

ChatGPT 기반 자동 요약을 활용한 연구 동향 분석 방법에 관한 연구*

A Study on Research Trend Analysis Methods Using Automatic Summarization Based on ChatGPT

조용신 (Yong Sin Jo) **
김용환 (Yong Hwan Kim) ***

초 록

특정 분야의 연구의 동향을 파악하고 특징을 포착하는 연구는 연구 분야의 미래를 예측하며 연구자들의 연구 방향을 설정하는데 주요한 역할을 한다. 최근 연구 동향 분석 연구들은 제목 및 초록 텍스트 또는 메타데이터 분석을 통해 연구 동향을 분석하고 있으며, 해당 데이터를 통한 동향 분석은 논문의 포괄적인 주제 혹은 핵심 연구자 및 기관 등의 분석에 초점을 두고 있다. 본 연구에서는 내용적인 측면인 방법론, 세부주제 및 결과, 논의점 등을 살펴볼 수 있도록 데이터를 제목 및 초록, 메타데이터가 아닌 전문에 대한 자동요약문을 사용하는 방법론을 제안하고자 한다. 정보격차를 주제로 수집된 논문을 대상으로 제목-초록 텍스트와 ChatGPT를 이용해 전문을 서론, 방법론, 결과, 논의로 구조화한 자동요약문을 구축하고, 단어 빈도 분석과 동시 출현 단어 빈도 분석, 키워드 네트워크 분석 방법을 사용하여 결과를 비교하였다. 그 결과, 제목-초록 텍스트와 서론 요약문의 분석 결과는 해당 영역의 주요 주제가 유사하게 나타나는 것을 확인하였다. 방법론, 결과, 논의에서는 기존에 파악하기 어려웠던 세부적인 내용이 파악되는 것을 확인할 수 있었다.

ABSTRACT

Studies that identify research trends and capture characteristics within a specific field play a crucial role in forecasting the field's future and guiding researchers' directions. Recent trend-analysis studies have relied on title-and-abstract text or metadata analysis, focusing on broad topics or on key researchers and institutions. In this study, we propose a methodology that employs automatic summaries of full texts instead of titles, abstracts, or metadata to examine substantive aspects such as methodology, detailed subtopics, results, and discussion points. For papers collected on the topic of the digital divide, we generated automatic summaries by structuring the full text into Introduction, Methodology, Results, and Discussion using ChatGPT and then compared outcomes through word-frequency analysis, co-occurrence analysis, and keyword network analysis. The findings confirm that analyses of title-and-abstract texts and Introduction summaries reveal similar major themes. In contrast, analyses of the Methodology, Results, and Discussion sections uncover detailed information that was previously difficult to identify.

키워드: 동향 분석, ChatGPT, 단어 빈도 분석, 네트워크 분석, 자동요약
research trends analysis, ChatGPT, term frequency analysis, network analysis, automatic summarization

* 이 논문은 2023~2025년도 청주대학교 연구장학 지원에 의한 것임

이 논문은 2022년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임
(NRF-2022S1A5C2A02093658).

** 청주대학교 문헌정보학과 박사과정(ys2495@cju.ac.kr) (제1저자)

*** 청주대학교 문헌정보학과 부교수(kimyonghwan@cju.ac.kr) (교신저자)

■ 논문접수일자 : 2025년 5월 23일 ■ 최초심사일자 : 2025년 6월 8일 ■ 게재확정일자 : 2025년 6월 19일

■ 정보관리학회지, 42(2), 229-258, 2025. <http://dx.doi.org/10.3743/KOSIM.2025.42.2.229>

※ Copyright © 2025 Korean Society for Information Management

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. 서 론

특정 분야의 연구 흐름이나 특징에 대한 파악은 해당 분야의 미래 연구를 예측하거나, 개별 연구자들의 연구 방향을 설정하는데 중요한 역할을 수행한다. 따라서 특정 분야의 연구 동향을 파악하기 위하여 기존의 연구 결과들을 분석하는 것은 효과적이라고 할 수 있다(오세훈, 2005). 따라서 기존에 출판된 논문들을 바탕으로 다양한 분야의 연구 동향을 분석하는 연구들이 지속적으로 진행되어 왔다.

문헌정보학 영역에서는 문헌정보학 분야에 대한 연구 동향 분석(구본진, 장덕현, 2023; 오세훈, 2005; 이재윤, 2023; 이해경 외, 2019)과 더불어 문헌정보학 영역의 세부 분야인 도서관 관련 연구동향(이용재, 김지은, 2021; 장로사, 2023; 최재황, 천세진, 2022), 교육 관련 분야(이성숙, 이지원, 2024), 큐레이션 및 기록관리 관련 분야(박민석, 이지수, 2024; 최재은, 2024; 홍덕용, 허준석, 2023), 리터러시 관련 분야(이다현, 신동희, 2024; 홍소람, 장우권, 2023) 등 다양한 세부 분야에 대한 동향 분석 연구가 이루어져 왔다.

연구 동향을 파악하는 연구들에서는 대부분 제목과 초록(김규환 외, 2009; 박옥남, 2011), 저자 키워드, 저자 정보(김희전, 조현양, 2010; 이종욱, 박혜린, 2016), 기관정보(이재윤 외, 2011), 인용 정보(오세훈, 2005; 이재윤, 2015)와 같은 데이터를 활용하여 분석을 수행해 왔다. 이러한 데이터를 통해 도출되는 결과들은 특정 연구 분야의 핵심적인 연구 주제(구본진, 장덕현, 2023; 신유미, 박옥남, 2019) 또는 해당 주제를 다루는 주요 연구자(감미아, 2023; 손혜

인, 남영준, 2016), 기관(강순애, 2013), 관련 연구 영역(한미경, 2017; 황금숙, 2005)에 대해 초점을 두고 있다. 그러나 연구동향을 파악한다는 것에는 핵심적인 주제뿐만 아니라 연구방법, 연구의 핵심결과, 논의점과 같은 내용적인 측면이 포함되기 때문에, 지금까지 사용되는 제목 및 초록, 메타데이터로는 이를 파악하는데 한계점을 가질 수 있다.

따라서 본 연구에서는 기존의 데이터와는 달리 ChatGPT(<https://chatgpt.com/>)를 이용하여 전문에 대한 요약문을 구성하고, 이를 통해 동향 분석을 수행하는 방법을 제안하고자 한다. 초록의 경우 저자에 의한 요약문이기 때문에 전문에 대한 요약문과 차이점이 나타나지 않는다고 판단할 수 있다. 하지만 초록의 대부분이 연구의 필요성 및 주제 등에 대한 내용만을 다루고 있어(이태영, 2005) 논문에 대한 체계적인 요약문을 이용한 동향 분석은 차이점을 보여 줄 수 있다. 따라서 ChatGPT를 통해 본문을 서론, 방법론, 결과, 논의라는 4개의 구체적인 구조로 요약한 뒤 분석하는 방법을 제안하고자 한다. ChatGPT를 활용한 요약문은 높은 평가와 선호를 받는 동시에(서종훈, 2024; Pu et al., 2023; Zhang et al., 2024) 일관성이나 신뢰성에 관한 우려 역시 제기되고 있다(Luo et al., 2024; Peters & Chin, 2025). 본 연구에서는 이러한 상반된 평가를 고려하여 동향 분석 연구에서 ChatGPT의 활용 가능성에 초점을 맞추어 실험을 진행하고자 한다.

본 연구에서는 정보격차 주제의 논문들을 대상으로 연구를 진행하고자 한다. 정보격차 문제는 단순히 정보의 비대칭에서 그치는 것이 아니라 사회 양극화를 초래하고 분리된 계층을

고착화하는 중대한 요인으로 보고 있다(강인서, 김혜진, 2020). 이에 따라 문헌정보학계에서도 정보격차에 대한 연구들이 진행되며 관심을 보이고 있다(강인서, 김혜진, 2020; 고정현 외, 2021; 이성신 외, 2019). 따라서 본 연구에서는 국내 문헌정보학계의 정보격차 문제를 다룬 연구들을 대상으로 하여 기존의 방법론과 제안한 방법론을 비교 분석하는 연구를 수행하고자 한다.

본 연구의 구성은 2장에서 선행연구, 3장에서는 방법론, 4장에서는 결과, 5장에서는 논의 및 결론으로 구성된다.

2. 선행연구

선행연구는 동향 분석 연구의 방법론을 중심으로 살펴보았다. 문헌정보학계에서 수행된 동향 분석 연구를 살펴보면 단순 통계적인 분석, 인용 기반 분석, 토픽모델링 분석, 빈도 및 네트워크 분석, 크게 3가지로 구분된다.

먼저 단순 통계적인 분석을 바탕으로 동향 분석을 수행한 연구이다. 남태우와 이진영(2009)은 기록관리학 학술지의 논문을 대상으로 기록 관리학 연구의 동향 및 향후 발전 방안을 제시하기 위해 기록관리 학문과 관련된 논문 총 399 편을 수집하여 주제별 분포, 간행 시기별 분포, 학회지별, 분포, 연구자별 분포를 통계 분석하였다. 김정현(2011)은 우리나라 문헌정보학 분야 전문 학회지에 수록된 논문 가운데 자료 조직 관련 논문 607편을 손정표(2003)의 연구에서 제시한 자료조직 분야의 주제 구분을 바탕으로 학회지별, 연도별 논문 수 현황을 계량적

으로 분석하여 자료 조직 분야의 연구 동향을 분석하였다. 손혜인과 남영준(2016)은 기록관리학 분야 국내 학술지의 연구 동향을 분석하기 위해 ‘한국기록관리학회지’와 ‘기록학연구’를 대상으로 각 학회지가 2015년까지 발행된 논문 데이터를 바탕으로 저자정보를 수집하고, 기록학 분야의 14가지 주제분야로 논문을 구분하여, 저자분석과 주제분석을 진행하였다.

다음은 인용정보를 활용한 연구이다. Zhao & Strotmann(2008)의 연구에서는 정보학 연구 영역의 소주제 형성과 변화 과정을 분석하기 위해 저자서지결합법과 저자동시인용분석을 활용하는 연구를 수행하였다.

다음은 토픽모델링을 활용한 연구이다. Sugimoto et al.(2011)는 1930년부터 2009년까지 북미 문헌정보학 박사학위 논문 3,121편의 제목과 초록을 LDA(Latent Dirichlet Allocation) 토픽모델링을 활용하여 분석하는 연구를 수행하였다. 분석 결과 도서관의 역사, 사서, 정보 행태, 인용 등의 핵심 주제를 포착하였다. 박준형과 오효정(2017)은 국내 기록관리학 연구동향을 분석하기 위해 기록관리학 학술지와 문헌정보학 학술지를 수집대상으로 선정하고, 논문명, 초록, 키워드를 활용하여 LDA, HDP 토픽모델링을 활용하여 분석하였다. LDA 토픽모델링 분석 결과 전자기록, 국가기록원, 아카이브 구축과 기록물 활용 방안 등의 주제가 나타났으며, HDP 토픽모델링에서는 민간기록의 아카이빙, 전자기록물과 디지털 기술, 기록관리에 관한 표준 및 규칙 등의 주제가 나타났다. 이세나 외(2024)는 법학사서 관련 연구의 연구 동향을 확인하는 연구를 수행하였다. 이를 위해 Scopus(<https://www.scopus.com>)와

ProQuest(<https://www.proquest.com>)의 문헌정보학 데이터 베이스에서 논문을 수집하였다. 초록 데이터를 이용한 토픽모델링 분석 결과 장서 및 정보자원 개발, 법률 정보 자원의 접근 도구, 법률 연구 환경의 변화, 법학사서의 역할 변화, 법학 사서의 역량 발전이 주요 주제로 도출되었다. 제목 및 초록 데이터를 바탕으로 토픽모델링을 활용한 연구들의 결과를 살펴보면 토픽을 중심으로 단어를 해석하기 때문에 연구 주제에 대한 내용만이 등장하는 것을 알 수 있다.

다음은 단어 빈도, 동시 출현 단어 빈도, 네트워크 분석 등을 수행한 연구이다. Milojević et al.(2011)는 문헌정보학 저널의 10,344편의 논문 제목을 활용하여 단어 빈도 분석과 동시 출현 단어 분석, 다차원 척도 분석을 수행하여 지적 구조를 규명하였다. 분석 결과 Library science, Bibliometrics and sciento metrics, Information science 가 주요 주제로 확인되었다. 김하진과 송민(2014)은 국내외 정보학 분야의 동향 분석을 위해 정보관리학회지와 JASIST의 논문 제목과 초록을 대상으로 동시 출현 분석을 수행하였다. 또한 세부적인 동향을 살피기 위해 네트워크 시각화 분석을 하였다. 분석 결과 국내의 주요 주제는 도서관, 정보 서비스, 정보 이용자, 정보 자료 조직로 나타났으며, 국외는 정보 검색, 정보 이용자, 웹 관련 정보학, 계량 정보학을 중심으로 연구가 진행된 것을 알 수 있었다. 이성신(2016)은 국내외 도서관마케팅 연구 경향을 비교 분석하는 연구를 수행하였다. 이를 위해 국내 문헌정보학분야 4대 학회지(한국문헌정보학회지, 한국도서관·정보학회지, 한국비블리아학회지, 정보관리학회지)의 도

서관마케팅 관련 연구 57건과 207개의 저자 키워드를 수집하였다. 이어 Scopus 데이터베이스를 활용하여 문헌정보학분야 학술지에 실린 도서관마케팅 관련 연구 550건과 1,283개의 저자 키워드를 수집하였다. 이어 수집된 저자 키워드를 바탕으로 빈도 분석과 네트워크 분석을 진행하였으며, 네트워크는 연결 중심성과 매개 중심성을 중심으로 분석하였다. 강인서와 김혜진(2020)은 문헌정보학 분야에서의 정보격차 연구 동향을 분석하였다. 연구 동향을 분석하기 위해 정보관리학회지, 한국도서관·정보학회지, 한국비블리아학회지, 한국문헌정보학회지에서 발행된 195편의 논문 데이터의 저자 키워드와 초록을 수집하였다. 이어 연구대상, 연구목적, 연구방법 별로 분류한 뒤, 단어 빈도 및 동시 출현 단어 분석, 패스파인더 네트워크 분석을 통해 연구 대상과 연구 목적, 연구 방법의 변화를 추적하였다. 박영희와 김성희(2024)는 저자 키워드를 중심으로 정보기술 및 인공지능 관련 연구 동향을 파악하고 연구 트랜드 변화를 분석하는 연구를 수행하였다. 이를 위해 Web of science DB를 활용하여 총 4,233개의 논문의 수집하고, 시기별 변화를 확인하고자 출판년도를 활용한 서지 매팅 분석을 수행하였다.

기준의 동향 분석 연구들에서 사용되는 데이터들은 제목, 초록, 저자, 기관, 저자키워드, 인용정보들이 주로 활용되는 것으로 보이며, 방법으로는 단순통계, 인용분석, 토픽모델링분석, 단어빈도 및 네트워크 분석이 활용되고 있다. 최근에는 주로 제목, 초록, 저자 키워드를 바탕으로 단어빈도 및 네트워크 분석이 많이 활용되는 것을 확인할 수 있다. 그러나 제목 및 초록, 저자키워드에서 파악할 수 있는 정보는 논

문의 전반적인 내용과 관련된 것이기 때문에 이를 활용한 동향 분석 연구에서는 포괄적인 주제 확인에만 초점을 두고 있다. 본 연구에서는 논문 전문을 AI를 활용하여 자동으로 요약한 뒤 요약문을 기반으로 최근 많이 사용되고 있는 단어빈도, 동시출현빈도 및 네트워크 분석을 수행해 보고자 한다. 그리고 이를 제목 및 초록을 활용한 분석과 비교 분석해 보고자 한다.

3. 연구 방법

정보격차에 대한 연구 동향을 분석하기 위해서 <그림 1>과 같이 연구 방법을 설계하였다.

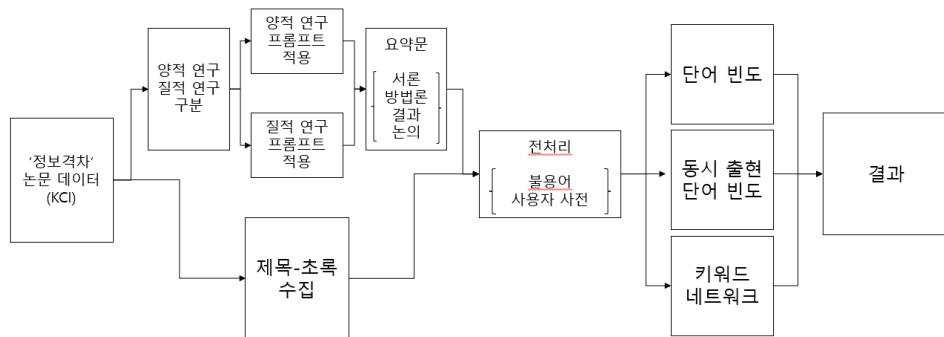
3.1 데이터 수집

정보격차에 대한 문헌정보학계의 연구 동향을 살펴보기 위해 한국학술지인용색인(KCI, <https://www.kci.go.kr>)을 통해 수집하였다. 검색조건으로 KCI 학문 분류는 문헌정보학 범위로 지정하고, 출판 기간은 2025년 3월까지로 설정하였다. 추가로 정보격차를 의미하는 단어

들인 ‘정보격차’, ‘정보불평등’, ‘정보빈곤’, ‘정보취약’을 ‘OR’ 연산자로 결합하여 질의를 구성하였다. 그 결과, 총 172건의 논문 중 초록 혹은 전문이 영어로 작성되어 있거나 파일의 내부 포맷이 이미지로 되어 있어 ChatGPT o4-mini를 통한 요약이 불가능한 논문을 제외한 총 149 건의 논문을 대상으로 제목과 초록 및 전문을 수집하였다.

3.2 ChatGPT를 통한 논문 요약

논문의 요약을 보다 체계적으로 작성하기 위해서 미국심리학회(American Psychological Association, APA)의 논문 작성 지침을 참고하여 논문의 구성요소를 결정하였다(American Psychological Association, 2019). APA에서는 논문을 양적 연구와 질적 연구로 구분하고, 서론, 방법, 결과, 논의의 네 가지 구성 요소로 세분화하여, 각 구성 요소 별 포함되어야 할 내용에 대해 설명하고 있다. 이를 바탕으로 논문의 구성요소를 서론, 방법, 결과, 논의로 설정하고, 세부 내용을 포함하도록 자동요약을 위한 프롬프트를 작성하였다.



<그림 1> 연구 방법

ChatGPT의 응답은 사용자에 따라 응답의 수준이 달라질 수도 있기 때문에 ChatGPT의 맞춤 설정 기능을 통해 요약을 요청하는 대상이 문헌정보학과 대학원생으로 설정하였다. 이후 ChatGPT를 통해 논문을 양적 연구와 질적 연구로 구분하고, 각각의 연구에 대해서 <그림 2>와 같이 논문의 구성요소 별로 요약문을 작성하도록 ChatGPT 프롬프트를 작성하였다.

<표 1>은 ChatGPT에 의한 논문 요약 결과에 대한 예시를 나타낸 것이다.

3.3 데이터 분석 방법

각 논문의 제목 및 초록 데이터와 ChatGPT

를 이용한 구성 요소별 요약문을 활용하여 전처리 작업을 진행하였다. 전처리 작업은 Python 기반의 Konlp 라이브러리의 모듈 중 불용어 처리와 이용자 사전을 활용할 수 있는 한국어형태소 분석기인 Komoran을 통해 전처리 작업을 진행하였다(박은정, 조성준, 2014). 논문 데이터의 특성상 자주 등장하는 ‘연구’, ‘논문’과 같은 단어들과 ‘이’, ‘것’ 등과 같이 의미를 특정할 수 없는 단어들을 불용어로 설정하였다. 그리고 ‘토픽모델링’, ‘자기효능감’, ‘저시력자’, ‘기초 생활수급자’ 등과 같은 복합어들을 대상으로 이용자 사전을 구축하였다.

전처리 단계에서 모든 데이터를 동일한 환경에서 분석하고자 동일한 불용어 사전과 이용자

<p>다음 양적 연구 논문을 아래의 보고 기준에 따라 요약해 주세요 .</p> <p>1. 서론</p> <ul style="list-style-type: none"> - 연구 문제 - 관련 연구 검토 - 가설 / 목표 / 세부 목표 <p>2. 방법</p> <ul style="list-style-type: none"> - 포함 및 배제 기준 - 참여자 특징 - 표집 절차 - 표본 크기, 검정력, 정확성 - 측정값 및 공변인 - 데이터 수집 절차 - 측정의 질 관리 방안 - 사용된 측정 도구 - 마스킹(블라인드) 여부 - 심리측정(신뢰도·타당도) 정보 - 연구 조건 및 설계 - 데이터 진단 절차 - 분석 전략 <p>3. 결과</p> <ul style="list-style-type: none"> - 참여자 흐름(flow chart) - 모집 현황 - 주요 통계 및 데이터 분석 결과 <p>4. 논의</p> <ul style="list-style-type: none"> - 원 가설 지지 여부 - 결과 간 유사성 비교 - 해석 - 일반화 가능성 - 실무적 / 이론적 시사점 	<p>다음 질적 연구 논문을 아래의 보고 기준에 따라 요약해 주세요 .</p> <p>1. 서론</p> <ul style="list-style-type: none"> - 연구 문제 혹은 질문 - 연구 목적 및 세부 목표 <p>2. 방법</p> <ul style="list-style-type: none"> - 연구 설계 개요 - 연구 참여자 또는 데이터 자원 - 참여자 모집 절차 - 데이터 수집 방법 - 분석 절차 및 기법 <p>3. 발견 내용 / 결과</p> <ul style="list-style-type: none"> - 주요 발견 내용을 세부 절(sub-section)로 구분하여 제시 <p>4. 논의</p> <ul style="list-style-type: none"> - 각 발견 결과에 대한 해석 - 기존 연구와의 연계성 - 일반화 가능성 - 현장 적용 및 이론적 시사점
--	--

<그림 2> ChatGPT 프롬프트

〈표 1〉 GPT 요약 결과 예시(이승민, 2021; 이제환, 윤유라, 2005)

구분	서론	방법론	결과	논의
양적 연구 (이승민, 2021)	스마트기기 활용이 고도화 된 환경에서 전통적 접근 중심 정보격차를 넘어서는 '스마트 디바이드' 형성 기제로서 미디어 리터러시의 역할을 실증적으로 분석하고자 한다. (중략) 본 연구의 목표는 미디어 리터러시의 다섯 요인인 스마트 디바이드 형성 요인에 미치는 영향을 검증하는 것이다.	온라인 설문조사 방식으로 2021년 3월 10~16일에 실시 하였으며, 20세 이상 스마트 기기 이용자 250명을 대상으로 포함·제외 기준에 따라 유효응답을 분석(N=250, Cronbach's $\alpha=.852$). (중략) SPSS 25.0을 사용해 Pearson 상관분석과 다중회귀분석을 수행하였다.	설문지 257부 배포 후 7부 제외, 최종 250부 유효응답 분석. 미디어 리터러시 요인과 스마트 디바이드 요인 간에 유의미한 양(+) 상관 ($r=.187 \sim .623$, $p<.01 \sim .05$) 관찰. (중략) 기기 보유에는 콘텐츠 다양성($\beta=.207$, $p<.01$), 어플리케이션 활용에는 콘텐츠 다양성($\beta=.293$, $p<.001$) 등이 유의하다.	미디어 리터러시의 핵심 요인들이 대체로 스마트 디바이드 형성 요인에 유의미한 영향을 미쳐 원 가설을지지. (중략) 스마트 디바이드 완화를 위해 미디어 리터러시 교육 강화와 정보 가치 판단 중심 정책 수립이 요구된다.
질적 연구 (이제환, 윤유라, 2021)	본 연구는 외국인 노동자 H의 삶을 사례로 하여, 사회적 약자로서의 외국인 노동자가 겪는 정보빈곤 현상을 탐구하고자 한다. 정보접근의 어려움, 언어·문화 장벽, 사회적 고립 등이 정보빈곤의 주요 요인으로 작용함을 문제의식으로 설정하였다.	심층면담을 통해 외국인 노동자 H의 생애사적 경험을 수집하고, 참여관찰과 문헌 분석을 보조 자료로 활용하여 질적 사례연구를 수행하였다. 수집된 데이터는 주제별로 코딩 및 범주화하여 분석되었다.	H는 한국 사회에서 언어·문화적 차이와 제도적 장벽으로 인해 주요 정보로부터 배제되었으며, 정보 취득 과정에서 중개인 의존, 구전 정보, 제한된 비공식 네트워크에 의존함을 발견하였다.	H는 한국 사회에서 언어·문화적 차이와 제도적 장벽으로 인해 주요 정보로부터 배제되었으며, 정보 취득 과정에서 중개인 의존, 구전 정보, 제한된 비공식 네트워크에 의존함을 발견하였다.

사전을 활용하여 전처리를 수행하였다. 그리고 전처리 결과를 바탕으로 단어 빈도 분석, 동시 출현 단어 빈도 분석, 네트워크 분석을 수행하였다. 단, 동시 출현 단어 빈도 분석에서는 '정보 격차'와 밀접한 주제를 파악하기 위하여 '정보 격차'와 동시 출현한 단어들을 분석하고자 하였다. 네트워크 분석에서는 Gephi 0.10.1를 활용하여 동시 출현 빈도 단어를 대상으로 빈도 순위 상위 300쌍을 기준으로 네트워크를 구축하였으며, 각각에 대한 중심성 분석 및 모듈리티 분석을 수행하였다.

4. 분석 결과

4.1 단어 빈도 분석 결과 비교

제목 및 초록 데이터와 ChatGPT를 이용한 구성 요소별 요약문에 대한 단어 빈도 분석을 수행하였으며 〈표 2〉는 단어 빈도 분석 결과 상위 20개 단어를 나타낸 것이다.

먼저, 제목 및 초록 텍스트에서는 '정보', '서비스', '도서관', '분석', '공공도서관' 순으로 단어들이 나타나는 것을 확인할 수 있다. 그리고 정보격차와 관련된 도서관 서비스에 대한 분석 ('정보', '서비스', '도서관', '공공도서관')과 주

〈표 2〉 단어 빈도 분석 결과 비교

순위	제목-초록	ChatGPT			
		구성요소별 요약문			
		서론	방법론	결과	논의
1	정보	정보	분석	서비스	필요
2	서비스	분석	서비스	정보	서비스
3	도서관	서비스	실시	도서관	강화
4	분석	도서관	문헌	이용	지원
5	공공도서관	공공도서관	도서관	분석	정보
6	장애인	방안	통계	자료	정책
7	정보격차	정보격차	활용	부족	제언
8	지역	제시	조사	제공	교육
9	제공	장애인	사례	프로그램	구축
10	방안	필요	공공도서관	활용	도서관
11	디지털	디지털	수집	교육	체계
12	정책	해소	정보	장애인	확대
13	활용	지역	자료	이용자	개발
14	이용	정책	정책	지원	제공
15	제시	격차	기반	지역	협력
16	교육	문제	설문조사	필요	프로그램
17	사회	사회	평가	요인	개선
18	격차	활용	검토	평균	해소
19	해소	문화	비교	비율	제안
20	지원	인식	장애인	공공도서관	맞춤

요 정보취약계층 중 장애인에 대한 연구(‘장애인’, ‘교육’, ‘지원’), 디지털 격차 해소(‘디지털’, ‘격차’, ‘해소’), 정보격차에 대한 정책적 접근(‘정책’, ‘제시’, ‘지원’)이 핵심적인 주제로 나타나는 것을 확인할 수 있다.

ChatGPT를 이용한 구성 요소별 요약문에 대한 결과를 살펴보면 다음과 같다. 서론의 경우 앞서 확인한 제목-초록 텍스트 분석 결과와 유사한 단어들이 등장한 것을 확인할 수 있었다. ‘정보’, ‘분석’, ‘서비스’, ‘도서관’ 등의 단어가 상위에 나타나고 있으며, 정보격차와 관련된 도서관 서비스에 대한 분석(‘정보’, ‘분석’, ‘서비스’, ‘도서관’, ‘공공도서관’)과 주요 정보취약계층인

장애인에 대한 방안 제시(‘방안’, ‘장애인’, ‘제시’, ‘필요’), 디지털 격차해소(‘디지털’, ‘격차’, ‘해소’), 정보 격차에 대한 정책적 접근(‘정책’, ‘지원’, ‘제시’)이 핵심적인 주제로 나타나 제목 및 초록 텍스트 대상의 분석 결과와 유사하게 나타나는 것을 확인 할 수 있었다.

방법론의 경우, ‘분석’, ‘서비스’, ‘실시’, ‘문헌’, ‘도서관’ 등과 같은 단어들이 상위에 나타나는 것을 확인할 수 있다. 도서관 서비스에 대한 분석(‘분석’, ‘서비스’, ‘도서관’)과 문헌 및 사례 조사(‘문헌’, ‘조사’, ‘사례’, ‘자료’), 설문조사 및 통계 자료 활용(‘통계’, ‘활용’, ‘설문조사’)이 핵심적인 내용으로 나타나는 것을 확인

할 수 있다. 특징적인 것은 정보격차와 관련된 연구들의 주요 분석 방법이 나타나며, ‘문헌조사’, ‘통계조사’, ‘사례조사’, ‘설문조사’ 등과 관련된 내용이 출현하는 것을 확인할 수 있다.

결과의 경우, ‘서비스’, ‘정보’, ‘도서관’, ‘이용’ 등의 단어들이 상위에 나타나는 것을 확인할 수 있다. 도서관 정보 서비스 이용에 대한 내용 ('서비스', '정보', '도서관', '이용', '분석')과 교육 자료 제공과 교육 프로그램 대한 연구('자료', '제공', '프로그램', '교육'), 이용자 및 장애인에 대한 지원 필요('장애인', '이용자', '지원', '필요')가 핵심적인 주제로 나타나는 것을 확인 할 수 있다. 특징적으로 ‘부족’이라는 단어가 다수 등장하는데 이는 장애인 서비스에서 시설, 장서, 인력 등 기본 자원의 부족이 시급한 문제라는 결과를 도출한 배경재(2016)의 연구나 시각장애인의 도서관 비이용 요인으로 시설·자료 부족을 주요 요인으로 확인한 오선경과 이지연(2011)의 연구와 같이 결과 부분에서 정보격차의 분석 결과로 특정 요인들의 결핍을 발견했기 때문으로 파악된다. 또한 ‘평균’과 ‘비율’이라는 단어들도 고빈도로 등장하였는데, 이는 앞서 방법론 파트에서 등장한 ‘통계’, ‘설문조사’의 영향으로 주요 결과들이 정량적 수치로 표현된 것으로 해석된다.

논의의 경우, ‘필요’, ‘서비스’, ‘강화’, ‘지원’ 등이 상위에 나타나는 것을 확인할 수 있다. 이는 정보격차 해소를 위한 제언에서 해당 용어들이 사용되기 때문인 것으로 해석된다. 또한 도서관 서비스 및 정책 개발('도서관', '서비스', '개발', '정책'), 협력 체계 구축('구축', '체계', '협력'), 맞춤 프로그램 개선 및 제공('제공', '프로그램', '개선', '맞춤')이 핵심적인 내용으로 나타나는

것을 확인할 수 있다.

각 텍스트들의 빈도 분석 결과 제목-초록 텍스트와 서론의 경우 유사한 단어들이 상위 빈도로 등장한 것을 확인할 수 있었으며, 방법론, 결과, 논의의 요약문에서는 각 구성 요소의 특징이 나타나는 단어들이 등장한 것을 확인할 수 있었다. 특히 방법론의 경우 다른 텍스트들에서 등장하지 않았던 ‘문헌’, ‘통계’, ‘사례’와 같은 단어들이 등장하여 논문의 구성 요소의 특징이 가장 강하게 나타나는 것으로 확인되었다.

4.2 동시 출현 단어 빈도 분석

제목 및 초록 데이터와 ChatGPT를 이용한 구성 요소별 요약문에 대한 ‘정보격차’와 동시 출현한 단어 빈도 분석을 수행하였으며 〈표 3〉은 동시 출현 빈도 분석 결과 상위 10개 단어를 나타낸 것이다.

먼저, 제목 및 초록 텍스트에서는 ‘해소’, ‘정보’, ‘공공도서관’, ‘분석’, ‘디지털’, ‘서비스’가 ‘정보격차’와 동시 출현하는 단어로 나타났다. ‘정보격차’와 밀접한 관계를 가지는 주제로 도서관 ('공공도서관', '도서관')의 역할과, 지역, 사회에서 나타나는 정보격차, 장애인에 관한 정보격차, 디지털 서비스와 관련된 정보격차가 주요한 주제로 나타나는 것을 확인할 수 있다. 특히 ‘해소’라는 단어가 높은 빈도로 동시 출현한 것으로 연구의 주요 목표가 정보격차의 해소를 위한 것으로 나타나고 있다.

ChatGPT를 이용한 구성요소 별 요약문에 대한 결과를 살펴보면 다음과 같다. 서론의 경우, 앞서 확인한 제목-초록 텍스트와 ‘정보격차’와 동시 출현하는 단어들과 유사한 것을 확인할

〈표 3〉 '정보격차' 키워드 동시 출현 분석 결과 비교

순위	제목-초록	ChatGPT			
		구성요소별 요약문		결과	논의
		서론	방법론		
1	해소	해소	분석	성별	해소
2	정보	분석	선정	연령	정책
3	공공도서관	방안	키워드	경제	제언
4	분석	정보	자료	수준	서비스
5	디지털	공공도서관	디지털	지적	지원
6	서비스	도서관	검색	영향	강화
7	도서관	서비스	실태	유의	필요
8	장애인	심화	문헌	정책	체계
9	사회	필요	검토	지표	정보
10	지역	사회	네트워크	실태	공공도서관

수 있다. 다만, '심화'의 경우에는 서론에서 문제 제기 또는 연구 문제 설정 과정에서 정보격차의 심화가 언급되는 것을 확인할 수 있다.

방법의 경우, '정보격차'와 동시 출현하는 단어로 '분석', '선정', '키워드', '자료', '디지털' 등과 같은 단어들이 상위에 나타나는 것을 확인할 수 있다. '정보격차'와 밀접한 관계를 가지는 주제로 디지털 정보격차, 디지털 격차에 대한 내용이 나타나는 것을 확인할 수 있다. 특징적으로 '키워드', '실태', '문헌', '네트워크'라는 단어들이 등장하며, 이는 정보격차 연구에서 사용된 주요 방법론이 키워드 분석, 실태 분석, 문헌 분석, 네트워크 분석이 사용된 것으로 확인 할 수 있다.

결과의 경우, '성별', '연령', '경제', '수준', '지적' 등의 단어들이 '정보격차'와 동시 출현하는 상위 단어로 나타나는 것을 확인할 수 있다. 이는 정보격차에 대한 분석에 성별, 연령층, 경제 수준, 지적 수준의 비교를 통해 분석 결과를 도출하는 것을 확인할 수 있다. 또한 '지표' 와 '실태'라는 단어가 등장하였다. 이는 각 계층의 실

태를 지표로 표현하여 서술한 결과인 것으로 확인할 수 있다. 이를 통해 주요 주제는 각 계층의 실태와 정책에 관한 내용인 것으로 나타난다.

논의의 경우, '해소', '정책', '제언', '서비스' 등이 단어들이 '정보격차'와 동시 출현하는 상위 단어들로 나타났다. 이는 정보격차의 해소를 위한 방안으로 정책 및 서비스의 지원 및 강화를 제언한 내용과 공공도서관의 정보격차 서비스 체계와 관련된 내용이 주요 논의 결과로 제안 된 것으로 확인할 수 있다.

정보격차와 동시에 출현한 단어들을 분석한 결과 제목-초록 텍스트의 경우 디지털이라는 단어가 특징적으로 등장하여, 디지털 정보격차 혹은 정보격차를 위한 디지털 서비스가 주요 주제에 포함된 것으로 보여진다. 구성 요소별 요약문의 경우 앞선 단어 빈도 분석의 결과와 동일하게, 서론의 경우 제목 및 초록 텍스트와의 결과와 유사하며 나머지는 각 요소 별 특징이 강조되는 단어들이 출현하는 것을 확인 할 수 있었다. 방법론의 경우 키워드 분석, 실태 분석, 네트워크 분석 등과 같이 구체적인 방법론을 의

미하는 단어들이 등장하였으며, 결과의 경우 정보격차가 발생하는 주요 계층들에 대한 지표의 비교 분석이 수행되었음을 확인할 수 있었다. 마지막으로 논의에서는 지원, 강화, 필요 등과 같은 단어들이 등장하여 정보격차 해소 정책 및 서비스에 대한 지원 및 강화가 필요하다는 제언을 나타내는 것을 확인할 수 있었다.

4.3 네트워크 분석 결과 비교

4.3.1 중심성 분석 결과 비교

네트워크 분석은 동시 출현 빈도 상위 300쌍을 기준으로 Gephi 0.10.1을 이용하여 네트워크를 구축하였으며, 각 노드들의 연결중심성, 매개중심성과 근접중심성을 분석하였다. 〈표 4〉는 제목 및 초록 텍스트를 통해 구성된 동시 출현 네트워크의 중심성 분석 결과이다.

중심성 분석 결과, 연결 중심성, 매개 중심성,

근접 중심성 모두에서 ‘정보’, ‘도서관’, ‘서비스’가 가장 중요한 단어로 확인되었다. 이는 국내 문헌정보학계의 정보격차 관련 논문들이 특히 ‘도서관 정보 서비스’와 관련된 연구들로 주를 이루기 때문에 이러한 결과가 나타나는 것을 확인할 수 있다. 이는 문헌정보학 영역의 정보격차에 대한 연구에서 빈도 분석 결과 중 ‘정보서비스’, ‘도서관 서비스’의 빈도가 높게 나타난 것과 유사한 결과를 보여준다(이성신 외, 2019).

〈표 5〉는 서론 요약문을 통해 구성된 동시 출현 네트워크의 중심성 분석 결과이다.

중심성이 높은 주요 단어로는 ‘분석’, ‘정보’, ‘도서관’으로 나타났으며, 이어 ‘방안’, ‘공공도서관’, ‘서비스’라는 단어가 높은 중심성을 갖는 것으로 나타났다. 이를 통해 정보격차와 관련하여 도서관 정보 서비스의 분석이 주요 연구 주제임을 확인할 수 있다. 또한, 앞서 분석된 제목 및 초록 텍스트의 결과와 유사한 단어들이

〈표 4〉 제목-초록 키워드 네트워크 중심성 분석

순위	연결중심성		매개중심성		근접중심성	
1	정보	51	정보	0.34	정보	0.7
2	도서관	51	도서관	0.27	도서관	0.68
3	서비스	51	서비스	0.25	서비스	0.68
4	공공도서관	37	분석	0.12	공공도서관	0.63
5	분석	35	공공도서관	0.12	분석	0.62

〈표 5〉 서론 요약문의 키워드 네트워크 중심성 분석

순위	연결중심성		매개중심성		근접중심성	
1	분석	51	분석	0.3	분석	0.78
2	정보	47	정보	0.24	정보	0.75
3	도서관	40	도서관	0.17	도서관	0.7
4	방안	33	공공도서관	0.08	방안	0.65
5	공공도서관	33	방안	0.08	공공도서관	0.64
6	서비스	30	서비스	0.05	서비스	0.64

등장하는 것을 확인 할 수 있다.

〈표 6〉은 방법론 요약문을 통해 구성된 동시 출현 네트워크의 중심성 분석 결과이다.

중심성이 가장 높은 주요 단어는 ‘분석’으로 타 단어들에 비해 매우 높은 중심성으로 보이고 있으며, 이어서 ‘통계’, ‘도서관’이 높은 중심성을 보이고 있다. 이는 국내 문헌정보학계의 정보격차 연구에서 주로 사용되고 있는 분석 방법이 통계 분석이라는 것을 보여주고 있다.

강인서와 김혜진(2020)의 연구에서는 정보격차 연구 동향 분석 결과 양적 연구는 전체 연구의 97%를 차지하고, 질적 연구는 3%만 수행되어 양적 연구 방법에 편중되어 있다고 언급한 것과 유사한 내용을 나타내고 있다.

〈표 7〉은 결과 요약문을 통해 구성된 동시 출현 네트워크의 중심성 분석 결과이다.

중심성이 높은 단어로 ‘서비스’, ‘정보’, ‘도서

관’, ‘프로그램’ 순으로 나타나고 있다. 이를 통해 정보격차 연구의 주요 결과로는 도서관 정보 서비스와 프로그램에 대한 내용이 주를 이루고 있다는 것을 보여주고 있다. 또한, 함께 등장하는 ‘부족’, ‘지원’, ‘제안’ 등의 단어는 정보격차 연구에서 주요 문제와 제언이 도서관 정보 서비스 및 프로그램의 부족 또는 지원의 부족이라고 살펴볼 수 있다. 조용완(2007)의 연구에서는 국외 도서관의 정보취약계층 도서관 서비스 관련 지침 등에서 언급된 내용들과 현재 국내 공공도서관에서 제공되고 있는 정보취약계층 관련 서비스의 내용을 비교하여, 우리 공공도서관의 전체적인 상황은 매우 미흡하다는 결과를 도출하였다. 또한 안인자 외(2018)의 연구에서는 한국에서 저소득계층을 대상으로 서비스하는 공공도서관의 사례를 분석한 결과, 저소득계층만을 위한 국내 공공도서관의

〈표 6〉 방법론 요약문의 키워드 네트워크 중심성 분석

순위	연결중심성		매개중심성		근접중심성	
1	분석	101	분석	0.86	분석	0.95
2	통계	24	통계	0.03	통계	0.56
3	도서관	23	활용	0.02	도서관	0.54
4	실시	23	장애인	0.02	실시	0.51
5	자료	21	시스템	0.02	자료	0.51
6	서비스	19	관찰	0.02	수집	0.50

〈표 7〉 결과 요약문의 키워드 네트워크 중심성 분석

순위	연결중심성		매개중심성		근접중심성	
1	서비스	72	서비스	0.43	서비스	0.64
2	정보	61	정보	0.37	정보	0.62
3	도서관	36	도서관	0.12	프로그램	0.53
4	프로그램	29	분석	0.07	도서관	0.52
5	교육	20	제안	0.06	교육	0.50
6	부족	18	프로그램	0.05	분석	0.50

서비스 사례는 전반적으로 많지 않은 것으로 결과를 도출하였다.

〈표 8〉은 논의 요약문을 통해 구성된 동시 출현 네트워크의 중심성 분석 결과이다.

중심성이 높은 주요 단어로는 ‘필요’, ‘강화’, ‘제언’으로 나타났다. 이어 ‘서비스’라는 단어도 높은 중심성을 갖는 것으로 나타났는데, 이는 정보격차에 대한 해소 방안으로 서비스의 필요성과 강화를 제언한 것으로 파악된다. 이제환과 윤유라(2005)의 연구에서는 외국인 노동자의 정보 빈곤을 해소하기 위한 방안으로 맞춤형 정보서비스, 언어 지원 프로그램, 공식적 정보채널 확대가 필요함을 제안하였으며, 이정연(2011)의 연구에서는 정보소외계층을 위한 도서관서비스는 통합적이고 전문적인 정책 개입과 협력 체계 강화가 필요하다 제언하였다.

네트워크 중심성 분석 결과, 각 요약문의 성격이 반영된 단어들이 높은 중심성을 보였다.

제목-초록 텍스트에서는 ‘정보’, ‘도서관’, ‘서비스’, ‘공공도서관’, ‘분석’이 높은 중심성을 보였다. 서론에서는 ‘분석’, ‘정보’, ‘도서관’, ‘방안’, ‘공공도서관’이 중심성이 높은 단어로 등장하여 제목-초록 텍스트의 결과와 순위의 차이만 보이고, 유사한 결과가 확인되었다. 방법론에서는 ‘분석’, ‘통계’가 주요 단어로 도출되어, 정보격차 연구에서 통계적 분석이 주요 방법으로 활

용되고 있음을 보여준다. 결과 파트에서는 ‘서비스’, ‘정보’, ‘도서관’이 높은 중심성을 보였다. 이는 제목-초록 텍스트와 서론에서 등장한 주요 단어들과 유사한 단어들이 등장한 것으로 보여진다. 그러나 ‘프로그램’, ‘부족’, ‘제언’이라는 단어들이 높은 중심성을 갖는 것으로 나타나, 도서관 정보 서비스 및 프로그램의 부족으로 해석되어 차이를 보인다. 논의 파트에서는 ‘필요’, ‘강화’, ‘제언’이 높은 중심성을 보여, 정보격차 해소를 위한 도서관 서비스의 필요성과 강화 방안에 대한 제언이 연구의 주요 결론으로 제시되고 있음을 확인할 수 있다.

4.3.2 모듈리티 분석 결과 비교

각각 네트워크의 모듈리티 분석을 통해 주요 주제를 파악하고자 하였다.

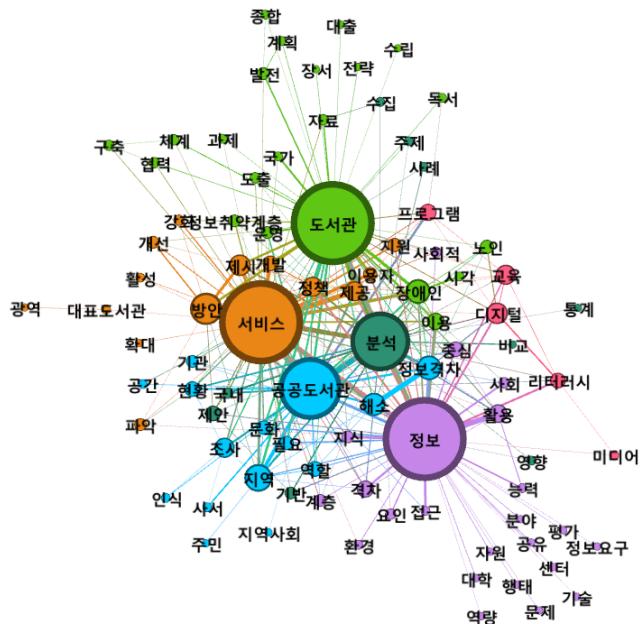
〈그림 3〉은 제목 및 초록 데이터를 통해 구축된 네트워크를 시각화 한 것이다. 네트워크를 구축한 결과 총 90개의 노드와 281개의 링크로 이루어져 있으며, 커뮤니티 검출 결과 6개의 class가 나타나는 것을 확인할 수 있다.

〈표 9〉는 제목과 초록 데이터 기반 네트워크의 모듈리티 분석 결과 나타난 상위 빈도 단어 5개를 나타낸 것이다.

각 Class의 상위 빈도 단어의 해석을 바탕으로 각 class의 주제를 표현하였다. Class 1은 ‘지

〈표 8〉 논의 요약문의 키워드 네트워크 중심성 분석

순위	연결중심성		매개중심성		근접중심성	
1	필요	48	필요	0.43	필요	0.83
2	강화	36	강화	0.11	제언	0.71
3	제언	35	제언	0.08	지원	0.67
4	지원	31	지원	0.08	서비스	0.67
5	서비스	31	서비스	0.06	강화	0.65



〈그림 3〉 제목 및 초록 텍스트의 키워드 네트워크

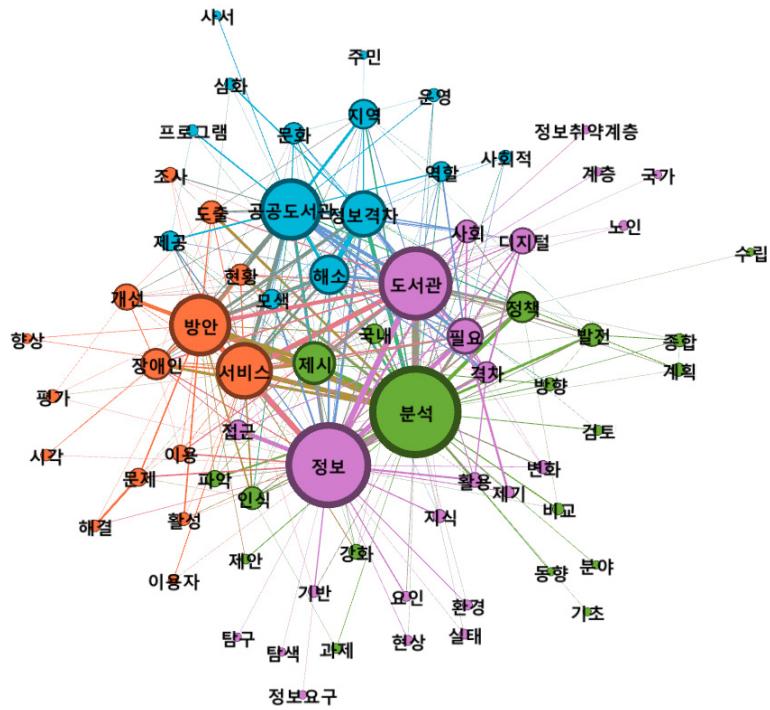
〈표 9〉 제목 및 초록 기반 키워드 네트워크의 모듈리티 분석 결과

순위	CLASS 1	CLASS 2	CLASS 3	CLASS 4	CLASS 5	CLASS 6
1	공공도서관	서비스	도서관	분석	정보	프로그램
2	지역	방안	장애인	제안	활용	리터러시
3	정보격차	제시	이용	기반	격차	디지털
4	해소	제공	노인	국내	중심	교육
5	조사	정책	발전	수집	지식	미디어

역 공공도서관의 정보 격차 해소'로 주제를 표현하였으며, 네트워크 내의 비중은 16.7%로 나타났다. Class 2(16.7%)는 '이용자 서비스 개발 및 제공과 정책 제안', Class 3(24.4%)은 '정보취약계층의 도서관 이용', Class 4(11.1%)는 '국내 사례 분석', Class 5(25.6%)는 '지식 정보 활용에 대한 접근성 격차', Class 6(5.6%)은 '디지털 리터러시 교육 프로그램'으로 주제를 표현하였다. 또한 네트워크에서 Class 5와 3의 비중이 높아 정보격차 연구에서 '지식 정보 활용에 대한 접근성 격차'와 '정보취약계층의 도서관 이용'이 핵심적인 주제로 나타나는 것을 확인할 수 있다.

〈그림 4〉는 ChatGPT 통해 구성 요소별로 작성된 요약문 중 서론을 기반으로 구축된 네트워크를 시각화한 것이다. 네트워크를 구축한 결과 총 71개의 노드와 308개의 링크로 이루어져 있으며, 모듈리티 분석 결과 4개의 class가 나타나는 것을 확인할 수 있다.

〈표 10〉은 서론 요약문 기반 네트워크의 모



〈그림 4〉 서론 요약문 기반 키워드 네트워크

〈표 10〉 서론 요약문 기반 키워드 네트워크의 모듈리티 분석 결과

순위	CLASS 1	CLASS 2	CLASS 3	CLASS 4
1	분석	정보	방안	공공도서관
2	제시	도서관	서비스	정보격차
3	정책	필요	장애인	해소
4	발전	디지털	개선	지역
5	인식	사회	도출	문화
6	국내	격차	현황	역할

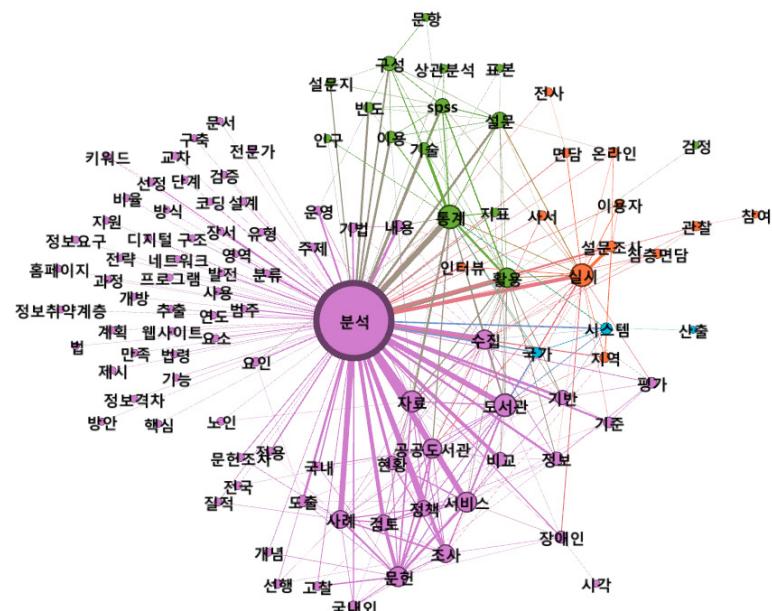
둘리티 분석 결과 나타난 상위 빈도 단어 5개를
나타낸 것이다.

각 Class의 상위 빈도 단어의 해석을 바탕으로 Class 1(26.8%)은 ‘국내 정책’, Class 2(32.4%)는 ‘디지털 정보 격차’, Class 3(21.1%)은 ‘장애인 서비스 개선’, Class 4(19.7%)는 ‘지역 공공도서관 정보격차 해소’로 주제를 표현하였다. 또한

네트워크에서 Class 2의 비중이 높아 정보격차 연구에서 '디지털 정보 격차'가 핵심적인 주제로 나타나는 것을 확인할 수 있다.

〈그림 5〉는 ChatGPT 통해 구성 요소별로 작성된 요약문 중 방법론을 기반으로 구축된 네트워크를 시각화한 것이다.

네트워크를 구축한 결과 총 108개의 노드와



〈그림 5〉 방법론 요약문 기반 키워드 네트워크

〈표 11〉 방법론 요약문 기반 키워드 네트워크의 모듈리티 분석 결과

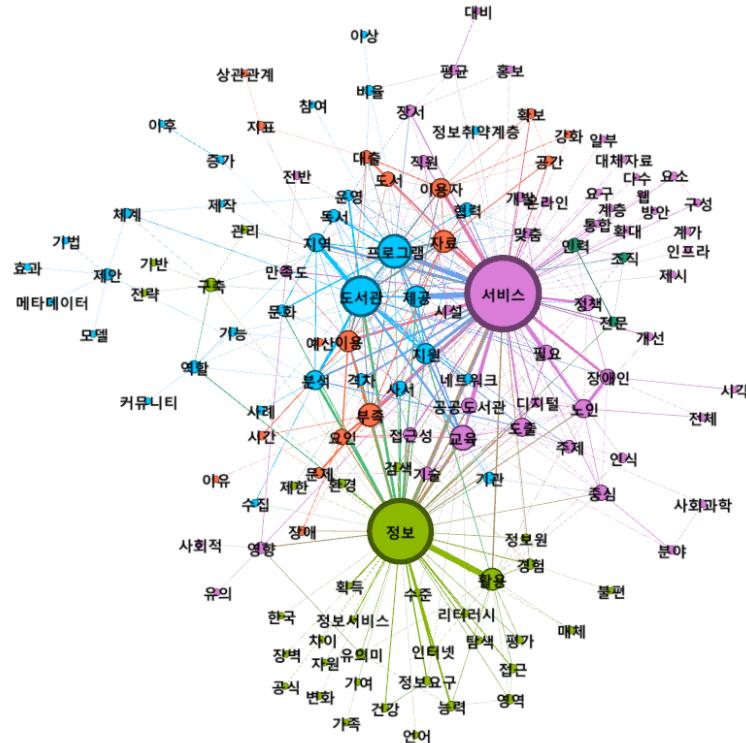
순위	CLASS 1	CLASS 2	CLASS 3	CLASS 4
1	분석	실시	통계	시스템
2	도서관	설문조사	활용	국가
3	자료	지역	설문	산출
4	문헌	관찰	구성	
5	서비스	실험면담	spss	

311개의 링크로 이루어져 있으며, 모듈리티 분석 결과 4개의 class가 나타나는 것을 확인할 수 있다. 〈표 11〉은 방법론 요약문 기반 네트워크의 모듈리티 분석 결과 나타난 상위 빈도 단어 5개를 나타낸 것이다.

각 Class의 상위 빈도 단어의 해석을 바탕으로 Class 1(72.2%)은 ‘문헌 자료 분석’, Class 2(11.1%)는 ‘설문조사와 면담’, Class 3(13.9%)은 ‘설문 통계’, Class 4(2.8%)는 ‘국가 시스템 산출’로 주제를 표현하였다. 네트워크에서 Class

1의 비중이 72.2%로 과반수 이상을 차지하고 있어, 정보격차 연구의 방법론으로 ‘문헌 자료 분석’이 주요 연구 방법으로 나타나는 것을 확인할 수 있다.

<그림 6>은 ChatGPT 통해 구성 요소별로 작성된 요약문 중 결과를 기반으로 구축된 네트워크를 시각화한 것이다. 네트워크를 구축한 결과 총 134개의 노드와 387개의 링크로 이루어져 있으며, 모듈리티 분석 결과 5개의 class가 나타나는 것을 확인할 수 있다.



〈그림 6〉 결과 파트 요약문 키워드 네트워크

〈표 12〉는 결과 요약문 기반 네트워크의 모듈러티 분석 결과 나타난 상위 빈도 단어 5개를 나타낸 것이다.

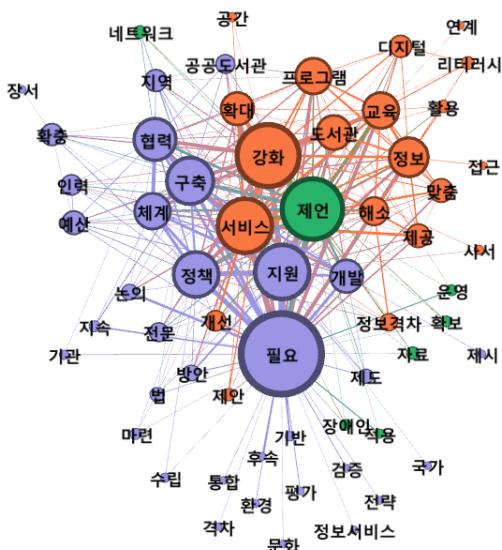
각 Class의 상위 빈도 단어의 해석을 통해
Class 1(26.1%)은 ‘인터넷 정보 활용 능력’, Class
2(35.1%)는 ‘정보 취약 계층 교육 서비스 필
요’, Class 3(23.9%)은 ‘도서관의 프로그램 제

공’, Class 4(12.7%)는 ‘정보격차 발생 요인’, Class 5(2.2%)는 ‘전문 인력’으로 주제를 표현하였다. 또한 네트워크에서 Class 2의 비중이 높아 정보격차 연구 결과로 ‘정보 취약 계층 교육 서비스 필요’가 핵심적인 주제로 나타나는 것을 확인할 수 있다.

〈그림 7〉은 ChatGPT 통해 구성 요소별로

〈표 12〉 결과 요약문 기반 키워드 네트워크의 모듈리티 분석 결과

순위	CLASS 1	CLASS 2	CLASS 3	CLASS 4	CLASS 5
1	정보	서비스	도서관	부족	인력
2	활용	교육	프로그램	자료	전문
3	구축	도출	제공	이용	조직
4	경험	노인	지원	이용자	
5	인터넷	장애인	분석	요인	



〈그림 7〉 논의 파트 요약문 키워드 네트워크

작성된 요약문 중 논의를 기반으로 구축된 네트워크를 시각화한 것이다. 네트워크를 구축한 결과 총 61개의 노드와 293개의 링크로 이루어져 있으며, 모듈리티 분석 결과 3개의 class가 나타나는 것을 확인할 수 있다.

〈표 13〉은 논의 요약문 기반 네트워크의 모듈리티 분석 결과 나타난 상위 빈도 단어 5개를 나타낸 것이다.

각 Class의 상위 빈도 단어의 해석을 바탕으로 Class 1(55.7%)은 '정책 지원 및 협력 체계 구축 필요', Class 2(32.8%)는 '도서관 정보 교

육 서비스 프로그램 강화', Class 3(11.5%)은 '장애인 자료 및 네트워크 운영'으로 주제를 표현하였다. 네트워크에서 Class 1의 비중이 과반수 이상을 차지 하는 모습을 보이고 있어, 정보 격차 연구의 논의점으로 '정책 지원 및 협력 체계 구축의 필요'가 핵심적인 주제로 나타나는 것을 확인할 수 있다.

〈표 14〉는 네트워크 모듈리티 분석 결과를 종합한 것이다.

네트워크의 모듈리티 분석 결과 제목-초록 텍스트에서는 총 6개의 클래스가 도출되었으

〈표 13〉 논의 요약문 기반 키워드 네트워크의 모듈리티 분석 결과

순위	CLASS 1	CLASS 2	CLASS 3
1	필요	강화	제언
2	지원	서비스	네트워크
3	구축	정보	자료
4	정책	교육	운영
5	협력	프로그램	확보

〈표 14〉 키워드 네트워크 모듈리티 분석 결과 비교

Class	제목-초록	ChatGPT			
		구성요소별 요약문			
		서론	방법론	결과	논의
1	지역 공공도서관의 정보격차 해소	국내 정책	문헌 자료 분석	인터넷 정보 활용 능력	정책 지원 및 협력 체계 구축 필요
2	이용자 서비스 개발 및 제공과 정책 제안	디지털 정보 격차	설문조사와 면담	정보 취약 계층 정보 교육 서비스 필요	도서관 정보 교육 서비스 프로그램 강화
3	정보취약 계층의 도서관 이용	장애인 서비스 개선	설문 통계	도서관의 프로그램 제공	장애인 자료 및 네트워크 운영
4	국내 사례 분석	지역 공공도서관 정보격차 해소	국가 시스템 산출	정보격차 발생요인	
5	지식 정보 활용에 대한 접근성 격차			전문 인력	
6	디지털 리터러시 교육 프로그램				

며, 구성요소별 요약문은 총 16개의 클래스가 도출되었다. 도출된 제목-초록 텍스트의 클래스는 연구의 큰 맥락을 파악할 수 있는 주제들로 구성되어 있는 반면, 구성요소별 요약문에서 도출된 클래스들은 상대적으로 구체적이고 세부적인 내용을 확인 할 수 있다. 서론에서는 연구주제, 방법에서는 구체적인 연구방법, 결론에서는 문제점, 논의에서는 해결책과 관련된 내용 및 각 부분의 핵심적인 내용에 대해 파악 할 수 있다는 특징을 가진다.

5. 논의 및 결론

5.1 GPT를 통해 생성된 요약문의 신뢰성 평가에 대한 논의

인공지능을 활용한 자동 요약 기법은 크게 추출 요약(Extractive Summarization)과 생성 요약(Abstractive Summarization)으로 구

분된다(Sharma & Sharma, 2022). 추출 요약은 원문에서 핵심이 되는 문장이나 구절을 그대로 발췌하여 요약문을 구성하는 방식이며, 생성 요약은 본문의 내용을 이해한 후 새로운 표현과 문장 구조로 요약문을 재구성하는 방법이다(윤세희, 신유현, 2024).

요약문의 평가는 정량적 평가와 정성적 평가로 나눌 수 있다. 정량적 평가는 자동 평가 지표를 활용하여 작성된 요약문과 정답 요약문을 비교하는 방식으로 진행된다. 대표적으로 ROUGE (Lin, 2004)와 BERTScore(Zhang et al., 2019)가 사용된다. ROUGE는 정답 요약문과의 N-gram 일치율을 기반으로 요약문의 품질을 평가하고, BERTScore는 사전 학습된 언어모델을 활용해 정답 요약문과의 문맥적 유사성을 측정한다. 정성적 평가는 인간 평가자가 요약문의 일관성, 유창성, 관련성, 응집성 등과 같은 기준을 바탕으로 직접 평가하는 방식이다.

본 연구에서 사용한 ChatGPT를 활용하여 생성된 요약문은 주로 생성 요약 기법을 통해

작성된 요약문이며, 생성 요약문의 경우 정량적 평가와 정성적 평가가 추가로 진행되어 결과의 신뢰도를 높이는 것이 보편적이다(윤세희, 신유현, 2024). 그러나 생성 요약문의 정량적 평가는 추출 요약문과 달리 요약문을 생성하는 단계에서 원문에 없는 단어나 새로운 문장 구조를 활용할 수 있기 때문에, 기존의 평가 방식처럼 정답 요약문과의 비교를 통한 평가는 생성 요약문의 특성을 정확하게 반영하기 어렵다(윤세희, 신유현, 2024; Yang et al., 2023). 또한 정성적 평가를 위한 인간 평가는 평가자들 간의 평가 일관성에 대한 문제와 시간과 비용이 많이 소요된다는 한계가 존재한다(최규리, 김한샘, 2024).

GPT를 활용한 요약문은 인간에 의한 요약문과 비교하여 높은 평가와 선호를 받는 동시에(서종훈, 2024; Pu et al., 2023; Zhang et al., 2024), 일관성이나 신뢰성 등에 대한 문제점이 지적되고 있다(Luo et al., 2024; Peters & Chin, 2025). GPT와 같은 모델의 생성 요약문을 평가할 때 명확한 평가 기준이 설정되지 않았기 때문에 상반되는 평가가 나타나는 것이라 볼 수 있다. 이를 위해 생성 요약의 평가에 대한 명확한 기준 또는 지표를 설정하는 연구가 필요할 것으로 보여진다.

본 연구에서는 앞서 언급한 바와 같이 GPT 기반 요약문의 신뢰성 평가와 관련된 한계를 충분히 인지하고 있었으며, 방법론과 결과 해석에 있어서 이를 고려하여 기존의 방식과 다른 형태의 결과가 나타날 수 있음에 초점을 두었다. 또한 GPT모델은 등장 이후부터 꾸준히 발전하고 있고(OpenAI, 2023), 시간의 흐름에 따라 한계들도 점차 개선될 것으로 보여 gpt 기

반 요약문을 활용한 방법론을 제안하는 것에 충분한 의의를 가진다고 생각한다.

5.2 결론

본 연구는 기존의 동향 분석 연구가 제목 및 초록 데이터를 기반으로 포괄적인 주제를 파악하는데 초점을 두고 있어, ChatGPT 기반 요약문을 활용하여 세부적인 내용을 파악할 수 있는 방법을 제시하였다.

연구 결과는 다음과 같이 요약할 수 있다. 첫째, 초록과 서론 요약문의 결과는 유사하게 나타난다는 점이다. 이는 초록의 내용이 주로 연구의 목적과 배경을 설명하는 문장들로 구성되어 있어, 결과 간의 유사성이 높게 나타난 것으로 해석할 수 있다. 이태영(2005)의 연구에 따르면, 초록은 목적 문장이 27%, 배경 문장이 27%, 결과 문장이 21%, 방법 문장이 7%, 결론 문장이 7%, 설명 문장이 11%로 구성되어 있는 것으로 나타났기 때문에 초록과 서론 요약문은 유사한 내용을 다룬다고 할 수 있다.

둘째, 각각의 구성요소 별 요약문에서는 보다 세부적인 내용이 파악된다는 점이다. 기존의 방식에서 파악하기 어려운 세부 방법론이나 중요하게 제기된 문제점, 해결 방안 등에 대한 파악이 가능하다는 것이다.

이러한 결과를 바탕으로 제시한 방법론은 다음과 같은 특성을 가진다. 첫째, 전문을 구조적으로 요약하였기 때문에 요약 파트에 따라 동일 단어의 등장에 대한 해석이 달라질 수 있다. ‘분석’이라는 단어는 모든 텍스트에서 등장하는 단어지만 방법론에서 등장할 경우 문헌 분석, 사례 분석, 통계 분석 등 분석 방법에 대한 내용

을 파악할 수 있다.

둘째, 각 요약문 별로 동시에 출현한 단어가 상이하기 때문에 요약문 별 다각적 분석이 가능하다. ‘정보격차’라는 단어와 동시에 출현한 단어들을 분석한 결과를 보면 방법론 파트와 결과 파트에서 특히 상이한 결과를 확인할 수 있었다. 방법론에서 등장한 단어들은 ‘키워드’, ‘실태’, ‘문헌’ 등의 단어가 등장하여 주요 방법론으로 키워드 분석, 실태 분석, 문헌 분석이 주요 분석 방법임을 나타내고 있으며, 결과 파트에서 등장하는 ‘성별’, ‘연령’, ‘경제’ 등의 단어들은 분석 결과 주요 계층에 따라 비교 분석하였다는 내용으로 해석할 수 있다.

셋째, 요약문 별 네트워크가 별개로 생성되기 때문에 네트워크 군집 분석과정에서 세분화된 주제의 도출이 가능하다. 제목-초록 텍스트의 키워드 네트워크의 모듈리티 분석을 통해 총 6개의 Class를 확인 할 수 있었으나, 각 요약문의 키워드 네트워크의 모듈리티 분석을 통해 나온 클래스는 총 16개의 클래스가 도출되어 보다 구체적이고 논문 구성요소별로 세분화된 분석이 가능해진다.

넷째, 각 분석 결과를 서론, 방법론, 결과, 논의로 이어지는 흐름에 따라 결과를 분석할 수 있

다. 방법론에서 등장하는 ‘통계’, ‘설문’과 같은 단어들은 결과 파트에서 등장하는 ‘평균’, ‘비율’과 같은 단어와 연결되어 주요 연구 방법이 양적 연구방법이며, 이에 따라 정량적 수치를 통해 주요 결과를 도출하였다는 해석이 가능해진다.

본 연구는 기준 제목-초록을 통해 텍스트 분석을 실행하였던 연구들과 달리 AI를 활용한 요약문을 활용하여 또 다른 방법론적 접근 가능성을 탐색하였다는 점에서 의의가 있다. 한편, 본 연구에는 앞서 논의에서 언급한 바와 같이 다음의 제한점이 존재한다. 첫째, 생성된 요약문에 대한 체계적 평가가 이루어지지 않고, 둘째, ChatGPT가 동일 프롬프트에 대해 일관되지 않은 답변을 제시하는 회발성 문제에 대한 보완이 부족하다는 점에서 한계를 가진다. 추가적으로 제안한 방법론을 적용하기 위해서는 논문 전문을 활용해야 하므로 저작권 문제에 대한 고려가 필요하다는 점이 한계점으로 지적될 수 있다.

향후 연구에서는 ChatGPT 요약문에 대한 체계적인 평가를 통해 데이터의 신뢰성을 확보하고, 시계열 분석을 추가함으로써 보다 상세하고 구체적인 분석 방법을 제시할 수 있을 것으로 기대한다.

참 고 문 헌

- 감미아 (2023). 「서지학연구」의 연구동향에 대한 계량서지학적 분석: 2002-2022 게재논문을 중심으로. *서지학연구*, 93, 51-75. <https://doi.org/10.17258/JIB.2023..93.51>
- 강순애 (2013). 기록관리의 역사에 관한 국내의 연구 동향: 회고와 전망. *한국기록관리학회지*, 13(3), 41-65. <http://doi.org/10.14404/jksarm.2013.13.3.041>

- 강인서, 김혜진 (2020). 문헌정보학 분야 정보격차 연구동향 분석. *정보관리학회지*, 37(2), 333-352.
<https://doi.org/10.3743/KOSIM.2020.37.2.333>
- 고정현, 강우진, 이종욱 (2021). 디지털 정보격차 관련 국내 연구 동향 분석. *한국도서관·정보학회지*, 52(4), 179-203. <http://doi.org/10.16981/kliss.52.4.202112.179>
- 구본진, 장덕현 (2023). 문헌정보학분야 해외 연구 동향 및 유망 주제 분석 연구. *한국문헌정보학회지*, 57(3), 71-96. <https://doi.org/10.4275/KSLIS.2023.57.3.071>
- 김규환, 장보성, 이현정 (2009). 우리나라 기록관리학 분야의 연구영역 분석: 논문제목의 구문 및 의미 구조를 중심으로. *한국문헌정보학회지*, 43(3), 417-439.
<https://doi.org/10.4275/KSLIS.2009.43.3.417>
- 김정현 (2011) 한국의 자료조직 분야 연구동향 분석: 1970~2010. *한국도서관·정보학회지*, 42(3), 149-164. <https://doi.org/10.16981/kliss.42.3.201109.149>
- 김하진, 송민 (2014). 동시출현단어 분석을 통한 국내외 정보학 학회지 연구동향 파악. *정보관리학회지*, 31(1), 99-118. <https://doi.org/10.3743/KOSIM.2014.31.1.099>
- 김희전, 조현양 (2010). 저자동시인용분석과 저자서지결합분석에 의한 지적 구조 분석: 사회복지학 분야 를 중심으로. *정보관리학회지*, 27(3), 283-306. <http://doi.org/10.3743/KOSIM.2010.27.3.283>
- 남태우, 이진영 (2009). 우리나라 기록관리학 연구 동향 분석. *한국도서관·정보학회지*, 40(2), 451-472. <https://doi.org/10.16981/KLISS.40.2.200906.451>
- 박민석, 이지수 (2024). 체계적 문헌고찰을 통한 국내 디지털 큐레이션 연구동향 분석. *한국기록관리학회지*, 24(2), 41-63. <https://doi.org/10.14404/JKSARM.2024.24.2.041>
- 박영희, 김성희 (2024). 서지 매핑을 이용한 도서관 정보기술 및 인공지능에 관한 연구동향 분석. *한국비블리아학회지*, 35(4), 45-65. <https://doi.org/10.14699/kbiblia.2024.35.4.045>
- 박옥남 (2011) 정보조직 지식구조에 대한 연구: 2000년~2011년 학술논문을 중심으로. *한국비블리아학회지*, 22(3), 247-267. <https://doi.org/10.14699/kbiblia.2011.22.3.247>
- 박은정, 조성준 (2014). KoNLPy: 쉽고 간결한 한국어 정보처리 파이썬 패키지. 제 26회 한글 및 한국어 정보처리 학술대회 논문집.
- 박준형, 오효정 (2017). 국내 기록관리학 연구동향 분석을 위한 토픽모델링 기법 비교: LDA와 HDP를 중심으로. *한국도서관·정보학회지*, 48(4), 235-258. <http://doi.org/10.16981/kliss.48.201712.235>
- 배경재 (2016). 도서관장애인서비스 품질요인의 실증적 평가에 대한 연구. *한국문헌정보학회지*, 50(2), 271-287. <https://doi.org/10.4275/KSLIS.2016.50.2.271>
- 서종훈 (2024). Chat GPT류 AI와 학습자 간 텍스트 요약 양상 고찰. *어문학*, 165, 5-47. <https://doi.org/10.37967/emh.2024.9.165.5>
- 손정표 (2003) 한국의 문헌정보학분야 연구동향 분석: 1957~2002. *한국도서관·정보학회지*, 34(3), 9-32.

- 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>
- 신유미, 박옥남 (2019). 장서개발관리 분야 최근 연구동향 분석에 대한 연구. *정보관리학회지*, 36(2), 105-131. <http://doi.org/10.3743/KOSIM.2019.36.2.105>
- 안인자, 노영희, 장로사 (2018). 저소득계층의 정보불평등 해소를 위한 도서관서비스 관련 연구. *한국비
블리아학회지*, 29(4), 113-137. <https://doi.org/10.14699/kbiblia.2018.29.4.113>
- 오선경, 이지연 (2011). 시각장애인의 공공도서관 비이용 요인에 관한 연구. *한국문현정보학회지*, 45(3), 141-162. <https://doi.org/10.4275/KSLIS.2011.45.3.141>
- 오세훈 (2005). 우리나라 문현정보학 학술지 논문 및 인용문헌 분석을 통한 연구동향 연구. *정보관리학
회지*, 22(3), 379-408. <https://doi.org/10.3743/KOSIM.2005.22.3.379>
- 윤세희, 신유현 (2024). 한국어 생성 요약 성능 평가 지표 분석 연구. *정보처리학회 논문지*, 13(12), 691-699.
<https://doi.org/10.3745/TKIPS.2024.13.12.691>
- 이다현, 신동희 (2024). 대학생 대상 정보 리터러시 연구의 키워드 네트워크 분석 및 토픽 모델링. *정보
관리학회지*, 41(3), 249-268. <https://doi.org/10.3743/KOSIM.2024.41.3.249>
- 이성숙, 이지원 (2024). 문현정보학 교육 분야 연구동향 및 지적구조 분석. *정보관리학회지*, 41(4), 209-234. <https://doi.org/10.3743/KOSIM.2024.41.4.209>
- 이성신 (2016). 키워드 네트워크 분석을 통한 도서관마케팅 연구 경향 분석: 우리나라와 국외연구의
비교분석. *한국문현정보학회지*, 50(3), 383-402. <http://doi.org/10.4275/KSLIS.2016.50.3.383>
- 이성신, 강보라, 이세나 (2019). 정보격차 연구 동향 분석: 문현정보학분야와 일반사회과학분야와의
비교. *한국도서관·정보학회지*, 50(3), 139-166. <https://doi.org/10.16981/kliss.50.201909.139>
- 이세나, 이성신, 백수민 (2024). 저자 키워드와 초록 분석을 통한 법학사서 연구동향 분석. *한국문현정보
학회지*, 58(2), 5-31. <https://doi.org/10.4275/KSLIS.2024.58.2.005>
- 이승민 (2021). 미디어 리터러시가 스마트 디바이드 형성에 미치는 영향. *정보관리학회지*, 38(2), 19-38.
<https://doi.org/10.3743/KOSIM.2021.38.2.019>
- 이용재, 김지은 (2021). 한국 도서관경영 연구의 동향과 과제-2014~2020 학술논문을 중심으로. *한국문
현정보학회지*, 55(1), 209-230. <https://doi.org/10.4275/KSLIS.2021.55.1.209>
- 이재윤 (2015). 문현동시인용 분석을 통한 한국 문현정보학의 연구 전선 파악. *정보관리학회지*, 32(4),
77-106. <https://doi.org/10.3743/KOSIM.2015.32.4.077>
- 이재윤 (2023). 저자집단 분석을 통한 한국 문현정보학의 학술커뮤니케이션 동향 연구. *한국문현정보학
회지*, 57(2), 409-434. <https://doi.org/10.4275/KSLIS.2023.57.2.409>
- 이재윤, 김관준, 강대신, 김희정, 유소영, 이우형 (2011). 계량서지적 기법을 활용한 LED 핵심 주제영역
의 연구 동향 분석. *Journal of Information Science Theory and Practice*, 42(3), 1-26.

- https://doi.org/10.1633/JIM.2011.42.3.001
- 이정연 (2011). 정보소외계층을 위한 도서관의 문화복지 정책에 관한 연구. *한국비블리아학회지*, 22(2), 27-39.
- 이제환, 윤유라 (2005). 외국인 노동자 H의 삶과 정보빈곤. *한국비블리아학회지*, 16(1), 181-202.
- 이종우, 박혜린 (2016). 문헌정보학 분야 국제 학술지 논문 계량분석을 통한 국내 연구자 특성 연구. *한국도서관·정보학회지*, 47(1), 217-242. https://doi.org/10.16981/KLISS.47.1.201603.217
- 이태영 (2005). 자동 발췌문/요약 시스템 구축에 관한 연구: 학술지 논문기사를 중심으로. *한국문헌정보학회지*, 39(3), 139-163.
- 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
- 장로사 (2023). 키워드 네트워크 분석을 활용한 국내 공공도서관 연구 동향 분석. *한국비бл리아학회지*, 34(4), 285-302. https://doi.org/10.14699/KBIBLIA.2023.34.4.285
- 조용완 (2007). 국내 공공도서관의 정보격차해소 활동에 대한 분석. *한국도서관·정보학회지*, 38(4), 167-186. https://doi.org/10.16981/kliss.38.4.200712.167
- 최규리, 김한샘 (2024). 생성형 AI모델의 요약 성능 평가 연구. *언어사실과 관점*, 62, 203-227. https://doi.org/10.20988/lfp.2024.62..203
- 최재은 (2024). 텍스트 마이닝을 활용한 국외 데이터 큐레이션 연구 동향 분석. *정보관리학회지*, 41(3), 85-107. https://doi.org/10.3743/KOSIM.2024.41.3.085
- 최재황, 천세진 (2022). 도서관 및 도서관학 분야 북한 논문의 관심 주제 및 용어 분석. *한국도서관·정보학회지*, 53(3), 377-404. https://doi.org/10.16981/KLISS.53.3.202209.377
- 한미경 (2017). 타 전공 연구자의 서지학적 연구 동향. *서지학연구*, 72, 225-250. https://doi.org/10.17258/JIB.2017..72.225
- 홍덕용, 허준석 (2023). 비정형 텍스트 데이터 분석을 활용한 기록관리 분야 연구동향. *한국기록관리학회지*, 23(4), 73-89. https://doi.org/10.14404/JKSARM.2023.23.4.073
- 홍소람, 장우권 (2023). 메타리터러시 연구동향 분석: 정보 리터러시와의 차이를 중심으로. *한국문헌정보학회지*, 57(2), 97-122. https://doi.org/10.4275/KSLIS.2023.57.2.097
- 횡금숙 (2005). 국내 독서치료 연구동향 분석 연구. *한국도서관·정보학회지*, 36(1), 117-130.
- American Psychological Association (2019). *Publication Manual of the American Psychological Association* (7th ed.). 강진령 옮김(2022). *APA논문작성법*(원서 7판). 서울: 학지사.
- Lin, C. Y. (2004). Rouge: A package for automatic evaluation of summaries. In *Text summarization branches out*, 74-81.
- Luo, Z., Xie, Q., & Ananiadou, S. (2024). Factual consistency evaluation of summarization in the Era of large language models. *Expert Systems with Applications*, 254, 124456.

<https://doi.org/10.48550/arXiv.2402.13758>

Milojević, S., Sugimoto, C. R., Yan, E., & Ding, Y. (2011). The cognitive structure of library and information science: Analysis of article title words. *Journal of the American Society for Information Science and Technology*, 62(10), 1933-1953. <https://doi.org/10.1002/asi.21602>

OpenAI (2023). Gpt-4 technical report. arXiv preprint arXiv:2303.08774.

<https://doi.org/10.48550/arXiv.2303.08774>

Peters, U. & Chin-Yee, B. (2025). Generalization bias in large language model summarization of scientific research. *Royal Society Open Science*, 12(4), 241776.

<https://doi.org/10.1098/rsos.241776>

Pu, X., Gao, M., & Wan, X. (2023). Summarization is (almost) dead. arXiv preprint arXiv: 2309.09558. <https://doi.org/10.48550/arXiv.2309.09558>

Sharma, G. & Sharma, D. (2023). Automatic text summarization methods: A comprehensive review. *SN COMPUT. SCI.* 4, 33. <https://doi.org/10.1007/s42979-022-01446-w>

Sugimoto, C. R., Li, D., Russell, T. G., Finlay, S. C., & Ding, Y. (2011). The shifting sands of disciplinary development: Analyzing North American Library and Information Science dissertations using latent Dirichlet allocation. *Journal of the American Society for Information Science and Technology*, 62(1), 185-204. <https://doi.org/10.1002/asi.21435>

Yang, X., Li, Y., Zhang, X., Chen, H., & Cheng, W. (2023). Exploring the limits of chatgpt for query or aspect-based text summarization. <https://doi.org/10.48550/arXiv.2302.08081>

Zhang, T., Kishore, V., Wu, F., Weinberger, K. Q., & Artzi, Y. (2019). Bertscore: Evaluating text generation with bert. <https://doi.org/10.48550/arXiv.1904.09675>

Zhang, T., Ladhak, F., Durmus, E., Liang, P., McKeown, K., & Hashimoto, T. B. (2024). Benchmarking large language models for news summarization. *Transactions of the Association for Computational Linguistics*, 12, 39-57. https://doi.org/10.1162/tacl_a_00632

Zhao, D. & Strotmann, A. (2008). Information science during the first decade of the web: An enriched author cocitation analysis. *Journal of the American Society for Information Science and Technology*, 59(6), 916-937. <https://doi.org/10.1002/asi.20799>

• 국문 참고문헌에 대한 영문 표기
(English translation of references written in Korean)

Ahn, Inja, Noh, Younghée, & Chang, Rosa (2018). A study on the library services for the solution of the information inequality of the low-income people in Korea. *Journal of the Korean*

- Biblia Society for Library and Information Science, 29(4), 113-137.
<https://doi.org/10.14699/kbiblia.2018.29.4.113>
- Bae, KyungJae (2016). A study on the empirical evaluation of quality factor on library service for the disabled. Journal of the Korean Society for Library and Information Science, 50(2), 271-287. <https://doi.org/10.4275/KSLIS.2016.50.2.271>
- Chang, Rosa (2023). An analysis of research trends on public libraries in Korea using keyword network analysis. Journal of the Korean Biblia Society for Library and Information Science, 34(4), 285-302. <https://doi.org/10.14699/KBIBLIA.2023.34.4.285>
- Cho, Yong Wan (2007). Analysis of the activities of public libraries for closing the information divide in Korea. Journal of Korean Library and Information Science Society, 38(4), 167-186. <https://doi.org/10.16981/kliss.38.4.200712.167>
- Choi, Gyu ri & Kim, Hansaem (2024). Summary evaluation performance studies with generative AI models. Language Facts and Perspectives, 62, 203-227.
<https://doi.org/10.20988/lfp.2024.62..203>
- Choi, Jae Hwang, & Cheon, Sejin (2022). An analysis of research areas and related terms in library and library science in North Korea. Journal of Korean Library and Information Science Society, 53(3), 377-404. <https://doi.org/10.16981/KLISS.53.3.202209.377>
- Choi, Jaeeun (2024). Analysis of research trends in data curation using text mining techniques. Journal of the Korean Society for Information Management, 41(3), 85-107.
<https://doi.org/10.3743/KOSIM.2024.41.3.085>
- Han, Mikyung (2017). Bibliographical research trend of researchers with different majors. Journal of Studies in Bibliography, 72, 225-250. <https://doi.org/10.17258/JIB.2017..72.225>
- Hoang, Gum-Sook (2005). An analytical study on research trends of bibliotherapy in Korea. Journal of Korean Library and Information Science Society, 36(1), 117-130.
- Hong, Deok Yong & Heo, Junseok (2023). Research trends in record management using unstructured text data analysis. Journal of Korean Society of Archives and Records Management, 23(4), 73-89. <https://doi.org/10.14404/JKSARM.2023.23.4.073>
- Hong, Soram & Chang, Woo Kwon (2023). Metaliteracy research trends analysis: Focused on the difference from information literacy. Journal of the Korean Society for Library and Information Science, 57(2), 97-122. <https://doi.org/10.4275/KSLIS.2023.57.2.097>
- Kam, Miah (2023). A bibliometric analysis of research trends in the journal of studies in bibliography: Articles from 2002 to 2022. Journal of Studies in Bibliography, null(93), 51-75.
<https://doi.org/10.17258/JIB.2023..93.51>

- Kang, In-Seo & Kim, Hea-Jin (2020). Analysis of 'digital divide' research trends in library and information science. *Journal of the Korean Society for Information Management*, 37(2), 333-352. <https://doi.org/10.3743/KOSIM.2020.37.2.333>
- Kang, Soon-Ae (2013). Research trends of records and archives management history in Korea: Retrospect and prospect. *Journal of Records Management and Archives Society of Korea*, 13(3), 41-65. <http://doi.org/10.14404/jksarm.2013.13.3.041>
- Kim, Gyuhwan, Jang, Bo Seong, & Yi, Hyun Jung (2009). A study on intellectual structure of records management and archives in Korea: Based on syntactic and semantic structure of article titles. *Journal of the Korean Society for Library and Information Science*, 43(3), 417-439. <https://doi.org/10.4275/KSLIS.2009.43.3.417>
- Kim, Ha Jin & Song, Min (2014). A study on the research trends in domestic/international information science articles by co-word analysis. *Korean Society of Information Management*, 31(1), 99-118. <https://doi.org/10.3743/KOSIM.2014.31.1.099>
- Kim, Heejeon & Cho, Hyun Yang (2010). A study on intellectual structure using author co-citation analysis and author bibliographic coupling analysis in the field of social welfare science. *Korean Society of Information Management*, 27(3), 283-306. <http://doi.org/10.3743/KOSIM.2010.27.3.283>
- Kim, Jeong-Hyen (2011). An analytical study on research trends of resource organization in Korea: 1970~2010. *Journal of Korean Library and Information Science Society*, 42(3), 149-164. <https://doi.org/10.16981/kliss.42.3.201109.149>
- Ko, Jeonghyeon, Kang, Woojin, & Lee, Jongwook (2021). Research trend analysis of digital divide in South Korea. *Journal of Korean Library and Information Science Society*, 52(4), 179-203. <http://doi.org/10.16981/kliss.52.4.202112.179>
- Koo, Bon Jin & Chang, Durk Hyun (2023). Research on overseas trends and emerging topics in field of library and information science. *Journal of the Korean Society for Library and Information Science*, 57(3), 71-96. <https://doi.org/10.4275/KSLIS.2023.57.3.071>
- Lee, Da-Hyeon & Shin, Dong-Hee (2024). Keyword network analysis and topic modeling in an information literacy study of undergraduate students. *Journal of the Korean Society for Information Management*, 41(3), 249-268. <https://doi.org/10.3743/KOSIM.2024.41.3.249>
- Lee, Hye-Kyung, Yang, Ki-duk, & Kim, Seon-Wook (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, Jae Yun (2015). Identifying the research fronts in Korean library and information science

- by document co-citation analysis. *Journal of the Korean Society for Information Management*, 32(4), 77-106. <https://doi.org/10.3743/KOSIM.2015.32.4.077>
- Lee, Jae Yun (2023). A study on scholarly communication trends in Korean library and information science studies through author group analysis. *Journal of the Korean Society for Library and Information Science*, 57(2), 409-434. <https://doi.org/10.4275/KSLIS.2023.57.2.409>
- Lee, Jae Yun, Kim, Pan Jun, Kang, DaeShin, Kim, Hee jung, Yu, So-Young, & Lee, Woo-Hyoung (2011). A bibliometric analysis on LED research. *Journal of Information Management*, 42(3), 1-26. <https://doi.org/10.1633/JIM.2011.42.3.001>
- Lee, Jae-Whoan & Youn, You-Ra (2005). Life and information poverty of foreign labor 'H'. *Journal of the Korean Biblia Society for Library and Information Science*, 16(1), 181-202.
- Lee, Jongwook & Bak, Hyerin (2016). Characteristics of Korean researchers through bibliometric analysis of papers published in international LIS journals. *Journal of Korean Library and Information Science Society*, 47(1), 217-242. <https://doi.org/10.16981/KLISS.47.1.201603.217>
- Lee, Jung Yeoun (2021). A study on the social welfare library policy for the information alienated groups journal of the Korean biblia society for library and information science, 22(2), 27-39
- Lee, Sena, Lee, Seongsin, & Baeck, Sumin (2024). An analysis of research trends in law librarians through author keywords and abstract analysis. *Journal of the Korean Society for Library and Information Science*, 58(2), 5-31. <https://doi.org/10.4275/KSLIS.2024.58.2.005>
- Lee, Seongsin (2016). A study on the library marketing research trends through keyword network analysis: Comparative analysis of Korea and other countries. *Journal of the Korean Society for Library and Information Science*, 50(3), 383-402.
<http://doi.org/10.4275/KSLIS.2016.50.3.383>
- Lee, Seongsin, Kang, Bora, & Lee, Sena (2021). A study on the information divide research trends: Comparative analysis of LIS fields and other social science fields. *Journal of Korean Library and Information Science Society*, 50(3), 139-166.
<https://doi.org/10.16981/kliiss.50.201909.139>
- Lee, Seungmin (2021). Effect of media literacy on the formation of smart divide. *Journal of the Korean Society for Information Management*, 38(2), 19-38.
<https://doi.org/10.3743/KOSIM.2021.38.2.019>
- Lee, Sung Sook & Lee, Ji Won (2024). A study on analysis of research trends and intellectual structure in the LIS education research. *Journal of the Korean Society for Information Management*, 41(4), 209-234. <https://doi.org/10.3743/KOSIM.2024.41.4.209>
- Lee, Tae Young (2005). A study on the construction of the automatic extracts and summaries:

- On the basis of scientific journal articles. *Journal of the Korean Society for Library and Information Science*, 39(3), 139-163.
- Lee, Yong Jae & Kim, Ji-Eun (2021). Trends and tasks of library management research in Korea: Focused on academic papers 2014~2020. *Journal of the Korean Society for Library and Information Science*, 55(1), 209-230. <https://doi.org/10.4275/KSLIS.2021.55.1.209>
- Nam, TaeWoo & 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>
- Oh, Se-Hoon (2005). A study on the research trends of library & information science in Korea by analyzing Journal articles and the Cited Literatures. *Journal of the Korean Society for Information Management*, 22(3), 379-408. <https://doi.org/10.3743/KOSIM.2005.22.3.379>
- Oh, Seon-Kyung & Lee, Jeeyeon (2011). A study on the non-use factors of the visually disabled in public libraries. *Journal of the Korean Society for Library and Information Science*, 45(3), 141-162. <https://doi.org/10.4275/KSLIS.2011.45.3.141>
- Park, Eun-Jeoung & Cho, Sung-Zoon (2014). KoNLPy: Korean natural language processing in Python, Proceedings of the 26th Annual Conference on Human & Cognitive Language Technology.
- Park, Jun-Hyeong & 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. <http://doi.org/10.16981/kliss.48.201712.235>
- Park, Minseok & Lee, Jisue (2024). A systematic review of trends of domestic digital curation research. *Journal of Korean Society of Archives and Records Management*, 24(2), 41-63. <https://doi.org/10.14404/JKSARM.2024.24.2.041>
- Park, Ok Nam (2011). Knowledge structures in knowledge organization research: 2000-2011. *Journal of the Korean Biblia Society for Library and Information Science*, 22(3), 247-267. <https://doi.org/10.14699/kbiblia.2011.22.3.247>
- Park, Younghhee & Kim, Seonghee (2024). Research trends on information technology and artificial intelligence for libraries using bibliographic mapping. *Journal of the Korean Biblia Society for Library and Information Science*, 35(4), 45-65. <https://doi.org/10.14699/kbiblia.2024.35.4.045>
- Seo, Jong-Hoon (2024). Exploration of text summarization patterns between ChatGPT-like AIs and learners. *Eomunhak*, 165, 5-47. <http://doi.org/10.37967/emh.2024.9.165.5>

- Shin, You Mi & Park, Oknam (2019). An analytical study on research trends of collection development and management. *Journal of the Korean Society for Information Management*, 36(2), 105-131. <http://doi.org/10.3743/KOSIM.2019.36.2.105>
- 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>
- Son, Jeong-Pyo (2003). Trend analysis of library and information science research in Korea: 1957~2002. *Journal of the Korean Library and Information Science Society*, 34(3), 9-32.
- Yoon, Sehwi & Shin, Youhyun (2024). A study on automatic metrics for Korean text abstractive summarization. *The Transactions of the Korea Information Processing Society*, 13(12), 691-699. <https://doi.org/10.3745/TKIPS.2024.13.12.691>