

# AI 기반 기록정보 검색 활용 고도화

지능형 검색 서비스 구축부터  
범국가 통합활용체계 확대로!

2026.05.20

Saltlux



# Contents

AI기반 기록정보 검색 활용 고도화 : 지능형 검색 서비스 구축부터 범국가 통합활용체계로의 확대까지

## Part 1

### ▶ AI기반 지능형 기록정보 검색 솔루션 개발 및 실증 사업

01 사업 개요

02 서비스 유즈케이스 개발

03 AI기반 검색 솔루션 개발

## Part 2

### ▶ AI 에이전트 기반 범국가 기록정보 통합 활용체계 모델설계 및 실증

01 사업 소개

02 AI 에이전트 서비스 모델 구성

03 PoC를 통한 서비스 활용 가능성 검토

전통기록문화  
한지와 목판인쇄

삼국시대~조선시대부터 발달된 목판인쇄술로 경전, 문집, 족보 등 다양한 서적 제작 및 지식 보급국가적 중요 유산인 건축물이나 역사적 사건을 기록하는 데 사용된 금석문, 현판 등의 기록물

근대적 기록관리 방식 도입  
(1945~1969)

1969년에는 정부기록보존소(현 국가기록원) 설립 이후 공공기록물의 체계적 관리 기반 마련

법적 기반과 현대화  
(1999~현재)

1999년: 기록관리 혁신의 출발점  
「공공기관의 기록물관리에 관한 법률」 제정으로 공공기록 관리의 법적 정비

2000년대 이후 국가기록관리 혁신의 본격적 추진  
디지털화와 전자기록 관리의 강조

디지털 전환과 클라우드  
(현대)

- 최근 클라우드 기반의 기록관리시스템(CRMS) 도입으로 기관 간 기록 공유 & 협업
- 디지털화된 기록으로 연구자 및 일반 시민의 활용도 증가

글로벌화

유네스코 세계기록유산으로도 인정받는 한국의 전통 기록문화 대표적으로  
- 훈민정음 해례본  
- 국채보상운동 기록물  
- 1983년 KBS 특별생방송 이산가족 찾기 관련 자료 등의 등재

우리나라의  
기록은 ...

전통적인 종이와 목판에서 시작해 현대적인 디지털 기술로 전환되며, 과거와 현재를 잇는 중요한 문화적 자산으로 자리 잡고 있습니다.

**그리고 국가기록원은  
연간 60만 건 이상의 기록을 제공하며  
국내 최대의 기록정보 기관으로  
성장하였습니다.**

데이터의 가치는

**기록** 을 넘어

활용될 때, 비로소

**자산** 이 됩니다.

수십 년간 정책과 행정, 그리고 역사를 정직하게 담아온 공공기록. 이제는 그 기록을 꺼내어 활용하고, 그 활용이 다시 새로운 기록이 되는

**지속 가능한 흐름을  
만들어야 할 때입니다.**

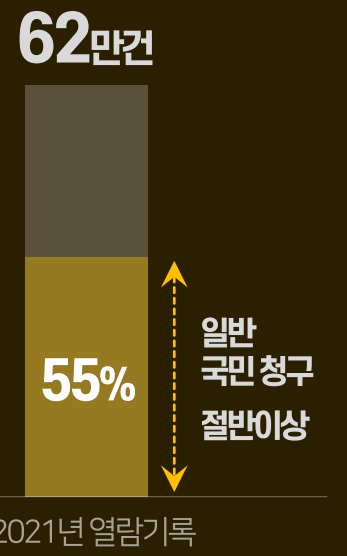


# 국가기록원 공공기록 1억 2천만 건!

## 쌓인 기록은 거대해 졌습니다. 하지만 국민이 필요한 '쓸 수 있는 기록'은 여전히 적습니다.



이제 국민이 주 이용자입니다.



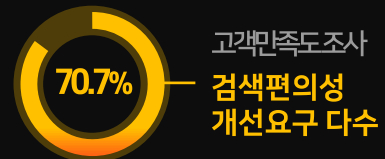
### 검색 문제점

기록을 찾기 어려운 구조입니다.

**키워드 중심의 검색 구조**

일반 국민은 원하는 기록을 찾기 어려운 구조와 통합성 부족

국민 과반수가 검색에 불편함을 느낍니다.



### 서비스 문제점

콘텐츠가 제한되어 있습니다.

열람 목적 중 70% 이상이 재산권·신분·행형 증빙에 집중

소송·보상·신원 확인 등 제한된 활용 목적에 편중

활용이 어렵습니다.

실제 활용과 동떨어지는 행정적 분류 중심의 기록 서비스

정보사용이 한정적 입니다.

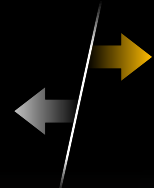
기타 열람(교육·연구·언론 등)은 전체의 약 30%로 다양한 정보 활용이 일어나지 못하고 있음

출처: 한국기록관리학회지 국가기록원의 기록제공서비스에 관한 탐색적 연구(2022), 기록의 발견과 이해를 위한 온라인 검색 가이드 연구(2015)

# 국민은 기록을 어떻게 사용하길 원할까요?

사용이 어렵고, 그저 행정적 정보를  
찾고만 싶어할까요?

NO



YES

활용이 가능한 맞춤기록을 찾아  
찾아서 쓰고 싶어 합니다.

단편적  
정보 검색이 아닌  
**‘문맥 기반 탐색 경험’**

단어가 아닌, 사건·사유·배경과  
데이터 특성까지 고려한 차세대 검색

- 필요한 정보를 빠르게!  
**대화형 크로스모달 통합 검색 플랫폼**
- 여러 번 검색하지 않아도!  
**Agentic RAG로 더 정교해지는 연관정보 제공**
- 한눈에 확인하는 필요한 정보!  
**의미 기반 기록물 요약과 사용자 맞춤형 추천**
- 보고 끝이 아니라 활용까지!  
**데이터 인용 & 기관 연계 API 확대**

‘보관’ 중심에서  
**‘활용’ 중심으로**  
바뀌는 기록

국민의 법적 증빙, 신원 확인,  
연구·교육 목적 활용 확대

진화된 검색 서비스 국가기록원의 **차세대 대화형 리서치 파트너!**

# Part 1

AI기반 지능형 기록정보 검색 솔루션 개발 및 실증 사업

01

▶ 사업 소개

02

▶ 서비스 유즈케이스 개발

03

▶ AI 기반 검색 솔루션 개발

## 01. 사업개요

## 2차년도

Business Scope 🔍

# NIPA AI기반 지능형 기록정보 검색 솔루션 개발 및 실증

국민들이 빠르고 편리하게 기록물을 검색하고 정보를 활용할 수 있도록 지원하는 AI 기반 지능형 기록정보 검색 솔루션 개발 및 실증

## 지원분야

## 국가기록분과

국민들이 빠르고 편리하게 기록물을 검색하고 정보를 활용할 수 있도록 지원하는 AI 기반 지능형 기록정보 검색 솔루션 개발 및 실증

## 과제사업내용

정부부처와 협력을 통해 다양한 산업 분야에서 활용 가능한 AI 솔루션을 개발·실증하여 부처 현장 적용 및 확산 추진

- ▶ 데이터 가공·활용 - 협업부처 보유 데이터 및 사업수행기관 자체 확보 데이터를 이용하여 학습데이터 구축
- ▶ AI 개발·실증 - 구축한 데이터를 기반으로 협업부처의 요구사항 등이 포함된 AI 솔루션 개발·고도화 및 실증, 현장적용 추진

## 사업기간

2026.1.1 ~ 2026.12.31.

\* 회계연도 단위로 협약하고, 1차년도('25.협약월~'25.12.31.), 2차년도('26.1.1.~'26.12.31.)

## 02. 1차년도 사업 추진 결과

### ● 국가기록원 학습데이터 구축 · 적용

항목	목표	구축 현황
데이터 수집	1차년도 750 만건 이상	<ul style="list-style-type: none"> <li>1차년도 수집 문서 (전자문서, OCR 데이터, 정부발행 간행물, 이미지, 업무매뉴얼)</li> </ul>
데이터 전처리	2017년 이관분까지 데이터 전처리	<ul style="list-style-type: none"> <li>2017년도 까지 이관된 전자문서 문서변환 및 데이터 추출 완료.</li> <li>OCR, 간행물 등 VLM 을 활용한 데이터 전처리 완료</li> </ul>
국가기록원 특화 LLM 학습용 데이터 셋 구축	기록원 특화 유즈케이스 구축	<ul style="list-style-type: none"> <li>전자문서, 매뉴얼, 업무지침, 정부 간행물 대상으로 유즈케이스 발굴 유즈케이스 별 QA 데이터셋 구축</li> </ul>

### ● 지능형 기록정보 검색 솔루션 개발 및 실증

항목	목표	구축 현황
AI 지능형 기록정보 검색 솔루션 개발	멀티턴 방식의 대화형 기록정보 검색 구현	<ul style="list-style-type: none"> <li>멀티턴 기반의 대화형 검색 API 구현 완료</li> <li>RAG 검색을 위한 검색 API 구성</li> <li>기록물 요약 / 주간 브리핑기능 구현 완료</li> </ul>
	지능형 기록정보 UI/UX 설계 및 개발	<ul style="list-style-type: none"> <li>사용자 UI 시안 3종 작성 국가기록원 직원 대상 투표를 통해 최종 A 안 선정.</li> </ul>
국가기록원 특화 모델 구축	기록원 특화 유즈케이스 적용	<ul style="list-style-type: none"> <li>유즈케이스별 시스템 프롬프트 / 사용자 프롬프트를 구성하여 상황별 검색에 사용하도록 적용</li> </ul>
	국가기록원 특화 모델 파인 튜닝	<ul style="list-style-type: none"> <li>모델 특성과 사용자 요구 분석을 바탕으로 프롬프트 설계를 정교화하여 다양한 작업 유형에서의 응답 일관성과 생성 품질 향상</li> <li>전처리·정제된 데이터를 기반으로 VLM 학습을 수행하여 VQA, Captioning 등 기본 성능 확보</li> </ul>

## 03. 2차년도 사업 추진 과제

### ● 국가기록원 학습데이터 구축 · 적용

항목	목표	과제 내용
데이터 수집	약 530만건 이상 수집	<ul style="list-style-type: none"> <li>공개팀에서 공개기록물로 확인된 전자문서 대상으로 재수집 및 전처리</li> <li>멀티모달 데이터인 사진 / 음성 및 영상 데이터의 수집 및 전처리</li> </ul>
데이터 가공	데이터 전처리 및 가공을 위한 파이프라인 구축	<ul style="list-style-type: none"> <li>국가기록원 포탈 시스템의 DB / 파일시스템에서 데이터를 수집하여 전처리 및 데이터 가공 및 색인을 위한 데이터 파이프라인 구축</li> </ul>
국가기록원 특화 LLM 학습용 데이터 셋 구축	기록원 특화 유즈케이스 구축	<ul style="list-style-type: none"> <li>대국민 서비스 구축을 위한 다양한 사용자군 및 이용환경/방법에 대한 유즈케이스를 발굴하고 설계 개발</li> </ul>

### ● 지능형 기록정보 검색 솔루션 개발 및 실증

항목	목표	구축 현황
AI 지능형 기록정보 검색 솔루션 개발	Multi Agent 기반 검색 시스템 고도화	<ul style="list-style-type: none"> <li>멀티모달 데이터를 포함하여 멀티턴 기반의 대화형 검색 API 고도화</li> <li>멀티 에이전트 기반 검색 시나리오 구성하여 스스로 계획하고, 추론하고 검증할 수 있는 구조 구성</li> <li>텍스트, 이미지, 음성 등 다양한 입력 데이터를 처리할 수 있는 LMM 기반의 검색 프레임워크를 제공</li> </ul>
	기록물 온톨로지 구축 및 서비스 구축	<ul style="list-style-type: none"> <li>기관명, 정책명, 작성자, 일자 등의 주요 정보를 중심으로 데이터 간 관계를 '소속됨', '관련됨', '작성함' 등 구성하여 의미적 연결을 제공</li> <li>모델링된 지식그래프를 토대로 데이터를 적재하여 LLM에 활용할 수 있는 체계 구축</li> </ul>
	LMM 학습용 데이터 고도화	<ul style="list-style-type: none"> <li>기록물 특화 LMM 학습용 데이터셋 보강</li> <li>LLM / VLM 모델 파인튜닝을 통한 기록원 특화 모델 구현</li> </ul>
현장적용 및 실증	국가기록원 내 검색 시스템 실증 환경 구성	<ul style="list-style-type: none"> <li>대국민 서비스 오픈을 위한 국가기록원 내 시스템 시범 오픈 및 운영 모니터링</li> </ul>

# 04. 구축 대상 데이터 범위



## 기록물 건 메타 데이터

21,548,810 건

생산년도 2003년 이후 기록원으로 이관된 모든 기록물의 메타 데이터

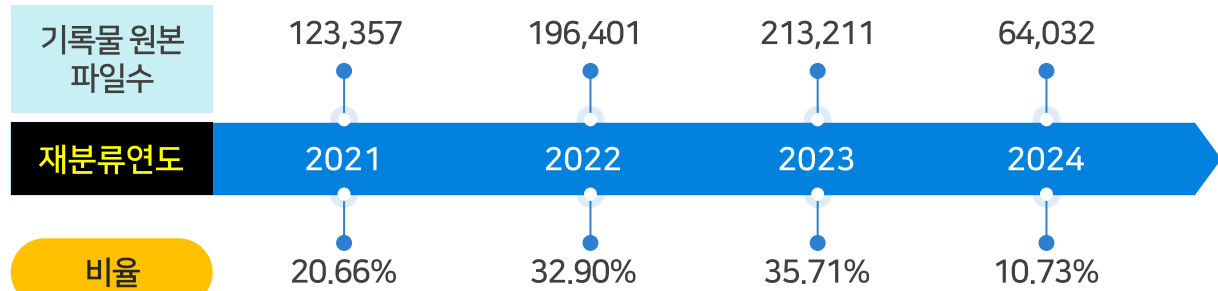
일반기록물 / 시청각기록물 / 정부간행물 / 해외기록물 / 민간기록물 포함

기록물형태	기록물 건수	비율
일반문서류	20,094,336	93.25%
도면류	23,400	0.11%
사진, 필름류	1,091,769	5.07%
녹음, 동영상류	78,809	0.37%
카드류	4,783	0.02%
정부간행물	254,395	1.18%
간행물	1,317	0.01%

기록물 유형	기록물 건수	비율
일반기록물	20,110,850	93.33%
시청각기록물	1,166,880	5.42%
정부간행물	254,037	1.18%
해외기록물	10,714	0.05%
민간기록물	6,329	0.03%

## 공개 재분류 완료된 전자기록 원문

2021년~2024년까지 공개팀에서 공개 재분류가 완료되어 원문 서비스가 가능한 기록물 원본으로 597,000 건



• 2003년 이후 출간된 정부 간행물로 한글로만 작성된 자료

## 온톨로지 활용을 위한 기록물 관계 정보



통합 기관 코드



기관변천 연혁정보



단위과제



정부별 정책과제



기록물 메타데이터

# 05. 구축 데이터의 AI 서비스 활용

## 기록원 보유 데이터에 대한 AI 서비스 활용

### 데이터 유형

소장 기록의 목록 데이터

내용 답변 가능한 원문

기록물 관계·맥락 (전거데이터)

### 데이터 종류

2003년 이후 생산된 전자 기록물로 모든 철/건 데이터의 메타데이터

공개팀에서 공개 재분류 완료된 기록물의 원본 데이터

전체 소장 기록의 목록과 연결

### 활용 데이터

존재여부   전체 규모   기관별/연도별 분포   기록군 구조

후속 열람/신청 안내를 담당

원문 있는 기록만 청킹   백터 DB   내용 기반 답변 근거

문단/원문 위치 연결

기관 변천   유사 개념   사건/정책/기관 관계

기록 간 상하/연관 관계

국가기록원의 장기적인 로드맵과 확장성까지 고려한 데이터 활용 UP!!



# 01. 유즈케이스 검증 위한 이용자 인터뷰 실시

- ▶ 총 2회(1차:3.23(월) 19~20시 / 2차:3.24(화) 19시~20시)
- ▶ 총 9명(국가기록포털 주요 사용자 7명\* / 우리원 열람담당자 2명)
  - \* 2025년도 국가기록포털 사용자 만족도 조사 응답자 중 성의있게 의견을 제시한 15명을 1차 후보군으로 선정하고 이 중 실제 인터뷰 가능한 7명을 최종 선정
- ▶ 온라인 회의방식으로 실시 하였으며, AI 검색서비스 유즈케이스 현실성, 유용성, 신뢰성 등 검증

## 유즈케이스 도출을 위한 시사점



### 이용자군별로 서비스 기능 및 활용방식 차이 확인

- 일반이용자 및 학생 : 기록물 내용의 정답형 답변 및 내용정리에 긍정적 반응
- 연구자 : 관련 기록을 폭넓게 탐색하고 구조적 확장 기능이 더욱 중요
- 행정실무 : 최신 법령 · 지침 확인, 비공개 · 열람 절차 안내 정확성 중시



### 검색결과 제시 방식과 탐색 구조에 대한 개선 요구

- 검색결과에서 보여주는 '관련기록'의 기준이 명확할 필요
- 또한 연구자는 정확도 높은 상위 몇 건의 기록만 제시되는 것보다 관련기록의 전체적 규모(생산기관·생산기간별 등) 파악이 우선 필요
- 유사어, 동의어, 표기변형 등을 포괄하는 검색확장 기능 필요성 제기



### 답변 신뢰성과 이용 안내의 중요성

- 답변 자체보다 기록을 근거로 생성 · 제시하는 방식이 중요
- 단순한 설명보다는 근거문서, 출처, 원문과 연결된 형태의 답변이 더 신뢰성 보장
- 비공개 절차 등은 적절하다고 평가하였으나, 실제 이용 시 더욱 구체적인 안내 필요
- 사진·영상 포함 여부, 저작권 및 초상권 문제, 2차 활용 가능 여부의 안내 필요



### 서비스 범위와 연계 확대에 대한 요구

- AI 검색서비스의 우선 대상이 전자문서 중심이라는 점은 수용하였으나, 이 범위가 충분히 설명되지 않으면 이용자가 실망으로 이어질 수 있음. 이에 종이기록, 디지털컬렉션 등으로 서비스 확장 필요
- 국가기록원에 자료가 없을 경우, 유관기관(국편, 한중연 등) 자료와 연계 수요 확인

## 02° 기록물 검색을 위한 사용자 유즈케이스 유형 ( 1/2)

### 이용자 의도분석 유즈케이스

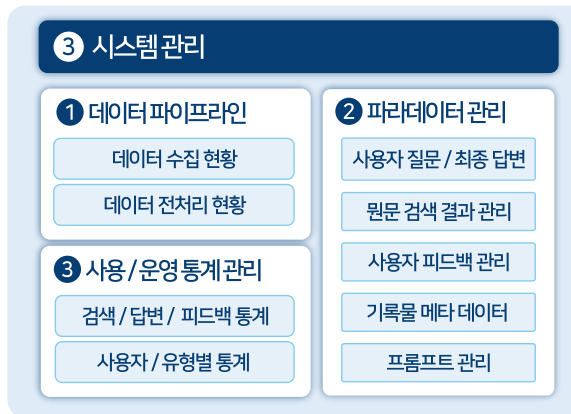
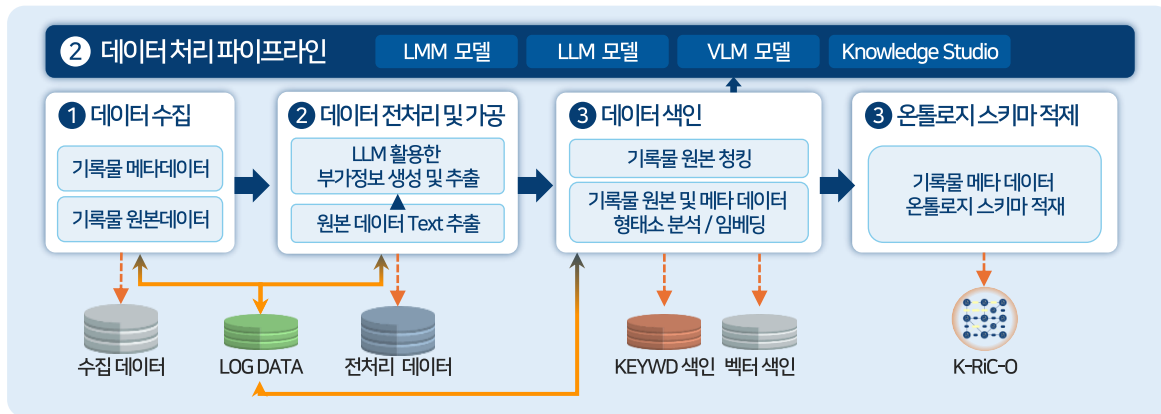
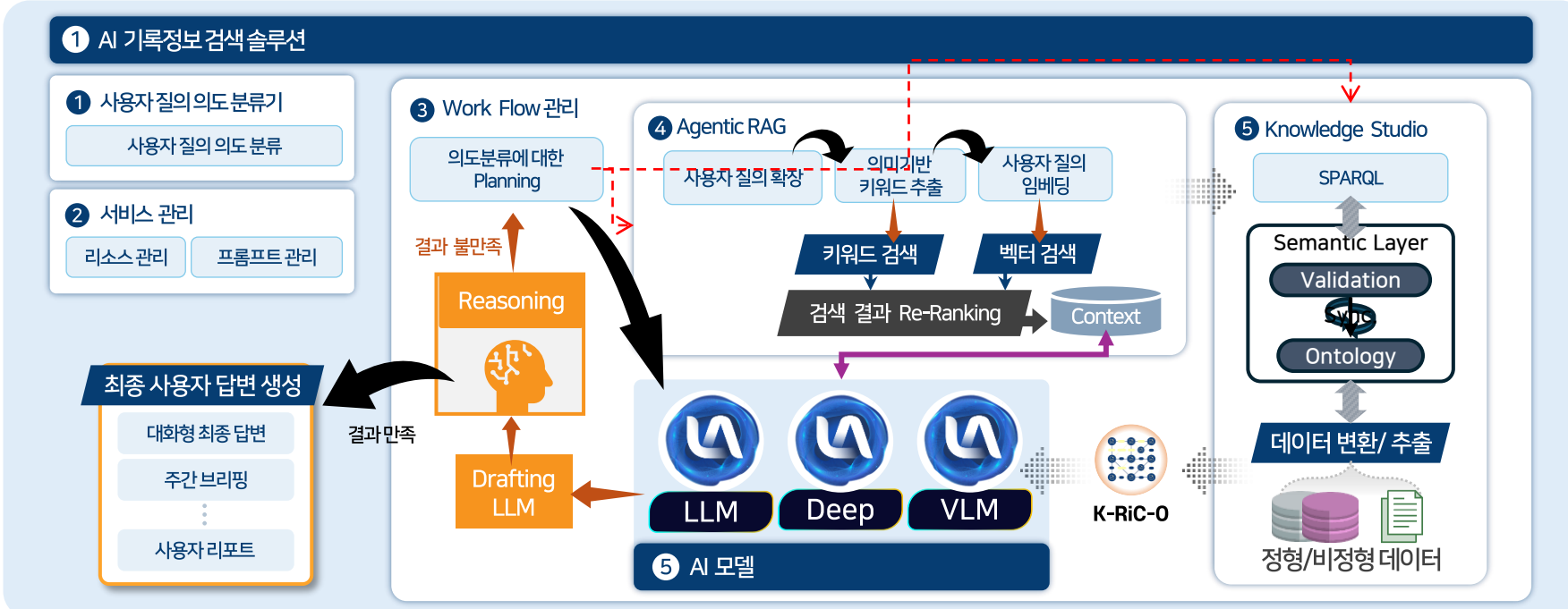
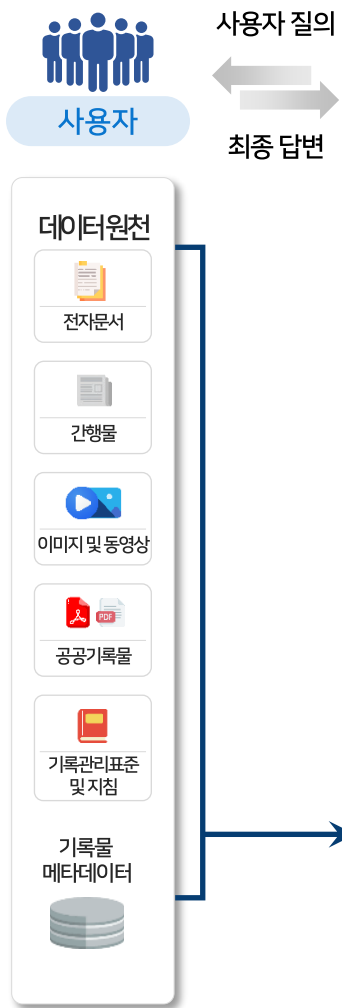
유즈 케이스 유형	목적	활용 방식	원칙
정답형 탐색	특정 사실, 수치, 날짜, 문구, 메타 정보 확인	<ul style="list-style-type: none"> <li>원문 있는 기록을 우선 검색해서 답변 생성</li> <li>목록만 있는 기록은 관련 기록 존재 여부나 추가 탐색 후보로 제시</li> <li>온톨로지는 동의어/기관/사건 연결 보조에 활용</li> </ul>	<ul style="list-style-type: none"> <li>내용 답변은 원문 근거 중심</li> <li>목록만 있는 기록은 "관련 기록 있음" 수준으로만 활용</li> <li>제목만 보고 내용을 추정하지 않음</li> </ul>
구조 탐색형	특정 주제 기록의 전체 규모, 흐름, 구성 파악	<ul style="list-style-type: none"> <li>먼저 주제 관련 기록 전체를 모음</li> <li>그 안에서 기획·지침 / 집행·조사 / 집계·보고 등으로 구조화</li> <li>원문 보유 여부를 함께 표시</li> <li>주요 기록은 시작점으로 제시</li> </ul>	<ul style="list-style-type: none"> <li>구조 탐색형은 목록-only 기록도 적극 활용</li> <li>원문 없는 기록도 전체 구조와 분포를 보여주는 데 포함</li> <li>원문 있는 기록은 구조를 대표하는 예시 문서로 제시</li> </ul>
연혁·기관 중심 탐색형	특정 기관/부서/기관 변천에 따라 어떤 기록이 생산됐는지 파악	<ul style="list-style-type: none"> <li>기관명, 전거 정보, 기관 변천 관계를 기반으로 관련 기록군 수집</li> <li>목록만 있는 기록을 포함해 생산 흐름 파악</li> <li>원문 있는 기록은 상세 확인용으로 연결</li> </ul>	<ul style="list-style-type: none"> <li>이 유형은 온톨로지 의존도가 가장 높음</li> <li>목록만 있는 기록도 기관 생산 흐름 파악에 매우 중요</li> <li>원문은 "깊이 보기" 역할</li> </ul>
주제 중심 탐색형	특정 정책/사건/주제 관련 기록을 모아서 보기	<ul style="list-style-type: none"> <li>주제 관련 기록을 폭넓게 수집</li> <li>원문 있는 기록은 설명 가능</li> <li>목록만 있는 기록은 관련 기록 확장</li> <li>온톨로지는 유사 개념, 관련 기관, 사건 연결에 활용</li> </ul>	<ul style="list-style-type: none"> <li>원문 = 깊이</li> <li>목록 = 넓이</li> <li>온톨로지 = 연결</li> </ul>
비교·맥락 탐색형	시기, 기관, 정책, 기록군 간 차이와 변화 파악	<ul style="list-style-type: none"> <li>비교 대상 기록을 수집</li> <li>원문 있는 기록은 차이점/공통점 설명 근거로 사용</li> <li>온톨로지는 비교 기준 설정과 관계 연결에 사용</li> <li>목록만 있는 기록은 비교 범위 확장에 활용</li> </ul>	<ul style="list-style-type: none"> <li>실제 차이 설명은 원문 중심</li> <li>목록만 있는 기록은 "더 비교할 후보" 역할</li> <li>온톨로지는 비교 축 정리 역할</li> </ul>
길잡이 탐색	어디서부터 시작해야 할지 모르는 이용자에게 출발점 제시	<ul style="list-style-type: none"> <li>주제 관련 전체 자료를 가볍게 훑음</li> <li>주요 기록, 기관별 보기, 연도별 보기, 전체 구조 보기 같은 시작점 제시</li> <li>목록만 있는 기록도 "관련 기록 존재 범위"를 보여주는 데 활용</li> </ul>	<ul style="list-style-type: none"> <li>길잡이형은 세 가지 정보를 모두 혼합 활용</li> <li>답을 닫지 않고 탐색 경로를 제시하는 데 중점</li> <li>주요 기록은 정답이 아니라 출발점</li> </ul>

## 02° 기록물 검색을 위한 사용자 유즈케이스 유형 (2/2)

### 안내형 유즈케이스

유즈케이스 유형	목적	활용 방식	원칙
비공개 기록물 신청 안내형	기록이 '비공개'상태인 경우, 정보공개 청구 프로세스 및 신청 웹페이지로 원활하게 이동할 수 있도록 안내	<ul style="list-style-type: none"> <li>원문 보유 기록과 목록만 있는 기록의 메타정보(공개상태: 비공개)를 식별하여 주 정보원으로 활용</li> <li>온톨로지는 해당 비공개 기록과 연관된 '공개' 기록이 있는지 탐색 확장</li> </ul>	<ul style="list-style-type: none"> <li>비공개 상태임을 명확히 고지하고, 행정적인 탐색 경로(정보공개청구 프로세스 및 신청 웹페이지)를 안내</li> </ul>
재검색 안내형	조건에 맞는 기록이 없을 때 범위 조정/키워드 변경을 안내하거나, 공개 기록은 없으나 비공개 기록이 N건 존재함을 알리고 정보공개 청구 절차로 유도	<ul style="list-style-type: none"> <li>목록만 있는 기록의 메타정보와 온톨로지를 활용하여 검색 범위를 조정하도록 유도</li> <li>공개상태: 비공개)를 식별하여 비공개 기록이 있는지 탐색 확장</li> </ul>	<ul style="list-style-type: none"> <li>이용자가 서비스를 이탈하지 않고, 지속적인 기록물 탐색이 가능하도록 재검색 유도</li> <li>비공개 기록물의 존재함을 안내하고, 행정적인 탐색 경로(정보공개청구 프로세스 및 신청 웹페이지)를 안내</li> </ul>
검색 결과 과다시 안내형	조건에 맞는 기록이 수백/수천 건일 때 바로 제시하지 않고, 연도/기관/문서유형 등으로 범위를 좁혀달라고 안내	<ul style="list-style-type: none"> <li>원문 보유 기록과 목록 only를 주 정보원으로 활용</li> <li>목록만 있는 기록의 메타정보(생산연도, 생산기관, 문서유형 등)를 식별하여 범위를 좁혀 안내</li> </ul>	<ul style="list-style-type: none"> <li>이용자가 확인하기 용이한 검색 결과가 제공할 수 있도록 범위를 좁혀 안내</li> </ul>

# 01. 최종 목표 서비스 구성도



# 02. 데이터 처리 파이프라인 구축

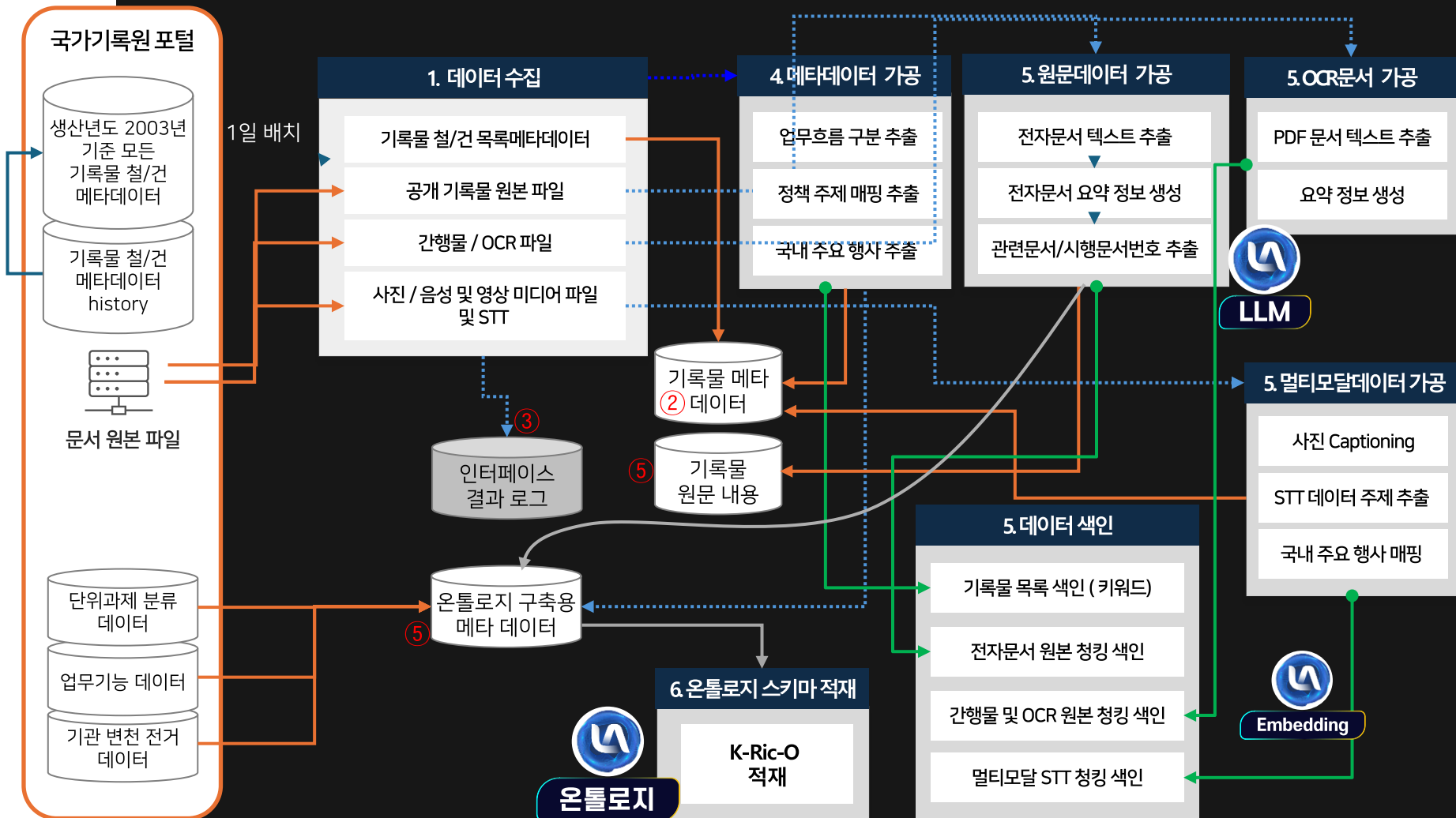
## 데이터 처리 파이프라인 흐름도

### 데이터 처리에 대한 자동화 프로세스 구축

CAMS → 국가기록포털 시스템으로  
데이터 연계 시 자동으로 AI 검색시스템으로  
적재되어 서비스에 활용할 수 있는  
자동화 프로세스 구축

### 구축방안

- 현재는 생산년도 기준 2003년부터 현재까지의 모든 기록물에 대한 목록을 대상으로 함
- 사업기간동안 국가기록원이 보유하고 있는 모든 기록물 목록에 대한 메타데이터를 처리할 수 있는 체계를 구축



## 03° 기록물 검색에 온톨로지의 필요성

### 기록물의 검색에서 온톨로지가 왜 필요한가?

온톨로지는 전거데이터와 기록물 메타데이터를 연결하여, LLM이 기록물의 명칭·기관·업무·시대적 맥락을 이해하고 정확하게 검색·추천·답변하도록 만드는 지식 기반입니다.

#### 1 기록물의 의미와 맥락을 구조화하기 위해

기록물은 단순 문서가 아니라 누가, 언제, 어떤 업무 과정에서, 어떤 사건과 관련하여 생산했는지가 중요

#### 2 전거데이터를 AI가 활용 가능한 지식으로 바꾸기 위해

전거데이터를 AI가 추론하고 탐색할 수 있는 지식 네트워크로 확장

#### 3 키워드 검색의 한계를 보완하기 위해

온톨로지가 있으면 서로 다른 표현을 업무기능·개념 관계로 연결하여, 사용자가 입력한 단어와 정확히 일치하지 않아도 의미 기반 검색이 가능

#### 4 RAG 검색 결과의 정확도를 높이기 위해

LLM 기반 RAG에서는 검색된 문서가 정확해야 답변도 정확해 집니다. 온톨로지를 활용하면 RAG 검색 전에 다음과 같은 보강이 가능

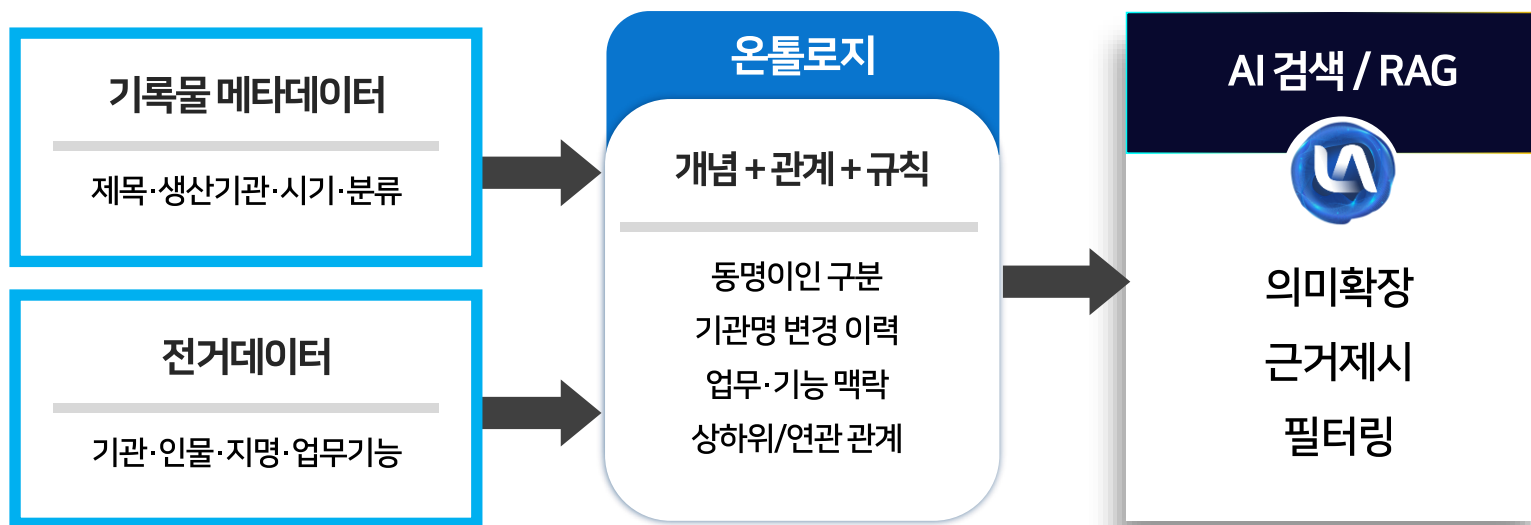
#### 5 LLM의 환각을 줄이고 설명 가능한 검색을 만들기 위해

온톨로지는 LLM에게 검증 가능한 기준을 제공함. 즉, 온톨로지는 AI 검색 결과의 신뢰성, 설명가능성, 검증가능성을 높입니다.

## 04. 기록물 검색에 맥락정보 활용을 위한 온톨로지 구축 방안

### ▶▶ 지능형 기록정보 검색시스템에서의 온톨로지의 역할 ◀◀

“단순 키워드”가 아니라 기록의 생산 맥락과 전거 관계를 따라 탐색



- 1 전거데이터의 의미 표준화**  
 동일 인물·기관·지명, 명칭 변경, 약어/이형표기를 하나의 식별자로 연결
- 2 검색 확장과 질의 이해**  
 사용자 질문을 관련 기관·업무기능·시대·상위/하위 개념으로 확장
- 3 RAG의 근거와 필터 제공**  
 LLM이 관련 기록을 잘 찾도록 메타데이터 필터와 관계 기반 후보군을 제공
- 4 설명가능성과 신뢰성 확보**  
 왜 이 기록이 검색되었는지 “기관-기능-기록물” 관계로 설명 가능

온톨로지는 선택적 부가기능이 아니라, 기록물 검색을 “문자열 검색”에서  
**“맥락 기반 지능형 탐색”**으로 전환하는 기반

# 05° 사용자 UI/UX 구성

행정안전부 국가기록원
공개 리포트 > 내가 만든 리포트 >
홍길동님, 안녕하세요. N | 로그아웃

AI 기반 지능형 기록정보 검색서비스

⊕ 새로운 대화

⊖ 나의 대화

주 5일은 언제부터 시작되었... ✎

IMF 경제 위기 극복을 위한... ✎

국가 외교의 발전 과정 ✎

국가기록원이 하는 일은 뭐야 ✎

페이 처리된 기록물 처리 방... ✎

주 5일은 언제부터 시작되었나요?

IMF 경제 위기 극복을 위한... ✎

국가 외교의 발전 과정

국가기록원이 하는 일은 뭐야

페이 처리된 기록물 처리 방식은...

⊖ 나의 기록물

지구촌 자연재해 Newsletter(20...

대통령기록물 관리에 관한 법률 시...

지구촌 자연재해 Newsletter(20...

대통령기록물 관리에 관한 법률 시...

대외무역법, 관세법 시행령 등

⊕ 주간 브리핑

⊕ 맞춤 리포트

대화 검색

제목 / 내용

TEXT

일시	대화 제목
26/02/15 13:00	주 5일 근무제는 언제 시작되었나요?
26/02/15 13:00	주 5일 근무제는 언제 시작되었나요?
26/02/15 13:00	주 5일 근무제는 언제 시작되었나요?
26/02/15 13:00	주 5일 근무제는 언제 시작되었나요?
26/02/15 13:00	주 5일 근무제는 언제 시작되었나요?
26/02/15 13:00	주 5일 근무제는 언제 시작되었나요?

**2002년 한일월드컵의 성과와 그 파급 효과**

102년 한일월드컵은 경기장과 인프라 확충, 치안 및 통신 지원 강화 등 정부 주도의 준비를 통해 성공적으로 개최되었으며, 경제적 파급 효과로 2조 382억 원 투자와 18만 5,000명 고용 창출을 이뤘습니다. 또한 거리 응원 문화와 '레드 스피릿'의 등장, 한류 확산, 동북아 경제 협력 강화라는 사회문화 국제적 성과를 남겼습니다.

**한일월드컵, 정부의 결의와 국민의 열기로 빛났던 2002년**

정부의 준비, 경기장과 인프라의 혁신

102년 한일월드컵의 개최를 위해 대한민국 정부는 역사상 유례없는 강도와 규모의 준비를 선보였습니다. 경기장 건설은 물론, 전국에 걸친 인프라 확충 및 통신 체계의 현대화가 추진되어 대회의 원활한 진행이 보장되었습니다. 특히 정부는 2조 원이 넘는 막대한 예산을 투입하여 서울, 부산, 대구, 주 등 10개 도시에 월드컵 경기장을 신축 및 보수하고, 주변 도로와 각종 기반시설을 현대화하였습니다. 또한 국가적 치안 및 질서 유지를 위해 경찰청 중심으로 15만 명에 달하는 경찰력이 동원되어 경기 기간 중 안전사고 방지와 질서 유지를 위한 전방위 노력이 이뤄졌습니다. 더불어, 방송 통신망 안정을 위해 전파관리 당국이 첨단 통신장비를 도입하고, 주파수 관리 및 불법 무선기기 단속 등 기술적 지원도 강화하였습니다. 이러한 다각적 준비는 단 한 경기 개최를 넘어, 한국 사회 전반의 인프라 수준과 운영 역량을 세계에 강하게 각인시키는 계기가 되었습니다. 1 2 3

**2. 경제적 파급, 투자와 고용의 성과**

102년 월드컵은 경제 측면에서도 매우 큰 성과를 거두었습니다. 경기장 건설 등 직접적인 사업을 통해 18만 5,000여 명에 달하는 고용 효과가 파급되었습니다. 정부 산하 연구기관의 조사 결과에 따르면, 월드컵 개최로 인해 대회 운영을 위한 인프라 확충, 치안 및 통신 지원 강화 등 정부 주도의 준비를 통해 성공적으로 개최되었으며, 경제적 파급 효과로 2조 382억 원 투자와 18만 5,000명 고용 창출을 이뤘습니다. 또한 거리 응원 문화와 '레드 스피릿'의 등장, 한류 확산, 동북아 경제 협력 강화라는 사회문화 국제적 성과를 남겼습니다.

**경제장관간담회(2001.9.14)회의결과 문서등록**

1. 월드컵의 활용 필요성 월드컵은 단순한 스포츠 경기를 넘어 경제, 문화, 외교 등 국가 전반에 지대한 영향을 미치며 국가 발전의 획기적 계기로 적용됨. 비교와 평가: 일본과의 공동 개최로 인해 대회 운영 능력뿐만 아니라 IT, 관광 산업 등 국가 경쟁력 전반에 걸친 비교와 평가가 이루어질 전망이다.

이뤄졌고, 이 과정에서 경기장 등 교통 등 관련 산업에도 소비 진작 효과가 기대됨. 생산유발 효과는 11조5,000억 원, 고용유발 효과는 11만 600명, 관광 수입 등 국가 경쟁력 전반에 걸친 비교와 평가가 이루어질 전망이다.

활성화 기반을 다졌다는 평을 받고 있습니다. 한국은 월드컵 개최를 계기로 현대적인 경제시스템 운영 능력과 대영 이벤트 인프라 조성 역량을 국제적으로 인정받게 되었습니다.

답변이 만족스러우셨나요?

👍 👎

⋮ 관련 기록물 더보기 클릭하신 후 기록물 제목을 클릭하시면 전체 내용을 확인하실 수 있습니다.

관련 기록물 더보기(18)

2002년 한일월드컵의 개최를 위해 정부에서는 어떤 노력을 했고, 개최 성과는 어땠나요?

관련 기록물 더보기 클릭하신 후 기록물 제목을 클릭하시면 전체 내용을 확인하실 수 있습니다. 관련 기록물 더보기(18)

다음으로 이런 내용도 궁금하실 수 있어요

- ✓ 2002년 한일월드컵 개최로 인한 한국 축구의 변화는 무엇이었나요?
- ✓ 한일월드컵 개최 당시 정부의 홍보 및 마케팅 전략은 무엇이었나요?
- ✓ 월드컵 개최가 한국의 국제적 위상에 미친 영향은 무엇이었나요?
- ✓ 한일월드컵 준비 과정에서 주요 인프라 개선 사례는 무엇이었나요?

대화 내용으로 온도표지가 생성되었습니다. 확대해서 보시면 관련 기록 탐색 및 핵심 주제를 중심으로 관련된 기록물을 확인하실 수 있습니다.

19

# Part 2

AI 에이전트 기반 범국가 기록정보 통합 활용체계 모델설계 및 실증

01

▶ 사업 소개

02

▶ AI 에이전트 서비스 모델 구성


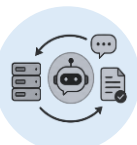

03

▶ PoC를 통한 서비스 활용 가능성 검토

# 국가기록원 지능형 검색에서 범국가 기록정보 통합 활용체계로

국가기록원 기록정보의 AI 검색 기반을 구축하는 단계

## 국가기록원 내부 검색 지능화

-  메타데이터·원문·관계정보 수집·가공·색인
-  대화형 검색, 요약, 추천, 근거 기반 답변 실증
-  온톨로지로 기관·업무·정책·기록물 맥락 연결

AI  
검색




기반  
구축

검색 고도화 성과를  
통합 활용체계의  
기반으로 확장

통합  
활용

국가전체 기록정보의 연결·통합·활용 체계로 확장하는 단계

## 범국가 기록정보 통합 활용체계

-  기관 시스템은 유지하고 필요한 범위만 경량 연계
-  메타데이터 표준화와 온톨로지 기반 용어·관계 통합
-  AI 에이전트가 기관별 기록정보를 탐색·연결·통합 응답

“국가기록원 내부 검색 지능화” ... 모든 국가 기록정보를 연결해 활용하는 **AI 기록정보 생태계**

# 기록정보 공유·활용의 정책·제도 기반이 마련되어 통합활용체계를 구체화할 시점



1억 2천만 여건 이상의 기록정보를 보유한  
**국내 최대 기록정보기관**입니다.

## 정책

123대 국정과제로  
**‘정보공개 및 기록물관리 강화’** 추진

정보공개 및 기록물관리 강화(국정과제 14번): 정보공개 기준과 절차의 합리적 개편으로 정보공개 확대, 공공기록물 등의 생산·지정·열람·공개 등 서비스 개선



## 법제도

공공기관이 공동 활용할 수 있는  
**‘법정부 생성형 AI 플랫폼’** 구축·운영 근거 마련

‘데이터기반행정활성화에 관한 법률’을 ‘인공지능 및 데이터기반행정활성화에 관한 법률(이하 ‘공공 AI법’),’로 개정(26.1.29)하고, ‘법정부 생성형 AI 플랫폼’ 구축·운영 등 규정

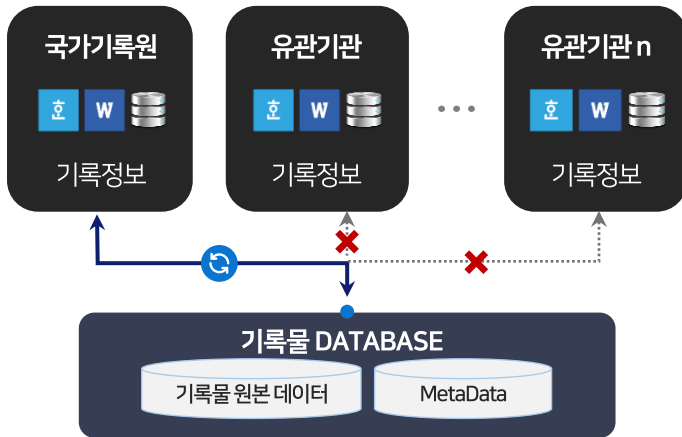
AI 기반 기록정보 통합 활용을 위한 정책·법제도적 근거 마련

기록물은 이제 단순 보존을 넘어, AI를 기반으로 융합·활용되는  
국가 지식자산으로의 전환이 필요한 시점입니다.

# 국민들이 원하는 기록정보를 한 곳에서 편리하고 쉽게 검색·열람 가능한 체계 마련 필요

## 현황및문제1

### 분산된 기록정보



- 기록정보 분산 관리·제공으로 인해 이용자가 원하는 기록정보를 검색하기 위해서는 **각 기관의 시스템을 일일이 방문**하는 불편함 발생
- 메타데이터 표준 불일치 등으로 **통합 관리 어려움**

기관 시스템 교체 없이 참여 가능한 **AI 에이전트** 기반 통합체계 구조 마련 필요

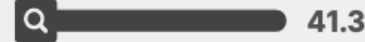
## 현황및문제2

### 시대에 맞지 않는 정보 검색 방식

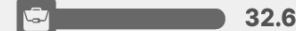
#### 생성형 인공지능 주요 이용 목적

\*출처: 2024년 한국미디어패널조사 (정보통신정책연구원, 2024)

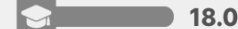
일상적인 정보검색을 위해



업무를 위해



과제 등 학업을 위해



취미 활동을 위해



시대에 맞지 않는 정보 검색 방식

사용자들은 기존의 키워드 중심 검색방식을 넘어 자신의 복잡한 질문 의도와 맥락을 이해하는 **대화형 검색방식**을 새로운 **표준으로 인식 및 기대**

기존의 키워드 검색 방식을 벗어나 **생성형 AI 대화 검색** 방식 필요

## 개선 방향

국민들이 원하는 기록정보!  
편리하고 쉽게 한 곳에서 검색·열람할 수 있는

범국가 기록정보  
통합·활용 체계  
모델 설계 필요

분산된 기록정보 통합

생성형 AI 대화형 검색

## 과제개요

## AI 에이전트 기반 범국가 기록정보 통합 활용체계 모델설계 및 실증

AI 에이전트를 활용해 국민과 공무원이 단일 창구에서 민관 기록정보를 통합 검색·열람할 수 있는 범국가 통합 활용체계 설계 및 실증

## 연구기간

계약일로부터 180일

## 연구 개발 내용



공공·민간 기록물관리기관이 참여 가능한 시 기반 통합활용서비스를 위해 연계기준(표준)과 운영체계를 설계하고 적용가능성을 검증

- ▶ 기관 간 데이터 연계 및 표준화 방안 - 연계기관의 메타데이터를 검토해 표준 항목 도출, 수집·동기화·자동 점검, 표준화·정제, 운영기준 정립
- ▶ 온톨로지(용어·관계 통합) 구축 방안 - 기관마다 다른 용어·분류·식별자 등을 하나로 통합 가능한 분류체계(안) 마련, 키워드가 아닌 맥락을 통한 탐색 방안 마련
- ▶ AI 에이전트 기반 통합 검색 및 열람서비스 방안 - 대화형 원스톱 검색 및 검색결과 출처 제시, 공개기준 및 개인정보 보호기준 내장
- ▶ 파일럿 구축 - 1~2개 기관과 파일럿을 운영하여 정확도, 응답시간, 만족도 등을 검토, 온보딩 키트 마련
- ▶ 통합 체계 구현을 위한 거버넌스 및 세부 방안 마련 - 법령 및 지침(가이드) 제개정 방향, 기관간 협력방안, 필수 H/W 및 S/W, 추진 로드맵 및 소요 예산 수립

## 02. 기존 통합 시도의 한계와 본 과업의 차별성



중앙집중형 통합 = 비용 부담 ↑ → 확산 한계



경량 연계 방식 = 비용 부담 ↓ → 확산 가능성 ↑

# 01. 기존 검색과 AI 에이전트 방식의 차이

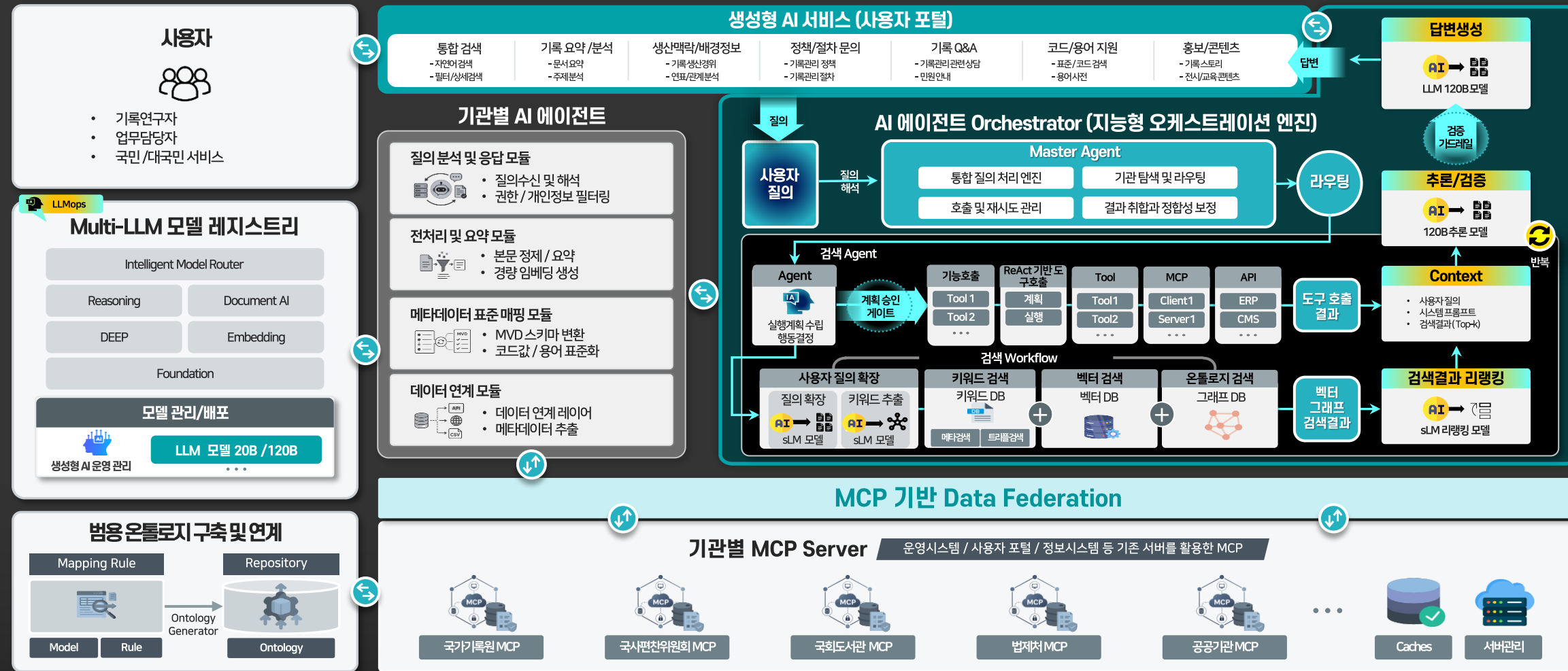


AI 에이전트는 단순 검색창이 아니라, **질문을 이해**하고 필요한 자료를 찾아 연결해 주는

**지능형 탐색 도구**

# 02. AI 에이전트 기반 통합 검색 및 열람서비스

## AI 에이전트 기반 통합 검색 및 열람서비스 목표 시스템 아키텍처



# 03. 질의 의도 정밀 분석 기반 AI 에이전트 서비스 모델 정의

## 수직적 통합: MCP 기반의 표준 인터페이스 구축

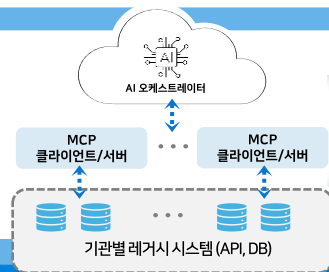
MCP는 중앙 AI 오케스트레이터가 각 기관의 레거시 시스템과 소통하기 위한 '수직적 통합의 표준'으로 사용

### 도입원칙 데이터 중앙 집중화 지양 및 분산형 연계 지향

- 독립성과 자율성 보장: 각 기관 내부 기록물과 시스템 통제권 유지
- 기술 표준화 기반 유연성 확보: 개방형 표준 프로토콜 도입

### 역할 MCP의 역할

- 기관별 상이한 구조 추상화
- 실시간 정보를 안전하게 가져오는 완충 역할



## 핵심 메시지

## 에이전트 서비스

MCP 도입 기반으로

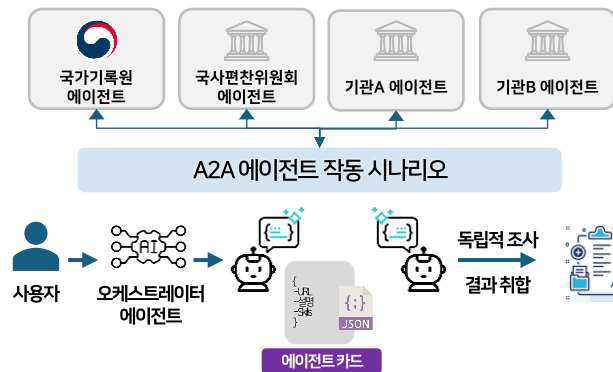
범정부 A2A 생태계 조성

✓ MCP 도입 일회성 개발이 아닌 표준 규격에 따른 유연한 연계 환경 제공

✓ 범정부 A2A 생태계 조성을 위한 초석 마련

## 수평적 통합: A2A 협업 생태계로의 진화

MCP를 통해 데이터 연계 표준이 정립되면, 각 기관의 자체 AI도입 시 자연스럽게 A2A 기반의 다중 에이전트 협업 체계로 전환 가능



# 01. [파일럿 시연 예시] AI 에이전트 검색 및 결과 제공 화면

**1** 최근대화

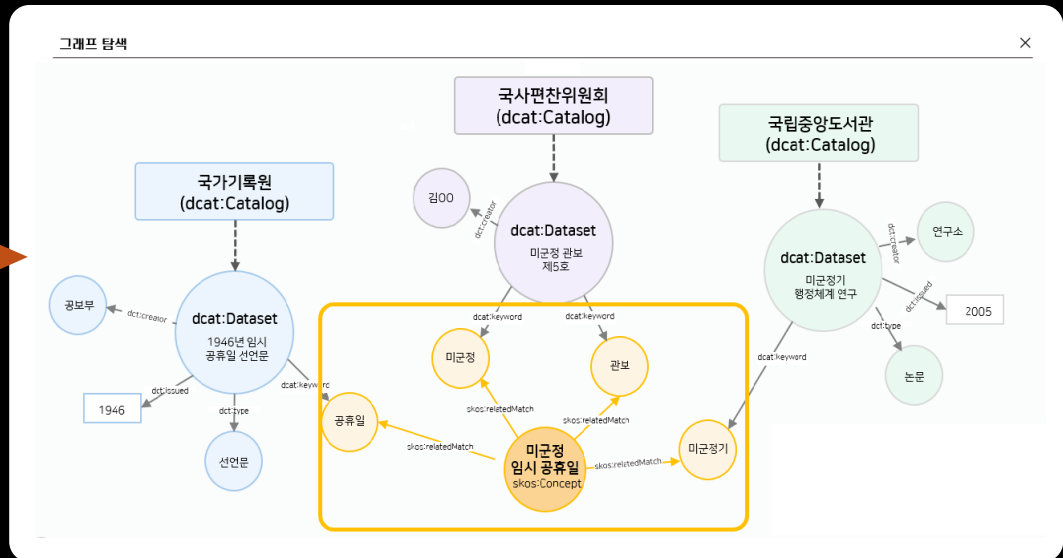
**2** 1946년 미군정 관보 제 5호에 실린 임시공휴일 선언 내용과 당시 생산 책임자가 누구인지 설명해주세요.

**3** 그래프 탐색

**4** 연결 링크

국가기록원 - 미군정 관보 및 행정기록  
국가데이터터치 - 공공기록 메타데이터

SAMPLE



## 사용자 질의에 대하여

- 1 과거 검색 이력을 통한 맥락 연속성 보장
- 2 사용자 의도를 이해하는 답변 제공
- 3 그래프 탐색 결과 제공
- 4 '연결 링크'로 답변의 근거 자료 제시

※ 본 화면은 연구개발용 파일럿 환경의 UI 구성안이며, 실제 구축 시 변경될 수 있습니다.



# 감사합니다

Saltlux