

개인 온톨로지 기반 지식 그래프 리랭킹이 PKMS 검색 품질에 미치는 영향

The Effect of Personal Ontology-Based Knowledge Graph Reranking on PKMS Search Quality

정혜지(Hyeji Jung)

E-mail: arche.jhj@gmail.com

컨텍스트에이 (ContextA) 대표



논문접수 2026-04-08

최초심사 2026-04-22

게재확정 2026-05-26

ORCID

Hyeji Jung
<https://orcid.org/0009-0003-1812-9830>

© 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.

초 록

본 연구는 개인지식관리시스템에서 생산자의 사고방식과 가치관을 반영한 온톨로지 기반 검색 리랭킹의 효과를 검증한다. 기록학의 출처주의가 생산 맥락으로 기록의 의미를 보존하듯, 생산자의 인지적 맥락이 검색 관련성을 결정한다는 가설 하에, 연구자 본인의 마크다운 기반 PKMS 환경 3,702건을 대상으로 시맨틱 검색 결과를 7개 온톨로지 요인(허브 부스트, 정체성 렌즈, 교차 도메인, 사고 패턴, 그래프 근접도, 가치 정렬, 저자성)으로 리랭킹하는 POR(v1.0~v2.1.0)을 설계하였다. 5,717 노드와 13,544 �지의 지식 그래프에서 10개 허브 노트 전체가 PageRank 상위 15위 이내에 위치하여 허브 선정의 타당성을 확인하였다. 10개 쿼리의 실시간 A/B 비교 결과, 평균 Kendall's tau 0.621(SD=0.252)로 쿼리 유형에 따라 차등적 재배열을 보였다. 허브 관련 쿼리(tau=0.300)에서 가장 큰 순위 변경이 발생하였고, 일반 쿼리(tau=0.786)에서는 기존 순위가 보존되었다. 정체성 렌즈 매칭(57건)과 교차 도메인 교집합(29건)이 가장 활발한 요인이었다. 이로써 개인 온톨로지가 시맨틱 유사도만으로는 포착할 수 없는 맥락적 관련성을 쿼리 유형에 따라 차등 제공함을 확인하였다.

ABSTRACT

This study investigates the efficacy of ontology-based search reranking that incorporates the producer's cognitive dispositions and value systems within a Personal Knowledge Management System (PKMS). Grounded in the archival principle of provenance, the study hypothesizes that integrating a producer's cognitive patterns and value systems can enhance the relevance of search results in personal records. To test this, POR (ver. 1.0-2.1.0), a system that reranks Gemini File Search RAG results using seven ontology scoring factors—hub boost, identity lens matching, cross-domain intersection, thinking pattern boost, graph proximity, value alignment, and authorship—was developed. The system was assessed on 3,702 markdown-based PKM notes, constructing a knowledge graph with 5,717 nodes and 13,544 edges, including 1,810 auto-linked HUB_CONNECTION edges. All 10 hub notes ranked within the top 15 by PageRank, supporting the validity of the hub selection methodology. In a live A/B comparison with 10 queries across 5 types, ontology-based reranking changed an average of 4.4 out of 10 positions, with a mean Kendall's tau of 0.621 (SD = 0.252). Hub-related queries showed the largest reranking (tau = 0.300), while general queries preserved existing rankings (tau = 0.786). Identity lens matching (57 activations) and cross-domain intersection (29 activations) were the most frequently activated factors. These results demonstrate that personal ontology modeling provides contextual relevance beyond what embedding-based semantic similarity alone can achieve, with effectiveness varying by query type.

Keywords: 온톨로지, RiC-O, 지식 그래프, 검색 리랭킹, 개인지식관리(PKM), RAG RAG ontology, RiC-O, Knowledge graph, Search reranking, Personal Knowledge Management (PKM), Retrieval-augmented generation

<https://jksarm.koar.kr>

1. 서론

1.1 연구 배경

개인지식관리(Personal Knowledge Management, PKM) 도구의 보급이 확산됨에 따라 개인이 축적하는 디지털 기록의 양이 급격히 증가하고 있다. 하나의 마크다운 기반 PKMS(Personal Knowledge Management System)에 수천 건 이상의 노트가 축적되면 키워드 검색이나 폴더 탐색만으로는 원하는 정보를 효율적으로 찾기 어렵다. 최근 대규모 언어 모델(Large Language Model, LLM) 기반 RAG(Retrieval-Augmented Generation) 기술이 시맨틱 검색을 가능하게 했지만, 이러한 검색은 텍스트의 의미 유사도만으로 결과를 정렬해 생산자의 맥락을 반영하지 못한다.

기록학에서 ‘맥락(context)’은 기록의 의미를 결정하는 결정적 요소이다. 출처주의(principle of provenance)는 기록을 생산자별로 분리 유지하여 생산 맥락을 보존하는 원칙이며, 원질서 존중(Original Order)은 기록의 원래 배열이 담고 있는 의미를 보존하는 원칙이다(Shepherd & Yeo, 2003). 최근 RiC(Records in Contexts) 표준은 기존의 계층적 기술에서 맥락 중심의 다차원 기술로 패러다임을 전환하며, 기록을 고립된 개체가 아닌 ‘맥락 속의 관계’로 기술한다(강주연 외, 2025). 이러한 기록학의 맥락 중심 사고는 개인기록의 검색에도 적용할 수 있다. PKMS에서 ‘출처’란 곧 생산자 자신의 사고방식, 가치관, 관심 영역이며, 이를 정형화하면 검색의 맥락적 관련성을 높일 수 있다.

1.2 연구 목적 및 연구 질문

본 연구의 목적은 개인 온톨로지 기반 지식 그래프 리랭킹(Personal Ontology Reranking, 이하 POR)이 PKMS 시맨틱 검색의 품질에 어떤 영향을 미치는지 실험으로 검증하는 데 있다. POR은 본 연구에서 설계한 검색 후처리 시스템을 가리키는 약어로서, 본 논문 전체에서 같은 의미로 사용한다. 구체적인 연구 질문은 다음과 같다.

- RQ1: 개인 온톨로지 기반 리랭킹이 PKMS 시맨틱 검색의 순위를 유의미하게 변경하는가?
- RQ2: 리랭킹된 결과가 사용자의 주관적 적합성 평가에서 더 높은 만족도를 보이는가?
- RQ3: 지식 그래프의 구조적 특성(허브 연결성, 그래프 근접도)이 리랭킹 효과에 어떤 영향을 미치는가?

RQ1은 온톨로지 리랭킹이 시맨틱 검색 결과의 순위를 실질적으로 변경하는지, 즉 리랭킹의 작동 여부 자체를 검증한다. RQ2는 순위 변경이 실제 사용자의 정보 요구에 부합하는지를 주관적 적합성 평가로 확인한다. RQ3는 지식 그래프의 구조적 속성이 리랭킹 효과의 차이를 설명하는 요인인지를 분석하여, 시스템 설계의 이론적 근거를 제공한다.

1.3 연구의 범위와 한계

본 연구는 연구자 본인의 마크다운 기반 PKMS 환경(3,702건, 2025.09 ~ 2026.03 축적)에 대한 사례 연구이며, 출처주의의 구조적 매핑을 단일 사용자 환경에서 1차로 실증하는 위상에 한정된다. 즉 1차 실증 위상의 자기 한정이다. 본 논문은 다음과 같이 구성된다. 2장에서 기록학적 맥락론, PKMS 검색, 온톨로지 기반 정보 검색의 선행 연구를 검토한다. 3장에서 시스템 아키텍처, 개인 온톨로지 설계, 지식 그래프 구축 방법을 기술한다. 4장에

서 6종의 주요 실험 설계와 결과를 제시한다. 5장에서 연구 질문에 대한 답변, 기록학적 함의, 일반화 가능성을 논의하고 결론을 맺는다. 단일 사용자가 자신의 기록 환경을 직접 설계하고 실험한 사례 연구의 성격으로 인해 결과의 일반화에는 한계가 분명하나, 3,702건의 실제 사용 데이터에 대해 정량적 실험을 수행한 점에서, 개인기록 검색이라는 구체적 맥락을 통한 온톨로지의 효과를 깊이 있게 분석할 수 있다.

2. 이론적 배경 및 선행 연구

2.1 기록학에서의 맥락과 검색

기록학의 주요 원칙인 출처주의와 원질서 존중은 기록의 생산 맥락을 보존하여 의미를 유지하는 것을 목표로 한다. ISAD(G)(ICA, 2000)로 대표되는 전통적 기술 체계는 기록을 Fonds→Series→File→Item의 계층 구조로 기술하며, Shepherd & Yeo(2003)는 이 구조가 생산 맥락을 반영하고, 검색도구(Finding Aids)가 이용자의 기록 접근을 돕는 메타데이터 체계로 기능한다고 설명하였다. 그러나 이 계층적 기술은 기록 간의 수평적 관계나 다차원적 맥락을 표현하는 데 한계가 있었다.

이러한 한계를 극복하기 위해 ICA의 EGAD(Expert Group on Archival Description)는 2023년 RiC 표준의 첫 버전을 발표하였다(ICA EGAD, 2023). RiC-CM(Conceptual Model)은 기존 ISAD(G), ISAAR(CPF), ISDF, ISDIAH의 네 가지 국제 표준을 통합하여 기록을 계층 구조 대신 맥락의 네트워크로 기술하는 패러다임 전환을 추구했다. RiC의 주요 개체인 Record, Agent, Activity, Place 등과 이들 간의 관계는 기록의 맥락을 다차원적으로 기술할 수 있게 한다(강주연 외, 2025). 특히 RiC-O(Ontology)는 OWL 2 온톨로지로 구현되어 링크드 오픈 데이터(LOD) 환경에서 기록 기술을 가능하게 한다. 이를 바탕으로 김현채 외(2024)는 RiC-O를 활용하여 생산기관 변천정보 서비스를 개선하는 방안을 제시하며 온톨로지가 기록의 맥락적 연결을 강화할 수 있음을 확인했다.

한국 기록학계에서 AI 기반 연구는 빠르게 늘고 있다. 김민지 외(2024)는 ChatGPT를 활용하여 기록관리 메타데이터를 자동 추출하는 사례를 연구하였고, 송치호(2024)는 자연어 처리의 개체명 인식(NER)을 통한 기록집합체의 메타데이터 추출 방안을 제시하였다. 최유리 외(2025)는 지능형 기록정보서비스 챗봇의 응답 논리 구조를 설계하였으며, 이수현과 오효정(2024)은 대통령기록포털의 검색어 로그를 분석하여 이용자 검색 행태를 연구하였다. 박하람 외(2024)는 FAIR 원칙을 디지털 아카이브에 적용하여 평가하였다. 그러나 이러한 연구들은 대부분 공공기록을 대상으로 하며, 개인, 기업, 단체 등 공공영역 이외의 기록에 대한 온톨로지 기반 검색 실험은 찾아보기 어렵다.

본 연구의 개인 온톨로지가 출처주의 및 원질서 존중과 맺는 관계는 원리의 직접 이식이라기보다 구조적 매핑에 가깝다. 두 원칙이 보존하려는 것(생산자별 분리, 원래 배열의 의미)을 검색 알고리즘의 가중치 구조에 매핑한 것이며, 생산자와 이용자가 분리되는 기관 기록의 전제는 본 연구의 PKMS 환경에서 충족되지 않는다. 다만 RiC-CM이 출처주의를 Agent 개체의 정체성·역할·가치 속성으로 확장 정의한 한도 내에서는, 본 연구도 그 확장된 정의에 부분적으로 동일한 선상에서 논의될 수 있다. PKMS 환경에서 생산자=이용자라는 출처주의의 원래 전제가 깨지는 문제는 본 연구의 한계로서 마지막 절에서 다룬다.

2.2 PKM과 시맨틱 검색

PKM은 개인이 지식을 수집, 조직, 검색, 활용하는 체계이다. 그 기원은 독일 사회학자 니클라스 루만(Niklas Luhmann)의 제텔카스텐(Zettelkasten)에서 찾을 수 있다. 루만은 1950년대부터 약 90,000장의 색인카드를 축적하며 70권의 저서와 400편 이상의 논문을 생산하였고, Schmidt(2016)는 이 카드 색인이 저장소에서 나아가 ‘사고

도구이자 커뮤니케이션 파트너'였다고 분석하였다. Ahrens(2017/2023)는 이를 현대적으로 재해석하여, 원자적 노트(Atomic Notes)와 노트 간 링크를 통한 지식 네트워크 구축 방법론을 제시하였다.

현대 PKM의 또 다른 축은 티아고 포르테(Tiago Forte)의 세컨드 브레인 방법론이다. 포르테는 PARA 체계(Projects, Areas, Resources, Archive)를 통해 디지털 정보를 실행 중심으로 조직하는 방법을 제안하였다(Forte, 2022/2023). PARA는 현재 진행 중인 프로젝트(Projects)를 최상위에 두고, 지속적으로 관리해야 할 영역(Areas), 인젠가 유용할 자원(Resources), 완료된 항목의 보관소(Archive)로 계층화한다. 이는 Fonds 계층 구조와 유사하지만, 생산 맥락이 아닌 '실행 가능성'을 기준으로 정보를 분류한다는 차이가 있다. 포르테는 또한 CODE 워크플로우(Capture, Organize, Distill, Express)를 제시하여, 정보의 수집에서 창작물 생산까지의 전 과정을 체계화하였다. 이 방법론은 노션(Notion), 에버노트(Evernote) 등 데이터베이스 기반 도구와 잘 결합되며, 특히 프로젝트 중심 워크플로우를 선호하는 지식 노동자들에게 널리 채택되었다.

이 외에도 다양한 PKM 방법론이 제안되었다. 에버그린 노트(Evergreen Notes)는 시간이 지나도 유효성을 유지하도록 지속적으로 갱신하는 영구 노트 개념을 강조하며, 노트 작성의 밀도와 연결성을 주요 원칙으로 삼는다. 닉 마일로(Nick Milo)의 LYT(Linking Your Thinking) 방법론은 MOC(Maps of Content)라는 개념을 도입하여, 특정 주제에 대한 노트들을 하나의 지도 노트와 연결하는 방식을 제시한다(Milo, 2020). 이는 허브 노트(hub notes) 개념과 유사하며, 본 연구의 개인 온톨로지에서 허브 노트가 수행하는 역할과 유사하다.

한국에서는 2020년대 초반부터 제텔카스텐과 마크다운 기반 PKM 도구에 대한 관심이 급증하였다. 특히 옵시디언(Obsidian)을 활용한 제텔카스텐 구현에 대한 한국어 커뮤니티 리소스가 활발히 생산되고 있다. PARA 체계를 조니 데시멀(Johnny Decimal) 번호 체계와 결합한 폴더 구조 템플릿, Dataview 플러그인을 활용한 자동 인덱스 생성 방법, Templater를 통한 일일 노트 자동화 등 구체적인 워크플로우가 공유된다. 또한 한글의 형태소 특성을 고려한 검색 최적화(예: 조사 제거, 어간 추출), 한글 문서에 대한 LLM 기반 임베딩 성능 비교 등의 실험도 커뮤니티 수준에서 진행되고 있다. 그러나 이러한 실천 사례에 대한 학술적 연구는 아직 초기 단계이며, 개인 온톨로지 기반 검색 최적화와 같은 심화 연구는 찾아보기 어렵다.

옵시디언, 로엠리서치(Roam Research), 로그시크(LogSeq) 등은 제텔카스텐을 디지털로 구현한 대표적 도구이다. 이들은 공통적으로 마크다운 파일과 위키링크로 노트 간 연결 구조를 형성하며, 특정 도구에 종속되지 않는 마크다운 기반 PKMS 환경을 구성한다.

이와 같이 PKM 도구의 발전으로 개인이 지식을 체계적으로 축적하고 연결하는 것이 용이해졌으나, 노트 수가 수천 건을 넘어서면 새로운 문제가 발생한다. 키워드 검색은 정확한 용어를 기억해야 하고, 폴더 탐색은 분류 체계를 숙지해야 하며, 위키링크 탐색은 이미 연결된 경로만 따라갈 수 있다. 생산자 자신도 과거에 작성한 노트의 존재를 잊거나 관련성 있는 노트 간의 연결을 발견하지 못하는 상황이 빈번해진다. 곧 PKM의 가치는 축적에 비례해 늘지만, 검색 효율성은 규모에 반비례하여 떨어진다.

이 검색 문제를 해결하기 위해 RAG 기술이 주목받고 있다. Lewis et al.(2020)이 제안한 RAG는 LLM의 생성 능력에 외부 문서 검색을 결합하여 응답의 정확성을 높이는 기술이다. Gao et al.(2024)은 Naive RAG(단순 검색-생성 파이프라인), Advanced RAG(사전/사후 검색 최적화), Modular RAG(구성 요소 교체 기능)의 3세대 패러다임으로 RAG의 발전 과정을 정리하였다.

그러나 RAG 기술의 강점인 시맨틱 검색은 텍스트의 임베딩 벡터 간 코사인 유사도만으로 결과를 정렬하므로 근본적 한계가 몇 가지 있다. 첫째, 임베딩은 개별 텍스트 청크의 의미를 포착하지만 문서 간 구조적 관계(위키링크, 폴더 계층, 태그 체계)를 반영하지 못한다. 둘째, 동일한 쿼리에 대해 모든 사용자에게 동일한 결과를 반환하여 생산자의 관심 영역, 사고 패턴, 가치관 등 개인적 맥락을 반영하지 못한다. 셋째, Collins와 Loftus(1975)가 제안한 확산 활성화 이론(spreading activation theory)이 인간의 기억 검색에서 연상적 연결을 통한 활성화 확산을 설명하는 것처럼, PKM 검색에서도 유사도 외에 개인의 인지 구조를 반영한 연상적 검색이 필요하다.

2.3 온톨로지 기반 정보 검색 향상

시맨틱 검색의 한계를 극복하기 위해 지식 그래프와 온톨로지를 활용한 검색 향상 연구가 진행되고 있다. Edge et al.(2024)은 Microsoft GraphRAG를 발표하며, LLM을 활용하여 소스 문서에서 개체 지식 그래프를 추출하고, 밀집 연결된 노드 집합을 커뮤니티로 감지한 후, 각 커뮤니티의 요약물 계층적으로 생성하는 접근을 제안하였다. 이 커뮤니티 요약 기반 검색은 ‘데이터셋의 주요 주제는 무엇인가?’와 같은 전역적 질의에서 기존 Naive RAG를 상회하는 포괄성과 다양성을 보였다. Peng et al.(2024)은 Graph RAG 분야에서 그래프 기반 인덱싱(G-Indexing), 그래프 가이드 검색(G-Retrieval), 그래프 강화 생성(G-Generation)의 3단계 분류 체계를 제시하였다. 그러나 GraphRAG는 범용적 지식 그래프를 구축하므로, 특정 사용자의 관점이나 가치체계를 반영하지 않는다.

Gutierrez et al.(2024)은 HippoRAG를 제안하며 뇌의 해마 기반 장기 기억 메커니즘을 RAG에 적용하였다. HippoRAG는 LLM이 대뇌피질의 역할을, 지식 그래프가 해마 색인의 역할을, Personalized PageRank(PPR)(Page et al., 1999)가 패턴 완성의 역할을 수행하는 3중 구조이다. 이 접근은 다중 단계 질의응답에서 기존 방법 대비 최대 20% 향상을 보였으며, 반복 검색 대비 10-20배 저렴하며, 6-13배 빠른 효율성을 달성하였다. 그러나 HippoRAG의 PPR은 그래프 구조에 기반한 범용적 패턴과 관련있으며 특정 개인의 사고방식이나 가치관을 반영하는 메커니즘은 포함되어 있지 않다.

개인 지식 그래프(Personal Knowledge Graph, PKG) 연구도 존재한다. Balog와 Kenter(2019)는 PKG의 개념을 ‘사용자에게 개인적으로 관련된 개체, 그 속성, 개체 간 관계에 대한 구조화된 정보 자원’으로 정의하며, 개인화된 서비스의 기반이 될 수 있다고 제안하였다. Chakraborty와 Sanyal(2023)은 PKG 연구의 포괄적 서베이에서 의학, 금융, 교육 분야의 활용 사례와 데이터 소스에 따른 구축 방법을 분류하였다. Skjæveland et al.(2024)은 PKG 생태계 서베이에서 구축, 활용, 관리의 3축 분류를 제시하며 PKG가 대규모 백과사전적 지식 그래프의 한계를 보완하는 사용자 중심 지식 표현임을 강조하였다. 그러나 기존 PKG 연구는 주로 사용자의 개체 수준 데이터(일정, 선호도, 인적 네트워크)를 구조화하는 데 초점을 맞추며, 사용자의 사고 패턴이나 가치 체계를 모델링하여 검색을 개선하는 접근을 찾기는 어렵다.

2.4 연구의 차별성

본 연구의 POR 접근은 기존 연구와 세 가지 측면에서 구분된다. 첫째, 개인화의 구체성이다. GraphRAG와 HippoRAG가 범용적 지식 그래프를 구축하고 PKG 연구가 개체 수준의 개인 데이터를 구조화하는 반면, 본 연구의 POR은 생산자의 정체성 렌즈, 사고 패턴, 가치 체계 등 인지 수준의 개인화를 구현한다(<표 1>의 개인화 수준 참고).

둘째, 검색 향상 방법이 다르다. GraphRAG의 커뮤니티 요약이나 HippoRAG의 PPR 기반 패턴 완성이 확률적·통계적 접근인 반면, POR의 7개 스코어링 요인은 규칙 기반(rule-based)이어서 각 요인의 기여를 해석 가능하고 검증 가능하다(<표 1>의 검색 향상 방법 참고).

<표 1> 관련 접근법 연구 비교

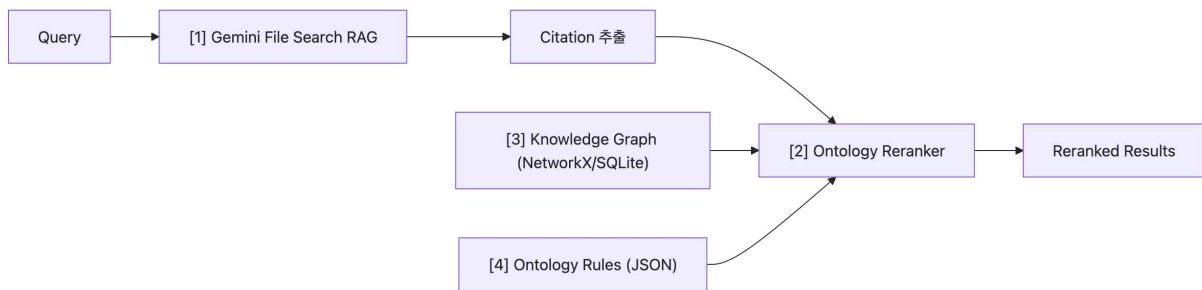
접근법	지식소스	개인화 수준	검색 향상 방법
Naive RAG	임베딩 벡터	없음	코사인 유사도
GraphRAG	범용 개체 KG	없음	커뮤니티 요약
HippoRAG	KG + PPR	없음(범용)	패턴 완성
PKG 연구	개인 개체 데이터	개체 수준	추천/QA
POR(본 연구)	개인 온톨로지 규칙	인지 수준	7요인 리랭킹

셋째, 기록학적 기반이 존재한다. 본 연구의 POR은 기록학의 출처주의 원리를 PKMS 검색에 그대로 옮긴 것이 라기보다, 출처주의가 보존하려는 생산자 맥락의 구조를 검색 가중치 구조에 매핑한 것이다(2.1 참고). 이 매핑의 매개는 RiC-CM의 Agent 개체이다. 개인 온톨로지가 생산자의 정체성·사고 패턴·가치 체계를 모델링하는 것은 RiC-CM이 Agent 속성으로 정의한 정체성·역할·가치를 개인 수준에서 구체화한 것이며, 이 매핑이 닿는 한도 내에서는 출처주의 원리의 부분 적용에 해당한다. 따라서 본 연구의 위상은 출처주의의 구조적 매핑이며, RiC-CM 과의 정합 한도 내에서는 부분적 본질 적용도 함께 이뤄진다. AI 시대에 기록학적 원리를 개인기록관리에 적용하는 새로운 시도라는 점에서 의의가 있으나, 본 연구에서의 시도가 곧 출처주의의 직접 이식이 아님을 분명히 한다.

3. 시스템 설계

3.1 파이프라인 아키텍처

POR v1.0은 다음 4개 구성 요소로 이루어진 파이프라인 아키텍처이다.



〈그림 1〉 POR 파이프라인 4개 구성 요소와 데이터 흐름

3.1.1 Gemini File Search RAG

Google Gemini 2.5 Flash 모델의 File Search API를 활용하여 마크다운 기반 PKMS의 노트를 시맨틱 검색한다. 7개 File Search Store에 961개 파일이 인덱싱되어 있으며(전체 3,702건 중 콘텐츠 밀도가 높은 노트를 선별, 시스템 파일, 일일 노트, 템플릿 등 제외), 3단계 우선순위(CRITICAL/HIGH/MEDIUM)로 분류된다. 실제 작업 환경은 마크다운 파일과 파인더 폴더로 구성되며, 오피시디언은 뷰어로, VS Code와 CLI는 편집 및 자동화 도구로 사용된다.

3.1.2 Ontology Reranker

Gemini File Search RAG가 1차로 반환한 상위 N건의 시맨틱 검색 결과를 입력으로 받아, 3.2에서 정의한 7개 온톨로지 요인의 점수를 산출하고 시맨틱 점수와 가중 합산하여 최종 순위를 재정렬하는 후처리 모듈이다. 가중 합산식은 식 (1)과 같으며, 가중치 조합은 4.4 가중치 시뮬레이션(실험 3)에서 검증한 임계값을 따른다.

$$\text{final_score} = \text{semantic_score} \times 0.6 + \text{ontology_score} \times 0.4 \dots\dots (1)$$

3.1.3 Knowledge Graph

NetworkX DiGraph와 SQLite로 구현된 지식 그래프이다. PKMS의 노트를 파싱하여 노트(note, hub, tag, category)

와 엣지(RELATED_TO, HUB_CONNECTION, HAS_TAG, IN_CATEGORY)를 생성한다.

3.1.4 Ontology Rules

JSON 형식의 개인 온톨로지 규칙 파일로, 허브 노트, 정체성 렌즈, 교차 도메인, 사고 패턴, 가치 체계를 정의한다.

파이프라인의 작동 위상으로 보면 Ontology Reranker는 Gemini File Search RAG가 1차로 검색한 상위 N건(본 실험 N=10)을 입력으로 받아 재정렬하는 후처리 모듈이다. 따라서 1차 검색이 누락한 노트는 본 시스템에서 구제되지 않으며, POR의 효과는 시맨틱 검색의 회수율(recall)에 종속된다. 이 종속 관계의 영향은 5장에서 후술한다.

3.2 개인 온톨로지 설계

본 연구의 개인 온톨로지는 기록학의 두 원칙을 개인기록의 검색 맥락에 매핑하여 7개 스코어링 요인을 도출하였다. 첫째, 출처주의(생산자 측 맥락 보존)에서는 생산자가 자신을 어떤 다중 정체성으로 규정하는지(정체성 렌즈), 어떤 인지 패턴으로 사고하는지(사고 패턴), 어떤 가치 체계를 가지는지(가치 정렬), 그리고 어떤 문서가 생산자의 직접 산출물인지(저자성)의 4개 요인이 도출된다. 이는 RiC-CM(ICA EGAD, 2023)의 Agent 개체가 가지는 정체성·역할·가치 속성, 그리고 hasCreator/hasContributor 관계에 매핑된다. 둘째, 원질서 존중과 검색도구 기능(기록 배열·구조 보존)에서는 어떤 노트가 생산자의 지식 체계에서 주요 시리즈로 기능하는지(허브 부스트), 노트들이 그래프 상에서 얼마나 가까운지(그래프 근접도), 어떤 노트가 도메인 경계를 가로지르는지(교차 도메인)의 3개 요인이 도출된다. 이는 ISAD(G)(ICA, 2000)의 계층 구조 인식과 RiC-CM의 관계 기술에 매핑된다.

7개라는 수는 두 원칙에서 도출된 후보군의 결과이며 최적값으로 주장하지 않는다. 본 연구는 단일 사용자 사례의 1차 실증이므로, 후속 다중 사용자 연구에서 요인의 가감 및 가중치 재조정이 가능하다. 또한 실증적 근거가 직접 학술 문헌에 있는 요인(정체성 렌즈·교차 도메인·사고 패턴·그래프 근접도·저자성)과, PKM 실천 커뮤니티 담론에 근거를 둔 요인(허브 부스트), 본 연구가 처음 제안한 요인(가치 정렬)이 혼재한다. 이 분포는 본 연구가 기록학적 원리와 PKM 실천을 잇는 1차 매핑이라는 위상에서 비롯되며, 직관 기반 요인이라 하여 임의로 설정한 값은 아니며 RiC-CM의 Agent 속성과 정렬되도록 설계하였다.

이하 각 요인의 정의, 작동 방식, 이론적 근거를 차례로 기술한다. 가중치는 4.4 가중치 시뮬레이션(실험 3)에서 검증한 임계값을 따른다.

▪ 요인 1. 허브 부스트(Hub Boost)

4-tier 계층 구조의 허브 노트에 해당하면 가산점을 부여한다. root(+3.0), tier1_pillars(+2.5), tier2_applied(+2.0), tier3_identity(+1.5, 단 허브별 차등 적용)의 가중치를 적용하며, 총 10개 허브 노트를 정의하였다. 본 요인은 Milo (2020)의 LYT(Linking Your Thinking)가 제시한 MOC(Maps of Content)와 Matuschak(n.d.)의 에버그린 노트가 강조하는 영구 노트의 연결성 개념을 가져와, 생산자가 자신의 지식 체계에서 주요 시리즈로 명시한 노트에 우선 순위를 부여한다. ISAD(G)의 Fonds → Series 인식과 RiC-CM의 Record/RecordSet 관계에 대응한다.

▪ 요인 2. 정체성 렌즈 매칭(Identity Lens Match)

생산자의 3가지 정체성 렌즈에 각각 10~13개의 키워드가 정의되어 있으며, 문서에서 키워드가 매칭되면 렌즈 가중치를 곱하여 가산한다. 상한은 2.0이다. 생산자의 정체성 렌즈는 각각 아키비스트, 솔로프러너(Solo-Entrepreneur),

기업가이다. 본 요인은 ISAD(G)의 Creator 개체와 RiC-CM의 Agent 개체가 단일 정체성으로 환원되지 않고 다중 역할을 가질 수 있음을 개인 수준에서 구체화한 것이다. 동일한 기록이라도 생산자의 어느 정체성 렌즈와 정합하는지에 따라 검색 가중치가 차등 부여된다.

▪ 요인 3. 교차 도메인 교집합(Cross-Domain Intersection)

AI, archive, human 3개 도메인의 키워드가 2개 이상 도메인에서 매칭되면 +1.5, 3개 이상이면 +3.0을 부여한다. 본 요인은 Collins와 Loftus(1975)의 확산 활성화 이론에서 개념 노드 간 활성화가 도메인 경계를 가로질러 이루어진다는 인지 모형을 검색 점수에 반영한 것이다. 생산자가 학제 교차점에서 형성한 문서가 상위에 노출되는 효과가 있다.

▪ 요인 4. 사고 패턴 부스트(Thinking Pattern Boost)

삼원대비 패턴(A×B×C) 감지 시 +0.5, 정의 지향 트리거 용어(‘~이란’, ‘정의하면’) 감지 시 +0.3, 연결 이유 명시(‘왜냐하면’, ‘이유는’) 감지 시 +0.2, 순서파괴 사고(‘역방향’, ‘미래에서’) 감지 시 +0.1을 부여한다. 본 요인은 Ahrens(2017/2023)가 강조한 원자적 노트의 정의, 연결, 이유 명시를 인지적 활동의 신호로 보고, RiC-CM의 Activity 개체(생산자의 인지적 활동) 속성에 매핑한 것이다. 생산자의 사고 흔적의 밀도가 높은 노트에 가산점을 부여한다.

▪ 요인 5. 그래프 근접도(Graph Proximity)

0-hop(허브 자신): +1.5, 1-hop: +1.0, 2-hop: +0.5, 3-hop: +0.3, 4-hop: +0.15. 본 요인은 Page et al.(1999)의 PageRank가 그래프 위상으로 노드 중요도를 수치화한 접근을 따라, 허브 노트로부터의 hop 거리를 맥락적 근접성의 대리 지표로 삼는다. 위키링크와 자동 허브 링킹(3.3)으로 형성된 그래프 위에서 작동한다.

▪ 요인 6. 가치 정렬(Value Alignment)

누적(weight 1.5), 자연스러움(1.3), 의미 있는 비합리성(1.2), 공유(1.0), 고유 철학(1.0)의 5개 가치에 대해 긍정 키워드 매칭 시 가치별 가중치를 적용하고, 단기 회피 키워드 매칭 시 -0.3을 적용한다. [-1.0, +1.0] 범위로 제한한다. 본 요인은 RiC-CM의 Agent 개체가 가지는 가치 체계 속성을 검색 알고리즘에 통합한 것으로, 본 연구가 7개 요인 중 처음으로 제안한 요인이다. 직접 학술 근거는 부족하나, 가치가 기록의 의미 부여에 영향을 미친다는 기록학의 일반 원칙에서 도출하였다.

▪ 요인 7. 저자성(Authorship)

v2.1.0에서 추가된 요인으로(버전 운영은 4.1 참고), 문서의 파일 경로를 기반으로 생산자가 직접 작성한 문서와 외부에서 수집한 자료를 구분한다. 저자성은 3단계로 분류된다. 직접 생산(created) 경로(20-created/, 40-periodic/41-daily/, 10-work/14-writing/ 등)는 +1.5, AI 협업(ai_collaboration) 경로(24-agentic-engineering/, 세션 로그 등)는 +0.5, 외부 수집(collected) 경로(30-collected/32-web-clip/, 30-collected/35-archives-academic/ 등)는 +0.0을 부여한다. ISAD(G)에서 생산자와 관련 기관을 구분하는 원칙, RiC-CM에서 hasCreator와 hasContributor 관계를 명시하는 것과 대응된다.

온톨로지 총점은 7.0 상한이 적용된다.

3.3 지식 그래프 구축

지식 그래프는 `graph_builder.py`가 PKMS의 마크다운 파일을 파싱해 구축한다. 빌드 프로세스는 다음과 같다.

1. 노트 파싱

각 마크다운 파일의 YAML frontmatter에서 태그, 상태, 생성일 등 메타데이터를 추출하고, 본문에서 위키링크 ([[]])를 정규식으로 추출한다. frontmatter가 없는 파일은 본문 전체를 파싱 대상으로 처리한다.

2. 노트 생성

파싱된 정보를 바탕으로 노트(note), 태그(tag), 카테고리(category) 유형의 노트를 생성한다. 허브 노트는 `is_hub=True` 플래그와 `hub_role(root/tier1/tier2/tier3)`을 부여하여 일반 노트와 구분한다.

3. 엣지 생성

노드 간 관계를 4가지 유형의 엣지로 생성한다. 위키링크 기반 RELATED_TO는 노트 간 명시적 참조를, 태그 기반 HAS_TAG는 노트와 태그 간 연결을, 폴더 기반 IN_CATEGORY는 노트와 분류 체계 간 소속을 나타낸다.

4. 자동 허브 링킹(v2.0)

허브 노트의 테마 키워드와 일반 노트의 본문을 대조하여, 일정 수 이상의 키워드가 매칭되면 해당 노트와 허브 사이에 HUB_CONNECTION 엣지를 자동 생성하는 메커니즘이다. 티어 기반 임계값(`root/tier1`: 3개 이상, `tier2/tier3`: 2개 이상)을 적용하여 과잉 연결을 제어한다. 이로써 명시적 위키링크가 없는 노트도 허브 네트워크에 편입될 수 있다.

4. 실험 설계 및 결과

4.1 실험 개요

실험 데이터는 연구자 본인의 마크다운 기반 PKMS 내 3,702건의 콘텐츠 노트(총 4,746개 파일 중 시스템 파일 제외, 2025.09~2026.03 약 7개월 축적)이다. 실험 환경은 macOS Tahoe 26.3.1, Python 3.12, Claude Code(Anthropic사의 Claude 모델 기반 CLI), Gemini 2.5 Flash의 File Search API이다.

본 연구는 POR을 세 단계로 운영하였다. POR v1.0은 6개 스코어링 요인(허브 부스트, 정체성 렌즈, 교차 도메인 교집합, 사교 패턴, 그래프 근접도, 가치 정렬)으로 구성되어 실험 1~3, 5의 베이스 시스템으로 사용되었다. POR v2.0은 v1.0의 그래프 구조 한계(허브 고립)를 보완하기 위해 자동 허브 링킹(3.3 참고)을 도입하였고, POR v2.1.0은 사용자 적합성 평가(실험 6) 결과로 도출된 저자성 요인을 7번째 요인으로 추가한 버전이다. 본문 결과는 v2.1.0 기준이며, 본문에서 'POR'이 별도 명시 없이 등장하면 v2.1.0을 가리킨다. v1.0과 v2.x를 비교하는 맥락에서는 버전을 명시한다.

지식 그래프 스냅샷도 시점에 따라 구분하여 사용하였다. v7 및 v8은 PKMS 그래프 빌드 시점이며, POR 알고리즘 버전과는 별개의 식별자이다. 본 연구에서 인용하는 그래프 스냅샷은 세 가지이다.

- v7 (2026.02 빌드): 콘텐츠 노트 3,155건, 4,976 노드, 12,357 엣지, 허브 10개. <표 3> 성장 현황 비교의 시점 1에 해당한다.
- v8 (2026.03 빌드): 콘텐츠 노트 3,702건, 5,717 노드, 13,544 엣지, 허브 10개. 본문 주요 결과(실험 1·2·5·6, 결론 5.5)의 분석 대상 그래프. v8의 엣지 유형별 분포는 HAS_TAG 10,782건, HUB_CONNECTION 1,810건 (자동 허브 링킹 산출물), IN_CATEGORY 890건, RELATED_TO 62건이며, 그래프 빌드 시간은 2.32초이다.
- 실험 K용 갱신본 (2026.04.22 빌드): 7,542 노드, 16,395 엣지, 허브 11개. 콘텐츠 노트 수는 별도 집계되지 않았다. v9 재심 작업 도중 자동 허브 링킹의 위상을 추가 검증하기 위해 별도로 빌드한 스냅샷으로, 메인 결과의 그래프와 구분된다. 4.2절 실험 K(자동 허브 링킹 위상 검증) 전용이다.

v8 그래프가 본 연구의 메인 그래프이며, 결론 5.5에서 인용하는 ‘5,717 노드와 13,544 엣지’가 이에 해당한다.

<표 2>에 본 논문에서 상세 기술하는 6종의 주요 실험을 요약하였다. 실험 1~3은 시스템의 구조적 기반을 검증하고, 실험 4는 자동 온톨로지 생성의 가능성을 탐색하며, 실험 5~6은 검색 품질의 정량적·정성적 효과를 평가한다.

<표 2> 실험 개요

실험	목적	방법	핵심 지표
1. 그래프 구조 분석	지식 그래프 건강도	NetworkX 통계, PageRank	노드/엣지 수, 허브 타당성
2. 허브 연결성	그래프 근접도 활성화	자동 링킹 전후 A/B	boost-eligible rate
3. 가중치 시뮬레이션	온톨로지 가중치 효과	6개 가중치 조합	순위 변동 건수
4. 클러스터 분석	자동 온톨로지 가능성	TF-IDF, SBERT, HDBSCAN	Silhouette score
5. 실시간 A/B 비교	검색 품질 비교	Gemini API 실제 호출	Kendall's tau
6. 사용자 적합성 평가	주관적 품질	4점 리커트 블라인드	평균 적합성

실험에 사용된 분석 도구 및 통계 방법은 다음과 같다. 그래프 분석에는 Python 기반 그래프 분석 라이브러리인 NetworkX(Hagberg et al., 2008)를 사용하였다. 지식 그래프의 구조적 타당성 검증(실험 1)에서는 웹 페이지 간 링크 구조를 기반으로 노드의 상대적 중요도를 산출하는 PageRank 알고리즘(Page et al., 1999)과, 모듈러리티 최적화를 통해 그래프를 커뮤니티로 분할하는 Louvain 알고리즘(Blondel et al., 2008)을 사용하였다. 모듈러리티(modularity)는 네트워크가 밀접하게 연결된 하위 집단(커뮤니티)으로 얼마나 명확하게 분리되는지를 나타내는 지표로, 0에서 1 사이의 값을 가지며, 일반적으로 0.3 이상이면 유의미한 커뮤니티 구조가 존재한다고 해석한다. 자동 온톨로지 생성 가능성 탐색(실험 4)에서는 문장 수준의 의미적 유사도를 포착하는 Sentence-BERT(Reimers & Gurevych, 2019) 임베딩과, 밀도 기반으로 임의 형태의 클러스터를 탐지하는 HDBSCAN(McInnes et al., 2017) 알고리즘을 사용하였다. 실시간 A/B 비교(실험 5)에서는 두 순위 목록 간의 일치도를 -1에서 +1 사이의 값으로 나타내는 비모수적 통계 지표인 Kendall's tau 순위상관계수를 사용하였다. tau 값이 낮을수록 온톨로지 리랭킹이 시맨틱 검색 대비 더 큰 순위 변동을 유발한 것으로 해석된다. 사용자 적합성 평가(실험 6)에서는 순위가 낮아질수록 적합성 점수를 할인하여 합산하는 DCG@5(Discounted Cumulative Gain at rank 5)(Järvelin & Kekäläinen, 2002)와, 대응 표본의 비모수적 검정 방법인 Wilcoxon 부호순위 검정을 사용하였다.

4.2 실험 1: 그래프 구조 분석

지식 그래프의 구조적 특성을 파악하기 위해 실험 1에서는 NetworkX 기반 분석을 수행하였다. 본 절의 분석

대상은 4.1에서 정의한 v8 그래프(2026.03 빌드, 5,717 노드 · 13,544 엣지 · 허브 10개)이며, 절 후반의 추가 검증인 실험 K는 별도로 빌드한 2026.04.22 갱신본(7,542 노드 · 16,395 엣지 · 허브 11개)을 대상으로 한다.

<표 3> 그래프 기본 통계

지표	v7 (2026.02)	v8 (2026.03)	변화
콘텐츠 노트	3,155	3,702	+547 (+17.3%)
총 노드	4,976	5,717	+741 (+14.9%)
총 엣지	12,357	13,544	+1,187 (+9.6%)
허브 인식	10/10	10/10	-
HUB_CONNECTION	1,559	1,810	+251
커뮤니티	-	35	-
모듈러리티	-	0.6696	-
빌드 시간	2.07초	2.32초	-

<표 3>은 v7(2026.02, 4,976 노드 · 12,357 엣지)과 v8(2026.03, 5,717 노드 · 13,544 엣지) 간 지식 그래프의 성장 현황이다. 약 1개월간 콘텐츠 노트가 547건(17.3%) 증가하였고, 총 노드와 엣지도 비례하여 증가하였다. 허브 인식률은 10/10을 유지하여, 그래프 규모 확장에도 허브 구조가 안정적임을 확인하였다.

PageRank 중심성 분석 결과, v8 그래프에서 수동 지정 허브 10개 전체가 상위 15위 이내에 위치하여 허브 선정의 타당성을 확인하였다. 1위는 root 허브인 ‘AI-아카이브-인간의-삼원구조’로, 889개의 1-hop 이웃 노드와 연결된다. Louvain 커뮤니티 감지 결과 35개 커뮤니티가 식별되었으며, 모듈러리티 0.6696은 노트들이 주제별로 뚜렷한 군집을 형성하고 있음을 보여준다.

위 PageRank 결과는 자동 허브 랭킹(3.3절)이 적용된 그래프에서 측정된 것이다. 자동 허브 랭킹이 허브 중심성을 인위적으로 형성한 것은 아닌지 검증하기 위해, 자동 허브 랭킹 이전 그래프(auto_linked=True 엣지 제거)에서 동일한 PageRank를 추가로 측정하였다(실험 K, 2026.05.06). 실험 K는 본 절 전반부의 v8 그래프와는 별개로, 자동 링킹 위상 검증을 위해 v9 재심 작업 도중 빌드한 2026.04.22 스냅샷(7,542 노드 · 16,395 엣지 · 허브 11개)을 대상으로 한다. 이 스냅샷 위에서 자동 링킹 적용 전후를 동일 그래프로 두고 비교하였다. 자동 링킹 적용 후에는 11개 허브 중 3개가 상위 15위, 6개가 상위 50위, 8개가 상위 100위에 위치한 반면, 자동 링킹 이전 그래프에서는 어떤 허브도 상위 50위에 진입하지 못하였고 상위 100위 내에 단 1개만 포함되었다. 1위였던 root 허브의 순위는 1위에서 89위로 하락하였다. 이 결과는 허브의 PageRank 중심성이 위키링크만으로는 형성되지 않으며, 키워드 기반 자동 허브 랭킹이 그래프 위상을 재편한 결과임을 의미한다. 따라서 본 절의 PageRank 결과는 키워드 의미 매칭이 반영된 그래프에서의 허브 중심성을 보여주는 것이며, 자동 허브 랭킹의 위상에 대해서는 4.3절과 5.3절에서 다시 기술한다. 실험 K의 코드와 결과 JSON은 본 연구의 연구 자료에 보관되어 있으며(코드: experiment_k_pagerank_no_autolink.py, 결과: k_pagerank_no_autolink_result.json), 후속 연구의 재현 검증에 사용 가능하다.

4.3 실험 2: 허브 연결성과 그래프 근접도

허브 노트의 연결성은 그래프 근접도 스코어링의 전제 조건이다. 초기 그래프에서 허브 대부분이 고립(위키링크 0건)되어 graph_proximity가 작동하지 않는 문제를 발견하고, 허브 노트의 테마 키워드와 일반 노트의 본문을 대조하여 HUB_CONNECTION 엣지를 자동 생성하는 자동 허브 랭킹(3.3 참고)으로 해결하였다.

<표 4> 허브 연결성 및 그래프 근접도

지표	수치
허브 발견율	10/10 (100%)
그래프 근접도 도달률	74/100 (74%)
평균 허브 거리	2.07 hops
거리 분포(1-hop)	34건
거리 분포(3-hop)	38건
거리 분포(5-hop)	1건
HUB_CONNECTION 엣지	1,810건

<표 4>는 자동 허브 링킹 적용 후 허브 연결성과 그래프 근접도 현황을 보여준다. 자동 허브 링킹으로 1,810건의 HUB_CONNECTION 엣지가 생성되었다. 100개 샘플 노트를 무작위 추출하여 가장 가까운 허브까지의 최단 경로를 측정된 결과, 74건(74%)이 4-hop 이내에서 허브에 도달하였다. 1-hop(34건)과 3-hop(38건)에 분포가 집중되어 있으며, 평균 거리 2.07 hops는 그래프 근접도 요인이 실제로 작동할 수 있는 연결성을 확보하였음을 보여준다.

자동 허브 링킹의 위상은 다음과 같다. 본 메커니즘은 허브 노트의 테마 키워드와 일반 노트 본문의 의미 일치율 그래프 엣지로 명시화하는 보완 도구이며, 위키링크가 부족한 PKMS 환경에서 그래프 근접도 요인이 작동하기 위한 인프라이다. 자동 허브 링킹은 그래프 위상에 영향을 미치는 의도된 설계 선택이며, 4.2절의 PageRank 결과는 이 보완 후 그래프에 한정된 해석이라는 점을 명시한다. 자동 링킹 없이 위키링크만으로 형성된 그래프에서는 허브의 PageRank 중심성이 형성되지 않음을 실험 K(4.2 참고)에서 확인하였으며, 본 연구는 허브의 그래프 중심성이 자동 링킹의 산물임을 인정한다. 다만 자동 링킹은 임의 엣지를 추가하는 방식과 달리 키워드 일치도 임계값(root/tier1: 3개 이상, tier2/tier3: 2개 이상)을 적용하는 의미 기반 보완이므로, 그 결과 형성된 허브 중심성도 의미적 근거가 있다.

4.4 실험 3: 가중치 시뮬레이션

온톨로지 가중치 변경이 검색 순위에 미치는 영향을 6개 가중치 조합으로 시뮬레이션하였다.

<표 5> 가중치 조합별 순위 변동

조합	semantic	ontology	평균 위치 변동(건/쿼리)
default(기준)	0.6	0.4	0.0
balanced	0.5	0.5	0.4
ont_heavy	0.4	0.6	1.2
max_ontology	0.3	0.7	1.2

<표 5>는 시맨틱 점수와 온톨로지 점수의 가중치 비율을 변화시켰을 때 검색 순위에 미치는 영향을 시뮬레이션한 결과이다. 기본값(semantic 0.6, ontology 0.4)에서 온톨로지 비중을 0.5, 0.6, 0.7로 점진적으로 높이며, 쿼리당 평균 위치 변동 건수를 측정하였다.

ont_heavy와 max_ontology 조합에서는 평균 1.2건의 위치 변동이 발생하였다. 온톨로지 가중치를 0.4에서 0.6으로 올릴 때 변동이 발생하지만, 0.7로 더 높여도 추가 변동이 없어, 0.6 부근이 변별력의 포화점임을 확인하였다.

4.5 실험 4: 클러스터 분석

PKMS의 자동 온톨로지 생성 가능성을 평가하기 위해 4가지 클러스터링 방법을 비교하였다. 비교 대상은 (1) TF-IDF (Raw)로 원시 텍스트의 단어 빈도-역문서 빈도 특징을 K-Means로 군집화한 방법, (2) TF-IDF (Korean NLP)로 Okt 형태소 분석기로 명사를 추출한 후 TF-IDF로 군집화하여 한국어 형태소 특성을 반영한 방법, (3) SBERT (K-Means)로 Sentence-BERT(Reimers & Gurevych, 2019) 다국어 임베딩 모델(paraphrase-multilingual-*ini*LM-L12-v2)로 문서를 벡터화한 후 K-Means로 군집화한 의미 기반 방법, (4) HDBSCAN (SBERT)로 SBERT 임베딩에 밀도 기반 클러스터링(McInnes et al., 2017)을 적용한 방법이다.

〈표 6〉 클러스터링 방법별 비교

방법	Silhouette	최적 k	특징
TF-IDF (Raw)	0.1049	5	URL 노이즈로 허위 양성
TF-IDF (Korean NLP)	0.0554	8	Okt 형태소 분석, 명사 추출
SBERT (K-Means)	0.0722	3	거시 3개 영역 분리 확인
HDBSCAN (SBERT)	0.2816	113	noise 51%, 자동 k

모든 방법에서 Silhouette score가 0.3 미만으로, PKMS 환경의 다학제적 특성상 전통적 클러스터링이 자동 온톨로지 생성에 적합하지 않다는 점을 확인했다. TF-IDF (Korean NLP)는 0.0554로, 한국어 형태소 분석을 적용했음에도 명확한 의미 군집 형성에 실패하였다. SBERT (K-Means)는 0.0722로 거시 3개 영역(아카이브, AI, 인간) 분리만 식별 가능한 수준이었다. HDBSCAN (SBERT)은 0.2816으로 최고치였으나, 밀도 기반의 자동 군집 수 결정으로 113개의 미세 군집이 생성되고 전체 노트의 51%(약 1,401건)가 노이즈로 분류되어, 실용적인 자동 온톨로지 추출 도구로 활용하기 어렵다는 한계를 확인했다.

4.6 실험 5: 실시간 A/B 비교

Gemini File Search API를 실제 호출하여 Condition A(순수 시맨틱 검색)와 Condition B(온톨로지 리랭킹)를 비교하였다. 5개 유형(단일 도메인, 교차 도메인, 허브 관련, 가치 관련, 일반)에서 각각 2개의 쿼리로 총 10개 쿼리를 사용하였다.

〈표 7〉 실험 쿼리 목록

ID	유형	쿼리
Q1	단일 도메인	기록학에서 출처주의 원칙
Q2	단일 도메인	RAG 파이프라인 설계와 구현
Q3	교차 도메인	아카이브와 AI의 교차점
Q4	교차 도메인	개인지식관리와 기록학의 접점
Q5	허브 관련	AI 아카이브 인간의 삼원구조
Q6	허브 관련	형성하는 존재로 살아남는다
Q7	가치 관련	데이터 주권과 로컬 우선주의
Q8	가치 관련	누적의 철학과 지속성
Q9	일반	창업 과정 기록
Q10	일반	A 프로젝트 회고

<표 7>은 실험에 사용된 10개 쿼리 목록이다. 5개 유형(단일 도메인, 교차 도메인, 허브 관련, 가치 관련, 일반)에 각 2개씩 배정하였다. 단일 도메인 쿼리는 기록학 또는 AI 등 한 분야에 한정된 주제를, 교차 도메인 쿼리는 두 분야 이상이 교차하는 주제를, 허브 관련 쿼리는 온톨로지에 정의된 허브 노트와 직접 연결된 주제를, 가치 관련 쿼리는 생산자의 가치 체계에 부합하는 주제를, 일반 쿼리는 특정 요인과 무관한 일상적 주제를 다룬다. 쿼리 선정의 절차는 5개 유형 정의가 쿼리 선정에 선행하였으며, 각 유형의 일반적 사용 시나리오를 대표하는 쿼리를 2개씩 선정하였다. 시스템에 유리한 쿼리만 의도적으로 선정하지 않도록, 유형별 비율을 사전에 균등하게 고정하여 표본 선택 편향을 통제하였다.

두 조건의 순위 일치도를 Kendall's tau 순위상관계수로 측정하였다. Kendall's tau는 두 순위 목록 간의 일치 정도를 -1에서 +1 사이의 값으로 나타내는 비모수적 통계 지표로, +1이면 두 순위가 완전히 동일하고, 0이면 순위 간 상관이 없으며, -1이면 순위가 완전히 역전되었음을 의미한다. 본 연구에서 tau 값이 낮을수록 온톨로지 리랭킹이 시맨틱 검색 대비 더 큰 순위 변경을 유발한 것으로 해석된다.

<표 8> 실시간 A/B 비교 결과

쿼리 유형	Kendall's tau	평균 순위 변경	Top-3 교체
단일 도메인	0.750	5.5	1.0
교차 도메인	0.752	4.5	0.5
허브 관련	0.300	3.5	0.5
가치 관련	0.517	4.5	0.5
일반	0.786	4.0	0.5
전체 평균	0.621 (SD=0.252)	4.4 (SD=1.56)	0.6

<표 8>은 5개 쿼리 유형별 Kendall's tau, 평균 순위 변경 건수, Top-3 교체 건수를 보여준다. Kendall's tau가 낮을수록 온톨로지 리랭킹에 의한 순위 변경이 크며, 허브 관련 쿼리(0.300)와 가치 관련 쿼리(0.517)에서 재배열이 두드러진다. Top-3 교체는 상위 3건 중 온톨로지 리랭킹으로 새로 진입한 문서의 수를 나타낸다.

전체 평균 Kendall's tau는 0.621(SD=0.252)로, 온톨로지 리랭킹이 쿼리 유형에 따라 차등적 순위 변경을 유발하였다. 한 쿼리당 상위 10건 결과 중 평균 4.4건의 위치가 변경되었다. 허브 관련 쿼리(tau=0.300)에서 가장 큰 재배열이 발생하여, 생산자의 주요 관심사에 해당하는 쿼리에서 온톨로지가 가장 적극적으로 개입함을 확인하였다. 반면, 일반 쿼리(tau=0.786)에서는 기존 순위가 대체로 보존되었다.

<표 9> 온톨로지 스코어링 요인 활성화 (10개 쿼리 합계)

요인	활성화 횟수
정체성 렌즈 매칭	57
교차 도메인 교집합	29
가치 정렬	16
사고 패턴 부스트	6
허브 부스트	3
저자성	0
그래프 근접도	0

정체성 렌즈 매칭(57건)과 교차 도메인 교집합(29건)이 가장 활발한 요인이었다(<표 9> 참조). v7에서 0건이었던 허브 부스트가 v8에서 3건 활성화되어, 리랭커의 4-tier 허브 인식 정합성 확보가 그래프 연결성과 결합되어

활성화로 이어졌음을 확인하였다. 가치 정렬은 16건 활성화되었다. 반면 7개 요인 중 2개(저자성, 그래프 근접도)는 RAG 인터페이스의 제약으로 실험에서 활성화되지 못하였다. 저자성은 구현은 끝났으나 Gemini RAG citation에 파일 경로(URI) 정보가 포함되지 않아 경로 기반 판별이 불가능하였다. 그래프 근접도 역시 citation title과 그래프 node title 간 정규화 불일치로 0건이었다.

이 0건 결과는 두 요인의 비활성화는 그래프 구조 자체가 무효하다는 뜻이 아니며, RAG 인터페이스가 노출하는 메타데이터의 제약에서 비롯된다. 즉 본 실험에서 그래프 기반 요인(그래프 근접도) 및 경로 기반 요인(저자성)은 측정되지 못한 것이지 작동하지 않은 것이 아니다. 이 두 요인의 효과 검증은 (1) Gemini File Search citation에 파일 경로·노드 식별자가 포함되도록 메타데이터를 확장하거나, (2) citation title을 그래프 node title로 매핑하는 별도의 정규화·매핑 모듈을 두는 후속 개발에서 가능하다(5.4 향후 연구 참고). 결과적으로 본 실험에서 정량 효과가 측정된 요인은 7개 중 5개(허브 부스트, 정체성 렌즈, 교차 도메인 교집합, 사고 패턴, 가치 정렬)이며, 이 5개 요인은 모두 텍스트 키워드 매칭만으로 산출 가능한 것들이다. 그래프 위상의 직접 활용 효과는 본 연구에서 검증되지 않았으며, 후속 검증 과제로 명시한다.

4.7 실험 6: 사용자 적합성 평가

실험 5에서 제시된 동일 10개 쿼리에 대해 사용자 적합성 평가를 수행하였다. 평가자는 연구자 본인이며, 본 연구가 단일 사용자 사례 연구로서 PKMS 적합성의 본질이 생산자에게 귀속된다는 위상에서 1차 실증으로 수행한 평가에 해당한다. 시맨틱 검색(조건 A)과 온톨로지 리랭킹(조건 B)이 반환한 상위 5건 결과를 블라인드 조건으로 제시하고, 각 결과의 적합성을 4점 리커트 척도(0=무관, 1=약간 관련, 2=관련, 3=매우 관련)로 평정하였다. 단일 평가자 구조의 객관성 약점과 후속 다중 평가자 설계는 5.3과 5.4에서 다룬다.

평가 지표로는 평균 적합성과 DCG@5(Discounted Cumulative Gain at rank 5, Järvelin & Kekäläinen, 2002)를 사용하였다. DCG@5는 상위 순위에 위치한 적합 문서에 더 큰 가중치를 부여하는 누적 이득 지표로, 식 (2)와 같이 산출한다.

$$DCG@k = \sum_{i=1}^k (rel_i / \log_2(i+1)) \dots\dots (2)$$

<표 10> 사용자 적합성 평가 결과

조건	평균 적합성	표준편차	DCG@5
A (시맨틱)	2.98	0.948	9.049
B (온톨로지)	3.06	0.947	9.371

<표 10>에서는 두 조건의 전체 평균 적합성과 DCG@5를 비교한다. 온톨로지 조건(B)이 평균 적합성 3.06, DCG@5 9.371로, 시맨틱 조건(A)의 2.98, 9.049보다 소폭 높았다. 차이는 작지만 두 지표 모두 온톨로지 조건이 우세하며, 특히 DCG@5에서 0.322의 차이는 적합한 문서가 상위 순위에 더 효과적으로 배치되었음을 가리킨다.

Wilcoxon 부호순위 검정 결과는 $p=0.544(W=57.0)$ 였다. $N=10$ 에서는 검정력이 부족하다. <표 11>은 5개 쿼리 유형별 선호 분석 결과이다. 허브 관련 쿼리(Q5, Q6)에서는 두 쿼리 모두 온톨로지 조건이 선호되었다. 특히 Q6 ‘형성하는 존재로 살아남는다’는 실험 5에서 가장 큰 순위 변동($\tau=0.300$)을 보인 허브 관련 쿼리 유형에 속하며, 온톨로지 리랭킹이 시맨틱 검색과 크게 다른 결과를 제시했음에도 사용자가 온톨로지 쪽을 선호한 것은 의미 있는 발견이다. 일반 쿼리(Q9 ‘창업 과정 기록’)에서는 시맨틱 검색이 선호되었다. 평가자는 ‘직접 작성한

문서가 우선적으로 검색되어야 한다'는 피드백을 남겼으며, 이 피드백은 7번째 요인인 저자성을 도출하는 직접적 근거가 되었다. 본 연구는 이를 POR v2.1.0의 설계에 반영하였다.

<표 11> 쿼리 유형별 선호 분석(블라인드 해제 후)

쿼리 유형	쿼리 수	B 선호	A 선호	동일	B의 평균 이점
단일 도메인	2	0	0	2	+0.10
교차 도메인	2	1	1	0	+0.10
허브 관련	2	2	0	0	+0.10
가치 관련	2	1	0	1	+0.20
일반	2	0	1	1	-0.10
전체	10	4	2	4	+0.08

4.8 실험 결과의 통합 해석

6종의 실험은 개별 발견을 제공하면서도 서로 연결되어 개인 온톨로지 기반 검색의 작동 메커니즘에 대한 통합적 이해를 제공한다.

첫째, 그래프 구조 검증(실험 1)과 허브 연결성 확보(실험 2)는 검색 단계의 전제조건이었다. v7에서 0건이던 허브 부스트가 v8에서 3건 활성화된 것은, 자동 허브링킹으로 확보한 그래프 연결성이 리랭커의 4-tier 허브 인식 정합성과 결합하여 검색 시점에 작동할 수 있는 조건을 형성했기 때문이다.

둘째, 실험 5의 정량 결과와 실험 6의 사용자 선호는 일관된 방향을 가리킨다. 큰 순위 변경이 발생한 허브 관련 쿼리에서 사용자 평가도 온톨로지 조건을 선호했다는 사실은 생산자의 주요 관심사에 해당하는 쿼리에서 시맨틱 검색만으로 포착하기 어려운 맥락이 존재한다는 가설을 뒷받침한다. 반면 순위 변경이 작았던 일반 쿼리에서 시맨틱 검색이 선호된 사실은 맥락 독립적 쿼리에서는 시맨틱 유사도만으로 충분하며 저자성과 같은 추가 맥락 신호가 필요함을 시사한다.

셋째, 실험 4의 클러스터링 실패는 이 일관된 방향을 강화한다. 모든 자동 클러스터링 방법의 Silhouette score가 0.3 미만이었다는 사실은 PKMS의 다주제적 특성상 표준화된 분류 체계로는 생산자 맥락을 포착할 수 없음을 의미한다. 노트의 51%가 노이즈로 분류된 결과는 생산자의 고유한 관심사와 연결 방식으로만 의미를 가지는 노트가 다수임을 보여주며, 원질서 존중이 외부 표준 분류 대신 생산자의 원래 배열을 보존해야 한다는 기록학적 원칙과 일맥상통한다.

요인별 기여도와 쿼리 유형별 차등 효과는 <표 8>과 <표 9>에 정리되어 있다. 정량 효과가 측정된 5개 요인 중 정체성 렌즈 매칭과 교차 도메인 교집합이 리랭킹효과를 가장 크게 설명하는 요인이었다. 두 요인은 모두 출처주의의 생산자 측 맥락(생산자가 자기를 어떻게 규정하는가, 어떤 학제 교차에서 사고하는가)에 대응한다. 그래프 기반 요인(그래프 근접도)과 경로 기반 요인(저자성)의 효과는 RAG 인터페이스의 메타데이터 제약(4.6 참고)으로 본 실험에서 측정되지 못하였고, 효과 검증은 후속 연구 과제로 남는다.

본 연구의 6종 실험과 추가 분석(실험 K)은 개인 온톨로지 기반 검색이 다음 세 조건에서 작동함을 시사한다. 첫째, 지식 그래프의 연결성이 자동 허브 링킹 등 키워드 기반 의미 매칭의 보완을 포함하여 확보되어야 한다(실험 1·2·K). 위키링크만으로 형성된 그래프에서는 허브의 그래프 중심성이 형성되지 않으며, 이는 PKMS 환경에서 자동 링킹이 보완 수준을 지나 그래프 위상을 결정짓는 설계 선택임을 의미한다. 둘째, 온톨로지 설계가 생산자의 인지 구조를 반영해야 하며, 자동 생성보다는 생산자의 명시적 모델링이 필요하다(실험 4·5·6). 셋째, 맥락의 차원(지식 구조, 가치 지향, 저자성)을 쿼리 유형에 따라 차등 활용해야 한다(실험 6). 검색 시스템은 단일 관련성

모델 대신 다차원 맥락 모델을 갖추어야 한다. 다만 본 연구에서 정량 효과가 검증된 차원은 텍스트 매칭으로 산출 가능한 5개 요인에 한정되며, 그래프 위상과 저자성의 직접 효과는 후속 연구의 검증과제로 남는다.

5. 결론

5.1 연구 질문에 대한 답변

6종의 실험에서 세 가지 연구 질문에 대한 답변을 도출하였다. 정리하면, 온톨로지 리랭킹은 쿼리 유형에 따라 차등적으로 작동하며, 생산자의 주요 관심사와 관련된 쿼리에서 가장 큰 효과를 보였다.

RQ1: 온톨로지 리랭킹은 시맨틱 검색의 순위를 쿼리 유형에 따라 차등적으로 변경한다. 평균 Kendall's tau 0.621(SD=0.252)은 전반적으로 중간 수준의 재배열을 보이지만, 허브 관련 쿼리($\tau=0.300$)에서는 매우 큰 순위 변경이 발생하고, 일반 쿼리($\tau=0.786$)에서는 기존 순위가 대체로 보존되었다. 이는 온톨로지 리랭킹이 균일하게 작용하지 않고, 생산자의 인지적 맥락이 관련된 쿼리에서 선택적으로 활성화됨을 확인한다.

RQ2: 온톨로지 리랭킹 결과가 사용자 적합성에서 통계적으로 유의미하게 더 높다고는 할 수 없다. 전체 평균 적합성은 Condition B(온톨로지) 3.06 vs Condition A(시맨틱) 2.98로 B가 소폭 높았으나, Wilcoxon signed-rank 검정 결과 $p=0.544$ 로 유의수준 0.05를 충족하지 못하였다. 이는 10개 쿼리 중 4건에서 두 조건의 결과가 동일하였고, 작은 표본 크기($N=10$)로 검정력이 부족하였기 때문이다. 그러나 쿼리 유형별 분석에서는 뚜렷한 패턴이 나타난다. 허브 관련 쿼리 2건에서 모두 온톨로지 조건이 선호되었고, 가치 관련 쿼리에서도 온톨로지가 우세하였다. 이는 개인 온톨로지의 주요 허브와 가치 체계가 리랭킹에 효과적으로 반영된 결과이다. 일반 쿼리에서 시맨틱 검색이 선호된 결과는 저자성 요인의 필요성을 시사하였으며, 본 연구는 이 피드백을 POR v2.1.0의 설계에 반영하였다.

RQ3: 지식 그래프의 구조적 특성은 리랭킹 효과에 직접적으로 영향을 미친다. 10개 허브 전체가 PageRank 상위 15위 이내에 위치하여(실험 1) 수동 허브 선정의 그래프 이론적 타당성을 확인하였다. 1,810건의 자동 허브 리랭킹과 74%의 그래프 근접도 도달률(실험 2)은 허브 부스트 활성화의 전제 조건이다. v7에서 0건이었던 허브 부스트가 v8에서 3건 활성화된 것은 그래프 연결성과 리랭커 코드 정합성이 동시에 확보되었을 때 효과가 발현됨을 보여준다. 다만 본 답변은 허브 부스트의 활성화로 한정되며, 그래프 근접도는 4.6에서 기술한 RAG 인터페이스 제약으로 본 실험에서 측정되지 못하였다. 그래프 위상의 직접 효과 검증은 후속 연구의 과제로 남으며, RQ3에 대한 본 답변은 본 실험에서 측정 가능했던 차원에 한정된 부분적 답이다.

5.2 기록학적 함의

5.2.1 출처주의의 구조적 매핑

전통 기록학의 출처주의는 기록을 생산자별로 구분하여 맥락을 보존한다. 본 연구의 개인 온톨로지는 출처주의를 PKMS 검색에 그대로 적용한 것이라기보다, 출처주의가 보존하려는 생산자 맥락의 구조에 검색 가중치 모델을 매핑한 것이다(2.1, 2.4 참고). 이 매핑의 매개는 RiC-CM의 Agent 개체이다. RiC-CM이 출처주의를 Agent의 정체성·역할·가치 속성으로 확장 정의한 만큼, 본 연구도 그 확장 정의 안에서 '생산자의 사고방식, 가치관, 관심 영역'을 정형화된 규칙으로 모델링한다. 따라서 본 연구의 위상은 출처주의의 직접 적용과 완전한 매핑 사이에 있으며, RiC-CM 매핑이 닿는 한도 내에서는 부분적 본질 적용, 그 너머에서는 구조적 매핑으로 이해할 수 있다.

정체성 렌즈는 생산자가 기록을 바라보는 다중 관점을 명시하며, 이에 따라 동일한 기록이라도 다른 맥락적 가치를 받는다.

다만 RiC-CM의 Agent는 기관 기록의 생산자를 전제하며, 자기-규정의 메타인지 함정은 보존소, 역사가, 동료 평가자라는 외부 권위에 의해 보정된다. 본 연구는 생산자=이용자=평가자가 동일인이므로 이 외부 보정 메커니즘이 작동하지 않으며, 따라서 본 연구의 RiC-CM Agent 매핑에는 그 외부 보정이 부재한 환경에서의 부분 적용이라는 추가 한정이 따른다.

실험 5의 실시간 A/B 비교는 이에 대한 실증적 근거에 해당한다. 허브 관련 쿼리에서 가장 큰 순위 재구성이 발생한 반면 일반 쿼리에서는 기존 순위가 보존되었다는 것은, 온톨로지 리랭킹이 모든 쿼리에 균일하게 작용하지 않고 생산자의 인지적 맥락이 관련된 쿼리에서 선택적으로 활성화됨을 보여준다. 이는 기록학의 Fonds에서 주요 기능에 대한 쿼리가 출처 정보로 효과적으로 해결되는 것과 유사하다. 정체성 렌즈 매칭이 57건으로 가장 활발한 요인이었다는 점은 생산자의 다중 정체성이 본 실험의 리랭킹 효과를 가장 크게 설명하는 요인임을 뒷받침한다.

5.2.2 검색도구 확장으로서의 온톨로지

기록학에서 검색도구는 사용자가 기록을 찾으려 하는 메타데이터 체계이다. 본 연구의 개인 온톨로지는 검색도구의 기능을 포함하면서 더 넓은 범위에 걸친다. 검색도구가 기록의 조직 구조와 메타데이터를 기술하는 데 초점을 맞추는 반면, 개인 온톨로지는 생산자의 정체성(정체성 렌즈), 사고 패턴(삼원대비, 정의 지향), 가치 체계(누적, 자연스러움, 의미 있는 비합리성)까지 모델링한다. 즉, 온톨로지는 ‘무엇을 어디서 찾을 수 있는가’에 더해 ‘왜 그것이 이 생산자에게 관련되는가’까지 답하는 의미적 프레임워크이다. 허브 노트가 주요 시리즈를 가리키고 그 래프 근접도가 맥락적 관련성을 수치화하는 것은 검색도구 기능에 해당하지만, 정체성 렌즈와 가치 정렬에 기반한 리랭킹은 검색도구의 범위를 초월한다.

5.2.3 맥락 차원에 따른 차등 적용

쿼리 유형별 차등 효과에서 맥락이 단일 차원에 그치지 않고 다차원 구조임을 실질적으로 확인했다. 허브 관련 쿼리(Q5, Q6)에서 온톨로지가 일관되게 선호된 것은 지식 구조의 맥락(생산자의 주요 관심사, 네트워크 중심성)이 활성화되었기 때문이다. 가치 관련 쿼리(Q7, Q8)에서는 ‘가치 지향적 맥락’(생산자의 철학과 원칙)이 작동하여 해당 가치를 지지하는 문서들이 상위로 배치되었다. 이는 RiC-CM의 Agent 개체가 가지는 ‘가치 체계’ 속성을 검색 알고리즘에 통합한 것으로 해석할 수 있다. 일반 쿼리에서 시맨틱 검색이 선호된 것은 생산자 맥락(누가 작성했는가), 곧 저자성 요인의 필요성을 실증하였다. 저자성 요인은 ISAD(G)에서 Creator와 관련 기관을 구분하는 원칙에 대응하며, 본 연구는 이 요인을 POR v2.1.0의 설계에 반영하였다. 다만 RAG 인터페이스 제약으로 실험적 검증은 후속 과제로 남아 있다.

5.2.4 개인 기록의 정량적 실험

기록학 분야의 AI 관련 연구는 대부분 공공기록 대상이거나 사례 연구 수준이다. 본 연구는 3,702건의 개인기록에 대해 정량적 실험을 수행한 시스템 평가로, 기록학 분야에서 드문 실험적 접근이다.

5.2.5 POR과 ‘예상 밖의 연결’의 양립 가능성

POR 리랭킹이 생산자의 기존 인지 구조에 부합하는 결과를 상위로 끌어올림으로써 확증 편향을 강화하고 새로운 통찰을 억제할 수 있다는 우려가 제기될 수 있다. 이는 Ahrens(2017/2023)가 PKM의 가치로 강조한 ‘예상 밖의 연결’과 본 연구의 방향이 양립 가능한지에 대한 질문이기도 하다. 본 연구의 응답은 다음 세 가지이다.

첫째, POR은 시맨틱 점수 0.6과 온톨로지 점수 0.4의 가중 합산으로 작동하며, 시맨틱 검색 결과를 제거하지 않는다. 임베딩이 잡아낸 의미 거리는 그대로 보존되고 그 위에 생산자 맥락 가중치가 더해지는 구조이므로, 시맨틱 검색이 제공하는 기존 인지 구조 외부의 후보들을 가산점만큼 낮출 위험은 있으나 차단하지는 않는다.

둘째, 7개 요인 중 교차 도메인 교집합은 단일 도메인 노트보다 도메인 경계를 가로지르는 노트에 가산점을 부여한다. 아렌스가 강조한 ‘예상 밖의 연결’을 적극적으로 점수화하는 요인이 본 온톨로지 안에 포함되어 있다는 의미이다. 실험 5에서 이 요인이 두 번째로 활발한 요인(29건)이었다는 사실은 POR이 확증 편향한 방향으로만 작동하지 않음을 가리킨다.

셋째, 실험 결과 자체가 POR의 작동이 일률적이지 않음을 보여준다. 일반 쿼리($\tau=0.786$)에서는 시맨틱 순위가 대체로 보존되었고, 사용자 적합성 평가에서도 일반 쿼리에서는 시맨틱 검색이 선호되었다. POR이 모든 쿼리를 자기 강화 방향으로 강제 재배열하는 도구였다면 이러한 결과가 나오기 어렵다. 다만 확증 편향이 0이라고 주장할 수는 없으며, 평가자 동일인 구조에서 비롯된 평가 단계의 확증 편향 가능성은 다음 절과 같이 다르다.

5.3 일반화 가능성과 한계

본 연구의 접근은 원리적으로 다른 마크다운 기반 PKMS으로 이식할 수 있다. 위키링크와 마크다운 파싱이 가장 중요한 입력이며, 이는 표준화된 형식이기 때문이다. 그러나 노선과 같은 데이터베이스 기반 시스템은 마크다운 기반으로 개발되었다 하더라도 API를 통한 접근이 필요하므로 추가 개발이 든다.

온톨로지 구축의 수동 작업은 순수 투입 시간 기준 약 20시간이었으며, Claude Code를 CLI에 설치한 IDE(Integrated Development Environment, 통합개발환경)에서 수행되었다. 온톨로지는 PKMS에 노트가 축적되는 과정과 함께 5개월(2025.11~2026.03)에 걸쳐 점진적으로 진화하였다. 초기 개념 설계, 코드화 및 4-tier 허브 재설계, 검증 및 정제, 저자성 요인 추가의 단계를 거쳤다. 자동 그래프 빌드는 2.32초이다. 수동 정의 부분(정체성 렌즈, 허브 선정, 가치 체계)은 본질적으로 자동화가 불가능하며, 이는 본 접근의 실용적 한계이자 온톨로지의 정의 단계가 AI 자동화에 적합하지 않음을 확인한 결과이기도 하다. 다만 PageRank 기반 허브 자동 선정(실험 1에서 타당성 확인)과 키워드 자동 추출 등 부분 자동화는 가능하다.

먼저 본 논문은 단일 사용자 연구라는 점에서 표본 한계는 분명하다. PKMS의 규모(3,702건), 구조(폴더의 심진 분류체계), 도메인(AI, 아카이브)이 고유하므로, 다른 사용자에게 동일한 효과를 보장하기 어렵다. 본 연구의 실험적 발견이 일반화 가능한 원리로 확립되기 위해서는 다중 사용자 연구가 필수적이다. 또한 본 연구는 연구자 본인의 PKMS를 대상으로 하는 자기 실험 구조의 단일 사용자 사례 연구로, 외부 인간 피험자가 포함되지 않으므로 IRB 심의 대상이 아니며, PKMS 데이터는 학회 게재용으로 외부에 공개되지 않는다. 향후 다중 사용자 연구로 확장될 때에는 외부 피험자 데이터에 대한 IRB 심의와 개인정보 처리 절차가 별도로 필요하다.

본 연구의 첫 번째 방법론적 한계는 자기 인지의 부분성이다. 개인 온톨로지의 정체성 렌즈, 사고 패턴, 가치 체계는 생산자가 자신의 인지 구조를 자각하고 형식화한 결과이며, 메타인지의 한계로 인해 모든 인지 패턴을 누락 없이 포착할 수는 없다. 본 연구는 7개월에 걸친 점진적 정제 과정(초기 → 4-tier 허브 재설계 → 검증·정제 → 저자성 추가)으로 이 한계를 보완하였으나, 보완은 점진적이며 완결될 수 없다. 후속 연구에서는 사용자가 작성하지 못한 사고 패턴을 LLM이 외부에서 추정하여 사용자 검증을 거쳐 온톨로지에 통합하는 반자동 정제 절차가 보완 방안이 될 수 있다.

두 번째 한계는 시간에 따른 온톨로지 정합성이다. 생산자의 사고방식과 가치 체계는 시간이 지나면서 변하며, 이에 따라 온톨로지도 갱신되어야 한다. 갱신 시점에 과거 노트와 새 온톨로지 사이에 불일치가 생길 수 있고(과거 노트가 더 이상 현재 정체성 렌즈에 부합하지 않는 경우 등), 이는 검색 결과의 시간적 일관성에 영향을 미친다. 본 연구의 v2.1.0 운영 기간(약 2개월) 내에서는 큰 정합성 문제가 발견되지 않았으나, 장기 운영에서의 정합성

보장은 본 연구가 다루지 못한 과제이다. 온톨로지 버전 관리·재검증 주기·과거 노트 재태깅 메커니즘을 후속 연구에서 설계해야 한다.

세 번째 단계는 RQ2의 평가 구조에서 비롯된다. 본 연구는 시스템 설계자, 온톨로지 정의자, 쿼리 작성자, 적합성 평가자, 결과 해석자가 모두 동일인이라는 구조 위에서 수행되었다. 정보검색 평가의 일반적인 역할 구조 ([쿼리 설계자] ≠ [시스템 개발자] ≠ [평가자])와 달리, 본 연구는 평가자가 자신이 설계한 시스템에 대해 작동하기를 기대하는 무의식이 평가에 개입할 수 있고, 적합성 4점 리커트의 경계 판단(적합/부분 적합)에 그 기대가 반영될 수 있다. 또한 평가자가 자신이 작성한 모든 노트의 맥락을 알고 있으므로 4.7절의 블라인드 조건(어느 세트가 시맨틱이고 어느 세트가 온톨로지인지 가려둠)이 충분하다고 보기 어렵다. 적합성 판단이 ‘이 문서가 이 쿼리에 적합한가’에서 사실상 ‘내가 이 쿼리를 던졌을 때 원하던 문서인가’로 변질될 가능성도 있다. PKMS의 적합성은 본질적으로 생산자만이 판단할 수 있다는 점에서 단일 사례 1차 실증의 위상에서는 부분적으로 정당화되지만, 이 구조 자체가 객관성의 약점이라는 점을 인정한다.

네 번째 단계는 파이프라인의 1차 검색 중속이다. POR은 Gemini File Search RAG가 1차로 검색한 상위 N건(본 실험에서 N=10)에서만 순위를 재배열하므로, 1차 검색이 누락한 노트는 POR이 구제할 수 없다. 즉 POR의 효과는 근본적으로 시맨틱 검색의 회수율에 중속된다. 1차 검색의 N을 늘리거나 별도 검색 후보 풀을 도입하는 등 추가 설계가 필요하다.

다섯 번째 단계는 PageRank 허브 입증의 위상이다. 4.2절에서 보고된 허브 PageRank 상위 15위 결과는 자동 허브 링킹이 적용된 그래프에서 측정된 것이며, 자동 링킹 이전 그래프에서는 어떤 허브도 상위 50위에 진입하지 못한다(실험 K). 즉 본 연구의 PageRank 결과는 키워드 의미 매칭이 반영된 그래프에서의 허브 중심성을 보여주는 것이지, 허브의 순수 그래프 이론적 중심성을 입증하는 것은 아니다. 자동 링킹은 위키링크 부재를 키워드 기반 의미 매칭으로 보완하는 의도된 설계 선택이지만, 그 결과 형성된 허브 중심성에는 자동 링킹의 작동이 함께 반영되어 있다.

여섯 번째 단계는 그래프 기반 요인의 미검증이다. 7개 스코어링 요인 중 그래프 근접도와 저자성은 RAG 인터페이스의 메타데이터 제약으로 본 실험에서 활성화되지 못하였다(4.6 참고). 따라서 본 연구가 정량 효과를 검증한 요인은 5개에 한정되며, 그래프 위상의 직접 효과는 후속 검증 과제로 남는다.

5.4 향후 연구 방향

첫째, 저자성 요인의 실험적 검증이 필요하다. v2.1.0에서 경로 기반 3단계 분류 체계를 구현하였으나, Gemini RAG citation 내 파일 경로(URI) 정보가 포함되지 않아 실시간 실험에서 활성화되지 못하였다. 그래프 근접도 역시 RAG citation title과 그래프 node title 간 정규화 불일치로 비활성 상태이다. 현재 7개 요인 중 2개가 RAG 인터페이스 제약으로 비활성이므로, citation 메타데이터를 확장하거나 별도의 경로 매핑 모듈을 개발해 해결하면 리랭킹 효과가 더 커질 것으로 예상된다.

둘째, 온톨로지 규칙의 반자동 생성 방법을 연구해야 한다. LLM을 활용하여 PKMS 내용으로부터 정체성 렌즈와 허브를 자동 제안하고, 사용자가 검증하는 방식이 가능하다. 나아가 이러한 반자동 온톨로지 구축 도구의 진입 장벽이 낮아진다면, 더 많은 개인이 자신의 지식 체계를 명시적으로 모델링할 수 있게 되어, 지식관리와 개인 아카이브의 실천적 저변이 확대될 것이다.

셋째, 다중 사용자·다중 평가자 연구로 일반화 검증을 진행한다. N=30~50명 규모의 다양한 PKMS 사용자 집단을 대상으로, 직업군(연구자, 작가, 개발자 등), PKM 규모, 도메인 특성(단일 전문 분야 대 다학제 교차), 링크 밀도를 독립변수로 설정해 실험을 반복한다. 이로써 정체성 렌즈 매칭과 교차 도메인 교집합이 보편적 요인인지, 특정 유형의 사용자에게만 유효한지를 확인하고, PKM 규모와 링크 밀도의 임계값 등 구조적 조건을 탐색할

수 있다. 또한 RQ2의 평가 구조 한계(5.3 참고)를 보완하기 위해, 동일 쿼리에 대해 (1) 본인 PKMS에서 본인이 평가한 결과, (2) 동일 PKMS를 7개월 이상 사용한 동료 평가자가 평가한 결과, (3) 외부 평가자가 자기 PKMS에서 같은 쿼리를 평가한 결과의 세 조건을 비교하는 설계가 효과적이다. 이는 PKMS 적합성이 본질적으로 생산자에게 귀속된다는 본질적 난점을 우회하지 못하지만, 평가자 동일인 구조에서 비롯된 확증 편향을 정량화하고 통제하는 절차로 기능할 수 있다.

넷째, 시간에 따른 온톨로지 정합성 관리도 과제이다. 생산자의 사고방식과 가치 체계는 변화하므로 온톨로지 갱신이 불가피하며, 갱신 시점에 과거 노트와 새 온톨로지 사이의 불일치를 어떻게 다룰지가 후속 과제이다. 온톨로지 버전 관리, 재검증 주기 정의(예: 6개월마다 정체성 렌즈 키워드 재검토), 과거 노트 재태깅 메커니즘(이전 온톨로지 버전의 점수를 보존하면서 새 버전 점수를 병기), 두 버전 간 일관성 평가 지표를 함께 설계하면, 장기 운영에서도 검색 품질의 안정성을 확보할 수 있다.

다섯째, 자동 허브 링킹 의존도가 어떤 일반화 패턴을 보이는지도 분석 과제이다. 본 연구는 단일 PKMS에서 자동 링킹이 그래프 위상을 결정짓는 설계 선택임을 확인하였으나(4.2, 5.3 참고), PKMS의 위키링크 밀도와 허브 정의 방식에 따라 자동 링킹의 비중은 달라질 수 있다. 다양한 위키링크 밀도(저밀도·중밀도·고밀도)의 PKMS에서 자동 링킹 전후 PageRank 허브 위상 변화를 비교하면, 자동 링킹이 어떤 PKMS 환경에서 필수이고 어떤 환경에서 보조 도구로 충분한지에 대한 일반화 패턴을 도출할 수 있다.

여섯째, 7개 스코어링 요인의 강건성 분석(요인 제거 분석)도 진행해야 한다. 본 연구의 4.4 가중치 시뮬레이션은 시맨틱 점수와 온톨로지 점수의 비율 변화에 한정되었으며, 7개 요인 중 일부의 가중치를 0으로 두는 요인 단위 분석은 수행되지 않았다. 후속 연구에서 정체성 렌즈 매칭(57건), 교차 도메인 교집합(29건) 등 활성도가 높은 요인을 단독으로 작동시켰을 때의 리랭킹 효과를 측정하면, 각 요인의 독립 기여도와 결합 기여도를 분리할 수 있다. 이는 본 연구가 정성적으로 추론한 요인별 기여 분포(4.8 참고)를 정량적으로 검증하는 절차이기도 하다.

일곱째, 다중 사용자 연구로 확장될 때 다중 비교 보정 절차도 도입해야 한다. 본 연구는 RQ1·2·3 각각에 대해 독립적으로 검정을 수행하였으나, 후속 연구에서 다수의 쿼리·평가자·요인을 동시에 비교하는 구조에서는 Bonferroni 보정, 거짓 발견율(False Discovery Rate, FDR) 등 통제 절차가 통계적 추론 엄밀성 확보에 필수적이다. 단일 사례 1차 실증의 위상에서는 작은 표본·탐색적 분석으로 수용된 약점이, 다중 사례로 확장될 때에는 명시적으로 통제되어야 한다.

여덟째, RiC 표준과 호환 가능한 개인 온톨로지 모델을 설계하여, 개인기록관리와 기관기록관리의 연속선에서 통합적 접근을 모색할 수 있다.

5.5 결론

본 연구는 개인 온톨로지 기반 지식 그래프 리랭킹(POR v1.0)을 설계하고, 3,702건의 마크다운 기반 PKMS에 대해 실험적으로 검증하였다. 주요 기여내용은 다음과 같다.

첫째, 개인의 사고방식과 가치관을 7개 스코어링 요인으로 정형화한 개인 온톨로지가 시맨틱 검색 결과의 순위를 쿼리 유형에 따라 차등적으로 변경함을 확인하였다. 허브 관련 쿼리($\tau=0.300$)에서 가장 큰 재배열이 발생하였고, 일반 쿼리($\tau=0.786$)에서는 기존 순위가 보존되었다.

둘째, 정체성 렌즈 매칭(57건)과 교차 도메인 교집합(29건)이 가장 활발한 스코어링 요인으로, 본 실험의 리랭킹 효과를 가장 크게 설명하는 요인이었다. v7에서 0건이었던 허브 부스트가 3건 활성화되어, 7개 요인 중 허브 부스트의 작동을 확인하였다.

셋째, 본 연구의 메인 그래프인 v8(2026.03 빌드, 5,717 노드와 13,544 엣지, 허브 10개)에서 10개 허브 전체가 PageRank 상위 15위 이내에 위치하여 수동 허브 선정의 타당성을 실증하였다. 1,810건의 자동 허브 링킹으로

74%의 그래프 근접도 도달률을 확보하였다. 본 메인 결과의 허브 중심성은 자동 허브 링킹의 그래프 위상에 한정된 해석임을 다루었다.

넷째, 사용자 블라인드 평가에서 온톨로지 조건이 허브 관련(2/2) 및 가치 관련(1/2) 쿼리에서 일관되게 선호되어, 쿼리 유형에 따른 차등 효과를 확인하였다. 전체적으로는 통계적 유의차가 없었으며($p=0.544$), 이는 7개 요인 중 2개(저자성, 그래프 근접도)가 RAG 인터페이스 제약으로 비활성 상태인 현 시스템의 한계에 기인한다.

본 연구는 기록학의 출처주의가 보존하려는 생산자 맥락의 구조를 개인기록의 검색 가중치 모델에 매핑한 1차 실증으로서, 생산자의 온톨로지가 검색도구의 기능을 포함하면서도 생산자의 사고방식과 가치관까지 반영하는 의미적 프레임워크로 기능할 수 있음을 확인했다. 본 매핑은 후속 다중 사용자 연구와 RiC-O, RiC-CM 정합의 추가 정렬을 거쳐 출처주의의 디지털 확장으로 발전할 가능성이 있으며, AI가 일반화된 시대에 개인기록의 조직과 검색에 기록학적 원리를 적용하는 한 방향을 제시한다.

참고문헌

- 강주연, 이진성, 윤선우, 오효정 (2025). RiC(Record in Context) 뜯아보기: 실무적·학술적 관점으로. 한국기록관리학회지, 25(2), 47-71. <https://doi.org/10.14404/JKSARM.2025.25.2.047>
- 김민지, 강성희, 이해영 (2024). 챗GPT를 활용한 기록관리 메타데이터 추출 사례연구. 한국기록관리학회지, 24(2), 89-112. <https://doi.org/10.14404/JKSARM.2024.24.2.089>
- 김현채, 강성희, 이해영 (2024). RiC-O(Records in Contexts - Ontology)를 활용한 국가기록원 기록물 생산기관 변천정보 서비스 개선방안. 한국기록관리학회지, 24(1), 47-72. <https://doi.org/10.14404/JKSARM.2024.24.1.047>
- 박하람, 강지우, 송채은, 안지은, 이정운, 장소정, 조하연, 차화준, 김학래 (2024). FAIR 데이터 원칙을 적용한 국내 디지털 아카이브 평가와 개선 방향. 한국기록관리학회지, 24(4), 113-134. <https://doi.org/10.14404/JKSARM.2024.24.4.113>
- 송치호 (2024). 자연어 처리의 개체명 인식을 통한 기록집합체의 메타데이터 추출 방안. 한국기록관리학회지, 24(2), 65-88. <https://doi.org/10.14404/JKSARM.2024.24.2.065>
- 이수현, 오효정 (2024). 대통령기록포털 검색어 로그 분석 기반 이용자 정보요구 분석. 한국기록관리학회지, 24(3), 23-44. <https://doi.org/10.14404/JKSARM.2024.24.3.023>
- 최유리, 이예은, 오효정 (2025). 지능형 기록정보서비스 챗봇을 위한 논항구조 의미역 분석 기반 응답 논리 구조 설계. 한국기록관리학회지, 25(3), 259-279. <https://doi.org/10.14404/JKSARM.2025.25.3.259>
- Ahrens, S. (2017). How to Take Smart Notes: One Simple Technique to Boost Writing, Learning and Thinking. 김수진 옮김 (2023). 제탈카스텐: 글쓰는 인간을 위한 두 번째 뇌. 서울: 인간회극.
- Balog, K. & Kenter, T. (2019). Personal knowledge graphs: A research agenda. Proceedings of the 2019 ACM SIGIR International Conference on Theory of Information Retrieval (ICTIR '19), 217-220. <https://doi.org/10.1145/3341981.3344241>
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment, 2008(10), P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>
- Brin, S. & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. Computer Networks and ISDN Systems, 30(1-7), 107-117. [https://doi.org/10.1016/S0169-7552\(98\)00110-X](https://doi.org/10.1016/S0169-7552(98)00110-X)
- Chakraborty, P. & Sanyal, D. K. (2023). A comprehensive survey of personal knowledge graphs. WIREs Data Mining and Knowledge Discovery, 13(6), e1513. <https://doi.org/10.1002/widm.1513>
- Collins, A. M. & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. Psychological Review, 82(6), 407-428. <https://doi.org/10.1037/0033-295X.82.6.407>

- Edge, D., Trinh, H., Cheng, N., Bradley, J., Chao, A., Mody, A., Truitt, S., Metropolitan, D., Osazuwa Ness, R., & Larson, J. (2024). From local to global: A graph RAG approach to query-focused summarization. *arXiv*.
<https://doi.org/10.48550/arXiv.2404.16130>
- Forte, T. (2022). Building a Second Brain. 서은경 옮김(2023). *세컨드 브레인*. 서울: 쌤앤파커스
- Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Wang, M., & Wang, H. (2024). Retrieval-augmented generation for large language models: A survey. *arXiv*. <https://doi.org/10.48550/arXiv.2312.10997>
- Gutiérrez, B. J., Shu, Y., Gu, Y., Yasunaga, M., & Su, Y. (2024). HippoRAG: Neurobiologically inspired long-term memory for large language models. *Advances in Neural Information Processing Systems* 38 (NeurIPS 2024).
- Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). Exploring network structure, dynamics, and function using NetworkX. *Proceedings of the 7th Python in Science Conference (SciPy 2008)*, 11-15.
- ICA (2000). ISAD(G): General International Standard Archival Description (2nd ed.). International Council on Archives.
- ICA EGAD (2023). Records in Contexts — Conceptual Model (RiC-CM) v1.0. International Council on Archives.
- Järvelin, K. & Kekäläinen, J. (2002). Cumulated gain-based evaluation of IR techniques. *ACM Transactions on Information Systems*, 20(4), 422-446. <https://doi.org/10.1145/582415.582418>
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems* 33 (NeurIPS 2020).
- Matuschak, A. (n.d.). Evergreen notes. Available: https://notes.andymatuschak.org/Evergreen_notes
- McInnes, L., Healy, J., & Astels, S. (2017). hdbscan: Hierarchical density based clustering. *Journal of Open Source Software*, 2(11), 205. <https://doi.org/10.21105/joss.00205>
- Milo, N. (2020). LYT Kit: Linking Your Thinking. Available: <https://www.linkinyourthinking.com/>
- Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The PageRank Citation Ranking: Bringing Order to the Web. *The Web Conference*.
- Peng, B., Zhu, Y., Liu, Y., Bo, X., Shi, H., Hong, C., Zhang, Y., & Tang, S. (2024). Graph retrieval-augmented generation: A survey. *arXiv*. <https://doi.org/10.48550/arXiv.2408.08921>
- Reimers, N. & Gurevych, I. (2019). Sentence-BERT: Sentence embeddings using Siamese BERT-networks. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 3982-3992. <https://doi.org/10.18653/v1/D19-1410>
- Schmidt, J. F. K. (2016). Niklas Luhmann's Card Index: Thinking Tool, Communication Partner, Publication Machine. *Library of the written word*, 53, 289-311 https://doi.org/10.1163/9789004325258_014
- Shepherd, E. & Yeo, G. (2003). *Managing Records: A Handbook of Principles and Practice*. London: Facet.
- Skjæveland, M. G., Balog, K., Bernard, N., Łajewska, W., & Linjordet, T. (2024). An ecosystem for personal knowledge graphs: A survey and research roadmap. *AI Open*, 5, 55-69. <https://doi.org/10.1016/j.aiopen.2024.01.003>

• 국문 참고자료의 영어 표기

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

- Ahrens, S. (2017). *How to Take Smart Notes: One Simple Technique to Boost Writing, Learning and Thinking*. Translated by Kim, Soo-jin (2021). *Zettelkasten: The Second Brain for Writing Humans*. Seoul: Ingan Huigeuk.
- Choi, Yuri, Lee, Yeeun, & Oh, Hyojung (2025). Logical answer structure design based on semantic role labeling (SRL) for an intelligent archival information service chatbot. *Journal of Korean Society of Archives and Records Management*, 25(3), 259-279. <https://doi.org/10.14404/JKSARM.2025.25.3.259>
- Forte, T. (2022). *Building a Second Brain: A Proven Method to Organize Your Digital Life and Unlock Your Creative Potential*.

- Translated by Seo, Eun-kyung (2023). *Second Brain*. Seoul: Sam & Parkers.
- Kang, Juyeon, Lee, Jinsung, Yoon, Sunwoo, & Oh, Hyojung (2025). Exploring record in context (RiC): From the perspectives of researchers and practitioners. *Journal of Korean Society of Archives and Records Management*, 25(2), 47-71.
<https://doi.org/10.14404/JKSARM.2025.25.2.047>
- Kim, Hyunchae, Kang, Sunghee, & Lee, Haeyoung (2024). Improving the National Archives of Korea's service for change information of records-creating agencies using Records in Contexts-Ontology (RiC-O). *Journal of Korean Society of Archives and Records Management*, 24(1), 47-72. <https://doi.org/10.14404/JKSARM.2024.24.1.047>
- Kim, Minji, Kang, Sunghee, & Lee, Haeyoung (2024). A case study on metadata extraction for records management using ChatGPT. *Journal of Korean Society of Archives and Records Management*, 24(2), 89-112.
<https://doi.org/10.14404/JKSARM.2024.24.2.089>
- Lee, Suhyun & Oh, Hyojung (2024). User information needs analysis based on search terms log of the Presidential Archives Portal. *Journal of Korean Society of Archives and Records Management*, 24(3), 23-44.
<https://doi.org/10.14404/JKSARM.2024.24.3.023>
- Park, Haram, Kang, Jiwoo, Song, Chaeun, An, Jieun, Lee, Jungyoon, Jang, Sojeong, Jo, Hayeon, Cha, Hwajun, & Kim, Hakrae (2024). The evaluation of domestic digital archives using the FAIR data principles and their improvement directions. *Journal of Korean Society of Archives and Records Management*, 24(4), 113-134.
<https://doi.org/10.14404/JKSARM.2024.24.4.113>
- Song, Chiho (2024). A method for metadata extraction from a collection of records using named entity recognition in natural language processing. *Journal of Korean Society of Archives and Records Management*, 24(2), 65-88.
<https://doi.org/10.14404/JKSARM.2024.24.2.065>