너무 이른 온톨로지.
지금은 메타데이터 RAG

공공부문 AI 신뢰 인프라 구축을 위한 현실적 접근
온톨로지 과잉 설계를 넘어
메타데이터 RAG 기반
실시간 현행화 · 출처 통제 · 정책 인텔리전스 (Policy Intelligence)

공공부문 AI 도입의 현주소

현재 활용 수준

- 단순 문서 요약
- 기초 민원 질의응답
- 초기 탐색 단계

목표 수준

- 복잡한 의사결정 지원
- 정책 분석 고도화
- 지능형 행정 구현

⚠ 핵심 장애요인: 데이터의 상태



- 파일명 규칙 미통합
- 버전 관리 미흡

정제된 데이터 부재
체계적 데이터베이스(DB) 미구축

AI

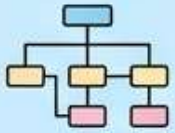
- AI 학습 제한
- AI 참조 한계
- 도입 확산 정체

공공 온톨로지 프로젝트의 4가지 현실적 어려움

기초 데이터 부재

- 엔티티(Entity) 정의 미흡
- 관계(Relationship) 구조 미정비
- 표준 용어 사전 부재
- 일원화된 데이터 구조 부족





행정 도메인 이해 장벽

헌법 · 법률 · 시행령 · 조례
규칙 · 행정 지침
위임 구조 (Delegation Structure)
책임 구조 (Accountability Structure)
외부 IT 중심 설계 한계
실무 활용도 저하



과잉 설계

- 하향식 설계 (Top-down)
- 전사 통합 설계
- 대규모 예산 투입
- 수년 단위 구축
- ROI 불확실
- 환경 변화 속도 > 구축 속도
- AI 기술 발전 속도 가속





AI 시대 부적합 구조

2009 빅데이터 아키텍처
정형 데이터 중심 (Structured Data)
문맥 이해 부족
의미 기반 처리 한계
거대언어모델(LLM) 비적합
AI 이전 시대 설계 관성

학술적 타당성 ≠ 현장 적용성　　이론 중심 설계 → 실무 단절

지금 공공 AI에 진정으로 필요한 것

신뢰 (Trust) • **책임** (Accountability)

무거운 시스템

전면적 온톨로지 구축
방대한 구조화, 장기, 고비용

현실적 접근

출처 통제 기반 AI
메타데이터 RAG (Metadata RAG)

정확히 찾는 것 (Accurate Retrieval)

- 많이 아는 것 ≠ 정확한 답변
- 공식 법령
- 공식 지침
- 명확한 출처(Source)
- 출성 기반 응답(Control)



실시간 현행화 (Real-time Update)

- 지침 개정
- 조례 변경
- 정책 수정
- 최신 데이터 반영
- 유연한 구조(Flexible Architecture)
- 정책 수정
- 구버전 차단



공공 웹사이트
국가법령정보시스템
→ 메타데이터 추출

검색 기반 생성

메타데이터 관리
(Metadata Layer)

검색 기반

메타데이터 RAG
(Retrieval-Augmented Generation)

- 출처 명시 응답
- 출처 가능 응답
- 책임 가능한 AI
- 행정 오류 최소화

온톨로지 보류 | 데이터 성숙도 고려 | 단계적 고도화 전략

공공 데이터 3층 구조와 스캐터엑스 AI 데이터 전달 체계
(3-Layer Public Data Architecture & AI Delivery Model)

1층

제도·근거 데이터
(Institutional & Legal Data)

법률, 시행령, 조례, 시행규칙, 행정 지침, 고시,
공식 홈페이지 안내문
비정형 텍스트 (Unstructured Text), 대외 공개 데이터

정확한 데이터 찾아주기
(Retrieval)

외부 신뢰 데이터 플랫폼,
국가법령정보시스템, 조문 검색,
조문 검색, 출처 기반 응답

2층

행정 운영 데이터
(Administrative Operational Data)

인허가 처리 내역, 민원 처리 이력, 세입·세출 예산 집행,
계약 관리 데이터, 관계형 데이터베이스(RDB),
정형 데이터 (Structured Data), 시스템별 상이한 구조

정확한 데이터 만들어주기
(Data Transformation)

ETL(추출·변환·적재),
데이터 카탈로그(Data Catalog)
데이터 정제, AI 이해 가능 구조

3층

현장·설비·센서 데이터
(Field & Sensor Data)

상수도 수압, 하수처리장 수질, 화재 감지 센서,
CCTV 메타데이터, 실시간 수치 데이터 (Real-time Data), 이기종 장비

정확하게 설명하기
(Contextual Explanation)

운영 매뉴얼, 정비 매뉴얼,
표준작업절차서(SOP), RAG(검색 증강 생성)
행정적 맥락 부여, 기술적 맥락 부여

제도 제정
(Legislation)

행정 집행
(Administration)

인프라 운영
(Infrastructure Operation)

스캐터엑스
AI 데이터 전달 체계
(AI Data Delivery Framework)

AI 성격 인지 기반 활용

계층별 맞춤 처리

유기적 연결
(Integrated Architecture)

데이터 무작위 통합 지양 | 계층별 특성 기반 설계 | 지자체 맞춤형 AI 도입 전략

메타데이터 RAG와 순도 높은 데이터 가공

일반 RAG

사용자 질문(Query)

↓

벡터 변환(Vectorization)

↓

벡터 데이터베이스(Vector DB)

↓

유사도 검색(Semantic Similarity)

↓

텍스트 조각(Chunk)

↓

LLM 응답 생성

위험 요소 (Risk Factors)

- ⚠ 의미 유사성 중심
- ⚠ 과거 폐기 지침 혼입
- ⚠ 타 지자체 조례 혼입
- ⚠ 비공식 출처 혼입
- ⚠ 행정적 유효성 구분 불가
- ⚠ 출처 위계 미반영
- ⚠ 환각(Hallucination) 위험

기술 중심 → 데이터 전달 체계 중심

- ✓ 신뢰성 (Trust)
- ✓ 감사 추적성 (Auditability)

원천 문서 + 메타데이터 결합(JSON)

- 발행 기관
- 담당 부서
- 문서 유형(조례·지침·보도자료)
- 시행일자
- 보안 등급

의미 기반 검색 + 메타데이터 통제
검색 범위 선제적 제한
행정 효력 중심 AI

범용 검색 AI ≠ 행정 신뢰 AI
메타데이터 기반 통제 아키텍처
공공 특화 RAG 전략

메타데이터 RAG

전달

↓

사용자 질문(Query)

↓

메타데이터 필터링(Metadata Filtering)

현재 시행 문서 | 우리 시 조례 | 최신 중앙부처 지침

↓

신뢰 구역(Trusted Search Space) 형성

↓

제한적 벡터 검색(Vector Search)

↓

재정렬(Re-ranking)

↓

LLM 전달

- 권한 외 문서 차단
- 폐기 문서 원천 배제
- 정확한 출처 명시(Source Citation)
- 시행일 명시
- 감사 대응 가능(Audit Ready)

메타데이터 RAG와 순도 높은 데이터 가공 체계
(Metadata RAG & High-Purity Data Pipeline)

신뢰성 (Trust) | **감사 추적성 (Auditability)** | **문맥 이해도 (Context Awareness)**

원천 데이터(Source Data)



법령



조례



지침


매뉴얼


공문서

4단계 순도 높은 데이터 가공 (4-Step Processing)

1단계: 보안 등급 분류 (Security Classification)

- 메타데이터 분석
- 내용 분석





- 공개 데이터 내부 민감 데이터
- 접근 권한 분리(Access Control)
- 데이터 풀(Pool) 분리
  - 대민 챗봇
  - 업무용 에이전트
- 정보 유출 방지

2단계: 부서 권한 및 도메인 분리 (Domain Segmentation)

- 소유 부서 태깅(Tagging)
- 업무 도메인 메타데이터



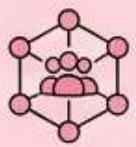
- 질문자 소속 인식



- 맥락(Context) 기반 검색
- 용어 의미 구분
  - 복지 지원금
  - 기업 지원금
- 맥락 혼선 차단

3단계: 핵심 업무 단위 (Clustering)

- 사업 단위 그룹화
- 상위 법령



- 지자체 조례
- 내부 시행 계획
- 관련 서식
- 정책 중심 묶음
- 통합 근거 구조

Garbage Data → 오류 증폭 | **High-Purity Data → 검색 정확도 극대화**

4단계: 공공 문서 구조 최적화 청킹 (Smart Chunking)

- 기계적 500자·1000자 분할 배제
- 문맥 소실 방지





- 법령 구조 인식 (장-절-조-항-호-목) 공문 구조 인식 (개요-목적-세부계획)
- 논리 단위 분할 상위 조문 제목 연결
- 문맥 꼬리표(Context Tag)
- 예외 규정 보존 단서 조항 보존

벡터 데이터베이스(Vector DB)

메타데이터 RAG

정확한 해석	단서 조항 인식
예외 규정 반영	출처 명시 응답
출서 명정 응답	감사 대응 가능

기존 공공 웹 구조를 활용한 메타데이터 RAG 전략

기술 선택 → 데이터 전달 구조 | 과잉 설계 배제 | 기존 정보 자산 재활용

Traditional Approach

신규 온톨로지 설계
(New Ontology)

- 지식그래프 구축
- 백지 상태 설계
- 대규모 예산
- 수년 단위 구축
- 구축 리스크 증가



텍스트 조각(Chunk) + 메뉴 경로 메타데이터

메타데이터 RAG

ScatterX Strategy

기존 정보 아키텍처 활용
(Reuse Architecture)

- 공식 홈페이지
- 국가법령정보센터
- 자치법규정보시스템
- 검증된 분류 체계
- 경량 온톨로지

구축 기간 단축 예산 절감
검증된 체계 활용

행정 흐름 일치 응답
업무 체계 기반 검색
논리적 답변 구조

공직자 친화적 구조 민원인 친숙 분류
행정 업무 흐름 일치 검색 신뢰성 강화

온톨로지 구축 없이, 기존 자산 최대 활용, AI 신뢰성·정확성 극대화

'실시간 현행화: 공공 특화 변경 감시 에이전트 (Watching Agent Architecture)

최신성 (Timeliness) | **무결성 (Integrity)** | **행정 신뢰성 (Administrative Trust)**

공공 데이터 소스

- 법령·조례
- 중앙부처 지침
- 공식 홈페이지
- 내부 행정망(새올 등)

변경 감시 에이전트 (Watching Agent)

자동 모니터링 → 변경 감지 → 변경 유형 분류

메타데이터 갱신

유효 기간 업데이트

검색 대상 재정렬

최신 규정 기반 AI 응답

고빈도 변동 데이터 (High-Frequency Data)

- 취업 공고
- 입찰 공고
- 새소식
- 내부 게시판
- 일단위 감지
- 실시간 감지
- 즉시 반영

이벤트성 변동 데이터 (Event-Based Data)

- 법령 개정
- 조례 제·개정
- 행정 지침 변경
- 공포일 기준 추적
- 고시일 기준 추적
- 이벤트 트리거 감시

저빈도 정적 데이터 (Low-Frequency Data)

- 조직도
- 기관 기본 정책
- 정합성 점검
- 장기 주기 모니터링
- 주기적 검증

과거 규정 참조 차단 | **변경 사항 자동 추적** | **실시간 현행화 기반 공공 AI**

구조 기반 정밀 변경 감지 기술

최신성 | 무결성 | 실질적 의미 변화 감지

비효율적 방식

문서 전체 비교



- 디자인 변경
- 오탈자 수정
- 사소한 수정에도 전체 재처리
- 서버 부하 증가
- 불필요한 재색인



조례 문서

조례 문서
- 제1조 → 각 조문별 해시(Hash) 값 생성
  Article-Level Hash
- 제2조
- 제3조
- 제4조
- 제5조 → 제5조 제2항 해시 값 변경 감지
  - 제2항

변경 발생

일부 개정 or 전부 개정

변경 조문만 업데이트
Selective Update

메타데이터 RAG 재색인

스캐터엑스 방식

문서 논리 구조 인식

- 장-절-조-항-호-목 구조
- 조문 단위 관리



공포일 자동 추출
시행일 자동 추출
유효 기간 갱신
Effective Date Control



현재 효력 문서만 검색 허용
Valid Documents Only

전체 재처리 배제 | 조문 단위 정밀 추적 | 현재 시점 효력 중심 통제 | 행정 사고 예방

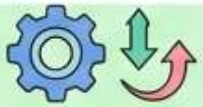
자동 현행화가 가져오는 실무적 가치 (Operational Value of Auto-Update)

최신성 (Timeliness) **무결성** (Integrity) **지속적 신뢰성** (Sustainable Trust)



변경 감시 에이전트
(Watching Agent)

자동 변경 추적
자동 메타데이터 갱신
자동 RAG 반영



운영 효율성 향상
(Operational Efficiency)

수작업 최소화
수동 재학습 불필요
일일 점검 업무 제거
관리자 업무 경감
운영 비용 절감



감사 대응 체계 강화
(Audit Readiness)

변경 시점 기록
AI 반영 시점 기록
자동 로그(Log) 생성
이력 추적 가능(Traceability)
근거 기반 행정 증명



적법성 객관적 증명
(Legal Accountability)

특정 시점 기준 규정 확인
당시 유효 규정 기반 처리
행정 사고 예방
책임 행정 구현

수동 관리 체계

- 상시 점검 부담
- 누락 위험 존재

자동 현행화 체계

- 실시간 반영
- 누락 원천 차단
- 온톨로지 재구축 불필요

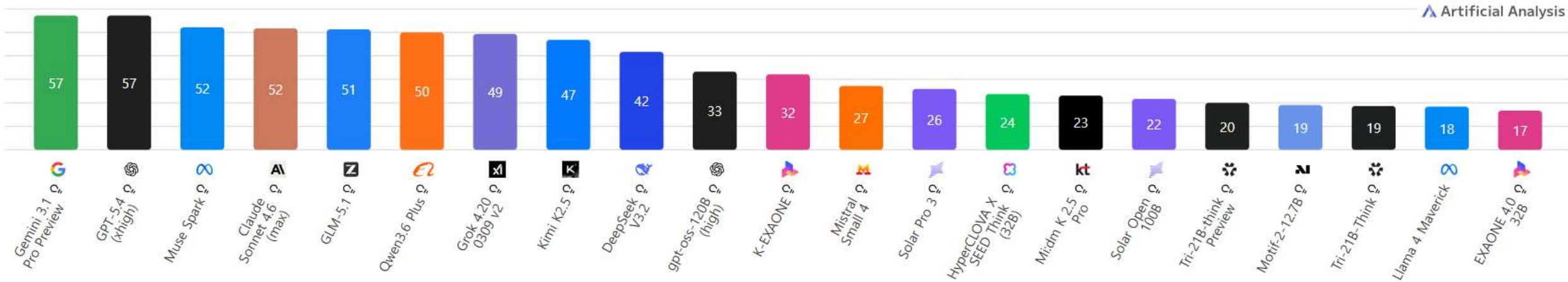
복잡한 재구축 없이, 지속적 신뢰 유지, 공공 특화 AI 운영 표준 모델

Artificial Analysis Intelligence Index

Artificial Analysis Intelligence Index v4.0 incorporates 10 evaluations: GDPval-AA, τ²-Bench Telecom, Terminal-Bench Hard, SciCode, AA-LCR, AA-Omniscience, IFBench, Humanity's Last Exam, GPQA Diamond, CritPt

21 of 473 models

+ Add model from specific provider



Reasoning models are indicated by a lightbulb icon.

Artificial Analysis Intelligence Index

Artificial Analysis Intelligence Index v4.0 includes: GDPval-AA, τ²-Bench Telecom, Terminal-Bench Hard, SciCode, AA-LCR, AA-Omniscience, IFBench, Humanity's Last Exam, GPQA Diamond, CritPt. See Intelligence Index methodology for further details, including a breakdown of each evaluation and how we run them.

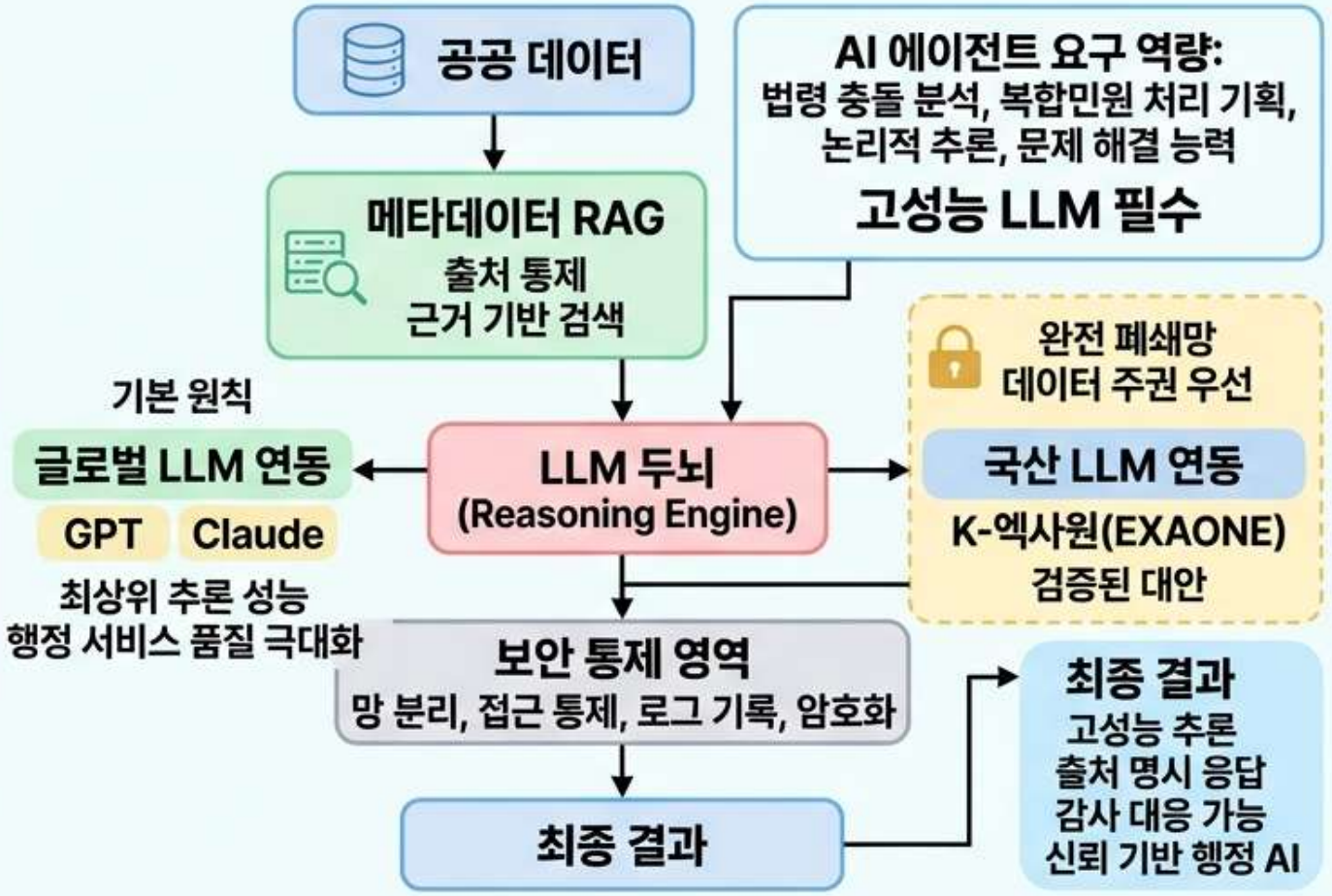
공공 AI 에이전트의 두뇌: 글로벌 LLM 도입 전략

(Global LLM Strategy for Public AI Agent)

메타데이터 RAG + 고성능 LLM | 복잡한 행정 추론 | 국가 보안 규정 준수

Artificial Analysis Intelligence Index v4.0

상위권 50~57점	제미나이(Gemini) GPT 클로드(Claude)	글로벌 최상위 모델
중위권 32점	K-엑사원(EXAONE)	국내 1위 글로벌 오픈 웨이트 7위
하위권 20점대 초중반	하이퍼클로바X 믿음(Mi:dm) 솔라(Solar)	
18점	라마 4(Llama 4) 범정부 공통 기반 검토 모델	



선택이 아닌 성능 기준 | 객관적 지표 기반 결정 | 유연한 LLM 연동 아키텍처

국정원 AI 보안 가이드 2.0 준수 아키텍처

(AI Security Guideline 2.0 Compliance)

국정원 가이드 2.0 핵심 요구사항

- 데이터 중요도 분류
- 민감도 등급화
- 망분리 일변도 지양
- 안전한 활용 권장

- 업무망 ↔ 인터넷망 연계 통제
- 개인정보 식별
- 비공개 자료 식별
- 마스킹(Masking)
- API 구간 암호화(Encryption)
- 접근 통제(Access Control)

데이터 등급별 통제 | **다층 보안 체계 (Multi-layer Security)** | **안전한 글로벌 LLM 활용**

기관 내부망 (폐쇄망, Intranet Zone)

공공 원천 데이터 | 행정 시스템 | 내부 결재 정보 | 주민등록번호 | 연락처

보안 전처리 구역 (Security Pre-processing Zone)

데이터 등급 판별 | 민감 정보 탐지 | 개인정보 마스킹 | 비공개 정보 차단 | 로그 기록(Log) | 전송 데이터 최소화

외부 연계 구간(API Secure Channel)

암호화 통신(TLS) | 접근 키 관리 | 전송 이력 기록

글로벌 LLM 또는 K-엑사원

지능 제공(Reasoning Only) | 원천 데이터 저장 없음 | 추론 결과 반환

공개 데이터 처리

- 공개 법령
- 공식 홈페이지 자료
- 대외 공개 문서

메타데이터 RAG

보안 검증 후 LLM 전달

경함과 민감 데이터 외부 반출 차단 | 지능만 활용 | 데이터는 내부 보관 | 감사 대응 가능 | 국정원 가이드 2.0 충실 이행

글로벌 성능 + 국내 보안 기준 준수 | **하이브리드 보안 아키텍처** | **공공 특화 안전한 AI 운영 모델**

RTX 4090 기반 내부망 전처리 아키텍처
(On-premise Pre-processing Architecture)

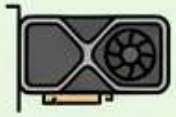
국정원 AI 보안 가이드 2.0 준수 | **고비용 GPU 대체** | **합리적 예산·고효율 구조**

고비용 GPU 랙 서버



- H100
- A100
- 수십억 원 예산
- 구축 복잡성
- 유지보수 부담

RTX 4090 워크스테이션 1장



- On-premise
- 합리적 도입 비용
- 경량 인프라
- 보안 전처리 전용 서버

기관 내부망 (폐쇄망)

원천 비정형 데이터

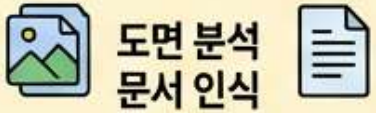


RTX 4090 서버

데이터 관문(Gateway)
데이터 정수기(Filter)
외부망 완전 차단

멀티모달 데이터 내부망 텍스트화
(Multimodal to Text)

AI OCR(광학문자인식)
이미지 → 텍스트 변환



STT(Speech-to-Text)
음성 → 텍스트 변환
회의 녹취 변환



- 보안 위반 차단
- 민감 데이터 외부 반출 금지
- 내부망 내 실시간 처리
- 고속 연산 RTX 4090

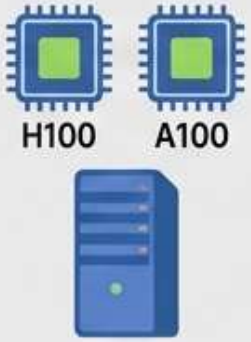
내부망앙 완전 차단

고비용 인프라 불필요 | **RTX 4090 1장으로 보안 전처리 완결** | **공공 특화 안전한 AI 연계 기반**

RTX 4090 기반 내부망 DLP·비식별화 아키텍처 (On-premise DLP & De-identification)

국정원 AI 보안 가이드 2.0 준수 | 고비용 GPU 대체 | RTX 4090 1장 기반 효율 구조

고비용 GPU 랙 서버



고비용 GPU 랙 서버

- 수십억 원 예산
- 구축 복잡성
- 운영 부담

RTX 4090 워크스테이션



RTX 4090 워크스테이션

- On-premise
- 합리적 예산
- 보안 전처리 전용
- 경량 고효율 파이프라인

기관 내부망 텍덱! Gateway 서버

기관 내부망 텍스트 데이터

OCR/STT 변환 완료 데이터

RTX 4090 보안 서버
데이터 관문(Gateway)
데이터 정수기(Filter)

외부 전송 데이터 = 비식별 데이터만
원본 데이터 외부 반출 0

RTX 4090 1장으로 실시간 처리
지연 최소화(Low Latency)

입력 텍스트
만원 내용
내부 보고 문서
업무 기록

sLLM 기반 보안 분석 (Security sLLM)
NLP 분석
문맥 이해(Context Awareness)
개인정보 탐지(PII Detection)
민감 정보 탐지(Sensitive Data Detection)

탐지 대상 예시
실명
주민등록번호
주소
전화번호
직책
내부 기밀 정보

완전 비식별화(De-identification)
마스킹(Masking)
홍길동 → OOO
010-1234-5678 → 010-****-****
주민등록번호 → *****-******

순도 높은 익명 텍스트
익명 데이터
(Anonymized Data)

문맥 기반 정밀 탐지 | 완벽한 비식별화 통제 | 글로벌 LLM 연계를 위한 안전한 관문

RTX 4090 기반 안전 연동·재식별화 아키텍처 (Secure Hybrid AI Architecture)

원천 데이터 외부 반출 0 | 지능만 활용(Reasoning Only) | 비용·성능·보안 최적 균형

내부망(폐쇄망, Intranet Zone)

- OCR/STT 완료 데이터
- 완전 마스킹 데이터
- 순도 높은 익명 데이터(Anonymized Data)

RTX 4090 전처리 서버
- 비식별화 완료
- 전송 데이터 최소화
- 관문(Gateway)

암호화 통신 구간
(Encrypted API Channel)
- TLS 암호화
- 접근 키 통제
- 전송 로그 기록

외부 LLM 영역

글로벌 LLM 동역

GPT | Claude or K-엑사원(EXAONE)

입력 데이터
- 익명 텍스트
- 메타데이터
- RAG 법령 근거

출력
- 추론 결과
- 보고서 초안
- 해결 방안

Reasoning Engine Only
원천 데이터 저장 없음

고비용 GPU 랙 서버
- H100
- A100
- 수십억 원

VS

RTX 4090 1장
- On-premise
- 합리적 예산
- 경량 보안 인프라

내부망 재유입

암호화 응답 수신

재식별화(Re-identification)
- 마스킹 정보 매핑 복원
- 원본 데이터 내부 결합

실무 공무원 화면 출력
- 최종 보고서
- 행정 처리 결과

글로벌 1위 수준 AI 지능 활용 | 국정원 보안 가이드 2.0 충족 | 데이터 외부 유출 0 | 공공 특화 하이브리드 AI 표준 모델

메타데이터 RAG 기반 공공 AI 에이전트

A 그룹: 근거·정책 검색형 에이전트 (Policy & Evidence Agents)

메타데이터 RAG | 변경 감시 에이전트 | LLM 추론(Reasoning) | 출처 통제(Source Control)

정책·조례 변화 감시 에이전트



사용 데이터
- 자치법규정보시스템
- 국가법령정보센터
- 중앙부처 고시
- 조문 단위 해시(Hash)

핵심 기능
- 상위 법령 제·개정 감지
- 메타데이터 즉시 갱신
- 현행 조례와 충돌 탐지
- 담당자 알림(Notification)

기대 효과
- 행정 오류 예방
- 입법 공백 최소화
- 법령 모니터링 자동화

근거 자동 첨부 보고서 에이전트



입력

검색 범위
- 법령
- 지침
- 과거 결재 문서
- 보도자료

메타데이터
- 시행일
- 소관 부서
- 현행 조문

출력
- 문단별 근거 자동 매핑
- URL 표시
- 조항 번호 주석

기대 효과
- 자료 조사 시간 단축
- 보고서 신뢰도 향상
- 의사결정 지원 강화

감사 대응 에이전트



입력
- 사업명
- 특정 기간

검색 기준
- 시행일
- 폐지일
- 조례 연혁
- 과거 지침
- 예산집행 내역

출력
- 과거시점 유효 법령
- 시계열 정리
- 소명 자료 초안

기대 효과
- 감사 준비 행정력 절감
- 적법성 객관적 입증
- 심리적 부담 완화

정책 충돌 감지 에이전트



입력
- 정책 기획안
- 조례 제·개정안

교차 검색
- 중앙부처 법령
- 타 지자체 조례
- 행정심판 재결례

LLM 분석
- 조항 간 의미 충돌
- 절차적 누락 위험
- 검토 의견서 자동 생성

기대 효과
- 법적 분쟁 예방
- 심사 기간 단축
- 정책 완성도 향상

법령 변화 실시간 감지 → 근거 자동 연결 → 과거 시점 소명 지원 → 정책 충돌 사전 차단

검색 보조 도구 → 실질적 업무 에이전트

공공 특화 지능형 행정 지원 체계

메타데이터 RAG 기반 공공 AI 에이전트
B 그룹: 민원·행정 처리형 에이전트 (Civil Service Agents)

메타데이터 RAG | RTX 4090 내부망 전처리 | DLP 비식별화 | LLM 추론(Reasoning)

국민신문고 민원 사전 분석 에이전트



입력 데이터	민원 텍스트, 수기 문서 사진, 현장 영상, 통화 녹음
내부망 처리	AI OCR, STT(Speech-to-Text), DLP 마스킹
검색 범위	관련 조례, 유사 민원 사례, 과거 처리 이력
출력	민원 요약, 적용 법규, 담당 부서 추천, 사전 분석 리포트
기대 효과	초기 검토 시간 단축, 이관 지연 최소화, 처리 형평성 강화

민원 대응 근거 에이전트



입력	민원 질의 내용
검색 데이터	공식 홈페이지 FAQ, 업무 매뉴얼, 민원 응답 이력, 현행 조례
메타데이터 필터	대국민 공개 가능, 최신 규정, 현행 시행 조문
출력	정중한 답변 초안, 조례 링크, 매뉴얼 출처 명시
기대 효과	반복 민원 처리 시간 단축, 오안내 예방, 대외 신뢰도 향상

인허가 자동 검토 보조 에이전트



입력	신청서, 첨부 도면, 스캔 문서
내부망 처리	AI OCR 텍스트 추출, 필수 서류 누락 점검
법규 대조	건축 조례, 지구단위계획 지침, 업종 제한 사항
출력	요건 충족 여부, 제한 사항 표시, 검토 보조 명세서
기대 효과	처리 기간 단축, Human Error 심층 검토 집중 가능

비정형 데이터 자동 처리 | 근거 기반 민원 대응 | 요건 검토 자동화 | 행정 효율성·형평성 동시 강화

검색 보조 도구 → 실무 처리형 AI 에이전트

메타데이터 RAG 기반 공공 AI 에이전트
C 그룹: 회의·업무 자동화형 에이전트 (Automation Agents)

개단 준조 | RTX 4090 내부망 처리 | DLP 비식별화 | 메타데이터 RAG | LLM 추론(Reasoning)

내부 회의 자동 비서 에이전트

입력 데이터
- 회의 녹음 파일
- 위원회 회의
- 부서 협의
- 주간 업무 보고

내부 네부망 처리 믈료

음성 파일 → 온프레미스 STT(Speech-to-Text) → AI DLP 마스킹 민감 정보 비식별화 → LLM 요약·분석

출력 결과
- 회의 요약
- 주요 결정 사항(Key Decisions)
- 부서별 후속 조치(Action Item)
- 문서 양식 자동 정리

마이크→체크리스트

기대 효과
- 회의록 작성 시간 절감
- 보안 규정 준수
- 부서 간 정보 공유 강화
- 실행력 향상

시장·군수 인사말 자동 생성 에이전트

입력 정보
- 행사 개요
- 참석자 정보
- 강조 키워드

검색 데이터
- 최근 보도자료
- 시정 백서
- 관련 조례 현황
- 지역 통계

메타데이터 RAG 검색
- 현행 정책 맥락 반영
- 역점 사업 연계

LLM 생성 결과
- 3~5분 연설문 초안
- 격식체
- 친근체

어조 조절
연단 마이크 + 문서 생성

기대 효과
- 연설문 작성 시간 단축
- 시정 메시지 일관성 확보
- 대외 신뢰도 향상
- 최신 통계 정확 반영
- 대외 신뢰도 향상

회의 자동 문서화 | 정책 맥락 기반 연설문 생성 | 보안 준수 + 업무 자동화 | 공공 행정 생산성 혁신

메타데이터 RAG 기반 공공 AI 에이전트

D 그룹: 정책 분석·인텔리전스형 에이전트 (Policy Intelligence Agents)

정형 데이터 분석(Structured Analytics) → 메타데이터 RAG → LLM 추론(Reasoning) → 근거 기반 정책 설계

생활인구 기반 정책 설계 에이전트



입력 데이터	• 통신사 유동 인구 데이터 • 시간대별 카드 매출 • 타 지자체 우수 정책 사례 • 연구 보고서
분석 흐름	• 인구 밀집 시간대 분석 • 소비 패턴 도출 • 유사 조건 지역 검색 • 상권 활성화 조례 매칭 • 공공시설 입지 사례 추천
출력	• 정책 옵션 제안 • 입지 후보지 도출 • 사업 우선순위 제안
기대 효과	• 데이터 기반 정책 설계 • 청년 유입 전략 고도화 • 예산 효율적 배분

경제 유발 효과 분석 에이전트



입력 데이터	• 행사 전후 매출 변화 • 대중교통 이용량 • 중앙부처 평가 지침 • 과거 행사 보고서
분석	• 증감률 계산 • 직접 경제 효과 추정 • 지침 기반 논리 구조 매핑
출력	• 행사 성과 분석 리포트 • 경제 유발 효과 수치화 • 보고서 초안 자동 생성
기대 효과	• 외부 용역 비용 절감 • 신속한 성과 보고 • 예산 확보 근거 강화

포트홀·재난 예방 에이전트



입력 데이터	• 위험 신고 위치 • 강수량 • 기온 데이터 • 도로 포장 이력 • 유지보수 매뉴얼 • 재난 대응 지침
분석	• 신고 빈도 분석 • 사고 발생 확률 예측 • 취약 구간 도출
메타데이터 RAG	• 표준 정비 절차 검색 • 안전 조치 가이드 제시
출력	• 위험도 우선순위 목록 • 정비 권고안 • 예산 배분 기준
기대 효과	• 사고 예방 • 시민 안전 강화 • 예산 집행 투명성 확보

온톨로지 vs 메타데이터 RAG
공공부문 도입 적합성 비교

비용(Cost), 기간(Duration), 하드웨어(Hardware), 정확도·보안(Security & Accuracy)

	온톨로지(Ontology)	메타데이터 RAG(Metadata RAG)
비용(Cost)	✓ 신규 개념 체계 설계 ✓ 도메인 전문가 투입 ✓ 데이터 재분류 ✓ 대규모 컨설팅	✓ 기존 홈페이지 메뉴 재활용 ✓ 법령 사이트 계층 구조 활용 ✓ 정보 아키텍처 흡수 ✓ 정보 아키텍처 흡수 ✓ 재설계 최소화 ✓ 예산 효율화
기간(Duration)	✓ 1년 이상 장기 프로젝트 ✓ 관계 정의·검증 반복 ✓ 법령 변경 시 재설계	✓ 수개월 내 구축 ✓ 메타데이터 태깅 ✓ 검색 엔진 연동 ✓ 검색 ✓ 변경 감시 에이전트 ✓ 실시간 현행화
하드웨어 요구사항 (Hardware Requirements)	✓ 대규모 연산 구조 ✓ 내부망 LLM 자체 구동, H100 ✓ 고사양 GPU 랙 서버 ✓ 인프라 관리 부담	✓ RTX 4090 워크스테이션 ✓ OCR ✓ STT ✓ 비식별화 전처리 ✓ 외부 LLM 하이브리드 연동 ✓ 도입 부담 최소화
정확도 및 보안 (Security & Accuracy)	✓ Hallucination 위험 ✓ 과거·현재 데이터 혼합 가능성 ✓ 출처 자동 통제 한계	✓ 시행일 기준 필터링 ✓ 발행 기관 기준 제한 ✓ 검색 범위 물리적 통제 ✓ 출처 명시 응답 ✓ 감사 추적성(Audit Trail) 확보

이상적 기술 ≠ 현장 적합성 | 현실적 예산·인력·데이터 수준 고려 | 공공 도입 1단계 최적 해법 | 메타데이터 RAG 중심 전략

메타데이터 RAG 도입의 5대 기대 효과 (5 Key Benefits)

예산 , 신뢰 , 업무 , 보안 , 미래

예산
(Budget Efficiency)



과잉 투자 방지
온톨로지 대규모 구축 유보
오버스펙 하드웨어 지양
기존 공공 웹 재활용
RTX 4090 기반 전처리
종량제 LLM API
ROI 최적화 매몰 비용 최소화

신뢰
(Trust & Auditability)



출처 기반 답변
공식 문서 한정 검색
조례·법령 링크 명시
시행일 필터링
감사 대응 자료 활용
행정 투명성 강화

메타데이터 RAG
12종 AI 에이전트
공공 행정 혁신

업무
(Work Productivity)



실시간 현행화
변경 감시 에이전트
민원 사전 분석
회의록 자동 작성
보고서 근거 자동 첨부
반복 업무 경감
기획 업무 집중

미래
(Future Readiness)



메타데이터 체계적 태깅
순도 높은 데이터 축적
노이즈 제거
검증된 데이터 자산화
고도화 AI 인프라 기반
온톨로지 확장 가능
지능형 행정 플랫폼 준비

보안
(Security & Compliance)



국정원 AI 보안 가이드 2.0 준수
내부망 RTX 4090 처리 비식별화 데이터만 전송
OCR AI DLP 마스킹 정보 유출 구조적 차단

IT 교체 → 행정 방식 혁신 | **현재 효율화 + 미래 기반 구축** | **안전하고 지속 가능한 공공 AI 전환**

정책 인텔리전스를 향한 단계적 진화 로드맵

무리한 도약 지양 → 안전하고 안정적인 계단식 진화→ 신뢰 기반 AI 전환

정책 인텔리전스
지능형 행정 플랫폼
데이터 기반 의사결정

Step 1



메타데이터 RAG 기반 구축
(Trust Foundation)

- 출처 통제 검색 환경
- 공식 문서 한정 검색
- 시행일·발행 기관 필터
- 내부망 전처리 아키텍처
- RTX 4090, DLP 비식별화
- 실시간 현행화
- 변경 감시 에이전트
- 시스템 신뢰 확보
- 보안 기반 확립
- 행정 오류 최소화

Step 2



12종 AI 에이전트 실무 투입
(Operational Expansion)

- 민원 분석
- 보고서 자동 근거 첨부
- 회의록 자동화
- 정책 설계 보조
- 인허가 검토 보조
- 일상 업무 효용 체감
- 반복 업무 경감
- 정제된 행정 데이터 축적
- 고품질 데이터 자산화

Step 3



상향식 온톨로지 자동 생성
(Bottom-up Ontology)

- 순도 높은 데이터 기반
- AI 개념 자동 추출
- 관계 자동 연결
- 지식그래프 자동 구성
- 현장 맥락 반영
- 하향식 설계 배제
- 실제 행정 데이터 기반 구조
- 정확도·실용성 극대화

데이터 없는 하향식 설계, 과잉 구조화 → 데이터 축적 기반 상향식 진화, 현실적·단계적 AI 고도화