그림 23 | 내부망 전용 AI시스템 개념도

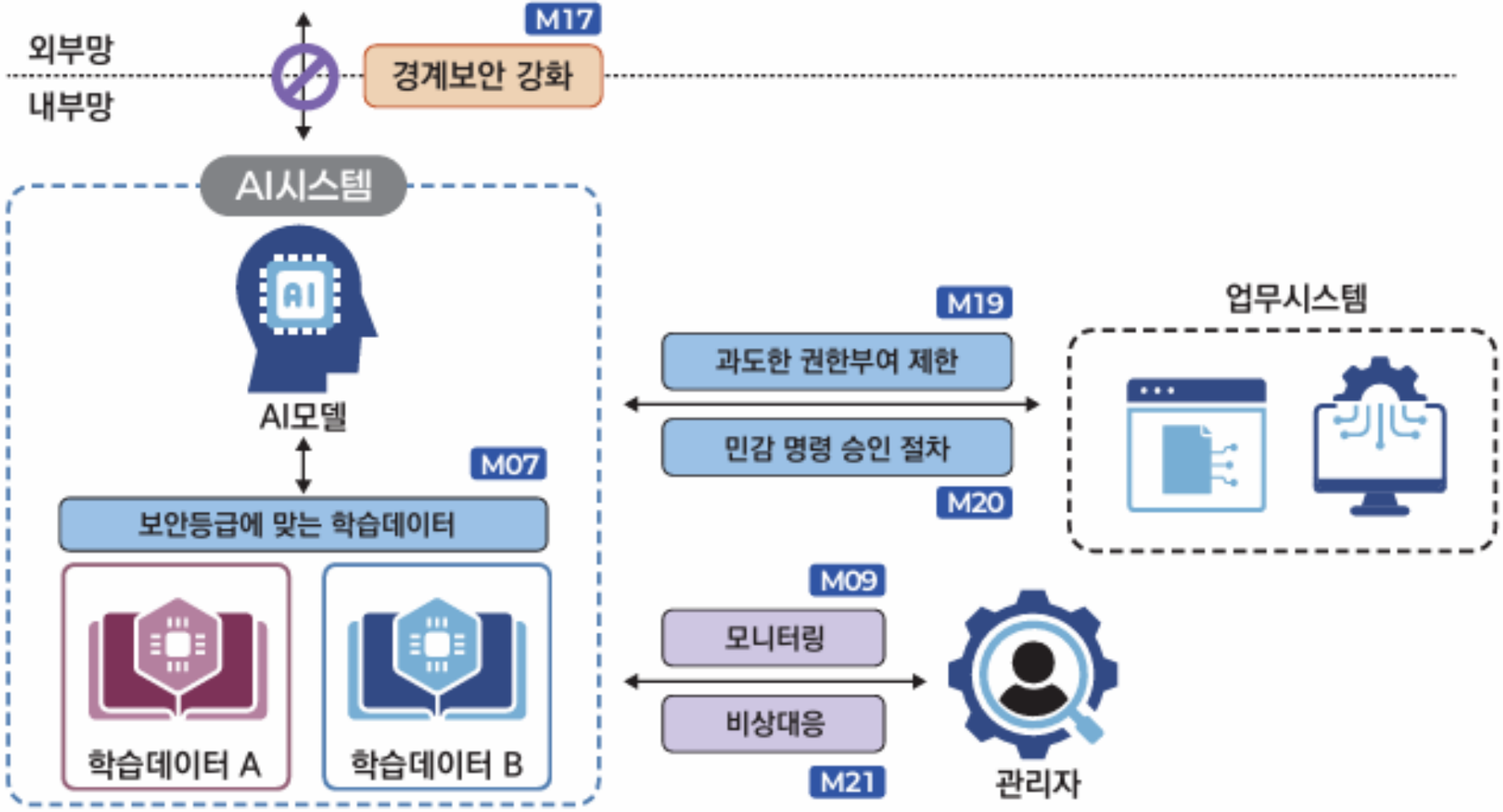


그림 26 내부업무용 AI시스템의 외부망 연계 개념도

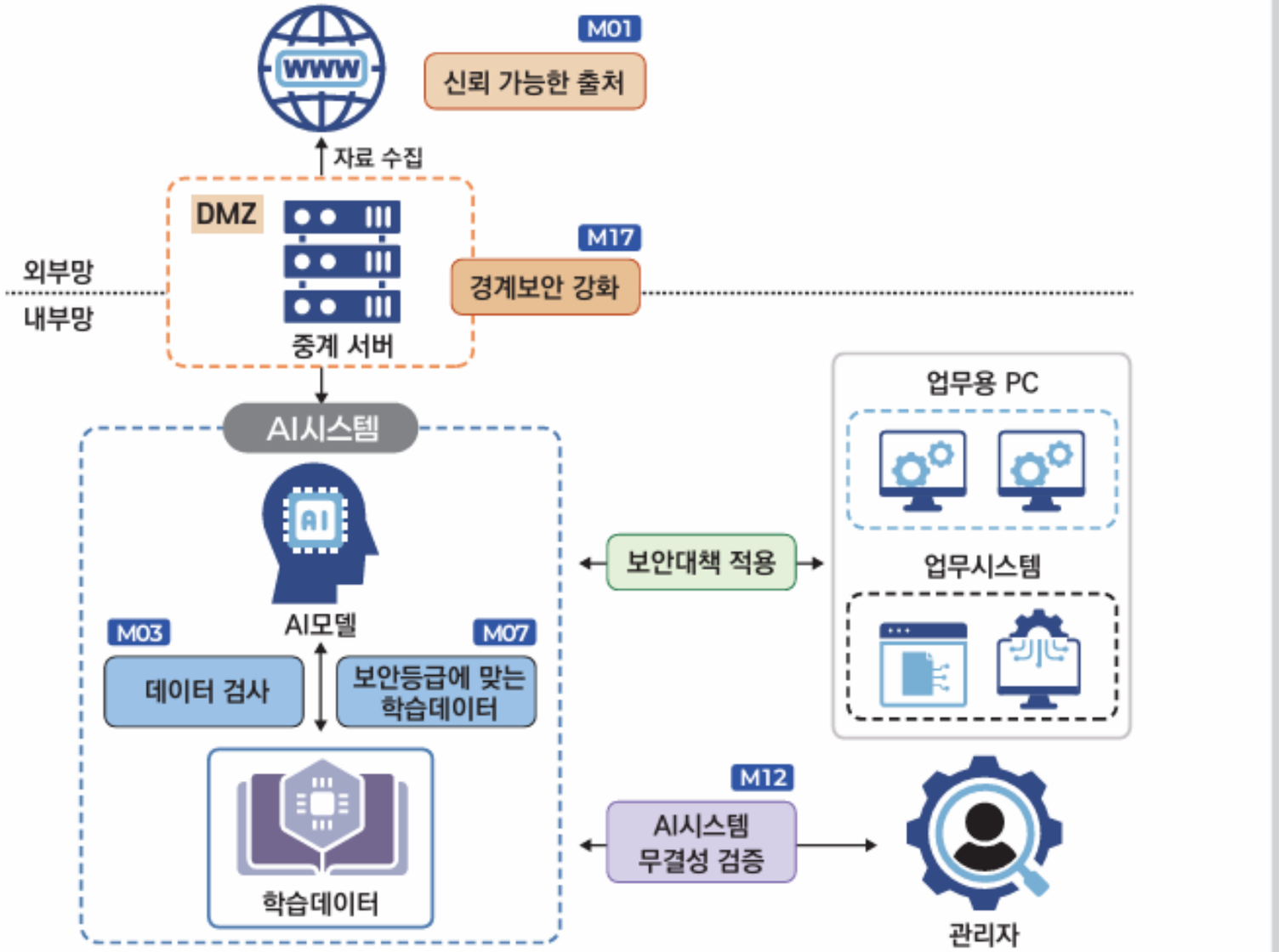


그림 28 대민서비스용 AI시스템의 내부망 연계 개념도

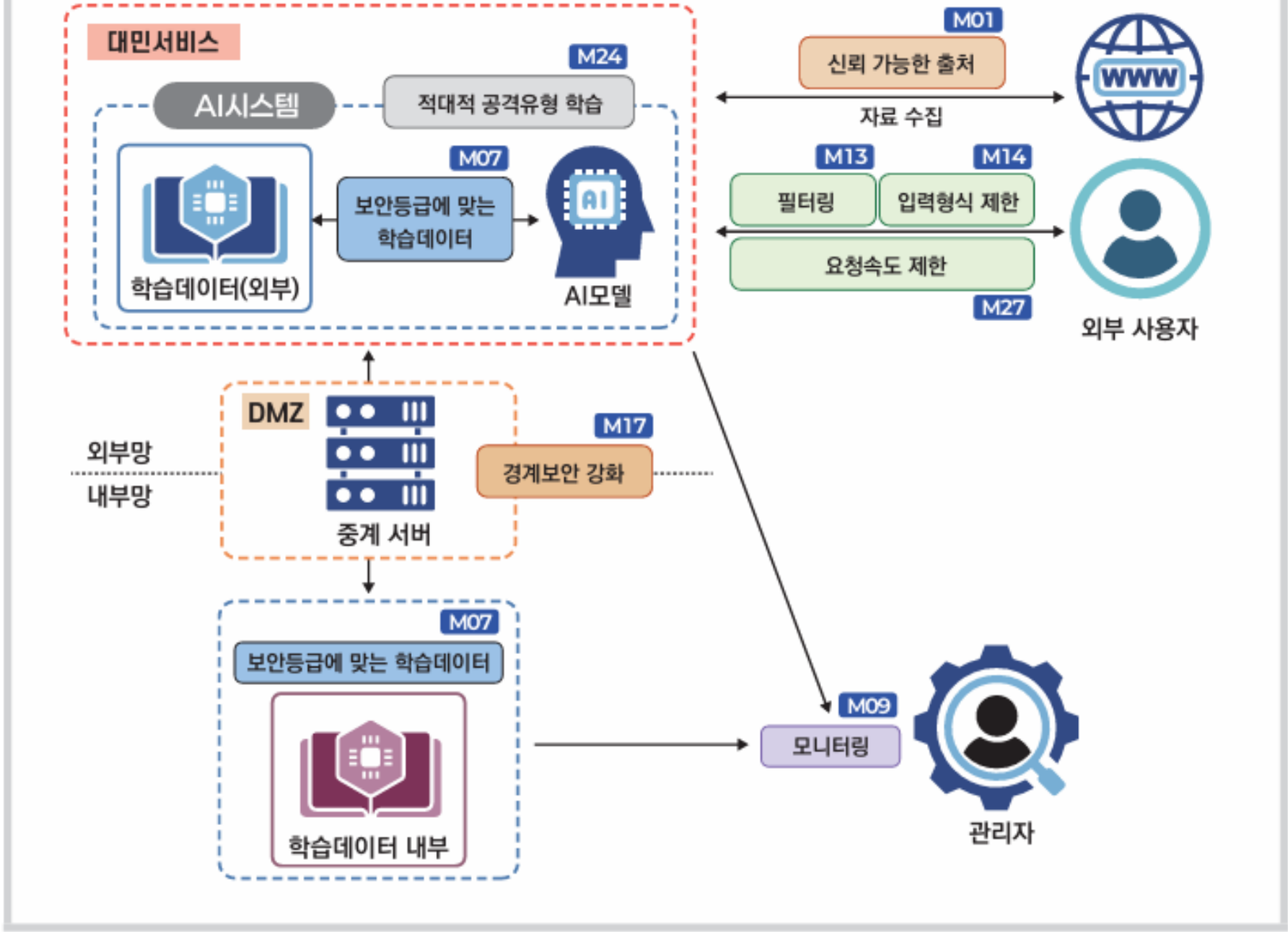


그림 30 상용 AI서비스 활용 개념도

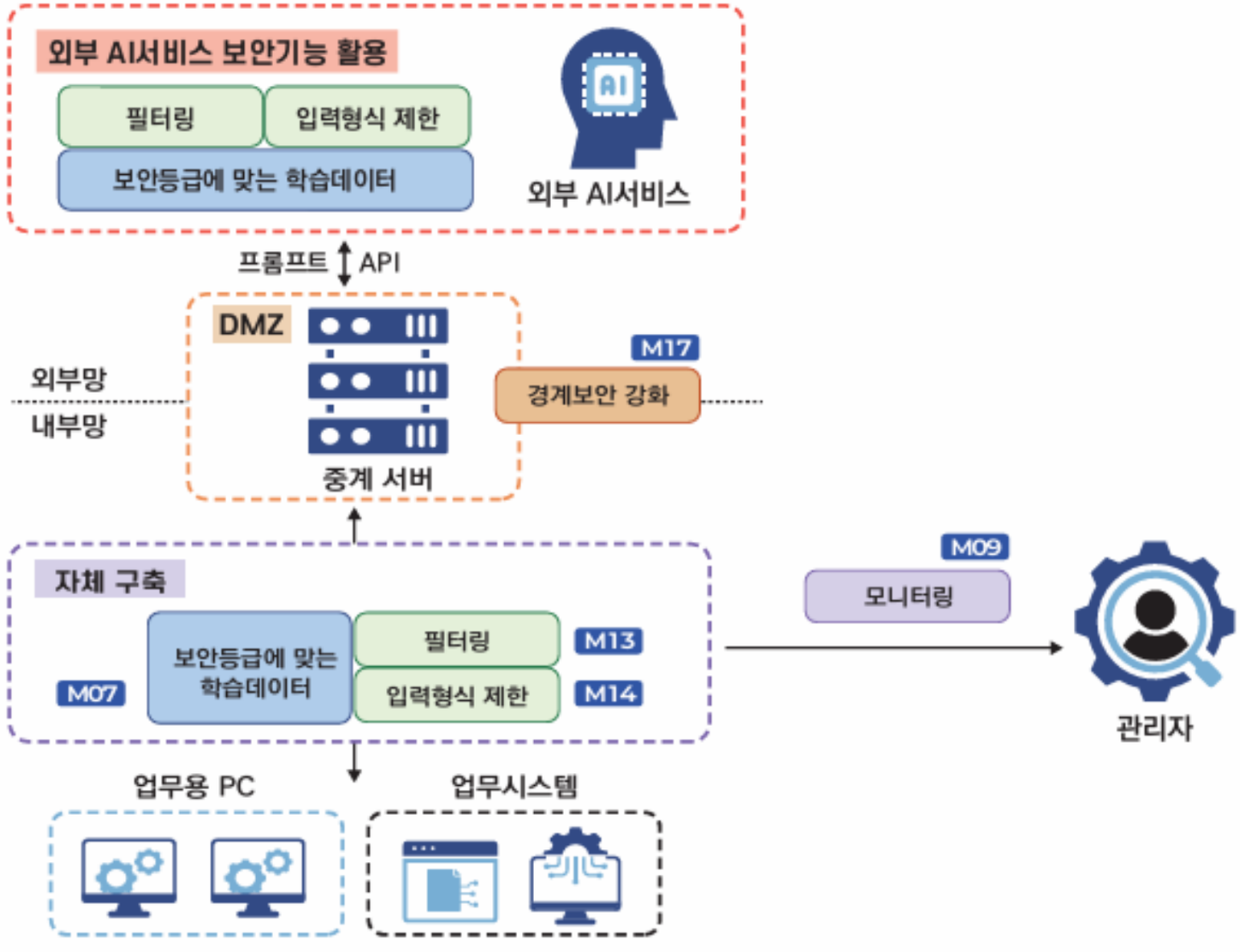
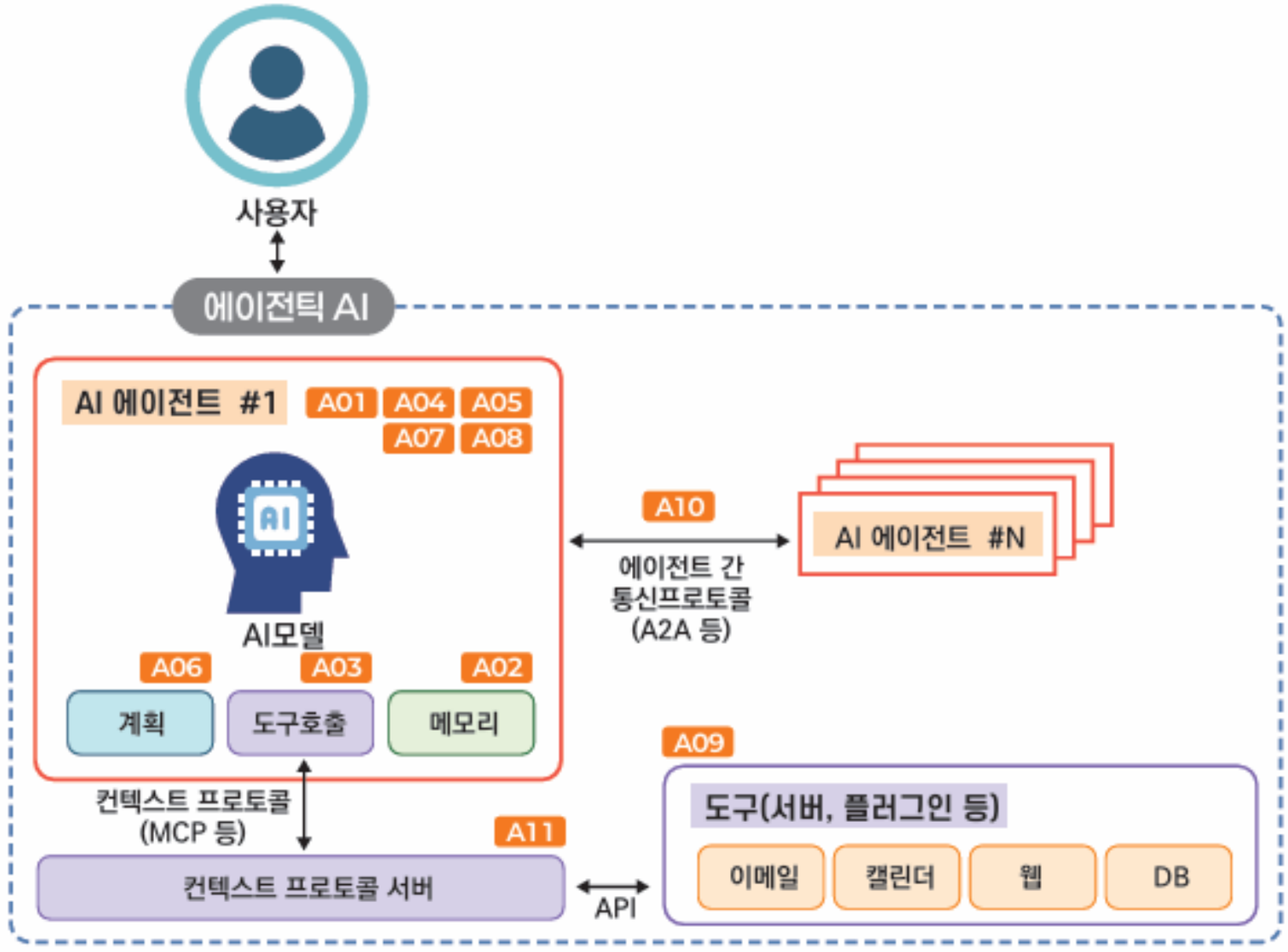


그림 32 | 에이전틱 AI 개념도



[참고] 범정부 AI 공통기반의 개념·제공 서비스·구축 방법

개요 범정부 AI 공통기반은 AI 모델·학습데이터·GPU를 중앙·지방정부가 행정 내부망에서 공동 활용하는 정부 전용 AI 플랫폼

■ 개념과 추진 배경

개념	추진 배경
AI 모델·데이터·개발 환경을 공통 제공하는 행정망 전용 AI 인프라	민간 AI는 보안 우려로 인터넷망에서만 쓸 수 있어 행정 업무 활용에 제약

■ 제공 서비스(제품 종류)

구분	제공 내용	특징
AI 플랫폼	삼성SDS·네이버클라우드 2종	내부망에서 쓰는 AI 챗서비스 제공
AI 모델	LLM·멀티모달 등 약 30종	국산 4개사 모델 탑재(아래 참고)
RAG	RAG API 7종(PaaS 형태)	법령·행정자료와 연계해 답변 근거 제시
GPU	공통기반이 GPU 일괄 관리	기관별 GPU 서버 구매·임차 불필요
연계 서비스	지능형 업무관리 플랫폼	메일·메신저·영상회의를 AI로 연계

■ 기관별 AI 서비스 구축 방법(5단계)

1 기획	AI 서비스 정의, 공통기반 활용 구현방안 검토
2 예산	비용 항목 확인, 공통기반 비용 시뮬레이션
3 계약	공통기반 이용 계약 체결
4 구축	RAG 우선 전략으로 서비스 구현
5 운영	성능 모니터링, 데이터 지속 업데이트

■ 구축 시 참고사항

- 구축 환경: 국가정보자원관리원 대구센터 PPP존
- '26.8.28 시행 법률에 공통기반 우선 이용 조항 반영
- 2차 사업(66억원)으로 AI 게이트웨이·외부망 연계 구축 추진

■ 탑재 국산 모델과 특징

업스테이지 솔라	삼성전자 삼성 LLM	네이버 HCX-GOV-Think	LG AI연구원 K-엑사원
한국어·문서 처리 강점, 근거 제시 삼성SDS 패브릭스로 제공	삼성전자 자체 개발 모델 삼성SDS 패브릭스로 제공	하이퍼클로바X 계열 자체 모델 클로바 스튜디오 포 거브로 제공	추론형 LLM 엑사원 계열 클로바 스튜디오 포 거브로 제공

빅테크 대비 성능 AAII 지수('26.8) 빅테크 61~63점 vs 국산 독자 모델 31~47점 · 정부 목표는 글로벌 최신 모델의 95%

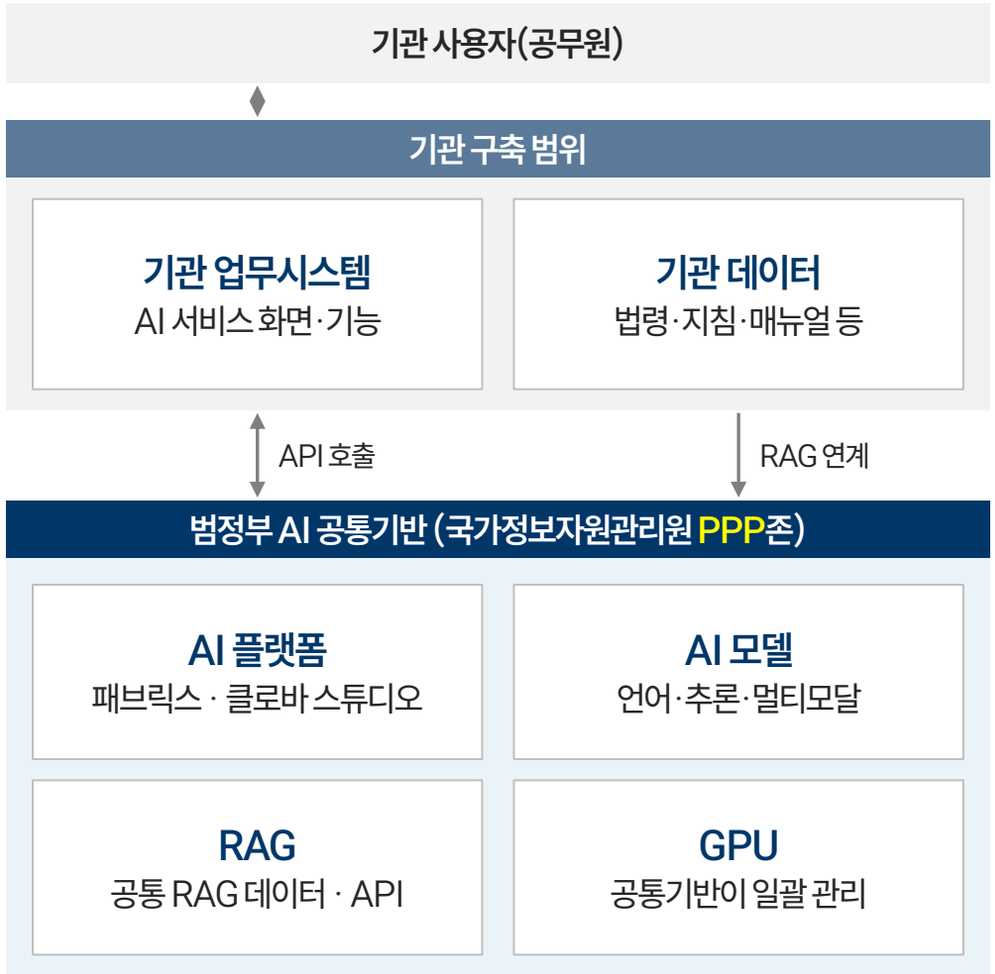
효과 중복 개발·투자 없이 내부 행정업무부터 공공서비스까지 AI를 업무 전반에 활용

※ 출처: 정책브리핑('25.11), NIA「공공부문 AI 도입·활용 가이드」('26.5), 전자신문·디지털데일리·연합뉴스('26.6), ZDNet Korea('26.8), 아이뉴스24('26.10), 서울경제('26.11), 이투데이('26.8)

[참고] 범정부 AI 공통기반 활용 AI 서비스 구축 개념도·아키텍처

개요 기관은 서비스와 데이터 연계를 구축하고, AI 플랫폼·모델·GPU는 공통기반을 이용

■ 개념도



■ 아키텍처(계층 구성)

■ 기관 구축 ■ 공통기반 제공

사용자	공무원이 업무시스템 화면에서 AI 기능 이용
AI 서비스	기관 업무 특화 AI 기능 (예: 법령안 검토 지원)
연계	모델·데이터 전처리 API, AI 게이트웨이(2차 사업)
AI 플랫폼	삼성SDS 패브릭스, 네이버 클로바 스튜디오 포 거브
AI 모델	한글 특화 언어·추론·멀티모달, 오픈소스 모델
데이터·RAG	공통 RAG 데이터(법령·지침·보고서) + 기관 데이터
인프라	GPU, 국가정보자원관리원 PPP존 클라우드

■ 구축 시 참고사항

- AI 게이트웨이·표준 API는 2차 사업('26.8 발주)에서 구축 추진
- 기관 보유 모델도 윤리성·안전성·신뢰성 평가 후 탑재 가능

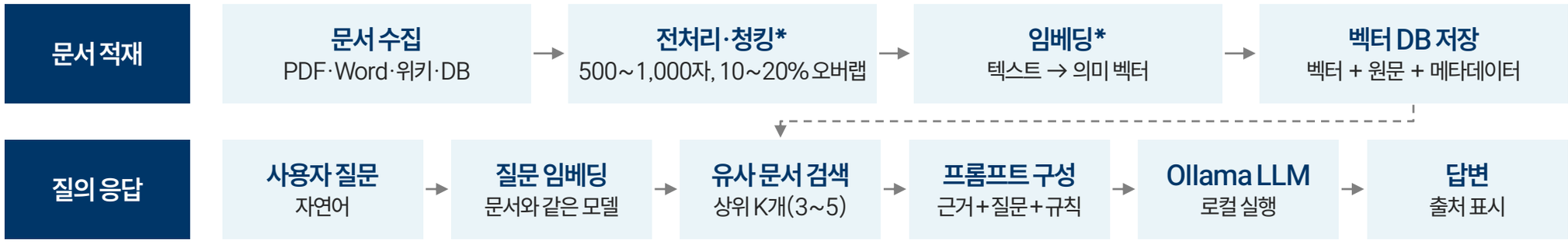
효과 플랫폼·모델·GPU 중복 투자 없이 기관은 업무 특화 기능과 데이터 연계에 집중

※ 공개 자료를 바탕으로 재구성한 개념 예시(공식 구성도 아님) · 출처: 정책브리핑('25.10), 디지털데일리('25.11), ZDNet Korea('26.8)

[참고] 로컬 LLM 기반 RAG 구축 개념도

개요 사내 문서를 벡터로 검색한 근거로 로컬 LLM(Ollama)이 답변, 문서·질문이 외부로 나가지 않음

■ 전체 구조



■ 구축 순서

① 수집·정제	표·스캔 PDF는 OCR 필요, 메타데이터 저장
② 청킹	500~1,000자, 조항·문단 경계 기준
③ 임베딩	다국어 모델, 문서·질문 동일 모델
④ 벡터 DB	소규모 Chroma·FAISS, 운영 Qdrant·pgvector
⑤ Ollama	사양에 맞는 모델 준비 (경량 3B~7B급)
⑥ 검색·답변	근거 없으면 확인할 수 없다고 답변

■ 품질을 높이는 방법

하이브리드	벡터 + 키워드(BM25) 검색 결합
리랭킹	상위 10~20개 재평가 후 3~5개 전달
출처 표시	답변에 문서명·페이지·링크 노출
권한 필터	부서·권한에 맞는 문서만 조회
평가셋	실제 질문 20~50개로 반복 측정
문서 갱신	변경 시 청크 삭제·재임베딩

■ 흔한 실패 원인

- 표·스캔 PDF를 단순 텍스트로 추출
- 청크 경계가 문단·조항을 무시
- 질문·문서 임베딩 모델 불일치
- 검색 결과 없이 LLM이 답을 생성
- 권한 없는 문서 검색으로 유출 위험

■ 권장 시작점

Ollama + Chroma + Python + 다국어 임베딩

문서 20~50개 규모로 검증한 뒤 Spring Boot, Qdrant/pgvector 구조로 확장

효과 문서와 질문을 외부로 보내지 않고 사내 문서 근거로 답변하는 지식검색 구현

※ 출처: 「로컬 LLM 기반 RAG 구축 가이드」(첨부 문서) 본문

[참고] RAG 전처리·청킹 개념도

개요 원문을 정리하고 메타데이터를 추출한 뒤, 검색하기 좋은 작은 단위(청크)로 나누는 단계

처리 흐름

원문 문서 PDF·Word·위키·DB → **전처리** 정리·메타데이터 추출 → **청킹** 500~1,000자 단위 분할 → **청크 + 메타데이터** 검색 단위 → **임베딩** 다음 단계

전처리

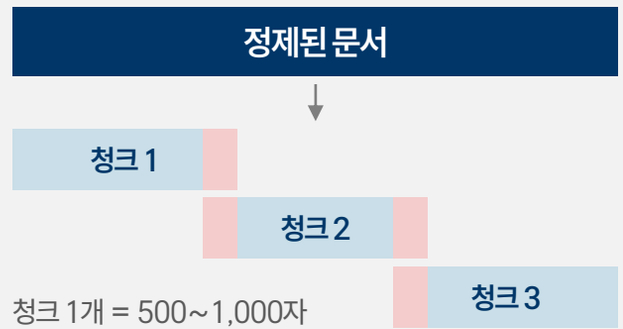
텍스트 추출	PDF·Word·HTML 등에서 본문 추출
OCR·레이아웃	표·스캔 PDF는 OCR 또는 레이아웃 인식 파서 필요
정리	표·이미지·중복·빈 줄 정리

메타데이터 (청크와 함께 저장)

문서 제목	원본 경로·URL
페이지 번호	작성일·개정일
부서·시스템·문서 유형	접근 권한

접근 권한은 검색 단계에서 사용자 부서·권한에 맞는 문서만 조회하는 필터로 사용

청킹 개념도



청크 크기

너무 크면	검색 정확도 저하
너무 작으면	문맥이 끊김

청킹 기준

크기	500~1,000자에서 시작
오버랩	10~20%
경계	제목·표·문단 경계 존중
규정 문서	조항 단위로 분할
API 문서	엔드포인트 단위로 분할

흔한 실패 원인

- 표·스캔 PDF를 단순 텍스트로만 추출해 정보 손실
- 청크 경계가 문단·조항을 무시해 문맥이 끊김

시사점 RAG 답변 품질은 LLM 자체보다 **검색 품질과 청킹**에 더 크게 좌우되는 경우가 많음

※ 출처: 「로컬 LLM 기반 RAG 구축 가이드」(첨부 문서) 구성 요소, 구축 순서 1~2. 청킹 그림은 오버랩 개념을 도식화한 것으로 비율은 설명용

[참고] RAG 임베딩 개념도

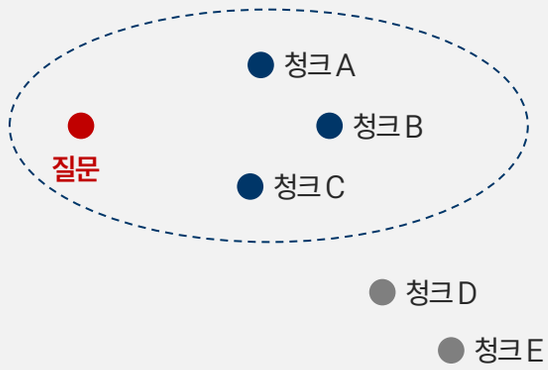
개요 청크와 질문을 같은 임베딩 모델로 의미 벡터로 바꿔, 벡터 DB에서 유사한 청크를 찾게 하는 단계

처리 흐름

문서 적재	**청크 + 메타데이터** 청킹 결과	**임베딩 모델** **청크를 벡터화**	**의미 벡터** 숫자 배열	**벡터 DB 저장** 벡터·원문 청크·메타데이터
질의 검색	**사용자 질문** 자연어 문장	**질문 임베딩** 문서와 동일한 모델	**유사도 검색** 상위 K개(3~5)	**관련 청크** 프롬프트에 전달

벡터 공간 예시

2차원으로 단순화한 설명용 그림



가까운 벡터 = 의미가 비슷한 청크
상위 K개(3~5)를 검색 결과로 반환

임베딩 모델

원칙	질문과 문서에 동일한 모델 사용
한국어	다국어 임베딩 모델을 우선 검토
예시	BGE-M3, multilingual-e5, sentence-transformers

검색 품질 보강

하이브리드	벡터 검색에 키워드 검색(BM25) 결합
리랭킹	상위 10~20개를 재평가해 3~5개만 LLM에 전달

벡터 DB 저장

벡터	원문 청크	메타데이터

소규모	Chroma, FAISS로 시작
운영	Qdrant, pgvector, MongoDB Vector Search

주의사항

- 질문 임베딩과 문서 임베딩에 다른 모델을 쓰는 것이 흔한 실패 원인
- 문서가 바뀌면 기존 청크를 삭제하거나 재임베딩

효과 키워드가 달라도 **의미가 비슷한 청크**를 찾아 LLM 답변의 근거로 제공

[참고] 로컬 LLM(Ollama)과 클라우드 LLM 비교

개요 Ollama는 모델이 아닌 로컬 실행 도구이며, 로컬 성능은 모델 크기와 서버 사양에 좌우

■ 항목별 비교 (7B~30B급 모델을 사내 서버에서 구동하는 경우의 정성 비교)

항목	Ollama 기반 로컬 LLM	빅테크 클라우드 LLM
복잡한 추론·에이전트	열세, 7B~14B급 소형일수록 격차 큼	우세, 어려운 작업일수록 격차 확대
요약·번역·문서 Q&A	실사용에 충분한 수준	우수하나 체감 차이는 작음
응답 속도	하드웨어에 좌우, VRAM에 모두 올라가야 빠름	대체로 빠르고 일정, 네트워크·호출 한도 영향
컨텍스트 길이	모델과 메모리 한도 안에서 제한적	최신 모델은 100만 토큰급 지원
데이터 보안	문서와 질문이 사내에 머묾	문서와 질문이 외부 서버로 전송
비용 구조	토큰 과금 없음, 하드웨어·운영 비용 부담	토큰 사용량에 따라 과금
오프라인 사용	가능	불가

■ 권장 기술 스택

빠른 프로토타입	Ollama + Chroma + Python + sentence-transformers
Spring Boot 서비스	Ollama + pgvector/Qdrant + Spring AI
대량·운영 서비스	문서 파서 + Qdrant/pgvector + Ollama/vLLM + 모니터링
완전 오프라인	Ollama + FAISS/Chroma + 로컬 임베딩 모델

■ 공개 벤치마크 동향 (첨부 문서 인용, 2026년)

- 시간 격차: 최상위 공개 모델이 약 4개월 뒤따름
- 종합 지수: 집계에 따라 3점~약 20점 차
- 소형 모델: HumanEval 로컬 7B 45~55%, GPT-5.6 90%
- 속도: GPU 전부 적재 약 140토큰/초, CPU 오프로딩 약 12.6토큰/초

시사점 RAG는 근거 요약 중심이라 로컬 모델이 비교적 강한 영역, 모델은 사내 평가셋으로 직접 측정

※ 출처: 「로컬 LLM 기반 RAG 구축 가이드」(첨부 문서) 권장 기술 스택, 부록. 벤치마크 수치는 문서가 인용한 공개 자료(Epoch AI, Artificial Analysis 등) 기준

[참고] 「공공부문 AI 도입·활용 가이드」 주요 내용

개요 행정안전부가 범정부 AI 공통기반으로 AI 서비스를 구축하는 전 과정을 5단계로 표준화한 실무 가이드를 배포

■ 배포 배경

법률 개정 대응	공통기반의 목적
'26.8.28 개정 시행 법률의 공통기반 우선 이용 조항에 선제 대응	기관별 중복 구축에 따른 예산 낭비와 기술 파편화 해결

■ 가이드 개요

구분	내용	세부
배포	행정안전부, '26.6.10 발표	전국 공공기관 대상
성격	실행 안내서	현장 담당자가 실무에 즉시 적용
범위	5단계 표준화	공통기반 활용 서비스 개발의 전 과정
기술	RAG 우선 전략	설계부터 적용까지의 과정 수록
설명회	6.10, 6.12 양일	중앙부처·지방정부·민간 AI 사업자 대상

■ 서비스 구축 5단계 표준 절차

1 기획	2 예산	3 계약	4 구축	5 운영

■ RAG(검색 증강 생성) 우선 전략

방식	기관 내부의 최신 문서(지식 데이터)를 먼저 찾고 그 내용을 바탕으로 답변
효과	환각 현상을 막고 행정 데이터에 기반한 정확한 답변 제공
수록	설계부터 적용까지의 과정을 가이드에 상세히 수록

■ 설명회와 현장 소통

- 가이드 핵심 내용 설명과 공통기반 활용 구축 사례 시연
- 질의응답으로 실무자 궁금증 해소
- 현장 의견을 수렴해 공통기반 이용체계 보완에 반영

기대 효과 예산 낭비 없이 국민이 체감하는 고품질 AI 서비스를 더 빠르게 실현

※ 출처: 행정안전부 보도자료 「공공 AI 구축, 더 쉽고 빠르게 「공공부문 AI 도입·활용 가이드」 배포」('26.6.10, 공공인공지능혁신과)

생성형 AI는 확률로 문장을 만들므로 답변 검증이 필요합니다

인공지능(AI) 활용 가이드 핵심 개념 ① · 인공지능(AI) 개념 이해

인공지능(AI)

인간의 지능적 활동을 모방하고 지원하기 위해 개발된 컴퓨터 시스템과 그 구현 기술

예: 콘텐츠 추천 · 자율주행 · 자연어 처리

생성형 AI

텍스트·오디오·이미지·동영상·프로그래밍 코드 등 새로운 콘텐츠를 만들어 내는 AI

예: 텍스트·이미지·동영상 생성, SW 개발

텍스트 생성 원리

다음 단어 예측: 입력과 가장 자연스럽게 이어질 확률이 높은 단어를 선택

예: 나는 → 너를 → 사랑해

생성형 AI의 한계

환각(Hallucination) 사실이 아니라 자연스러운 문장이 될 확률로 답변을 생성하므로, 존재하지 않는 정보를 사실처럼 답할 수 있음. 완벽하게 제거하는 기술적 해결책은 현재 없음

환각을 줄이는 프롬프트 작성

- 사용할 지식의 범위를 명확하게 명시
- 충분한 맥락(Context) 정보 제공
- 구체적이고 명확한 지시, 단계별 분석 요청
- 출처를 명시하고, 불확실하면 추가정보를 요청하도록 지시

AI 업무 활용 원칙

- 사실 기반 과제와 창의적 과제를 구분
- AI는 보조 도구로만 활용
- 중요한 정보는 뉴스·논문 등 외부 소스로 검증
- 의사결정 사항은 반드시 추가 검증 (교차검증, 출처 확인)

출처: 인사혁신처, 업무 효율성 향상을 위한 인공지능(AI) 활용 가이드, Ⅰ. 인공지능(AI) 개념 이해

프롬프트는 8개 구성요소와 5개 작성 원칙으로 설계합니다

인공지능(AI) 활용 가이드 핵심 개념 ② · 프롬프트 작성(Prompting)

프롬프트 AI가 원하는 작업을 수행할 수 있도록 제공하는 질문, 명령 또는 설명. 좋은 프롬프트는 좋은 결과를, 나쁜 프롬프트는 나쁜 결과를 가져옴

프롬프트 구성요소

Instruction (지시)
AI가 수행해야 하는 작업·행동

Context (문맥)
작업에 필요한 배경·상황 정보

Persona (역할)
AI가 수행해야 할 역할

Example (예시)
원하는 결과물의 샘플

Format (출력 형식)
생성해야 하는 출력의 형식

Style (스타일)
답변의 구조·구성·표현방식

Tone (톤)
태도·감정·뉘앙스

Input Data (입력 데이터)
첨부파일 등 작업대상 데이터

작성 기본원칙

① **구체적이고 명확하게 지시**

② **충분한 맥락(Context) 정보 제공**

③ **지시는 간결하게**

④ **지속적·반복적으로 프롬프트 개선**

⑤ **복잡한 작업은 단계별 접근**

출처: 인사혁신처, 업무 효율성 향상을 위한 인공지능(AI) 활용 가이드, Ⅱ. 프롬프트 작성(Prompting)

8개 구성요소는 지시부터 입력 데이터까지 역할이 나뉩니다

프롬프트 구성요소 상세 · 요소별 역할과 예시

구성요소	역할	예시
Instruction (지시)	AI가 수행해야 하는 작업 또는 행동	다음 텍스트를 번역해 주세요
Context (문맥)	작업에 필요한 배경이나 상황 정보	대상은 신규직원입니다
Persona (역할)	AI가 수행해야 할 역할	당신은 데이터 분석 전문가입니다
Example (예시)	원하는 결과물의 샘플	'행복해요' : 긍정, '기분 나빠요' : 부정
Format (출력 형식)	AI가 생성해야 하는 출력의 형식	결과를 표 형태로 보여주세요
Style (스타일)	답변의 구조, 구성, 표현방식	대화형으로 작성해 주세요
Tone (톤)	결과에서 느껴지는 태도나 감정, 뉘앙스	정중하고 친절한 톤으로 작성해 주세요
Input Data (입력 데이터)	분석·요약·번역 등을 위해 제공되는 데이터	이미지, 업로드한 PDF, 소스코드 등

출처: 인사혁신처, 업무 효율성 향상을 위한 인공지능(AI) 활용 가이드, Ⅱ-3. 프롬프트 구성요소

프롬프트는 단계별로 작성하고 반복해서 개선합니다

프롬프트 작성 단계와 작성 기본원칙 상세

작성 단계

문제 정의 → 해결 방향 설정 → 초기 프롬프트 작성 → 반복 개선

작성 기본원칙

① 구체적이고 명확하게 지시	• 모호성이 발생하지 않도록 작업 지시는 구체적이고 명확하게 • 모호하면 AI가 의도를 파악하지 못해 원하는 답변을 주지 못함
② 충분한 맥락(Context) 제공	• 목적, 목표, 배경, 최종 독자, 사례, 제약 조건 등을 제공 • 충분한 정보를 제공하되 불필요한 정보는 제외
③ 지시는 간결하게	• 복잡한 프롬프트는 AI를 혼란스럽게 하므로 가능한 단순하게 • 간결한 문장 구조와 명확한 지시어로 작게 나누어 제공
④ 지속적·반복적으로 개선	• AI 모델이 업그레이드되면 기존 프롬프트 품질이 저하될 수 있음 • 반복적인 테스트, 분석, 조정으로 기존 프롬프트를 개선
⑤ 복잡한 작업은 단계별 접근	• 한 번에 모든 것을 요구하지 않고 작업을 단계별로 나눔 • 예: 브레인스토밍 → 개요작성 → 섹션별 작성 → 결론 도출

출처: 인사혁신처, 업무 효율성 향상을 위한 인공지능(AI) 활용 가이드, Ⅱ-4. 프롬프트 작성 단계와 작성 기본원칙

패턴·기법·프레임워크로 프롬프트를 구조화해 업무에 씁니다

인공지능(AI) 활용 가이드 핵심 개념 ③ · 프롬프트 작성 방법과 업무 활용

패턴 (Pattern)

15개 소개

자주 사용되는 표준화된 문장 구조

- **페르소나** AI에게 역할·전문성 부여
- **청중 페르소나** 최종 독자의 특성에 맞춘 출력
- **그 외** 템플릿, 레시피, 사실 확인 리스트 등

기법 (Technique)

12개 소개

AI의 사고·행동방식을 전략적으로 제어

- **프롬프트 체이닝** 작업을 서브 작업으로 나눠 연결
- **N-Shot** 예시를 제공해 의도를 전달
- **생각의 사슬(CoT)** 단계적으로 추론하도록 유도

프레임워크 (Framework)

22개 소개

패턴과 기법을 종합한 프롬프트의 구조적인 뼈대

- **RTF** Role · Task · Format
- **그 외** ERA, TAG, BAB, CARE, RACE, STAR 등
- 본인에게 가장 잘 맞는 구조를 활용

업무 적용

업무분야별 프롬프트 활용

요약 및 정리 · 업무 도우미 챗봇 · 글쓰기 · 언론 보도자료 · 계획(정책) 실행결과 시뮬레이션 · 아이디어 발굴 · 이미지 인식 및 처리 · 데이터 분석

업무자동화 사례

- **ChatGPT GPTs** 채용시험 상담 챗봇
- **프로그램 코드(Script)** VB Script·엑셀 VBA 자동화
- **AI Tool** PPT 작성

출처: 인사혁신처, 업무 효율성 향상을 위한 인공지능(AI) 활용 가이드, Ⅱ~Ⅳ장. 패턴·기법·프레임워크 개수는 가이드 목차 기준

프롬프트 패턴 15개는 지시 문장의 구조를 표준화합니다

프롬프트 작성 패턴 상세 · AI와 대화할 때 자주 사용되는 문장 구조

01 페르소나
AI에게 역할·인격·전문성을 부여

02 청중 페르소나
최종 독자의 특성에 맞춰 출력

03 인지 검증자
하위 질문으로 나눠 답한 뒤 종합

04 질문 다듬기
AI 도움으로 프롬프트를 개선

05 템플릿
플레이스홀더로 출력 형식 지정

06 레시피
절차·체크리스트를 제공하거나 요청

07 대안적 접근
문제 해결 대안을 찾도록 요청

08 개요 확장
개요를 작성한 뒤 점진적으로 확장

09 메뉴 Action
메뉴별 작업을 정의하고 선택해 대화

10 사실 확인 리스트
답변 근거가 되는 사실 포함을 강제

11 꼬리 생성
답변 끝에 후속 질문·문구를 추가

12 시맨틱 필터
특정 정보를 삭제·대체·추출

13 플립
AI가 사용자에게 질문하며 해결

14 입력 요청
정해진 질문으로 필요한 정보 요청

15 메타 언어 생성
단어에 새로운 의미(작업)를 부여

프롬프트 기법 12개는 AI의 사고·행동방식을 제어합니다

프롬프트 기법(Prompting Technique) 상세 · 프롬프트를 구조화하고 최적화하는 방법

01 프롬프트 체이닝
작업을 서브 작업으로 나눠 처리한 뒤 체인처럼 연결

02 Shot 프롬프트 (N-Shot)
예시를 제공해 의도를 전달 (Zero-Shot, Few-Shot)

03 생각의 사슬 (CoT)
작업을 분할하고 단계적으로 추론한 뒤 답변

04 생각의 나무
문제를 트리 구조로 분해하고 단계마다 해결책 선택

05 자기 일관성
여러 추론을 수행하고 가장 많이 나온 결과를 채택

06 산파술
질문을 통해 복잡한 문제의 해결 방안을 찾음

07 복잡성 기반
프롬프트의 난이도와 복잡도를 점진적으로 높임

08 생성된 지식
먼저 관련 지식을 생성한 뒤 이를 바탕으로 해결

09 최소 최대
간단한 문제부터 풀고 더 복잡한 문제로 진행

10 자기 개선
생성한 답변을 스스로 재검토해 개선점을 찾음

11 방향 자극 (DSP)
힌트나 단서로 특정 방향의 사고를 유도

12 리액트 (React)
추론 → 행동 → 관찰을 반복해 문제를 해결

프레임워크 22개는 업무 유형에 맞춰 골라 씁니다

프롬프트 프레임워크 상세 · 패턴과 기법을 종합한 프롬프트의 구조적인 뼈대

프레임워크	주요 활용분야	프레임워크	주요 활용분야
RTF	간단한 질문, 글쓰기, 콘텐츠 생성	STAR	업무프로세스·고객불만 개선
ERA	개인학습	CO-STAR	말씀자료, 홍보자료 작성
TAG	자동화 프로그램·보고서 작성	RISEN	데이터 분석, 매뉴얼 작성
BAB	설득력 있는 문장, 아이디어 발굴	CRISPE	맞춤형 학습, 컨설팅
APE	일상 대화, 간단한 작업요청	POWER	세부계획 수립, 코드작성
RICE	요구사항·구조가 복잡한 콘텐츠	PRESS	사실중심 보도자료 작성
CARE	현황 문제 분석·해결·개선	NEWS	실적홍보 보도자료, 정책홍보
RACE	사업계획, 정책 전략, 리스크평가	TASK	프로그래밍 개발, 데이터분석
ROSES	실무자 교육, 프로세스 개선	CODE	프로그래밍 코드 작성
COAST	정책전략·시행계획 수립	AIDA	말씀자료, 연설문, 홍보글
TRACE	업무자동화, 시스템개발	Six Thinking Hat	현황분석, 의사결정

가이드 권고: 다양한 구조화 방식을 이해하는 차원에서 학습하고, 실전에서는 본인에게 가장 잘 맞는 프롬프트 구조를 활용

출처: 인사혁신처, 업무 효율성 향상을 위한 인공지능(AI) 활용 가이드, Ⅱ-8. 프롬프트 프레임워크 (활용분야는 원문에서 일부만 발췌)

가이드는 8개 업무분야와 3가지 자동화 사례를 제시합니다

업무분야별 프롬프트 활용과 업무자동화 사례 상세

업무분야별 프롬프트 활용 사례

요약 및 정리
- 논문(자료) 요약
- 벌칙 조항 정리

업무 도우미 챗봇
- 엑셀 도우미
- 관련 규정 학습·검색

글쓰기
- 기고문·말씀자료
- 공문서·외국어 공문서
- 문장 교정

언론 보도자료
- 성과 홍보용
- 정보(사실) 전달용

실행결과 시뮬레이션
- 새 방안 적용 시 예상 시나리오 분석

아이디어 발굴
- 가상 전문가 토론

이미지 인식·처리
- 영수증 스캔 정보 추출

데이터 분석
- 수강생 기본 현황 분석
- 주관식 답변 분석

업무자동화 사례

ChatGPT GPTs
- AI 채용시험 상담 챗봇
- 법령·지침·FAQ 파일을 올려 답변

프로그램 코드(Script)
- 행정전자서명 자동복사
- 엑셀 시트 이름 일괄 변경
- 엑셀 파일 취합·제출자 리스트

AI Tool
- AI Tool을 활용한 PPT 작성
- ChatGPT·Claude로 스크립트 작성
- GAMMA로 슬라이드 디자인

출처: 인사혁신처, 업무 효율성 향상을 위한 인공지능(AI) 활용 가이드, Ⅲ. 업무분야별 프롬프트 활용, Ⅳ. 업무자동화 사례

가이드북은 15개 보안위협에 30개 보안대책을 제시합니다

국가·공공기관 AI보안 가이드북 · 개요

15개
보안위협 유형
T01~T15

30개
보안대책
M01~M30 · 권고사항

3가지
구축 유형
유형별 중점 대책 제시

2종
새 유형 AI 대책
에이전틱 AI · 피지컬 AI

AI보안의 범위

사이버보안
네트워크 · 정보통신시스템 · 소프트웨어

+

AI보안에서 추가되는 보호대상
AI모델 · 학습데이터 · AI 입출력
모델·데이터 변조가 시스템 전체에 영향

AI시스템 발전과 공공분야 도입 추세

예측형 AI
분류 및 예측
2023~2024 도입 다수

생성형 AI
콘텐츠 생성
2025 도입 급증 (34% → 77%)

에이전틱 AI
자율적 목표 달성
2026~2027 도입 추진

피지컬 AI
물리적 환경과 상호작용
2026~2027 도입 추진

도입 추세는 2025년 국가정보원 공공분야 AI시스템 조사(300개 기관) 결과. 출처: 국가·공공기관 AI보안 가이드북 2.0 (2025.12.)

AI시스템 보안위협은 모델·데이터까지 포함해 15개 유형입니다

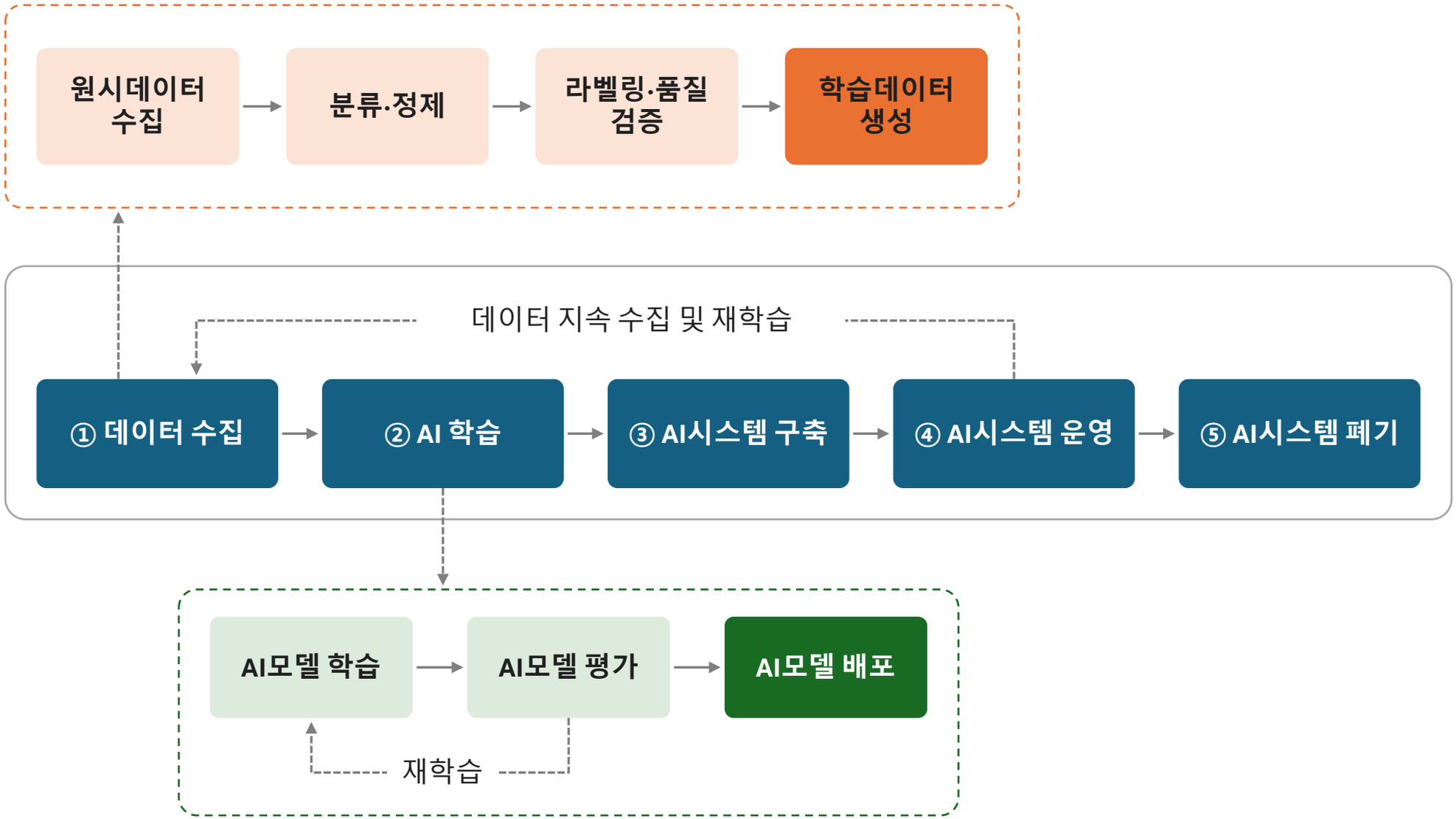
AI시스템 보안위협 · T01~T15

번호	보안위협	내용
T01	**학습데이터 오염**	학습·참조 데이터를 변조·삽입해 오염된 AI모델 생성
T02	**비인가 민감정보 학습**	학습데이터에 비인가 민감정보가 삽입·학습
T03	**AI 백도어 삽입**	학습데이터·오픈소스 라이브러리 등에 AI 백도어 은닉
T04	**학습데이터 추출**	반복 질의로 학습데이터를 재구성·추출
T05	**학습데이터 비인가자 접근**	비인가자가 학습데이터에 무단 접근
T06	**AI모델 추출**	AI모델 구조·가중치 등 주요 정보 추출
T07	**민감정보 입력·유출**	사용자가 AI에 민감정보를 입력해 유출
T08	**프롬프트 인젝션**	악의적인 프롬프트로 AI 출력·동작 변경
T09	**회피 공격**	입력정보를 조작해 AI의 오판 유도
T10	**통신구간 공격**	사용자-AI시스템 통신구간에서 정보 탈취
T11	**서비스 거부 공격**	AI에 과도한 입력을 발생시켜 과부하 유발
T12	**사고·이상행위 모니터링 체계 부재**	실시간 공격 탐지·모니터링 부재
T13	**AI시스템 권한관리 부실**	과도한 권한을 가진 AI가 임의로 타 시스템 조작
T14	**공급망 공격**	외부 AI모델·데이터, 업데이트를 통해 악성코드 유입
T15	**용역업체 보안관리 부실**	용역업체를 통해 수명주기 전 단계에서 위협 발생

내용은 원문 정의를 한 줄로 요약. 답변의 편향·환각은 가이드북 범위에서 제외. 출처: 국가·공공기관 AI보안 가이드북 2.0 (2025.12.)

데이터 수집과 AI 학습 단계는 세부 절차로 나뉩니다

AI시스템 수명주기 · 5단계와 세부 흐름



출처: 국가·공공기관 AI보안 가이드북 2.0 (2025.12.) 그림 19 AI시스템 수명주기

① 데이터 수집: 원시데이터를 모아 학습데이터로 만듭니다

AI시스템 수명주기 · 1단계 데이터 수집

① 데이터 수집 | ② AI 학습 | ③ AI시스템 구축 | ④ AI시스템 운영 | ⑤ AI시스템 폐기

정의

AI시스템이 해결해야 할 문제와 목표를 설정하고, 이를 해결하기 위해 원시데이터를 수집·정제하는 일련의 과정

목적

학습에 필요한 외부 데이터를 가지고 오거나, 내부 데이터를 선별하는 과정에서 수집·분류·정제 등을 수행

특징

내부 민감정보 포함, 외부 공개 데이터 사용, 서드파티 상용데이터 구입, 인터넷 크롤링 등 다양한 경로로 대규모 데이터가 유입

주요 내용

원시데이터 수집: 내부 데이터 생성, 외부 데이터 구매, 크롤링 등

학습데이터 생성: 분류·정제, 라벨 부여, 품질 검증을 거쳐 AI모델이 학습할 수 있는 형태로 구성

주요 보안위협

T01 학습데이터 오염: 외부 유입 데이터에 변조된 정보 포함

T02 비인가 민감정보 학습: 학습데이터에 민감정보 삽입

T03 AI 백도어 삽입: 학습데이터에 AI 백도어 포함

T14 공급망 공격: 외부 AI모델·벡터DB용 데이터에 악성코드 은닉

T15 용역업체 보안관리 부실: 용역업체를 통한 보안위협 발생

출처: 국가·공공기관 AI보안 가이드북 2.0 (2025.12.) 제1장 제3절 수명주기별 보안위협 · 주요 내용과 위협 설명은 요약

② AI 학습: 학습데이터로 모델을 학습·평가해 배포합니다

AI시스템 수명주기 · 2단계 AI 학습

① 데이터 수집 | **② AI 학습** | ③ AI시스템 구축 | ④ AI시스템 운영 | ⑤ AI시스템 폐기

정의

AI모델에 학습데이터를 학습시켜 성능을 최적화하고, 테스트를 수행하는 과정

목적

AI모델이 데이터의 패턴을 인식하고 결과를 추론·예측할 수 있도록 준비된 학습데이터를 학습

특징

GPU·NPU 및 대용량 서버 등 고성능 연산 자원이 필요하며, 학습에 사용된 데이터의 특성이 AI모델에 반영되어 결과로 도출

주요 내용

AI모델 학습: 데이터의 패턴·규칙 학습, 가중치 조정
AI모델 평가: 출력 품질·성능 평가, 재학습으로 최적화
AI모델 배포: 구축·운영에 활용할 수 있는 형태로 전달

주요 보안위협

T01 학습데이터 오염: 변조된 외부 데이터 유입
T02 비인가 민감정보 학습: 삽입된 민감정보 학습
T03 AI 백도어 삽입: 데이터에 포함된 백도어 학습
T05 학습데이터 비인가자 접근: 무단 접근
T14 공급망 공격: 외부 AI모델·벡터DB용 데이터에 악성코드 은닉
T15 용역업체 보안관리 부실: 용역업체를 통한 위협 발생

출처: 국가·공공기관 AI보안 가이드북 2.0 (2025.12.) 제1장 제3절 수명주기별 보안위협 · 주요 내용과 위협 설명은 요약

③ AI시스템 구축: 인프라를 갖추고 기존 시스템과 연계합니다

AI시스템 수명주기 · 3단계 AI시스템 구축

① 데이터 수집 | ② AI 학습 | **③ AI시스템 구축** | ④ AI시스템 운영 | ⑤ AI시스템 폐기

정의

시스템을 실제 운영할 수 있게 인프라를 구축하고 기존 정보통신시스템과 연계

목적

실제 서비스 제공을 위해 학습된 AI모델을 토대로 하드웨어·소프트웨어를 구성하고 기존 정보통신시스템과 연계

특징

AI모델을 실제 서비스 운영에 활용하고자 관련 소프트웨어를 개발·도입하고, 필요한 정보통신시스템과 API 등을 통해 연결

주요 내용

인프라 구축: CPU·GPU·스토리지 등 하드웨어 확보, RAG·MCP 구성과 운영에 필요한 도구 설치

기존시스템 연계: 기존 정보통신시스템과 AI모델을 API 등으로 연결

주요 보안위협

T03 AI 백도어 삽입: 오픈소스 라이브러리 등에 AI 백도어 은닉

T14 공급망 공격: 외부 AI모델·벡터DB용 데이터에 악성코드 은닉

T15 용역업체 보안관리 부실: 용역업체를 통한 구축단계 보안위협 발생

출처: 국가·공공기관 AI보안 가이드북 2.0 (2025.12.) 제1장 제3절 수명주기별 보안위협 · 주요 내용과 위협 설명은 요약

④ AI시스템 운영: 15개 보안위협이 모두 해당됩니다

AI시스템 수명주기 · 4단계 AI시스템 운영

① 데이터 수집 → ② AI 학습 → ③ AI시스템 구축 → **④ AI시스템 운영** → ⑤ AI시스템 폐기

정의

AI시스템을 목적에 맞게 운영하고 유지보수

목적

사용자에게 AI 기능을 챗봇, 분석, 추천, 업무지원 등 다양한 서비스로 제공하며, 주로 사용자 요청과 AI시스템 응답 형태로 운영

특징

사용자에게 인터페이스를 제공하여 공격 노출면이 늘어나고, 시스템 기능·학습데이터를 업데이트하며 모든 보안위협 발생 가능

주요 내용

시스템 운영: 정상 동작 여부를 로깅·모니터링하고 상황 발생 시 대응·관리

유지보수: 기대 성능 유지를 위한 데이터 수집·재학습, 최신 하드웨어·소프트웨어·데이터 업데이트

주요 보안위협 · T01~T15 전 위협 해당

이 단계에서만 발생 (T07~T13): 민감정보 입력·유출, 프롬프트 인젝션, 회피 공격, 통신구간 공격, 서비스 거부 공격, 모니터링 체계 부재, 권한관리 부실

다른 단계와 공통 (T03~T06, T14, T15): AI 백도어 삽입, 학습데이터 추출·비인가자 접근, AI모델 추출, 공급망 공격, 용역업체 보안관리 부실

재학습 시 (T01, T02): 학습데이터 오염, 비인가 민감정보 학습에 지속 노출

출처: 국가·공공기관 AI보안 가이드북 2.0 (2025.12.) 제1장 제3절 수명주기별 보안위협 · 주요 내용과 위협 설명은 요약

⑤ AI시스템 폐기: 구성요소를 복구할 수 없게 삭제합니다

AI시스템 수명주기 · 5단계 AI시스템 폐기

① 데이터 수집 → ② AI 학습 → ③ AI시스템 구축 → ④ AI시스템 운영 → **⑤ AI시스템 폐기**

정의

시스템의 운영 종료에 따라 AI모델, 인프라, 데이터 등 불필요한 구성요소를 삭제·파기

목적

AI시스템을 삭제·폐기하고 새로운 시스템으로 교체하거나 운영을 중단

특징

AI모델·학습데이터 등에 대한 폐기 절차가 부실할 경우 새로운 보안 취약요소가 발생하며, 재사용 등 악용 가능

주요 내용

시스템 폐기: AI시스템의 모든 구성요소를 복구 및 재사용이 불가능하도록 완전삭제 소프트웨어 등을 이용하여 삭제·폐기

주요 보안위협

T04 학습데이터 추출: 폐기단계 보안관리 부실로 학습데이터 추출

T06 AI모델 추출: 폐기단계 보안관리 부실로 AI모델 추출

T15 용역업체 보안관리 부실: 용역업체를 통한 폐기단계 보안위협 발생

출처: 국가·공공기관 AI보안 가이드북 2.0 (2025.12.) 제1장 제3절 수명주기별 보안위협 · 주요 내용과 위협 설명은 요약

운영 단계에는 15개 보안위협이 모두 해당됩니다

AI시스템 수명주기별 보안위협 · ● 해당

번호	보안위협	① 수집	② 학습	③ 구축	④ 운영	⑤ 폐기
T01	학습데이터 오염	●	●		●	
T02	비인가 민감정보 학습	●	●		●	
T03	AI 백도어 삽입	●	●	●	●	
T04	학습데이터 추출				●	●
T05	학습데이터 비인가자 접근		●		●	
T06	AI모델 추출				●	●
T07	민감정보 입력·유출				●	
T08	프롬프트 인젝션				●	
T09	회피 공격				●	
T10	통신구간 공격				●	
T11	서비스 거부 공격				●	
T12	사고·이상행위 모니터링 체계 부재				●	
T13	AI시스템 권한관리 부실				●	
T14	공급망 공격	●	●	●	●	
T15	용역업체 보안관리 부실	●	●	●	●	●
	해당 위협 수	**5개**	**6개**	**3개**	**15개**	**3개**

운영 단계에서 데이터를 계속 수집·재학습하면 T01·T02에도 지속 노출. 출처: 국가·공공기관 AI보안 가이드북 2.0 (2025.12.) 표 3

15개 위협에 대응하는 보안대책은 30개입니다

AI시스템 보안대책 · M01~M30

번호	보안대책	번호	보안대책
M01	신뢰할 수 있는 출처의 데이터 활용	M16	AI모델 구조·가중치 유출 방지
M02	신뢰할 수 있는 출처의 AI모델·라이브러리 활용	M17	AI시스템 경계보안 강화
M03	데이터 검사	M18	AI시스템 통신구간 보호
M04	데이터 암호화	M19	과도한 권한 부여 제한
M05	데이터 접근통제	M20	민감 명령 승인 절차 마련
M06	민감정보 사용 사전 승인	M21	비상대응 체계 마련
M07	보안등급에 맞는 학습데이터 구성·활용	M22	설명 가능한 AI 구성
M08	데이터 로깅·모니터링	M23	AI모델 대상 적대적 모의공격 수행
M09	AI시스템 로깅·모니터링	M24	AI모델에 적대적 공격유형 학습
M10	데이터 수집 명세서 관리	M25	구성요소 취약점 점검 및 보안업데이트
M11	AI시스템 구성요소 명세서 관리	M26	AI모델 복구
M12	AI시스템 구성요소 무결성 검증	M27	요청 속도 제한
M13	입·출력 필터링	M28	AI시스템 구성요소 완전 삭제
M14	입력 길이·형식 제한	M29	용역업체 보안관리
M15	가드레일 다중화	M30	사용자 교육 및 보안정책 수립

보안대책은 권고사항으로, 기관이 가능한 범위에서 유연하게 적용. 출처: 국가·공공기관 AI보안 가이드북 2.0 (2025.12.) 표 4

[참고] 시소러스와 온톨로지의 차이

개요 시소러스는 용어 간 관계를 정리한 사전, 온톨로지는 개념·개체·속성·관계·규칙까지 형식적으로 정의해 **기계 추론**을 가능하게 하는 지식 모형

■ 핵심 차이

구분	시소러스	온톨로지
목적	검색·색인에서 용어·개념 관계 통일	지식을 형식적으로 표현하고 추론 지원
다루는 대상	주로 용어·개념	클래스·인스턴스·속성·관계·규칙
관계 표현	동의어, 상위어·하위어, 관련어	클래스 계층, 속성, 제약, 역관계, 논리 규칙
추론 능력	제한적	규칙 기반으로 새 사실·모순 추론
대표 활용	도서관 검색, 논문 DB, 주제 분류	지식그래프, 데이터 통합, 의료·법률·산업 지식 시스템
구축 부담	상대적으로 낮음	높음 (표현력·자동화 가능성 큼)

■ 쉬운 비유

시소러스 = 관계 사전

"자동차 = 승용차, 상위 개념은 탈것, 관련 단어는 운전·연료"

온톨로지 = 지식 설계도

"자동차는 탈것의 한 종류로 엔진을 가지며, 운전자가 운전하고 도로에서 주행한다"

■ 시소러스의 특징

상위어	기준어	하위어	관련어
탈것	자동차	승용차	운전·연료

- 우선어·비우선어, 상위어·하위어, 관련어로 용어 관계를 정리
- 검색어 표현 차이를 줄이고 관련 개념으로 확장해 **검색 누락 감소**

■ 온톨로지의 특징 (법률 도메인 예시)

클래스	법률, 조문, 개정안
개체	개인정보보호법, 제15조
속성	법률의 시행일, 조문의 번호
관계	법률은 조문을 포함한다
규칙	개정안이 의결되면 기존 조문은 개정 이력을 가진다

추론 예시 "이 개정안은 어떤 조문에 영향을 주는가", "이 법률과 관련된 다른 법률은 무엇인가"에 컴퓨터가 논리적으로 답을 도출

적용 포인트 법조문 연계 판례 데이터에서 시소러스는 **일상용어→법령용어 검색 확장**에, 온톨로지는 **조문·판례·개정 이력 간 관계 추론**에 활용

[참고] 입력 통제(AI-DLP)의 개념

개요 AI-DLP(AI Data Loss Prevention)는 생성형 AI에 입력되는 프롬프트·첨부파일의 민감정보를 전송 전에 탐지해 차단·기록하는 보안 체계

■ 기존 DLP와 AI-DLP 비교

구분	기존 DLP	AI-DLP
통제 대상	파일 업로드·반출 중심	프롬프트 텍스트·첨부파일·클립보드
감시 지점	메일·단말·파일 전송 채널	생성형 AI 입력 단계
탐지 방식	정규식·키워드 패턴	패턴에 문맥 분석(NLP·딥러닝) 결합
탐지 정보	형식이 정해진 개인정보	개인정보와 소스코드·영업기밀
한계·보완	텍스트 프롬프트 탐지 곤란	AI 사용 현황 파악·응답 권한 통제 병행

■ 필요한 이유

프롬프트 유출

임직원이 생성형 AI 입력창에
소스코드·개인정보를 직접 입력

입력 후 회수 불가

입력한 정보가 AI 학습 데이터로 들어가면
사실상 회수할 수 없음

■ 주요 기능

1 프롬프트 입력 통제	개인정보·소스코드·영업기밀을 전송 전에 검사 차단 시 사용자 화면에 안내
2 첨부파일 통제	AI에 올리는 문서의 민감정보 포함 여부 검사 프롬프트와 같은 정책으로 차단
3 로깅 및 관리	누가 어떤 AI에 무엇을 보냈는지 로그로 기록 중요 기록은 관리자에게 알림

■ 처리 단계

1 입력 → 2 탐지 → 3 판단 → 4 차단 → 5 기록

- 탐지 대상: 주민등록번호 등 개인정보, 소스코드, 영업기밀·대외비
- N2SF의 C·S·O 등급별 데이터 탐지를 지원하는 제품도 출시

효과 생성형 AI 사용을 막지 않으면서 민감정보 유출을 사전 차단하고, 전송 로그로 증적 확보

※ 출처: 파수 AI 'AI DLP' 제품 소개, ZDNet Korea('26.6), 스마트비즈('26.9), Teleskope