서의 정보 관리 및 접근성 측면과 연계되어 있다는 것을 추측해볼 수 있다.

좌측의 클러스터 4(기관 리포지터리)를 중심으로 다른 클러스터들과의 연결이 활발하게 이루어지고 있었다. 특히, 클러스터 4와 클러스터 7(메타데이터), 클러스터 6(dcollection) 간의 연결 강도가 두드러지게 나타났다. 이는 기관 리포지터리와 관련하여 메타데이터 구축 및 dcollection 시스템 활용에 관한 연구가 긴밀하게 연계되어 수행되어 왔음을 보여준다. 이러한 네트워크 구조는 리포지터리 연구에서 기관 리포지터리를 중심으로 메타데이터와 시스템 구축 등 기술적 측면과의 연계 연구가 활발히 이루어져 왔음을 의미한다.

반면 공동보존서고 관련 연구는 상대적으로 독립적인 연구 영역을 형성하고 있는 것으로 나타났다. ‘공동보존서고’는 가장 높은 페이지랭크 값과 연결정도 중심성을 보인 키워드로(〈표 5〉 참조), 이 키워드가 해당 연구 영역에서 핵심적인 연구 주제로 다루어져 왔으며, 관련 연구들의 중심점 역할을 하는 것임을 알 수 있다.

〈표 5〉 클러스터 1의 상위 10개 키워드 및 페이지랭크

키워드	페이지랭크	네트워크
공동보존서고	0.03827	
지역대표도서관	0.02058	
자료보존	0.01973	
보존서고	0.01841	
공동보존도서관	0.01703	
대학도서관	0.01576	
공공도서관	0.01465	
자원공유	0.01248	
공동보존	0.01148	
보존도서관	0.01083	

다음 〈표 6〉은 클러스터 2(디지털 리포지토리)의 상위 10개 키워드를 보여준다. 페이지랭크 분석 결과, 클러스터 2 내에서 ‘디지털 리포지토리’(0.01490)가 가장 높은 중심성을 보였으며, ‘빅데이터’(0.01311), ‘오픈엑세스 저널’(0.01299) 순으로 높은 값을 보였다. 전체적으로 클러스터 1(공동보존서고)과 유사하지만, 주로 클러스터 2 내의 키워드끼리 연결된 경우가 많았다. 주목할 만한 점은 ‘디지털 리포지토리’가 ‘라이선스’를 매개로 하여 클러스터 4의 주요 키워드들(‘기관 리포지터리’, ‘오픈엑세스’, ‘셀프 아카이빙’)과 연결되어 있다는 점이었다. 이러한 연결은 디지털 리포지토리 연구에서 저작권 관리와 오픈엑세스가 밀접하게 연관되어 있음을 나타낸다. 특히 ‘라이선스’가 매개 키워드로서 중요한 역할을 하고 있는데, 정보자원의 저작권 관리가 디지털 리포지

토리와 오픈엑세스 연구를 연결하는 핵심 요소임을 추측해볼 수 있다.

〈표 6〉 클러스터 2의 상위 10개 키워드 및 페이지랭크

키워드	페이지랭크	네트워크
디지털 리포지토리	0.01490	
빅데이터	0.01311	
오픈엑세스 저널	0.01299	
영구기록물관리시스템	0.01083	
장서관리	0.01025	
전자자원	0.01025	
클라우드 컴퓨팅	0.01025	
자연어 처리	0.01016	
라이선스	0.00982	
장서관리	0.00945	

클러스터 3(데이터 리포지터리)에서는 ‘데이터 리포지터리’(0.03650)와 ‘연구데이터’(0.03001)가 높은 페이지랭크 값을 나타냈다(〈표 7〉 참조). 이는 데이터 리포지터리가 연구데이터의 수집, 관리, 보존을 위한 아카이브로서의 기능적 연관성을 반영한 결과로 해석된다. 클러스터 3 내의 다른 키워드들(‘re3data’ 0.01776, ‘coretrustseal’ 0.01541 등)은 상대적으로 낮은 페이지랭크 값을 보여, 이 분야의 연구가 주로 데이터 리포지터리와 연구데이터를 중심으로 이루어지고 있음을 확인할 수 있었다. 클러스터 간의 연결 구조를 살펴보았을 때, ‘연구데이터’는 클러스터 4의 ‘기관 리포지터리’와 클러스터 7의 ‘메타데이터’와 직접적인 연결을 형성하고 있었다. 반면에 중심성 값이 가장 높게 나타난 ‘데이터 리포지터리’는 상대적으로 제한된 외부 연결을 보였다.

〈표 7〉 클러스터 3의 상위 10개 키워드 및 페이지랭크

키워드	페이지랭크	네트워크
데이터 리포지터리	0.03650	
연구데이터	0.03001	
re3data	0.01776	
coretrustseal	0.01541	
연구데이터 리포지터리	0.01353	
연구데이터 관리	0.01183	
데이터 관리 계획	0.01163	
연구데이터 공유	0.00916	
functional requirement	0.00809	
데이터 공유	0.00516	

다음으로 클러스터 4(〈표 8〉 참조)의 주요 노드를 살펴보면, ‘기관 리포지터리’가 0.10060으로 전체 네트워크에서 가장 높은 페이지랭크 값을 나타냈으며, ‘오픈엑세스’가 0.06253으로 두 번째로 높았다. 특히 ‘기관 리포지터리’는 전체 네트워크 내에서 다른 키워드들과의 동시출현이 가장 많았던 노드로, 가장 높은 연결중심성을 보였다. 클러스터 4 내의 다른 키워드들(‘셀프 아카이빙’ 0.03841, ‘대학교수’ 0.00955 등)은 상대적으로 낮은 페이지랭크 값을 보였다. 네트워크의 구조적 측면에서는, 클러스터 4가 높은 중심성을 가진 키워드들을 포함하고 있음에도 불구하고, 클러스터 5(xml)와는 연결관계가 형성되어 있지 않았다.

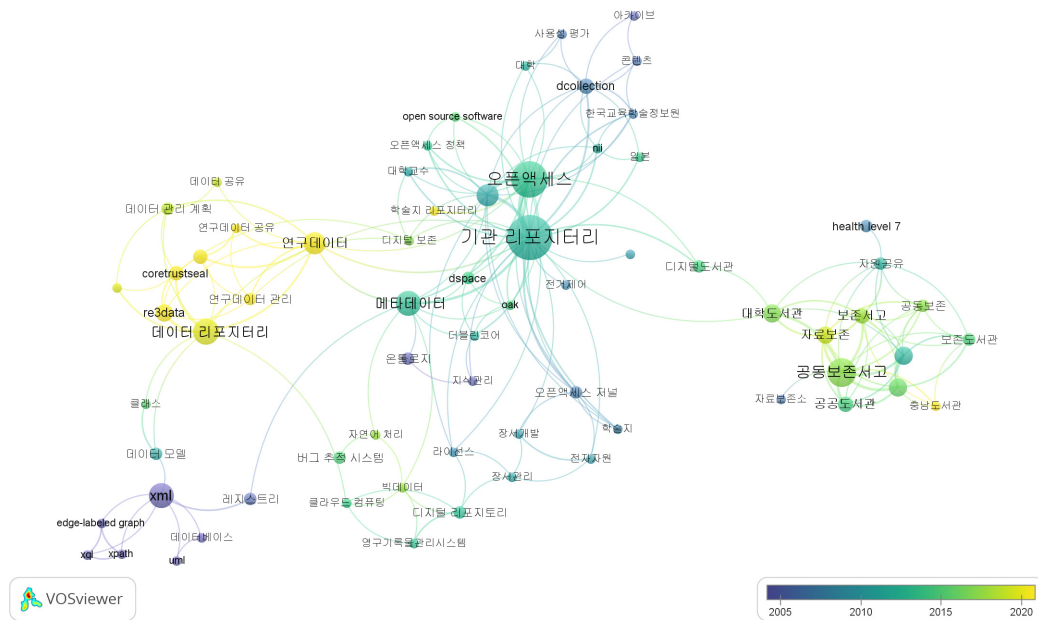
〈표 8〉 클러스터 4의 상위 10개 키워드 및 페이지랭크

키워드	페이지랭크	네트워크
기관 리포지터리	0.10060	
오픈엑세스	0.06253	
셀프 아카이빙	0.03841	
대학교수	0.00955	
오픈엑세스 정책	0.00945	
디지털 보존	0.00876	
open source software	0.00823	
학술지 리포지터리	0.00564	
대학 기관리포지터리	0.00337	
전자제어	0.00337	

이외에도 클러스터 5의 ‘xml’은 ‘레지스트리’를 매개 키워드로 클러스터 7의 ‘메타데이터’와 연결되어 있었으며, ‘데이터 모델’을 매개로 하여 클러스터 3의 ‘데이터 리포지터리’와 연결되어 있었다. 그러나, 이를 제외한 클러스터 5의 대부분은 타 클러스터 및 키워드와는 연결되어 있지 않았다. 클러스터 6(dcollection)에서는 특정한 대학의 dcollection과 직접적으로 연관된 키워드들로 구성되어 있었으며, 주로 오픈엑세스와 관련된 클러스터 4와의 연결이 많았다. 클러스터 7(메타데이터) 또한, 해당 클러스터 내의 키워드 중 온톨로지를 제외하고 모두 클러스터 4의 기관 리포지터리와 연결되어 있었다.

나. 키워드의 등장 시기 분석

다음으로는 〈그림 4〉를 통해 해당 시기에 빈번하게 등장한 키워드를 색으로 표현하였다. 〈그림 4〉의 범례에 표시되어 있듯이, 짙은 보라색의 경우 초반에 빈번하게 등장한 단어를 의미하며, 밝은 노란색으로 표시된 경우 최근에 사용되기 시작한 단어를 의미한다.



〈그림 4〉 시기별 리포지터리 네트워크 지도

시기별 키워드 네트워크 분석을 통해 리포지터리 연구의 시간적 변화와 주제적 진화 양상을 살펴보면 초기 연구 단계인 2005년경에는 ‘xml’, ‘xpath’, ‘dcollection’ 등 기술적 측면의 키워드가 주로 등장하였다. 이는 리포지터리 구축을 위한 기반 기술에 초점을 맞춘 연구가 진행되었음을 의미한다. 2010년부터 2020년 사이는 리포지터리 연구가 가장 활발하게 이루어진 시기로, 특히 이 시기에는 ‘기관 리포지터리’와 ‘오픈엑세스’가 핵심 키워드로 나타났다. 이는 2002년 Budapest Open Access Initiative(BOAI) 선언 이후, 학술지 구독료 급등과 복잡한 심사과정에 대한 대안으로서 오픈엑세스 패러다임이 확산된 것과 같은 맥락에서 이해할 수 있다(Budapest Open Access Initiative, 2002). 가장 최근에는 ‘연구데이터’, ‘데이터 리포지터리’, ‘공동보존서고’ 등의 키워드가 새롭게 부상한 것으로 나타났다. 이러한 변화는 과학기술정보통신부(2023)의 ‘국가연구데이터 관리 및 활용 촉진에 관한 법률’ 제정안 입법예고와 같은 정책적 변화의 맥락으로 볼 수 있다. 특히 연구데이터의 가치 증대와 함께 체계적인 공유 체계의 필요성이 대두되면서, 데이터 리포지터리와 연구데이터 관련한 연구가 증가하는 추세를 보이고 있었다.

요약하면, 리포지터리 연구가 기술적 기반 구축에서 시작하여, 기관 리포지터리와 오픈엑세스를 거쳐, 최근에는 연구데이터 관리와 보존으로 연구의 초점이 진화해 왔음을 보여준다.

V. 결론 및 제언

도서관, 대학, 공공기관을 포함한 모든 기관에서 정보자원을 효율적으로 관리하고 활용하기 위해서 리포지터리의 필요성은 점차 강조되고 있었으나, 국내의 리포지터리 도입 현황과 진행된 연구의 수는 부족한 실정이다. 따라서 본 연구는 KCI에서 수집한 리포지터리 관련 논문들을 대상으로 키워드 빈도 분석과 네트워크 분석을 활용하여 국내 리포지터리 연구의 동향을 체계적으로 분석하고자 하였다. 주요 연구 결과는 다음과 같다.

첫째, 키워드 빈도 분석 결과, 기관 리포지터리(39회), 오픈엑세스(26회), 공동보존서고(16회), 데이터 리포지터리(13회), xml(12회) 순으로 높은 빈도를 보였다. 시기별 분석에서는 초기의 기술 중심 키워드(xml, 관계정보)에서 최근에는 데이터 관리 중심 키워드(연구데이터, 데이터 리포지터리)로 연구 관심사가 변화하는 것으로 나타났다. 또한 꾸준한 빈도로 시기별 주요 키워드에 포함되어 있는 키워드로 메타데이터, 셀프 아카이빙 등이 있었다.

둘째, 네트워크 분석을 통해 총 7개의 클러스터가 도출되었으며, 공동보존서고 클러스터와 나머지 6개 클러스터가 뚜렷이 구분되는 구조를 보였다. 특히 ‘기관 리포지터리’와 ‘오픈엑세스’ 키워드가 가장 높은 연결 중심성을 보여, 리포지터리 연구의 핵심 주제로 자리잡고 있음을 확인하였다. 시기별 키워드 네트워크 분석을 통해 과거의 연구는 xml, dcollection과 같은 키워드가 중점적이었지만, 최근의 연구는 연구데이터, 데이터 리포지터리, 공동보존서고와 같은 키워드가 주를 이루고 있었다. 시기별 키워드의 빈도 분석과 네트워크 분석에서도 유사한 결과를 도출할 수 있었다.

셋째, 리포지터리의 다양한 형태로 기관 리포지터리, 데이터 리포지터리, 디지털 리포지토리, 연구 데이터 리포지터리 등이 있었으며, 최근의 연구 동향에서는 데이터 리포지터리와 연구데이터 관리에 대한 관심이 급증하고 있었다. 이는 데이터의 공정한 이용과 자연과학 분야를 중심으로 연구데이터의 체계적 관리와 공유에 대한 필요성이 증가하고 있는 경향이 반영되었음을 유추해볼 수 있다.

이러한 분석 결과를 바탕으로 다음과 같은 제언을 하고자 한다.

첫째, 리포지터리의 유형별 특성화 연구가 필요하다. 분석 결과에서 나타난 것처럼 기관 리포지터리, 데이터 리포지터리, 디지털 리포지터리 등 다양한 유형의 리포지터리가 존재하며, 각각의 목적과 기능이 다르다. 특히 네트워크 분석에서 도출된 7개 클러스터는 각 리포지터리 유형별로 서로 다른 특성과 주요 관심사가 상이함을 보여준다. 이러한 유형별 특성을 고려한 차별화된 발전 방안 연구를 통해 리포지터리 도입 기관의 확대에 기여할 수 있을 것이다.

둘째, 시기별 분석에서 나타난 연구 동향의 변화를 고려할 때, 데이터 관리 측면(연구데이터, 데이터 리포지터리)으로의 연구 확대가 필요하다. 과거에는 클러스터 5의 xml, 데이터 모델, 레지스트리와 클러스터 6의 dcollection 등 기술적 요소에 초점을 맞추었으나, 현재는 클러스터 3, 4에서 보이듯 데이터 리포지터리와 기관 리포지터리를 중심으로 한 정보 공유와 오픈엑세스 실현이 더욱

중요해지고 있는 것을 알 수 있다. 이는 선행연구에서 언급한 Ammarukleart와 Kim(2017)의 연구에서도 나타나는 결과로, 해외와 유사한 연구경향을 보여가는 것으로 볼 수 있다. 공동보존서가 보존 기능에 초점을 맞추는 것과 달리, 데이터 리포지터리는 정보자원의 공유와 활용이라는 측면에서 더 큰 잠재력을 가지고 있으므로, 데이터 관리 측면의 연구 확대가 더욱 필요한 시점이라고 할 수 있을 것이다.

이러한 제언점에도 불구하고 본 연구는 다음과 같은 한계점을 가진다. 분석 대상을 KCI 등재 논문으로 한정하여 학술대회 발표논문, 연구보고서 등 다양한 유형의 연구 성과물을 포함하지 못하였다. 또한 국내 리포지터리 연구만을 대상으로 분석을 수행하여, 글로벌 연구 동향과의 비교가 이루어지지 못했다는 한계를 가진다.

따라서 향후 연구에서는 Web of Science나 Scopus 같은 국제 인용 데이터베이스를 활용하여 해외 리포지터리 연구 동향과의 비교 분석이 필요하며, 이를 통해 국내 리포지터리 연구의 특성과 발전 방향을 보다 객관적으로 파악할 수 있을 것이다. 그럼에도 불구하고 본 연구는 국내 리포지터리 연구의 전반적인 동향을 체계적으로 분석했다는 점에서 의의가 있으며, 향후 리포지터리 연구와 도입 확대를 위한 기초 자료로 활용될 수 있을 것이다.

참 고 문 헌

과학기술정보통신부 (2023. 9. 26.). ‘국가연구데이터 관리 및 활용 촉진에 관한 법률’ 제정안 입법예고, 과학기술정보통신부. 출처:

<https://www.msit.go.kr/bbs/view.do?sCode=user&mId=113&mPid=238&pageIndex=&bbsSeqNo=94&nttSeqNo=3183524&searchOpt=ALL&searchTxt=>

김규환, 김지현 (2024). 국내 기관 리포지터리 운영 정책 수립을 위한 내부 환경 분석. 한국도서관·정보학회지, 55(3), 43-64. <https://doi.org/10.16981/kliss.55.3.202409.43>

김예현, 김지현 (2023). 과학기술분야 기관 연구데이터 리포지터리 운영 활성화 방안 연구. 한국비블리아학회지, 34(3), 109-134. <https://doi.org/10.14699/kbiblia.2023.34.3.109>

김주섭, 강효숙, 김선태 (2023). 생태 분야 데이터 리포지터리 운영 현황 분석 및 EcoBank 서비스 제안. 한국문헌정보학회지, 57(4), 289-310. <https://doi.org/10.4275/KSLIS.2023.57.4.289>

김주섭, 양성준, 김선태 (2022). 데이터 리포지터리 인증 체계 분석 및 인증 전략에 관한 연구 - Coretrustseal을 중심으로 -. 한국문헌정보학회지, 56(2), 209-229.

<https://doi.org/10.4275/KSLIS.2022.56.2.209>

김관준 (2021). 프로파일링에 기초한 키워드 유형별 지적구조 분석에 관한 연구 - 국외 오픈액세스

- 분야를 중심으로 -. 한국문헌정보학회지, 55(4), 115-140.
<https://doi.org/10.4275/KSLIS.2021.55.4.115>
- 김현희, 장우권, 박병희, 안태경 (2005). 국가지식정보의 효율적인 유통체제 구축을 위한 대학 리포지토리의 운영 모형 개발: dCollection을 중심으로. 정보관리학회지, 57, 103-127.
<https://doi.org/10.3743/KOSIM.2005.22.3.103>
- 김현희, 정경희, 김용호 (2006). 대학 기관 리포지토리의 운영 현황 분석 및 개선 방안에 관한 연구 - dCollection을 중심으로 -. 정보관리학회지, 23(4), 17-39.
<https://doi.org/10.3743/KOSIM.2006.23.4.017>
- 박형주 (2024). 데이터 스칼라십: 데이터 저널과 데이터 리포지토리를 중심으로. 문화기술의 융합, 10(1), 443-451. <https://doi.org/10.17703/JCCT.2024.10.1.443>
- 백종명, 이은지 (2013). 국내 기관 리포지터리 콘텐츠 현황분석 및 개선방안에 관한 연구. 한국정보관리학회 2013년도 제20회 학술대회 논문집, 83-87.
- 서선경, 정은경 (2013). 동시출현단어 분석 기반 오픈 액세스 분야 지적구조에 관한 연구. 한국비블리아학회지, 24(1), 207-228. <https://doi.org/10.14699/kbiblia.2013.24.1.207>
- 서혜란, 김선애, 강은영 (2015). 공공도서관 자료보존의 현황과 과제. 한국문헌정보학회지, 49(2), 111-134. <http://dx.doi.org/10.4275/KSLIS.2015.49.2.111>
- 신은자 (2017). 국내 대학 리포지터리의 세계적 영향력에 관한 연구. 정보관리학회지, 34(1), 197-218.
<https://doi.org/10.3743/KOSIM.2017.34.1.197>
- 윤희윤 (2013). 공공도서관 공동보존서고 건립의 타당성 분석. 한국도서관·정보학회지, 44(2), 5-26. <https://doi.org/10.16981/kliss.44.2.201306.5>
- 윤희윤 (2014). 대학도서관 공동보존서고 설립·운영모형 연구. 한국도서관·정보학회지, 45(3), 37-61. <https://doi.org/10.16981/kliss.45.3.201409.37>
- 윤희윤, 김신영 (2007). 국내외 문헌정보학 학술지의 오픈 액세스 동향 분석. 한국도서관·정보학회지, 38(1), 257-276. <https://doi.org/10.16981/kliss.38.1.200703.257>
- 이다현, 신동희 (2024). 텍스트 마이닝을 활용한 대학도서관과 공공도서관의 장서개발 정책 키워드 분석. 한국문헌정보학회지, 58(1), 285-302. <https://doi.org/10.4275/KSLIS.2024.58.1.285>
- 이재민, 손청기, 이미영 (2008). 국내 기관 레포지터리 구축 및 활성화 방안 - KAIST의 KOASAS 사례를 중심으로. 정보관리연구, 39(3), 179-204. <https://doi.org/10.1633/JIM.2008.39.3.179>
- 이재윤, 정경희 (2020). 해외 인문사회과학 학술지 오픈엑세스 전환 동향 연구. 정보관리학회지, 37(3), 301-326. <https://doi.org/10.3743/KOSIM.2020.37.3.301>
- 이종덕, 신규용, 유진철 (2013). 교육자료 아카이빙을 위한 리포지터리 설계 및 구축 방안. 컴퓨터교육학회 논문지, 16(6), 95-102. <https://doi.org/10.32431/kace.2013.16.6.009>

- 이혜림, 엄정호, 신영호, 임형준, 한나은 (2024). CoreTrustSeal 인증 획득을 통한 데이터 리포지토리의 신뢰성 향상을 위한 연구. 정보관리학회지, 41(2), 245-268.
<https://doi.org/10.3743/KOSIM.2024.41.2.245>
- 정영미 (2014). 일본의 리포지터리 확산과 성장 배경에 관한 연구. 일본근대학연구, 43, 427-448.
<http://doi.org/10.16979/jmak.43.201402.427>
- 정영미, 이상기 (2010). 성공적인 리포지터리의 운영정책에 관한 연구. 정보관리학회지, 27(4), 131-152. <https://doi.org/10.3743/KOSIM.2010.27.4.131>
- 정재민, 김완중 (2022). 토픽 성장 분석을 통한 오픈엑세스 분야 연구 동향 분석. 정보관리학회지, 39(4), 75-97. <https://doi.org/10.3743/KOSIM.2022.39.4.075>
- 추기숙, 남영준 (2017). 개인기록 전자출판 리포지토리 구축 방안에 관한 연구. 한국도서관·정보학회지, 48(4), 351-374. <http://dx.doi.org/10.16981/kliss.48.201712.351>
- 한상우 (2023). 키워드 네트워크 분석을 이용한 연구데이터 관련 국내 연구 동향 분석. 한국도서관·정보학회지, 54(4), 393-414. <http://doi.org/10.16981/kliss.54.4.202312.393>
- 황혜경, 이지연 (2009). 오픈엑세스기반 기관리포지터리 성공에 미치는 요인 분석. 정보관리학회지, 26(4), 35-57. <https://doi.org/10.3743/KOSIM.2009.26.4.035>
- Ammarukleart, S. & Kim, J. (2017). Institutional repository research 2005-2015: a trend analysis using bibliometrics and text mining. Digital Library Perspectives, 33(3), 264-278. <https://doi.org/10.1108/DLP-07-2016-0027>
- Budapest Open Access Initiative (2002, February 14). BUDAPEST OPEN ACCESS INITIATIVE. BOAI. Available: <https://www.budapestopenaccessinitiative.org/read/>
- Jain, P. (2011). New trends and future applications/directions of institutional repositories in academic institutions. Library Review, 60(2), 125-141.
<https://doi.org/10.1108/00242531111113078>
- OAK 국가리포지터리 (2024. 2. 20.). 2024년 오픈엑세스코리아(OAK) 리포지터리 보급 공모. OAK 국가리포지터리. 출처:
https://oak.go.kr/board/viewBoard.do?menuSeq=47&board_kind=2&pageUrl=selectBoardArticle&bbsId=BBSMSTR_000000000011&nttId=1564&pageIndex=1
- Pinfield, S., Salter, J., Bath, P. A., Hubbard, B., Millington, P., Anders, J. H. S., & Hussain, A. (2014). Open-access repositories worldwide, 2005-2012: Past growth, current characteristics, and future possibilities. Journal of the Association for Information Science and Technology, 65(12), 2404-2421. <https://doi.org/10.1002/asi.23131>
- Ranking Web of Repositories (2024, August). TRANSPARENT RANKING: All Repositories.

Ranking Web of Repositories. Available:

<https://repositories.webometrics.info/en/node/30>

VOSviewer (2024, August 3). VOSviewer Manual. VOSviewer. Available:

<https://www.vosviewer.com/>

Waltman, L., van Eck, N. J., & Noyons, E. C. M. (2010). A unified approach to mapping and clustering of bibliometric networks. *Journal of Informetrics*, 4(4), 629-635.

<https://doi.org/10.1016/j.joi.2010.07.002>

Zibani, P., Rajkoomar, M., & Naicker, N. (2022). A systematic review of faculty research repositories at higher education institutions. *Digital Library Perspectives*, 38(2), 237-248.

<https://doi.org/10.1108/DLP-04-2021-0035>

• 국한문 참고문헌의 영문 표기

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

Chu, Kisook & Nam, Youngjoon (2017). A study on the building self-publishing repository for the personal digital records. *Journal of Korean Library and Information Science Society*, 48(4), 351-374. <http://doi.org/10.16981/kliss.48.201712.351>

Chung, Jaemin & Kim, Wanjong (2022). Understanding research trends of open access via topic growth analysis. *Journal of the Korean Society for Information Management*, 39(4), 75-97. <https://doi.org/10.3743/KOSIM.2022.39.4.075>

Han, Sangwoo (2023). An analysis of domestic research trend on research data using keyword network analysis. *Journal of Korean Library and Information Science Society*, 54(4), 393-414. <http://doi.org/10.16981/kliss.54.4.202312.393>

Hwang, Hyekyong & Lee, Jeeyeon (2009). Analyzing the factors affecting the successful deployment of the open source based institutional repositories. *Journal of the Korean Society for Information Management*, 26(4), 35-57. <https://doi.org/10.3743/KOSIM.2009.26.4.035>

Jung, Youngmi (2014). A study on the current status and its policy background of repositories proliferation in Japan. *The Japanese Modern Association of Korea*, 43, 427-448. <http://doi.org/10.16979/jmak.43.201402.427>

Jung, Youngmi & Lee, Sanggi (2010). A study on the operational policies and best practice of repository. *Journal of the Korean Society for Information Management*, 27(4),

- 131-152. <https://doi.org/10.3743/KOSIM.2010.27.4.131>
- Kim, Gyuhwan & Kim, Jihyun (2024). Analyzing the internal environment to establish policies for domestic institutional repository operations. *Journal of Korean Library and Information Science Society*, 55(3), 43-64.
<https://doi.org/10.16981/kliss.55.3.202409.43>
- Kim, Hyunhee, Chang, Wookwon, Kawk, Byeongheui, & Ahn, Taekyoung (2005). Implementing the management model of the institutional repositories for the efficient distribution infrastructure of national knowledge focusing on the dCollection system. *Journal of the Korean Society for Information Management*, 57, 103-127.
<https://doi.org/10.3743/KOSIM.2005.22.3.103>
- Kim, Hyunhee, Joung, Kyounghee, & Kim, Yongho (2006). A study on the improvement of university institutional repositories (dCollection) based on its current state. *Journal of the Korean Society for Information Management*, 23(4), 17-39.
<https://doi.org/10.3743/KOSIM.2006.23.4.017>
- Kim, Juseop, Kang, Hyosuk, & Kim, Suntae (2023). Analysis of Ecological Data Repository Operation Status and EcoBank Service Proposal. *Journal of the Korean Society for Library and Information Science*, 57(4), 289-310.
<https://doi.org/10.4275/KSLIS.2023.57.4.289>
- Kim, Juseop, Yang, Seongjun, & Kim, Suntae (2022). A study on data repository certification scheme analysis and certification strategy: focused on Coretrustseal. *Journal of the Korean Society for Library and Information Science*, 56(2), 209-229.
<https://doi.org/10.4275/KSLIS.2022.56.2.209>
- Kim, Panjun (2021). A study on the intellectual structure analysis by keyword type based on profiling: focusing on overseas open access field. *Journal of the Korean Society for Library and Information Science*, 55(4), 115-140.
<https://doi.org/10.4275/KSLIS.2021.55.4.115>
- Kim, Yehyeon & Kim, Jihyun (2023). A study on strategies to promote the activation of institutional research data repositories in the field of science and technology. *Journal of the Korean Biblia Society for Library and Information Science*, 34(3), 109-134.
<https://doi.org/10.14699/kbiblia.2023.34.3.109>
- Lee, Dahyeon & Shin, Donghee (2024). A keyword analysis of collection development policies of university and public libraries using Text Mining. *Journal of the Korean*

- Society for Library and Information Science, 58(1), 285-302.
<https://doi.org/10.4275/KSLIS.2024.58.1.285>
- Lee, Jaemin, Son, Chungki, & Lee, Miyoung (2008). The institutional repository development and revitalization methods in Korea: focused on the KAIST's KOASAS. *Journal of Information Science Theory and Practice*, 39(3), 179-204.
<https://doi.org/10.1633/JIM.2008.39.3.179>
- Lee, Jaeyun & Joung, Kyounghee (2020). The current state of foreign transition to open access journal publishing in the field of HSS. *Journal of the Korean Society for Information Management*, 37(3), 301-326. <https://doi.org/10.3743/KOSIM.2020.37.3.301>
- Lee, Jongdeog, Shin, Kyuyong, & Yoo, Jincheol (2013). Design and implementation of the repository for archiving educational resources. *The Journal of Korean Association of Computer Education*, 16(6), 95-102. <https://doi.org/10.32431/kace.2013.16.6.009>
- Ministry of Science and ICT (2023, September 26). Announcement of the Draft Act on the Promotion of Management and Utilization of National Research Data. Ministry of Science and ICT. Available:
<https://www.msit.go.kr/bbs/view.do?sCode=user&mId=113&mPid=238&pageIndex=&bbsSeqNo=94&nttSeqNo=3183524&searchOpt=ALL&searchTxt=>
- OAK National Repository (2024, February 20). 2024 Open Access Korea (OAK) Repository Distribution Announcement. OAK National Repository. Available:
https://oak.go.kr/board/viewBoard.do?menuSeq=47&board_kind=2&pageUrl=selectBoardArticle&bbsId=BBSMSTR_00000000011&nttId=1564&pageIndex=1
- Paek, Jongmyung & Lee, Eunjee (2013). A study on the contents of institutional repositories in Korea. *Proceedings of the Korean Society for Information Management Conference*, 20, 83-87.
- Park, Hyoungjoo (2024). Data Scholarship: data journals and data repositories. *The Journal of the Convergence on Culture Technology*, 10(1), 443-451.
<https://doi.org/10.17703/JCCT.2024.10.1.443>
- Rhee, Healim, Um, Jungho, Shin, Youngho, Yim, Hyungjun, & Han, Naeun (2024). A study to improve the trustworthiness of data repositories by obtaining CoreTrustSeal certification. *Journal of the Korean Society for Information Management*, 41(2), 245-268. <https://doi.org/10.3743/KOSIM.2024.41.2.245>
- Seo, Sunkyung & Chung, Eunkyung (2013). Domain analysis on the field of open access

- by co-word analysis. Journal of the Korean Biblia Society for Library and Information Science, 24(1), 207-228. <https://doi.org/10.14699/kbiblia.2013.24.1.207>
- Shin, Eunja (2017). Global impact of institutional repositories in South Korean university. Journal of the Korean Society for Information Management, 34(1), 197-218. <https://doi.org/10.3743/KOSIM.2017.34.1.197>
- Suh, Hyeran, Kim, Sunae, & Kang, Eunyeong (2015). The status quo and tasks of the collection preservation of public libraries in Korea. Journal of the Korean Society for Library and Information Science, 49(2), 111-134. <http://dx.doi.org/10.4275/KSLIS.2015.49.2.111>
- Yoon, Heeyoon (2013). Feasibility analysis of establishment of the collaborative repository for public libraries in Korea. Journal of Korean Library and Information Science Society, 44(2), 5-26. <https://doi.org/10.16981/kliss.44.2.201306.5>
- Yoon, Heeyoon (2014). A model for the establishment and operation of the collaborative repository for academic libraries in Korea. Journal of Korean Library and Information Science Society, 45(3), 37-61. <https://doi.org/10.16981/kliss.45.3.201409.37>
- Yoon, Heeyoon & Kim, Sinyoung (2007). Trends analysis of open access for foreign and domestic scholarly journals in the field of library and information science. Journal of Korean Library and Information Science Society, 38(1), 257-276. <https://doi.org/10.16981/kliss.38.1.200703.257>