공공부문
AI 도입·활용 가이드

행정안전부
NIA 한국지능정보사회진흥원
NATIONAL INFORMATION SOCIETY AGENCY

CONTENTS

1 AI 이해

1. AI 개념

○ 인공지능(AI)은 방대한 데이터를 바탕으로 고유한 규칙과 패턴을 학습함으로써 예측, 분류, 추천, 생성 등 지능적 과업을 수행하는 기술 체계를 의미한다.

- 초기 AI는 통계 및 규칙 기반의 자동화 방식에서 출발하였으나, 최근 머신러닝(기계학습)과 딥러닝 기술의 비약적 발전을 통해 복잡한 데이터 처리 역량과 예측 정확도를 지속적으로 고도화하고 있다.

○ 파운데이션 모델은 대규모 데이터와 고성능 연산 자원을 투입하여 수십억 개 이상의 매개변수(Parameter)를 사전 학습한 범용 AI 모델이다. 이는 텍스트, 이미지, 음성 등 다각적인 데이터를 통합적으로 처리하며, 광범위한 영역에서 추론과 지식 수행이 가능하다는 특징이 있다.

○ 생성형 AI는 초거대 AI, 특히 대규모 언어모델(LLM)을 토대로 텍스트, 이미지, 코드 등 새로운 콘텐츠를 능동적으로 생성하는 기술이다. 이를 통해 문서 작성 및 요약, 질의응답 등 지식 기반의 행정 업무를 효율적으로 지원할 수 있다.

구분	AI	파운데이션 모델	생성형 AI
핵심 목적	분석, 예측, 자동화	범용 추론, 지식 처리	생성, 대화, 문서화
주요 기능	분류, 탐지, 예측	언어 / 지식 전반 처리(범용)	문서 작성, 요약, Q&A
기술 기반	ML, DL	대규모 딥러닝 모델	LLM 기반
공공 활용	민원 분류, 이상 탐지	정책, 지침 이해 지원	행정 문서, 대국민 응대

[표 1] AI 개념

2. 관련 법령 및 지침

○ **인공지능 관련 법령**

소관부처	자료명	핵심 내용	시행일
개인정보보호위원회	개인정보보호법	개인정보 처리 동의 절차, 안전성 확보 조치 및 오남용 금지 등 제반 사항을 규정함	2024. 09
과학기술정보통신부	국가인공지능전략위원회의 설치 및 운영에 관한 규정	AI 정책 심의·조정을 위한 위원회 설치·운영 사항과 범정부 추진 체계 및 역할을 명시함	2026. 01
산업통상부	산업 디지털 전환 및 인공지능 활용 촉진법	산업 전반의 디지털 전환 및 AI 활용 촉진을 위한 공공·민간의 역할과 국가 지원 체계를 규정함	2026. 07
과학기술정보통신부	인공지능 발전과 신뢰 기반 조성 등에 관한 기본법	국민 권익 보호와 국가 경쟁력 강화를 위한 AI 기술 범위 및 관련 주체별 책임과 권리를 명시함	2026. 01
과학기술정보통신부	지능정보화 기본법	국가 및 지방자치단체의 AI 추진 책무와 지능정보사회 기반 조성을 위한 법적 근거를 명시함	2026. 01

[표 2] 인공지능 관련 법령

○ 인공지능 관련 가이드라인

발행처	자료명	핵심 내용	발행일
과학기술정보통신부/ 한국정보통신기술협회(TTA)	2024 신뢰할 수 있는 인공지능 개발 안내서	AI 서비스 및 제품의 개발·운영 단계별 기술적 신뢰성 확보 방안을 제시함	2024. 02
과학기술정보통신부/ 한국지능정보사회원(NIA)	AI 데이터 품질관리 가이드라인 v3.5	AI 학습용 데이터의 수집, 가공, 검증 등에 관한 국가 표준 및 품질 기준을 제시함	2025. 05
국가정보원/ 과학기술정보통신부	AI 보안 가이드북	적대적 공격 등 AI 시스템 특유의 보안 위협에 대한 기술적 대응 방안을 수록함	2025. 12
디지털플랫폼 정부위원회	공공부문 초거대 AI 도입·활용 가이드라인 2.0	공공부문 AI 도입 시 준수해야 할 행정 절차, 기준 및 주요 고려 사항을 안내함	2025. 04
금융위원회	금융분야 AI 보안 가이드라인	AI 모델 개발 및 도입 시 반영해야 할 특화된 보안 요구사항을 규정함	2023. 04
	금융분야 AI 운영 가이드라인	금융분야 AI 시스템의 기획, 설계부터 운영까지 전 생애 주기별 관리 지침을 제공함	2021. 07
	금융분야 AI개발/활용 안내서	기획·설계 → 개발 → 평가·검증 → 도입·운영·모니터링을 단계별로 점검함	2022. 08
문화체육관광부/ 한국저작권위원회	생성형 AI 활용 저작물의 저작권 등록 안내서	생성형 AI 활용 저작물의 등록 기준, 법적 효력 및 유형별 사례 등을 상세히 안내함	2025. 06
개인정보보호위원회	생성형 인공지능(AI) 개발·활용을 위한 개인정보 처리 안내서	생성형 AI 수명주기별 개인정보 처리 이슈를 체계화하고 법적 준수 사항과 안전 조치를 명시함	2025. 08
인사혁신처	업무효율성 향상을 위한 인사혁신처 인공지능(AI) 활용 가이드	공직자의 업무 효율 증대를 위한 생성형 AI 프롬프트 작성법 및 실무 사례를 제공함	2025. 03
과학기술정보통신부/ 한국인터넷진흥원(KISA)	인공지능 (AI) 보안 안내서	AI 생애주기 전 단계에서의 보안 위협 분석 및 단계별 필수 보안 수칙을 가이드함	2025. 12
국가정보원/ 국가보안기술연구소	챗 GPT 등 생성형 AI 활용 보안 가이드라인	생성형 AI 서비스의 접속, 질의, 결과 활용 등 전 과정에서 준수할 보안 지침을 안내함	2023. 06

[표 3] 인공지능 관련 가이드라인

3. 국내외 기술 동향

■ 국내외 기술 동향

○ 생성형 AI는 실험 단계를 넘어 공공 및 민간 부문 전반에서 상시 활용되는 운영 단계로 전환되고 있으며, 도입 기준 또한 운영 안정성, 비용 효율성, 품질 관리 중심으로 변화하고 있다.

○ 글로벌 시장에서도 AI가 업무 프로세스에 상시 적용되는 단계에 진입함에 따라, 투자 영역이 단순 모델 도입에서 운영 인프라와 서비스 통합 중심으로 확대되는 추세이다.

○ 소버린 AI 정책과 AI 기본법 추진을 통해 데이터, 인프라, 운영 통제를 포함한 신뢰 기반 AI 운영 체계 구축 요구가 확대되고 있으며, 주요 국가들은 GPU, 데이터센터, 파운데이션 모델을 결합한 국가 단위의 AI 생태계 구축을 추진하고 있다.

○ 검색 증강 생성(RAG) 기반 구조가 내부 데이터 통제와 근거 제시 측면에서 표준 아키텍처로 확산되고 있으며, AI 서비스는 오케스트레이션 기반의 조합형 구조로 고도화되고 있다.

○ AI 활용 범위는 단순 질의응답을 넘어 보고서 작성, 업무 분류, 민원 대응 등 실제 행정 자동화를 수행하는 에이전트 중심으로 확대되고 있으며, 온프레미스 및 프라이빗 AI 도입이 늘어남에 따라 하이브리드 구조가 확산되는 상황이다.

○ 기술 선택 기준은 성능 중심에서 신뢰성, 데이터 거버넌스, 책임성 중심으로 이동하고 있으며, 이에 따라 평가·검증 체계 구축과 지속적인 품질 모니터링이 핵심 요소로 자리 잡고 있다.

■ 해외 공공 AI 거버넌스 사례

○ EU: AI 법(AI Act)을 통해 공공부문 AI 도입을 법적 거버넌스 체계로 제도화하고 있으며, 문서화, 검증, 로그 관리, 책임 체계를 운영 과정에 통합하는 모델을 제시하고 있다.

○ 싱가포르: 정부기술청(GovTech)을 중심으로 생성형 AI를 범정부 공통 서비스로 제공하며, LaunchPad 및 AIBots와 같은 플랫폼을 통해 기술의 신속한 배포를 지원하고 있다.

○ 영국: 대국민 서비스와 내부 행정·정책 지원 서비스를 분리하여 운영함으로써 효율성을 높이고, 레지스트리 등록과 기록 관리를 통해 운영의 투명성을 강화하고 있다.

○ 프랑스: 디지털부(DINUM)를 중심으로 오픈소스 기반의 소버린 생성형 AI인 '알베르(Albert)'를 개발하여 공무원의 내부 업무를 중심으로 확산하고 있다.

○ 미국: 국방 및 안보 분야를 중심으로 보안 등급에 따른 AI 운영 환경 분리 전략을 추진하고 있으며, 공통 기능 모듈의 재사용을 통해 기술 확산 속도를 높이고 있다.

2 공공부문 AI 도입

1. 계획 수립 시 고려사항

○ (목적 명확화) AI 도입은 기술 중심이 아니라 업무 목적과 성과 창출 관점에서 추진되어야 하며, 적용 대상과 기대 효과를 명확히 정의하는 것이 중요하다.

- 적용 대상 업무의 특성, 기대 성과, 활용 범위를 사전에 구체화해야 하며, 생성형 AI의 기술적 한계(정확성 편차, 환각 현상, 보안 위험 등)에 대한 충분한 이해와 대응 방안을 마련해야 한다.

○ (단계적 추진) AI 기술의 급격한 발전 속도를 고려하여, 무분별한 도입보다는 단계적 검증과 확산 전략을 수립하여 추진해야 한다.

- 단기 성과 도출이 용이한 업무를 대상으로 시범 적용(PoC)을 우선 수행하고, 성능 검증 및 보완을 거쳐 고부가가치 분야로 확대하는 단계적 접근이 필요하다.

○ (최신 기술 활용) 동일한 목적 달성이 가능한 경우, 민간의 최신 AI 기술을 적극 활용하여 서비스의 품질과 혁신성을 확보한다.

- 다만 데이터 보안 수준과 업무 특성을 고려하여적용 범위를 결정하고 공공 환경에 적합한 통제, 보안, 책임 관리 체계를 함께 구축해야 한다.

○ (범정부 AI 공통기반 우선) 정부 정책 방향에 따라 AI 도입 시 범정부 AI 공통기반 및 국가 AI 인프라 활용을 우선 고려한다.

- 외교, 국방, 안보 등 업무 특수성이 있거나 높은 보안 수준이 요구되는 경우에 한해 자체 AI 인프라 구축을 검토할 수 있다.
- 도입 대상 업무 데이터가 공개정보(O) 등급인 경우에는 민간 클라우드 기반의 AI 서비스 활용이 가능하다.

○ (데이터 공동활용) 데이터 활용 전략과 공동 활용 가능성을 고려하여 AI 도입 효과를 극대화해야 한다.

- 기관 간 데이터 공동 활용, 업무 표준화, 공통 플랫폼 활용 등을 통해 중복 구축 비용을 절감하고 기술 적용 범위를 효과적으로 확대할 수 있다.

○ (신뢰성 및 안전성 확보) AI 도입 시 보안, 개인정보 보호, 책임성 확보 등 신뢰 기반의 운영 체계를 반드시 함께 구축해야 한다.

- 접근 권한 관리, 로그 기록 관리, 결과 검증 절차, 인간 개입(HITL, Human-In-The-Loop) 체계 등을 포함한 운영 통제 방안을 사전에 마련해야 한다.

○ (적절한 계약 방식 선정) 업무 환경에 따라 최적의 도입 방식을 선택해야 하며, 인터넷망 기반 상용 서비스는 '구매 방식'으로, 내부 정보를 활용하는 업무망 서비스는 '용역계약(조달발주) 방식'으로 추진한다.

- 서비스 구매 방식의 경우 디지털서비스 이용지원시스템(www.digitalmarket.kr) 또는 디지털서비스몰(digitalmall.g2b.go.kr)을 적극 활용한다.

◈ 참고: AI 도입 사업비 산정 절차

○ AI 서비스 도입 사업비 = 서비스 총이용료 + 커스터마이징 작업 비용 + 구축 개발비용

절차	추가활동	설명	산출물
1. 사전 준비	○ 도입 대상 서비스 식별 - 세부 도입 서비스 항목과 서비스 도입 유형 결정	- 도입 대상 AI 서비스를 식별하고 해당 서비스에서 필요한 기능 수준을 파악함 - 대상 서비스의 도입 유형을 결정하고, 도입 유형에 따른 추가 작업에 대한 요구사항을 정의함	대상 서비스 및 추가 활동 항목
2. 서비스 이용료 계산	○ 서비스 특성을 고려하여 사용기간 결정 - 해당 서비스 가격표 또는 견적서를 참고하여 이용료 계산	- 필요한 기능 수준, 도입 기간, 규모(사용 인원, 데이터 사용량 등)를 설정한다. 해당 서비스의 공식 가격표 또는 견적서를 참고하여 총이용료를 산정함	서비스 이용료
3. 커스터마이징 작업비용 계산	○ 서비스 도입 시 필요한 커스터마이징 작업 비용 항목 식별 - 해당 서비스 가격표 또는 견적서, 유사 서비스 가격표 또는 견적서를 참고하여 커스터마이징 작업 비용 계산	- 이용료 범위 내에서 제공되는 기본 지원 활동과는 별도로 비용이 소요되는 커스터마이징 작업을 식별함 - 작업 비용은 서비스 가격표 또는 기업 견적가를 참고하여, 비용 항목별 단위(시간, 일, 주, 월 등)와 단위당 단가를 기준으로 산정함 - 서비스 기업에 커스터마이징 작업비 기준이 없는 경우, 유사 서비스 가격표 또는 견적서 등을 참고하여 투입공수 방식으로 산정함	커스터마이징 작업 비용
4. 구축·개발 비용 계산	○ 서비스 도입 시 필요한 소프트웨어 개발 및 시스템 통합 작업 비용 항목 식별 - 기존 소프트웨어 개발비 산정 방식에 따라 기능점수 또는 투입공수 방식으로 구축·개발 비용 계산	- 시스템 통합 작업 시 요구되는 통합시스템 구축 및 추가 기능 개발, UI/UX 개선 등의 비용항목을 식별함 - 구축·개발 비용은 'SW사업 대가산정 가이드'를 준용하여 기능점수 방식 또는 투입공수 방식으로 산정함	구축·개발 비용
5. AI 서비스 도입 사업비 산정	○ 서비스 도입 사업비 산정 - 서비스 이용료+커스터마이징 작업 비용+구축·개발 비용	- AI 서비스 도입 사업비는 서비스 총이용료와 커스터마이징 작업비용, 구축·개발 비용의 합으로 산정함	서비스 도입 사업비

출처: 한국인공지능·소프트웨어산업협회, SW사업 대가산정 가이드, 2025, 160-164면.

2. 공공 AI 서비스 도입 절차

구분	내용
계획 수립	AI 도입의 필요성과 타당성을 검토하고, 도입 효과가 기대되는 업무를 중심으로 추진 목적과 범위를 구체화함
보안 등급 설정	업무 중요도에 따라 기밀(Classified), 민감(Sensitive), 공개(Open) 등 3단계 등급으로 분류하여 차등화된 보안 정책을 적용함
AI 모델 검토 및 데이터 구성	파운데이션 모델, 파인튜닝된 모델, 사후 학습된 모델, RAG(검색증강생성) 기반 모델로 구분
클라우드 서비스 구성	클라우드 영역 및 규모 선정, 클라우드 도입유형 결정 등 클라우드 서비스 구성
운영 및 유지보수 방안	데이터 준비, 모델 구축, 배포, 모니터링 등 생애주기별 관리 방안을 수립하고 지속적인 최적화를 위한 거버넌스 체계를 마련함
성과관리	AI 도입 과제의 실질적인 효용성을 검증하기 위해 정량적·정성적 성과 지표를 설정하고 사후 관리를 수행함

[표 4] AI 서비스 도입 절차

가. 계획 수립

○ 업무 특성, 정책 목표, 서비스 수요 등을 종합적으로 고려하여 AI 도입의 필요성과 타당성을 우선 검토한 후, 도입 효과가 명확히 기대되는 업무를 중심으로 계획을 구체화한다.

○ 신기술 적용에 따른 기대 성과, 위험 요소, 정책 부합성 등을 체계적으로 점검하여 AI 서비스 도입 여부를 합리적으로 결정하고, 추진을 위한 기본 방향을 설정한다.

① 대상 업무 발굴 및 도입 필요성 검토

- 기관의 업무 현황 및 서비스 환경을 분석하고, 당면한 문제점과 개선 과제를 식별한다.
- AI 도입이 실제 문제 해결 및 업무 혁신에 실질적으로 기여할 수 있는지 면밀히 검토한다.
- 단기적 성과와 중장기적 파급효과를 동시에 고려하여 도입 우선순위를 설정한다.
- 내부 행정 업무, 대민 서비스, 민원 처리 등 구체적인 적용 가능 분야를 구체화한다.

② 도입 목적 및 기대효과 정의

- 행정 효율성 제고, 국민 체감 서비스 개선, 정책 품질 향상 등 도입 목표를 명확히 수립한다.
- 정량적·정성적 성과 지표를 설정하고, 이에 따른 기대 효과를 사전에 분석한다.
- 기관의 정책 전략, 국정과제, 정부 AI 정책 방향과의 정합성을 검토한다.

③ 서비스 범위 및 추진 방식 방향 설정

- 시범 적용 여부 및 단계적 확산 계획을 포함한 전체적인 적용 범위를 설정한다.
- 범정부 AI 공통 기반 활용, 자체 인프라 구축, 민간 클라우드 활용 등 최적의 방향을 검토한다.
- 데이터 보안 등급 설정, 클라우드 구성, 모델 유형 결정 등 후속 단계를 위한 기본 원칙을 수립한다.
- 기관 내 추진 체계를 구축하고 부서별·담당자별 역할을 명확히 정의한다.
- 사업 총괄 부서 및 전담 조직을 지정하고, 정보 보안·개인정보 보호·AI 윤리 등 관련 부서 간 협업 체계를 마련한다.
- 신속한 의사결정 구조와 명확한 책임 체계를 설정한다.

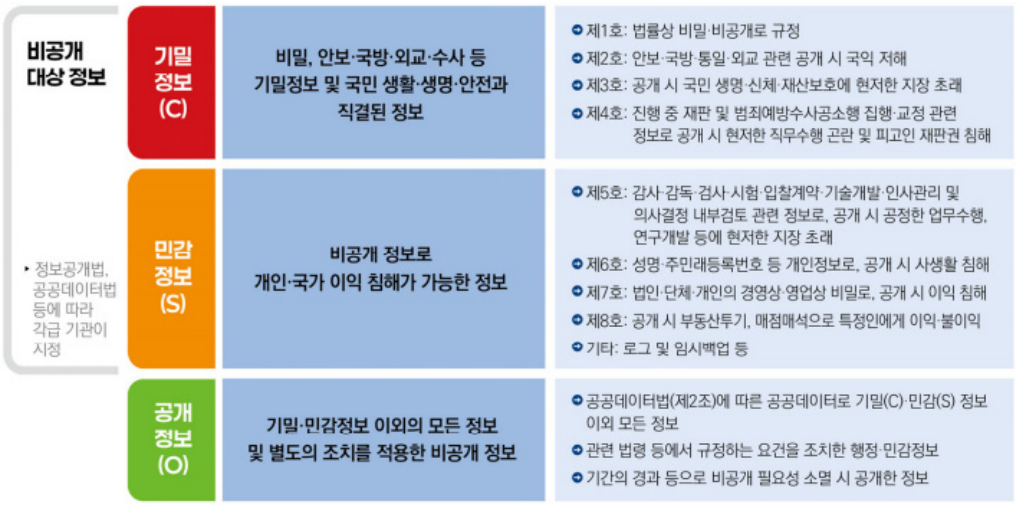
④ 리스크 및 사전 고려 요소 점검

- AI의 기술적 특성을 고려하여 편향성, 권리 침해, 환각 현상(Hallucination) 등 잠재적 위험 요소를 검토한다.
- 개인정보 보호, 디지털 소외계층 배려, 지역 불균형 해소 등 사회적 수용성 이슈를 사전에 점검한다.
- 서비스 운영의 지속 가능성, 성능 유지 방안, 예산 대비 효율성 등을 종합적으로 검토한다.
- 고영향 AI*의 경우, 관련 법령에 따라 체계적인 리스크 관리를 수행한다.

* 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」 제2조 제4호에 따른 고영향 인공지능을 의미

나. 보안 등급 설정

○ 기관의 업무 정보 및 정보 서비스 현황 식별·분석 등 기초 정보를 토대로, 도입 대상 업무에 대한 데이터 보안 등급을 중요도에 따라 구분한다.



[그림 1] 업무 정보에 대한 C/S/O 분류 기준 출처:국가정보원, 국가망 보안체계 보안 가이드라인 1.0, 2025. 23면

① 업무 중요도에 따라 기밀정보(Classified), 민감정보(Sensitive), 공개정보(Open) 등 3개 등급으로 분류함

② 데이터 보안 등급 분류에 따른 위험을 식별하고 보안 대책 대상을 선정하며, 보안 통제 항목의 선택 및 구현 계획 등 구체적인 보안 대책을 마련함

- 참고 자료 : 국가망 보안체계 보안 가이드라인, 국가 공공기관 AI 보안 가이드북

다. AI 모델 검토 및 데이터 구성

■ LLM 유형 결정

LLM 사용 목적 명확화	학습데이터 확보	작업유형 이해	학습 시간·컴퓨팅 자원 확보
• 사용 목적 정의 • 기능 및 수준 검토	• 도메인·전문데이터 유무 • 데이터 수집 기간 확보	• 사용할 작업유형 분류 • 작업유형별 LLM 활용 가능성 검토	• 자원 활용 기간 설정 • 운영·학습 인프라 비용, 규모 검토

[표 5] LLM 유형 결정 과정

① 제공하고자 하는 서비스에 따라 적합한 LLM 유형을 선정

- 국내 AI 서비스 시장에는 다양한 규모와 성능, 특성을 지닌 LLM 및 sLLM이 존재한다.
- 모델 선택을 위해서는 도입 목적을 명확히 정의하고, 기관의 업무적 특성과 가용 자원 규모를 종합적으로 고려해야 한다.
- LLM과 sLLM은 양자택일의 요소가 아니며, 공통 영역은 LLM을 활용하고 특화 분야는 sLLM을 적용하는 하이브리드 형태의 구성도 가능하다.

② 학습을 위한 데이터셋 구축 준비

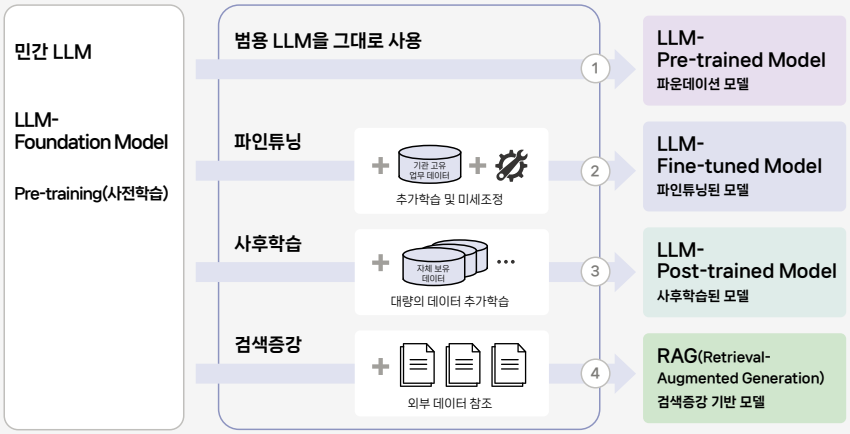
- 말뭉치(Corpus), 지도형 미세조정(SFT, Supervised Fine Tuning), 보상 모델(RM, Reward Model) 등의 데이터셋을 구축하여 사전 학습 및 미세 조정을 진행한다.
- 내부 데이터 학습 및 사용자 질의 과정에서 비공개 정보가 유출되지 않도록 적절한 보안 대책을 수립해야 한다.
- 모델 성능 향상과 최신 정책 여건 반영을 위해 주기적인 추가 학습 진행을 고려해야 한다.
- 개정 법률과 같이 시간 경과에 따라 정보가 변하는 데이터는 기존 학습 데이터와 모순될 수 있으므로, 시간 정보와 연동한 학습 체계가 필요하다.
- 학습 시에만 요구되는 고성능 연산 자원과 운영을 위한 상시 자원을 구분하여 효율적인 인프라를 구성하고 운영한다.

③ RAG 구축 준비를 위한 고려사항

- RAG(검색 증강 생성) 모델의 성능은 벡터 데이터베이스(Vector DB)의 구축 품질에 큰 영향을 받으므로, 업무 특성과 검색 패턴을 분석하여 최적의 DB를 선정해야 한다.
- 구축 단계부터 '데이터 수집-전처리-인덱싱-배포' 프로세스를 체계화하여 운영 단계의 유지 관리 용이성 확보가 필요하다.
- 데이터 변경 주기가 짧거나 최신성이 중요한 정책 문서 등은 업데이트 주기를 사전에 정의하고, 증분 업데이트 방식을 통해 시스템 성능 저하를 방지한다.
- 특화된 전문성이 높은 경우 학습이 필요하고, 데이터 변경이 낮고 최신성이 중요한 경우 RAG 구축이 필요하므로 구축하고자 하는 AI 서비스의 성격에 따라 데이터 구성 방안을 선택한다.

■ LLM 유형

○ 데이터 학습 방식에 따른 LLM 유형은 크게 파운데이션 모델, 파인튜닝된 모델, 사후학습된 모델, RAG(Retrieval-Augmented Generation, 검색 증강 생성) 기반 모델로 구분할 수 있다.



[그림 2] 학습 방식에 따른 LLM 유형

구분	내용	
파운데이션 모델	사용 목적	일반적 목적으로 사용하는 유형
	학습 데이터	자체 보유 데이터가 없어도 사용 가능
	작업 유형	텍스트 생성, 문장 완성 등 질의 응답 능력을 기반으로 하는 작업
	학습시간·컴퓨팅자원	추가적인 데이터 수집과 학습 과정이 없기 때문에 빠르게 구축 및 사용 가능
파인튜닝된 모델	사용 목적	문장 분류, 감정 분석 등 구체적인 작업
	학습 데이터	특정 작업과 관련된 소량의 데이터 확보 필요
	작업 유형	특정 작업과 관련된 문맥이나 패턴 이해를 기반으로 하는 작업
	학습시간·컴퓨팅자원	주기적인 추가 학습 인프라가 필요하므로 파운데이션 모델보다 상대적으로 많은 자원을 소요
사후학습된 모델	사용 목적	최신 정보와 성능을 필요로 하는 전문 작업
	학습 데이터	최신성, 전문성과 관련된 대량의 데이터 확보 필요
	작업 유형	파운데이션 모델을 최신화, 전문화한 다양한 작업
	학습시간·컴퓨팅자원	대규모의 추가적인 학습 인프라가 필요하므로 파운데이션 모델이나 파인튜닝보다 상대적으로 많은 자원을 소요
RAG 기반 모델	사용 목적	최신 정보를 검색·활용하여 답변 제공
	학습 데이터	모델 자체 학습 없이 보유 데이터를 활용하여 최신 데이터 반영
	작업 유형	최신 데이터를 활용한 최신 정보 생성, 질의응답, 요약, 검색 결과 기반의 문서 작성 등
	학습시간·컴퓨팅자원	모델 자체를 학습하지 않으므로 학습 시간이 소요되지 않으며, 보유 데이터의 신뢰성 확보가 중요

[표 6] LLM 유형별 주요 내용

라. 클라우드 서비스 구성

① 시스템을 운영할 클라우드 영역 및 컴퓨팅 리소스 규모를 선정함

- 데이터 및 시스템의 특성을 종합적으로 고려하여 클라우드 영역과 LLM 도입 위치를 명확히 구분한다.
- 데이터 학습 방식과 컴퓨팅 리소스(CPU, GPU, 메모리, 스토리지 등) 요구량을 분석하여 최적의 인프라 규모를 결정한다.

② 클라우드 도입 유형에 따라 자원의 활용 방법을 검토하여 결정함

- 도입하는 클라우드 자원을 기관 단독으로 사용하거나 타 기관과 공유하는지에 따라 자원 공유 방식을 결정한다.
- 클라우드 자원의 통제 수준을 도입기관이 직접 물리적으로 통제하거나 위탁을 통해 관리하는지에 따라 통제 방식을 구분한다.

구분	내용
퍼블릭 클라우드 (Public Cloud)	• 클라우드 서비스 공급자가 서버 및 클라우드 리소스 제공 • 모든 기관 또는 사용자가 자원을 공유
프라이빗 클라우드 (Private Cloud)	• 단일 조직에서 독점적으로 사용되는 컴퓨팅 리소스 제공 • 기관이 클라우드 자원의 통제권을 보유
멀티 클라우드 (Multi Cloud)	• 여러 퍼블릭 클라우드를 함께 쓰는 방식 • 안정성 확보를 위한 클라우드 분산 운영이 필요한 조직에 적합

[표 7] 클라우드 서비스 형태 분류

③ 이용 중인 정보 자원이 있을 경우 전환 및 재구축 방안을 검토하고, 이를 신규 시스템에 필요한 정보 자원 규모 산정에 반영하여 종합적으로 고려함

④ CSAP 인증 등에 기반한 보안성, 비용 효율성, 안정성, 향후 확장성 등을 고려하여 클라우드 컴퓨팅 서비스를 선정함

클라우드 서비스 도입·운영 유형별 고려사항

■ 민간 클라우드 서비스

○ 이미 개발된 민간 클라우드 서비스를 활용함으로써 구축 비용을 절감하고 도입 기간을 단축하는 효과가 있다.

○ 도입 절차 및 고려사항

- 클라우드 영역 및 컴퓨팅 리소스 규모 선정: 데이터 특성과 업무 중요도에 따라 클라우드 영역 및 LLM 도입 위치를 구분하고, CPU·GPU·메모리 등 컴퓨팅 자원 규모를 산정한다.
- 보안·연계 구조 설계: 내부 시스템과의 연계, 외부 인터넷 환경 사용 여부, 데이터 전송 구조 등을 검토하여 보안 위험을 최소화하도록 설계하며, 외부 환경으로의 데이터 유출 방지에 유의한다.
- 보안 인증 확인: 활용하려는 민간 클라우드 서비스의 클라우드 보안인증(CSAP 등) 충족 여부를 확인하고, 데이터 보안 등급 결과에 따른 보안 대책을 적용한다.

■ 국가정보자원관리원 민관협력(PPP) 클라우드 서비스

○ 보안성, 안정성, 책임성 측면에서 강화된 운영 환경이 제공되므로 국민적 영향도가 크거나 핵심 행정 서비스 등에 적용하는 것이 효과적이다.

○ 도입 절차 및 고려사항

- 네트워크 연계 환경 검토: 정부 업무망 및 기존 정보자원 시스템과의 안정적인 연계 가능성을 사전에 확인한다.
- 책임 체계 명확화: 장애 발생 및 보안 사고에 대비한 대응 체계를 구축하고, 기관과 운영 주체 간의 역할과 책임을 명확히 정의한다.
- 국가 망 보안 체계 적합성 확인: 국가 보안 통제 기준과의 적합성 여부를 면밀히 검토한다.

■ 기관 자체 민관협력형 클라우드 서비스

○ 국가정보자원관리원의 구축 환경을 활용하기 어려운 경우, 별도의 민관협력형 클라우드 환경을 구축하여 운영할 수 있다.

○ 도입 절차 및 고려사항

- 협력 모델 및 역할 정의: 기관과 민간 기업 간의 협력 구조, 운영 책임, 관리 범위, 거버넌스 체계 등 기능별 책임 관계를 구체적으로 설정한다.
- 기술 아키텍처 설계: 네트워크 연계 여부, 외부망 사용 여부, 데이터 흐름, 로그 관리, 접근 통제 등 전반적인 보안 체계를 설계한다.
- 운영 승인 및 관리 기준 설정: 운영 승인(ATO) 등 기관 내부 승인 절차를 마련하고, SLA(서비스 수준 협약)·장애 대응·성능 관리 기준을 계약을 통해 명확히 규정한다.
- 협약 관리: 민간 기업의 수익 구조와 협약의 구체적인 이행 사항 등을 설정하여 관리한다.

마. 운영 및 유지보수 방안

○ LLM을 효율적으로 운영하기 위한 데이터 관리, 유지보수 및 상세 운영 방안 수립

- 주요 내용에는 추가 학습, 교육, 배포, 모니터링, 최적화 등이 포함되어야 하며, 데이터 보호 및 환각 현상(Hallucination) 예방 조치 등을 병행한다.
- 또한 LLM 운영 전 과정에서 사용자 피드백을 통해 파라미터를 재조정하고, 강화 학습 등을 활용하여 모델 성능 개선을 지속적으로 추진한다.
- 지속적인 모델 관리와 유지보수를 위해 필요한 운영 예산을 사전에 확보한다.
- 효과적인 성능 개선과 기술 고도화를 위해 민간 전문가와의 협력 체계 구성을 검토한다.

○ **AI 서비스를 효율적으로 활용하기 위해서는 데이터를 지속해서 관리할 조직과 이를 위한 거버넌스적 체계 마련**

- 내부적으로는 언어 모델의 규모 및 비용 대비 성능 최적화를 위해 운영 중 발생하는 신규 데이터(사용자 로그, 질의응답, 오류 이력 등)를 체계적으로 수집·분석하여 모델 개선에 반영한다.
- 외부적으로도 AI 서비스 활용 과정에서 발생할 수 있는 편향성, 할루시네이션(환각 현상), 악용 등 사회적·윤리적 문제에 대응하기 위한 관리 방안을 수립한다.
- 서비스 사용자와 제공자 간의 책임 소재를 명확히 하고 서비스 품질을 보장하기 위해 구체적으로 측정 가능한 목표치를 설정한다.

○ **도입 대상 AI 서비스가 「인공지능기본법」에 따른 '고영향 AI'인 경우, 안전성 및 신뢰성 확보를 위해 필요한 조치 수행**

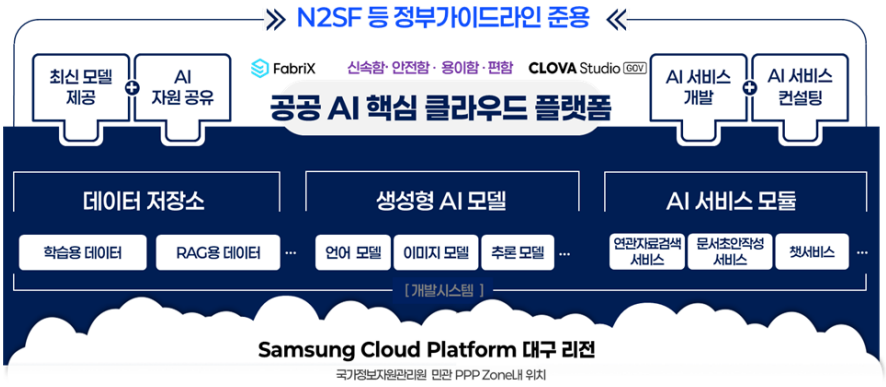
- 위험 관리 방안의 수립 및 운영, 결과 도출의 주요 기준과 학습 데이터 개요에 대한 설명 방안 수립·시행, 이용자 보호 방안 운영, 사람에 의한 관리·감독, 관련 문서 작성 및 보관 등을 철저히 이행한다.

3 범정부 AI 공통기반

1. 범정부 AI 공통기반 소개

■ 정의

○ 중앙 및 지방정부가 인공지능을 신속하고 안전하게 도입할 수 있도록 지원하는 '정부 전용 AI 핵심 인프라'이다. AI 모델, 학습 데이터, 컴퓨팅 자원 등을 공동 활용함으로써 기관별 중복 개발과 투자를 방지하고, 각 기관의 수요에 최적화된 AI 서비스 개발 및 운영을 지원한다.



[그림 3] 범정부 AI 공통기반 개념도

■ 목적

○ (안전한 공공 AI 서비스 구현 기반 제공) 국가정보자원관리원 대구센터 PPP Zone 내에 안전성이 검증된 생성형 AI 개발 환경을 제공한다.

○ (국가 자원 활용 효율화를 위한 공공 AI 자원 공동 활용) 중앙 및 지방정부가 AI 기획·구현·운영 시 공통으로 활용하는 GPU, AI 모델, 학습 데이터, RAG DB, 서비스 모듈 등 공공 AI 자원의 공동 활용을 지원한다. 이를 통해 예산 중복 투자를 방지하고 국가 자원 활용의 효율성을 제고한다.

○ (공공 AI 서비스 혁신 가속화) 범정부 차원에서 공통기반에 축적된 AI 서비스 및 모델 정보, 학습 데이터 등 AI 자산과 노하우를 공유할 수 있는 창구를 제공한다. 이를 통해 중앙 및 지방정부 전반의 공공 AI 서비스 혁신 주기를 단축한다.

■ 기대효과

○ (범정부 차원의 AI 서비스 및 데이터 구축 비용 절감) AI 플랫폼 구축 및 데이터(학습용, RAG용) 공유 체계를 통해 AI 서비스 개발 기간을 단축하고 관련 데이터 구축 비용을 절감한다.

○ (AI 시장 생태계 활성화) 범정부 AI 영역에 AI 모델, 클라우드, 학습 전문 기업 등 민간 기업의 참여 기회를 확대한다. 대기업과 중소기업 간 상호 협력을 통한 AI 서비스 개발의 장(場)을 제공한다.

○ (AI 데이터 기반 업무 혁신) 다양한 AI 서비스를 사전에 테스트하고, 검증된 서비스를 신속하게 현업에 적용할 수 있는 기반을 제공한다.

○ (정부 민감 정보 보안성 강화) 정부가 활용하는 AI 모델을 플랫폼에서 통합 관리하고 정부 보안망 내에서만 사용하도록 제한함으로써, 민감 데이터의 외부 유출을 원천적으로 방지한다.

2. 범정부 AI 공통기반 구성요소

○ 범정부 AI 공통기반의 주요 구성요소는 ① AI 플랫폼, ② AI 모델, ③ 학습 데이터, ④ RAG, ⑤ AI 서비스, ⑥ 개발 시스템으로 이루어진다. AI 서비스 구축 시 서비스 구현 단계별로 필요한 자원을 유연하게 선택하고 활용할 수 있다.

주요 구성요소	내용
① AI 플랫폼	– RAG 구축 환경 제공(데이터 업로드·정제, 파싱, 청킹, 임베딩을 통한 DB 구축) – AI 서비스 개발·운영 환경 제공(에이전트 설계, 개발, 평가, 배포 등) – FabriX, CLOVA Studio for GOV 등 플랫폼 지원
② AI 모델	– 다양한 생성형 AI 모델 30종 제공 (LLM, 멀티모달, 추론 모델 등) – 최신 AI 모델 업데이트 및 고도화 지원
③ 학습데이터	– 공동 활용 가능한 고품질 학습 데이터를 공통기반 개발 시스템에서 제공 • 지시 학습용 126,000 건 / 공통 서비스 학습용 8,000 건 데이터셋 포함
④ RAG	– AI 답변 성능 향상을 위한 행정 공통 RAG 10종* 제공 및 공동 활용 지원 * 보도 / 연구 / 현행 법령 / 현행 행정규칙 등 행정 업무 시 필요한 약 40만 7천 건의 데이터를 기반으로 구축
⑤ AI 서비스	– 기관 자체 시스템에 즉시 활용 가능한 공통 AI 서비스(연관 정보 검색, 문서 초안 작성 등) 제공 – AI 질의응답이 가능한 기본 챗(Chat) 서비스 지원
⑥ 개발시스템	– AI 플랫폼을 이용한 시스템 개발 및 운영을 지원하는 통합 창구 제공 – 요금 체계에 따른 정보시스템별 사용량 및 과금 현황 모니터링 기능 지원

[표 8] 범정부 AI 공통기반 구성요소

○ AI 공통기반 이해관계자

이해관계자		설명
관리·운영자	공통기반 관리자	범정부 AI 공통기반의 행정 및 관리 체계 전반을 총괄하여 관리하는 주체
	플랫폼 운영자	공통기반에서 제공하는 AI 플랫폼을 관리·운영하는 주체 : FabriX 운영자(삼성SDS), CLOVA Studio for GOV 운영자(네이버클라우드)
사용자	정보시스템 담당자	공통기반을 활용하는 이용 기관(중앙·지방정부, 공공기관 등)의 정보시스템을 담당하는 주체
	정보시스템 개발자	공통기반을 활용하여 이용 기관(중앙·지방정부, 공공기관 등)이 필요로 하는 AI 서비스 개발 또는 RAG 구축 등의 업무를 수행하는 주체
	일반사용자	AI 플랫폼사와 계약하기 전 카탈로그 조회 등을 활용하는 중앙·지방정부 공무원 및 공공기관 직원
	Chat서비스 이용자 (공무원, 공공기관)	AI 플랫폼에서 제공하는 '범정부 AI 공통기반 chat 서비스(chat.ai.go.kr)'이용하는 중앙·지방정부 공무원

[표 9] 범정부 AI 공통기반 이해관계자

2.1. AI 플랫폼

○ AI 플랫폼은 데이터 수집·전처리, RAG DB 구성, AI 모델 서빙, AI 서비스 개발 및 테스트, 배포와 운영에 이르기까지 AI 서비스 구현 전 과정에 필요한 자원을 제공하고 안정적으로 관리하는 통합 환경이다.

○ 현재 범정부 AI 공통기반에서 제공하고 있는 AI 플랫폼은 FabriX와 CLOVA Studio for GOV이다.

■ AI 플랫폼 구성요소 및 특징

○ AI 플랫폼 구성요소

[그림 4] AI 플랫폼 구성도

○ AI 플랫폼 특징

AI 서비스 개발운영 지원 도구 제공	민간 AI 최신 기술 도입·활용
⟫ GPU·모델 등 AI 자원 공동 활용 ⟫ 서비스 공통 개발환경 제공	⟫ 다수의 민간 플랫폼, 모델 도입 ⟫ 공공·민간 AI 생태계 구성

[그림 5] AI 플랫폼 특징

- (AI 서비스 개발·테스트·운영 지원 도구 제공) 플랫폼 내 제공된 자산을 활용하여 로코드(Low-Code) 또는 노코드(No-Code) 방식으로 손쉽게 AI 서비스를 개발할 수 있는 환경을 제공한다. 화면 구성, 비즈니스 로직 설계, 데이터 연결 등을 직관적으로 처리하여 신속한 AI 서비스 구축을 지원한다.

- (멀티 플랫폼 구조, 민간 최신 기술 도입·활용) 유연한 서비스 아키텍처를 확보하여 기술적 위험을 분산하고 서비스 지속성을 보장하는 멀티 플랫폼 구조를 지향하고, 기관의 수요와 목적에 따라 최적의 AI 플랫폼 및 AI 모델을 자유롭게 선택하여 활용할 수 있다.

◈ 참고: 공통기반 AI 플랫폼을 활용한 AI 서비스 개발 프로세스

구현 단계	단계별 To-Do	
❶ LLM 모델 선택	LLM 모델 선택 파라미터 설정	• AI 플랫폼에서 제공하는 LLM 모델 선택 • 답변 수준 위한 설정값(TopK, TopP, Max tokens 등 Parameter) 지정
❷ 시스템 지시문 작성	지시문 작성 지시문 테스트	• AI 서비스 문제 처리 방식, 답변 제시 형태를 지정하는 시스템 지시문(System prompt) 작성 – 작성 시 AI 플랫폼 제공 지시문 템플릿을 사용하거나 기관별 맞춤화해 사용 가능 • 작성 지시문에 따른 결과 확인 후 수정
❸ 기반지식 (RAG) 선택	기반지식(RAG) 생성 기반지식(RAG) 선택	• (RAG생성) 필요 데이터 수집/정제/저장 과정 거쳐 기반지식(RAG) 생성 • (RAG선택) 기관 생성 RAG와 공통기반 제공 RAG들 중에 사용기관이 구현하고자 하는 AI 서비스에 적합한 RAG 선택
❹ 플러그인 선택	플러그인 선택	• AI 플랫폼에서 기본 제공되는 플러그인 선택해 기능 확장 가능 (ex. fabriX에서 제공하는 코드 인터프리터 플러그인 선택 시 자연어로 Python코드 생성 가능)
❺ 테스트 및 배포	테스트 배포	• AI 구성요소(모델, 지시문, 기반지식, 플러그인) 선택 및 변경 통한 테스트 진행 • 각 기관 담당자가 사용할 수 있도록 AI 서비스 제공(배포)

2.2. AI 모델

○ 범정부 AI 공통기반에서는 LLM, VLM, 추론 모델 등 다양한 생성형 AI 모델을 제공한다.

○ 정보시스템 담당자 및 개발자가 최신 AI 모델을 안정적으로 활용할 수 있도록 지속적인 모델 최신화와 업데이트를 지원하고, 개별 기관 요청 시 검토 및 평가를 거쳐, 기관에서 필요로 하는 새로운 AI 모델을 추가로 서빙할 수 있도록 지원할 예정이다.

※ 추후 독자 AI 파운데이션 모델 추가 예정

○ 범정부 AI 공통기반의 독파모 AI 모델 공급 일정

- **Naver Cloud**: HCX-SEED-Think32B(HCX-GOV-Think로 제공)(2026.03.16.)
- **LG AI Research:** K-Exaone-236B(2026.04.30.)
- **Upstage:** Solar-open-100b(2026.2Q)
- **SK Telecom:** A,X K1-519B(공급 시기 미정)
- **NC AI:** VAETKI-112B(공급 시기 미정)

■ 공통기반 서빙 모델

○ AI 플랫폼(FabriX, CLOVA Studio for GOV)에 검증이 완료된 30종의 모델 풀(Pool)을 제공하며, 이 중 10종을 우선 탑재하여 서빙 중이다.

회사	모델명
FabirX	SamsungLLM 37B, Llama 3.3 70B, Gemma 3 27B, GPT-OSS 120B
CLOVA Studio for GOV	HCX-GOV-Think 24B, HCX-GOV-Think 32B, LLM42(Gemma 3 12B), LLM42(Gemma 4 31B), Translate Gemma 27B, GPT-OSS 120B, K-Exaone 236B

* 삼성SDS FabriX 및 네이버클라우드 CLOVA Studio for GOV는 2026년 2분기 중 학습을 완료할 예정이다.

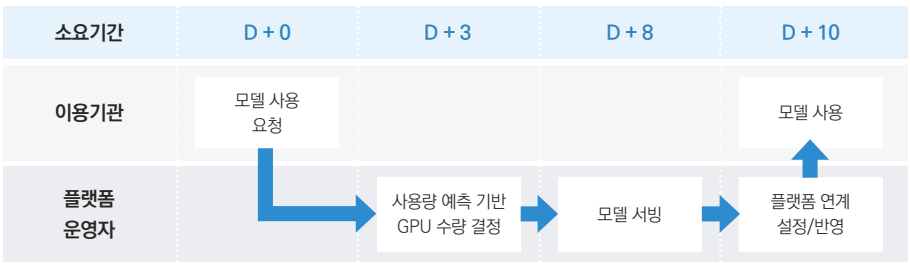
** [부록1]'공통기반 서빙 모델' 참고
공통기반에서 활용 가능한 최신 AI 모델 정보는 '개발시스템(dev.ai.go.kr) 〉 카탈로그 관리 〉 AI 모델 정보'에서 확인할 수 있다.

○ 각 기관은 구현하고자 하는 AI 서비스의 특성과 목적에 따라 적합한 모델을 선택하여 서비스 개발에 활용할 수 있다.

○ 현재 공통기반에 탑재된 10종을 제외하고, 플랫폼사 검증이 완료된 20종의 모델을 사용하고자 하는 기관은 플랫폼 운영자에게 이메일로 이용을 요청한다.

* 연락처: fabrix.cs@samsung.com, dl_clovastudio_gov@navercorp.com

- 모델 서빙이 완료되면 플랫폼 운영자는 이용 기관에 모델 적용 완료 여부를 유선 또는 메일로 통보하며, 해당 모델은 즉시 사용할 수 있다.



[그림 6] 검증 완료된 모델 반영 프로세스

- 특별한 도입 불가 사유가 없을 경우 처리 기간은 최소 2주* 소요된다.

* 소요 기간은 워킹데이 기준이며 GPU 자원 확보, 모델 크기 등에 따라 검토 기간이 추가 소요될 수 있음

■ 개별 기관 요청 AI 모델의 등록 및 서빙

○ 공통기반에서 제공하는 기본 AI 모델 외에 개별 기관의 서비스 구축에 특화된 모델이 필요한 경우, 행정안전부와 협의를 거쳐 새로운 AI 모델을 공통기반에 탑재할 수 있다.

2.3. 학습데이터

○ 범정부 AI 공통기반에서는 개별 기관이 특화 AI 모델을 훈련할 수 있도록 양질의 학습 데이터를 제공한다.

* 제공되는 학습 데이터는 '공통기반 〉 개발 시스템 〉 카탈로그 관리 〉 데이터 공유' 메뉴에서 다운로드 가능

- 공통기반에서 제공하는 AI 학습 데이터는 기관별 특화 LLM 학습 시 기초 자료로 활용 가능하다.
- 이를 통해 각 기관은 고유한 업무 도메인에 특화된 성능을 확보함과 동시에, 공공·행정 분야의 방대한 문서와 지식을 폭넓게 이해하는 활용성 높은 LLM을 구축할 수 있다.

◈ 참고: 학습데이터, RAG 비교

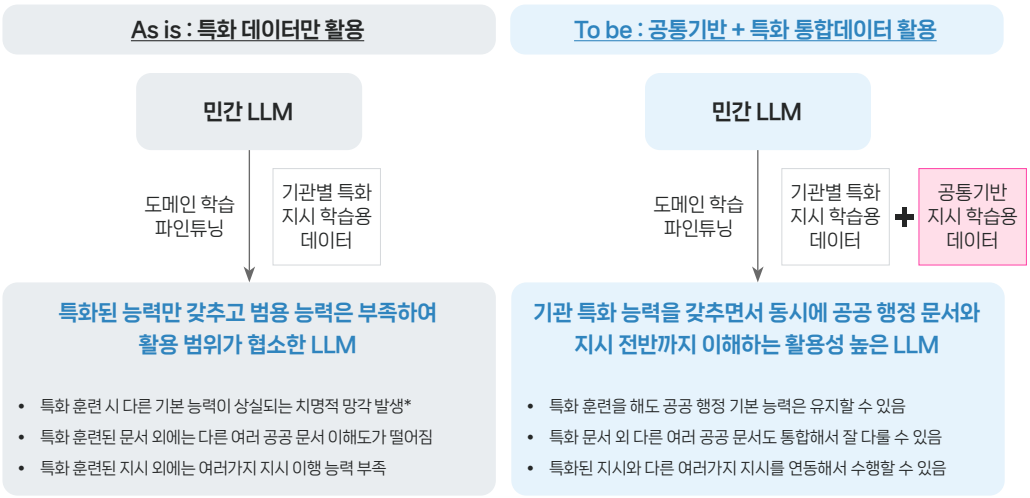
	학습데이터	RAG 데이터
목적	▸ AI 모델 자체에 언어 능력·지식·패턴 등을 학습시키기 위해 학습데이터를 활용	▸ AI 모델 추론 시 외부지식 데이터를 검색하여 답변결과를 생성하기 위해 RAG데이터(Vector DB)를 활용
데이터 준비과정	▸ 수집 → 데이터 전처리(정제/가공) → ▸ 포맷화 → **AI모델 학습에 활용**	▸ 수집 → 데이터 전처리(정제/파싱/청킹/임베딩) → 벡터 DB 구축 → **AI 서비스 구축 시 벡터DB 활용**

데이터 형태	▸ Instruction 포맷 형태 (JSON, CSV 등)
	json
	{
	"instruction": "서울의 수도는 어디니?",
	"response": "서울의 수도는 서울특별시 입니다."
	}
		▸ 벡터로 변환(임베딩)한 데이터를 벡터 DB에 저장한 형태
		[0.1 0.7 0.4 ...]
		[0.2 0.5 0.8 ...]
		[0.1 0.2 0.1 ...]
		[0.7 0.1 0.9 ...]
		[0.5 0.6 0.3 ...]
		...
		저장 → 벡터 DB

데이터 활용시점	▸ AI 모델 학습 단계에서 적용 예:) AI 모델 Fine-tuning, AI모델에 특정 task 학습	▸ 특정 AI 서비스 구성 시 RAG (벡터 DB) 활용 예:) 기업 내부 지식 챗봇 구성 시 최신 매뉴얼이 포함된 벡터 DB(RAG)를 서비스 구성요소로 포함
답변 생성방식	▸ 대규모 학습데이터로 미리 학습된 **모델 내부 지식**을 활용해 답변을 생성함	▸ 질문과 관련된 정보를 외부 데이터베이스(RAG DB)에서 검색하여 답변을 생성함

○ 공통기반 제공 학습데이터 활용 방법

- 기관 특화 AI 서비스를 구축하기 위해 AI 모델 학습 시, 기관의 특화 학습데이터와 공통기반 학습데이터를 병행하여 학습한다.
  - 데이터 통합: 공통기반으로부터 제공받은 지시 학습용 데이터를 기관 특화 데이터와 통합하여 고도화된 학습 데이터셋을 구성한다.
  - 모델 파인튜닝: 통합된 데이터셋을 기관 특화 LLM에 적용하여 해당 도메인에 최적화된 미세 조정을 진행한다.



*모델이 새로운 도메인(예: 법률, 의료)만 집중 학습하면 기존에 가지고 있던 일반 지식 또는 언어 능력을 잊는 현상

[그림 7] 학습데이터 활용 시 AS-IS, TO-BE

○ 공통기반이 제공하는 AI 학습데이터

가공데이터 항목	내용	구축 규모
지시학습용 데이터 ('25. 12월 기준)	- (용도) 민간에서 민간 말뭉치와 민간 지시 학습 데이터로 이미 훈련시켜 놓은 AI를 범정부 AI가 수행해야 할 지시들에 적응·훈련시키는 용도 - (구성) 문서 생애주기 전체에 걸친 다양한 지시 유형(약 40개*)에 대해 지시문과 모범 응답문 형태의 데이터로 구축됨 - (규모) 각 유형별로 학습 효과를 기대할 수 있는 최소 수량을 1,000건, 특정 유형에만 편향적인 학습을 방지하기 위해 한 유형당 최대 수량은 5,000건으로 제한함	126,000 건
평가·검증용 데이터* ('25. 12월 기준)	- (용도) 범정부 AI가 수행해야 할 지시를 얼마나 제대로 학습했는지를 평가하는 용도 - (구성) 지시학습용 데이터와 동일한 기준인 40가지 지시유형을 따라 구축되며 객관식 평가 데이터(지시문-제시문-선택지-정답) 와 주관식 평가 데이터(지시문-모범 응답문 10개)로 구성됨 - (규모) 40가지 지시유형* 하나당 100점 만점 점수로 3번을 평가할 수 있도록 최소 300건 이상으로 하되, 특정 유형에 편향되지 않도록 한 유형당 최대 1,000건으로 제한하여 구축함	30,000 건
공통 서비스 학습용 데이터 ('25. 12월 기준)	- (용도) 범정부 AI가 공통 서비스가 입력하는 데이터를 잘 이해하고, 공통 서비스가 원하는 형태로 응답하는 것을 적응·훈련시키는 용도 - (구성) 지시문과 모범 응답문 형태의 데이터로 구축됨 - (규모) 공통 활용 서비스에서 LLM이 사용되는 대표적인 2가지 유형(검색어 확장, 검색 문서 기반 질의응답)에 맞춰 평균 4,000건씩 총 8,000건 구축함	8,000 건

[표 10] 공통기반 제공 AI 학습데이터

- 상기 가공 데이터 항목 중 공통기반 플랫폼에서 다운로드하여 기관별 AI 모델 학습에 활용할 수 있는 학습 데이터는 지시 학습용 126,000 건과 공통 서비스 학습용 데이터 8,000건이다.

  • 평가·검증용 데이터는 모델 성능을 객관적으로 비교·검증하기 위한 기준 데이터이다.

  • 해당 데이터가 사전에 공개되거나 외부로 제공될 경우, 평가 결과의 객관성과 신뢰성이 저하될 우려가 있으므로 다운로드 대상에서 제외한다.

지시유형 대분류(7개)	세부 지시유형(40개)
1. 기획단계 (Planing Phase-A)	1.1. 주제·제목 제안 1.2. 목차·구조 설계 1.3. 핵심 메시지·논거 정의 1.4. 아이디어·키워드 제시 1.5. 초록·개요 작성
2. 작성 단계 (Drafting Phase-B)	2.1. 문서 초안 생성 2.2. 현재 문장·단락 완성 2.3. 다음 단락 작성 2.4. 특정 파트 작성
3. 수정/검토 단계 (Revision/Review Phase-C)	3.1. 문법·맞춤법·외래어·띄어쓰기 교정 3.2. 어휘·표현 개선 및 윤문 3.3. 공공행정 형식에 맞게 내용 수정 3.4. 내용 일관성·논리성 검토 3.5. 가독성 진단 및 개선 3.6. 특정 요구사항·피드백 기반 수정 3.7. 문서 용도·대상에 맞게 피드백 기반 수정
4. 활용 단계 (Utilization Phase-D)	4.1. 문서 요약 4.2. 문서 내용 검색 4.3. 문서 분석·분류 4.4. 문서 기반 질의응답

지시유형 대분류(7개)	세부 지시유형(40개)
5. 멀티모달 (Multimodal Phase-E)	5.1. 이미지 포함 문서 텍스트 추출 5.2. 이미지 내용 설명 및 해석 5.3. 도표·그래프 설명 및 해석 5.4. 표 데이터 분석·요약 5.5. 표 데이터 기반 질의 및 응답
6. LLM 종합 능력 (LLM Capability Evaluation Phaase-F)	6.1. 언어 이해 6.2. 언어 생성·편집 6.3. 소통·대화 관리 6.4. 정보 검색·추출 6.5. 지식 적용·정합성 검증 6.6. 논리·추론 6.7. 창의 아이디어 생성
7. 윤리/안전 (Ethics/Safety Phase-G)	7.1. 편향 최소화 7.2. 근거 제시·설명 가능성 7.3. 개인정보비식별화 7.4. 규정 준수 7.5. 유해·불법 콘텐츠 차단 7.6. 환각 최소화 7.7. 프롬프트 공격 대응

[그림 8] 40개의 지시유형

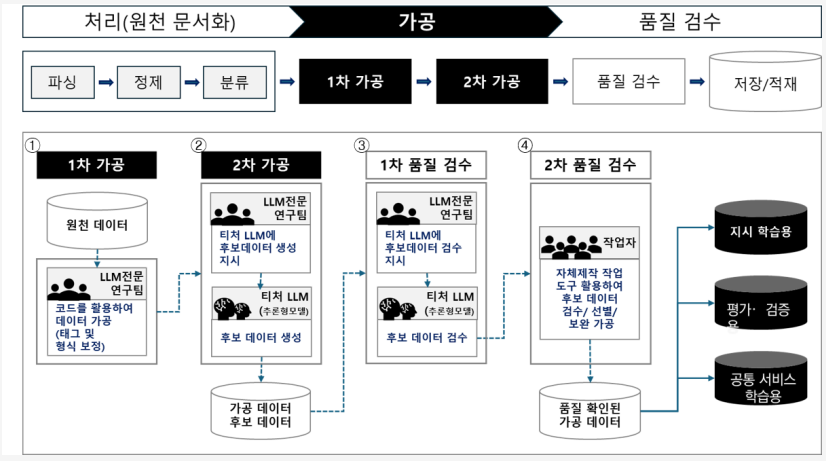
○ 학습데이터 가공 단계별 결과물과 수량

- **(원시데이터)** 범정부 학습데이터 구축을 위해 수집한 기초 데이터로 '공동활용데이터 등록관리시스템'에 등록된 행정 문서와 정부·공공기관에서 개방한 공공 데이터 등으로 구성된다.
- **(원천데이터)** 확보된 원시데이터를 분류, 추출(파싱), 텍스트 정제, 처리(저품질·중복 데이터 제외)하는 과정을 거쳐 저장한 데이터이다.
- **(가공데이터)** 원천데이터를 AI 모델 학습 용도에 맞게 가공하고 선별한 최종 데이터로 이는 지시 학습용 데이터, 평가·검증용 데이터, 공통 서비스 학습용 데이터로 구분된다.

원시데이터 확보 규모 ('25. 9월 기준)	원천데이터 구축 규모 ('25. 12월 기준)	가공데이터 구축 규모 ('25. 12월 기준)
282,414건	170,000건	164,000건

[표 11] 학습데이터 가공 단계별 결과물과 수량

◈ [참고] 공통기반 학습데이터 가공 방법 및 절차



① 확보된 원천데이터를 기반으로, LLM 전문연구팀이 앞서 정의한 40개의 지시유형별 지시문과 응답문으로 구성된 데이터 세트를 구성함

- 가공 절차에 따라 최종 가공 완료한 학습데이터의 형식은 JSON 형태이며, 지시(Instruction), 입력(Input), 응답(Output)으로 구성되어 있음

② LLM 연구팀이 LLM을 활용하여 문서 생애주기 전 단계에 다양한 지시유형을 학습할 수 있는 데이터 후보를 가공함

③ LLM 연구팀이 LLM을 활용하여 후보 데이터에 대해 1차 품질 검수를 진행함

④ LLM 검수를 통과한 후보 데이터에 대해 별도의 검수 작업자가 선별/보완하여 최종 데이터를 도출함

2.4. RAG

○ 범정부 AI 공통기반에서는 연구·보도자료, 행정 업무자료, 법령 자료 등 행정 업무에 활용할 수 있는 약 40만 7천건의 데이터를 RAG* 형태로 제공한다.

* 검색 증강 생성(Retrieval-Augmented Generation): AI가 외부 데이터베이스(DB)에서 관련 정보를 검색하여 답변을 생성함으로써 정보의 정확성과 최신성을 높이는 기술이다.

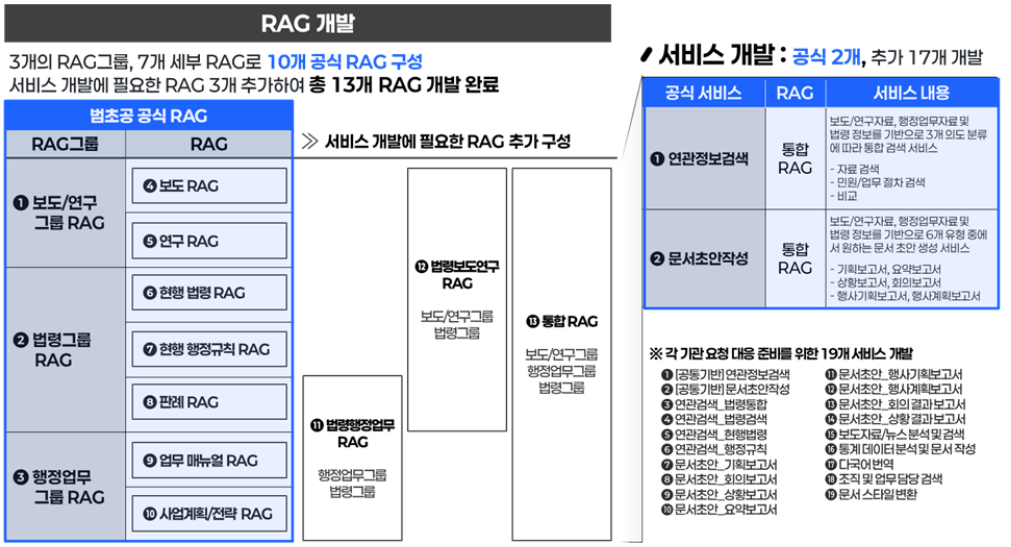
○ 개방데이터*와 공동활용데이터** 약 407,000 건을 수집하여 공통 RAG를 구축하였다.

* 중앙부처 홈페이지 자료 공유 사이트를 통해 개방된 데이터

** 국가공유데이터 플랫폼 〉 공동활용데이터 등록관리시스템에 각 공공기관들이 행정망 내에서 공유 활용 가능하도록 등록한 데이터

○ 2026년 5월 기준, 총 10개의 RAG 구축을 완료하였으며 공통기반에 등록된 RAG는 즉시 활용이 가능함

* 활용 경로: FabriX(Store 〉 Asset 〉 Knowledge), CLOVA Studio for GOV(지식저장소) 메뉴를 통해 공통 RAG 기능을 이용할 수 있다.



[그림 9] 범정부 AI 공통기반 RAG 구성도

2.5. AI 서비스

○ 범정부 AI 공통기반에서는 크게 두 가지 유형의 AI 서비스를 제공한다. 사용자는 행정 업무 전반에 활용할 수 있는 공통 AI 서비스(연관정보검색, 문서초안작성)와 Chat 서비스를 즉시 이용할 수 있다.

○ 해당 서비스들은 API 형태로도 제공되어 별도의 복잡한 개발 없이 활용이 가능하다. 다만, 제공되는 API 상세 사양은 플랫폼별로 상이하므로 사용자 매뉴얼을 통해 확인해야 한다.

[그림 10] 공통기반 AI 서비스 범위

■ 공통 AI 서비스

공통 AI 서비스는 공통기반의 10종의 RAG를 활용하여 공공행정 분야의 지식 문서를 검색하고, 이를 바탕으로 신뢰성 있는 답변을 생성하는 서비스로 '연관정보검색 서비스'와 '문서초안작성 서비스' 2종으로 구성되어 있다.

* ① 보도 RAG, ② 연구 RAG, ③ 현행 법령 RAG, ④ 현행 행정규칙 RAG, ⑤ 판례 RAG, ⑥ 업무 매뉴얼 RAG, ⑦ 사업계획/전략 RAG

- 활용 방식: API 형태로 호출하여 기관 특화 서비스 개발에 활용할 수 있으며, 호출 기관은 사용량에 따른 사용료를 지불해야 한다.

○ 연관정보검색 서비스

- 보도자료, 법령 정보 등 다양한 행정 지식 자료를 통합 검색하는 서비스이다.
- 질문 의도를 분석하여 관련성이 높은 자료를 출처와 함께 신속하게 답변한다.

○ 문서초안작성 서비스

- 사용자의 입력 내용을 바탕으로 사전에 정의된 양식에 맞춰 문서 초안을 자동 생성하는 서비스이다.

* 문서 초안 양식(6종): 기획보고서, 행사기획보고서, 행사계획보고서, 요약보고서, 상황보고서, 회의보고서

■ Chat 서비스

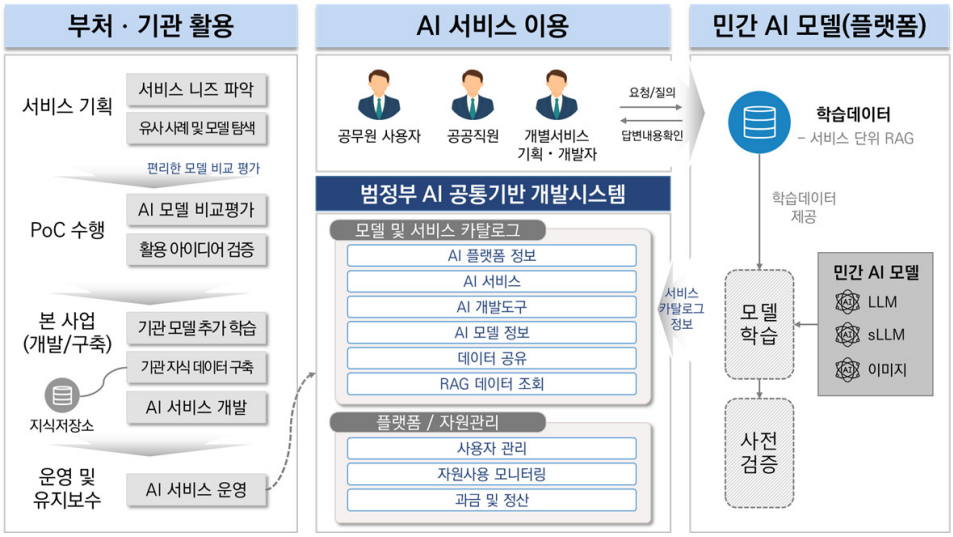
Chat 서비스는 범정부 AI 공통기반에서 기본 제공되는 기능으로, 사용자에게 친숙한 대화형 인터페이스를 통해 자연어 기반의 질의응답을 제공하는 서비스이다.

- 이용 방법: 범정부 AI 공통기반 Chat 포털(https://chat.ai.go.kr)에 접속하여 원하는 플랫폼(FabriX 또는 CLOVA Studio for GOV)을 선택한 후 이용한다.
- 활용 방식: API 형태로 호출하여 기관별 특화 서비스 개발에 활용할 수 있으며, 호출 기관은 사용료를 지불해야 한다.

2.6. 개발시스템

○ 범정부 AI 공통기반을 각 이용 기관에서 효율적으로 활용할 수 있도록 통합 창구를 제공한다.

- AI 플랫폼의 자원 및 기능 모듈 정보를 제공하여, AI 서비스의 개발과 확장에 필요한 기반 정보를 지원한다.
- AI 플랫폼 요금 체계에 따라 정보시스템별 사용량과 과금 현황을 모니터링하며, 전일까지의 누적 금액 및 월별 집계 금액을 실시간으로 조회할 수 있다.



[그림 11] 범정부 AI 공통기반 개발시스템 구성도

■ 개발시스템 주요 기능

○ AI 플랫폼 이용 현황 관리

- AI 플랫폼과의 연계를 통해 각 기관 정보시스템의 이용 현황을 체계적으로 관리할 수 있는 기능을 제공한다.

○ 사용자 등록

- 기관 내 정보시스템에서 AI 플랫폼을 활용할 수 있도록 사용자 정보를 등록하고 적절한 권한을 부여한다.
- 이용 현황을 실무적으로 관리할 담당자를 지정할 수 있는 기능을 지원한다.

○ **대시보드 조회**

- 정보시스템별 API 호출 횟수 및 토큰 사용량을 조회하여 월별 사용 현황과 일별·월별 변동 추이를 그래프 형태로 제공한다.
- 이용 약정 정보를 기반으로 산정된 사용 요금을 표시하며, AI 플랫폼의 LLM 모델별 상세 토큰 사용량을 확인할 수 있다.
- 플랫폼에서 제공하는 AI 서비스·모델·개발도구에 대해 사용자가 평가한 별점 순위(Top 5) 정보를 제공한다.

○ **카탈로그 제공**

- AI 플랫폼에서 제공하는 AI 서비스, 모델, 개발도구, RAG 데이터 정보를 체계적으로 등록하여 관리한다.
- 범정부 공통 AI 서비스 2종과 각 부처·기관에서 자체 구축한 AI 서비스까지 통합하여 조회할 수 있는 카탈로그 기능을 제공한다.

○ **데이터 공유**

- 공통기반에 구축된 학습데이터를 유형별(원시·원천·가공)로 구분하여 조회할 수 있는 기능을 제공한다.
- RAG 데이터 구축 등 필요 시 데이터를 다운로드하여 활용할 수 있도록 공유 기능을 지원한다.

○ **이용 요금 시뮬레이션(엑셀) 제공**

- AI 서비스 개발을 검토하는 기관의 예산 편성을 돕기 위해, 플랫폼별 이용 요금을 사전에 산정해 볼 수 있는 시뮬레이션 엑셀 파일을 제공한다.
- 이를 통해 서비스 기획 단계에서 비용 구조를 검토하고, 예산 산정을 위한 공신력 있는 참고 자료로 활용할 수 있다.

3. 공통기반을 활용한 AI 서비스 구축 프로세스

[그림 12] 공통기반 활용 프로세스

3.1. AI 서비스 컨셉 정의

○ 도입 예정 AI 서비스 정의

- 기관에서 도입하고자 하는 AI 서비스의 명칭, 이용 대상 및 규모, 주요 기능을 구체적으로 정의한다.
- 서비스 이용 대상은 사용자가 공무원인지 일반 국민인지 등을 명확히 구분하며, 서비스 규모는 1일 예상 사용자 수 또는 사용자 1인당 일평균 이용 횟수 등을 바탕으로 산정한다.

○ 기관 특화 RAG용 데이터 정의

- 기관의 업무 특성에 맞춘 전용 RAG(검색 증강 생성) 구축 필요 여부를 사전에 판단한다.
- 특화 RAG가 필요한 경우 세부 내용을 정의하고, 연계 대상 서비스와의 관계를 맵핑(Mapping)한다.
- 데이터별 수집 출처를 파악하고 목표 수집 건수, 방식, 주기를 설정하여 현실적인 데이터 확보 계획을 수립한다.

○ 공통기반을 활용한 서비스 구현 방안 검토

- 기관의 AI 서비스가 공통기반과 연계되어 동작하는 구조를 데이터, 모델, 서비스의 흐름이 포함된 아키텍처(Architecture)로 정의한다.
- 공통기반의 AI 에이전트 빌더(FabriX: AI Lab, CLOVA Studio for GOV: Agent Builder)를 활용하여 AI 모델, 시스템 지시문, 기반 지식(RAG), 플러그인의 4가지 요소를 조합한 최적의 구현 방안을 검토한다.

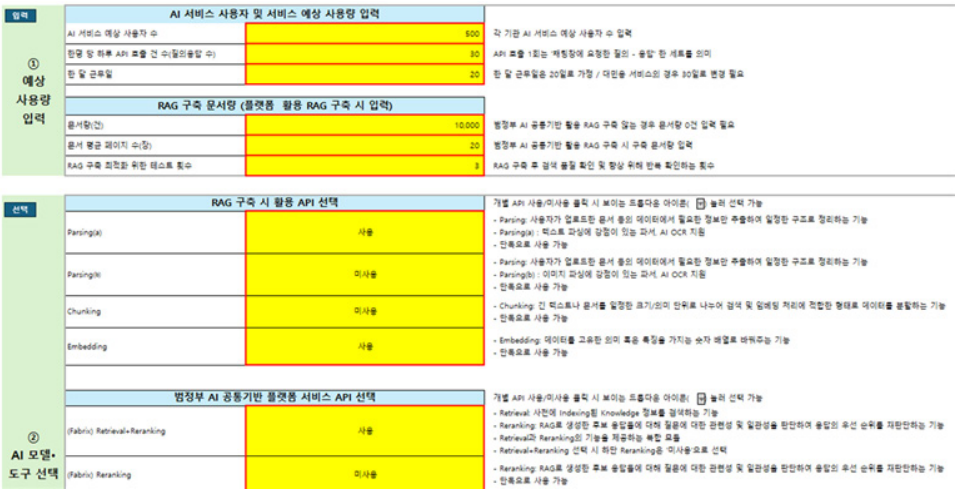
3.2. 예산 및 계획 수립

○ 공통기반 활용 예산 시뮬레이션

- 기관에서 활용할 공통기반 AI 자원(AI 모델, RAG 구축 여부, 보안 서비스 적용 여부 등)을 구체적으로 정의한다.
- 구현하고자 하는 AI 서비스의 예상 이용자 수, 호출 횟수, RAG용 데이터 규모를 설정한다.
- 정의된 조건을 토대로 시뮬레이션을 진행하여 API 이용료와 RAG 구축 비용에 대한 예산(안)을 확인한다.

※ 예산 시뮬레이션에 필요한 엑셀 파일은 '공통기반 개발시스템 〉 이용 안내 〉 요금 시뮬레이션 [범정부 AI 공통기반_예산 시뮬레이터.xlsx]'를 다운로드하여 사용(사용 방법은 파일 내 포함)할 수 있다.

[예시] 범정부 AI 공통기반 플랫폼(FabriX) 예상 요금 산정



[그림 13] 예산 시뮬레이션 진행 예시

○ **기관이 고려하여야 하는 AI 관련 구축 예산**

- 데이터 전처리, RAG 구축 및 품질 최적화 업무 수행을 위한 개발자 인건비를 편성해야 한다.
- 지속적인 RAG 데이터 최신화를 위한 운영 인건비를 고려해야 한다.
- 정보시스템 연계가 필요한 경우, API 호출 및 처리를 위한 애플리케이션 서버 등 인프라 구축 비용을 반영한다.
- 기관 특화 서비스를 위한 추가 모델 도입이나 모델 학습 및 운영에 따른 제반 비용을 산정한다.

○ **예산 확보 및 활용 계획 수립**

- 기재부 예산 신청, 자체 예산 편성, 또는 공모 사업 참여 등을 통해 필요한 예산을 확보한다.

○ **공통기반 활용 계획 수립**

- 기관 특화 RAG DB 구축을 위한 가공 도구 활용, 서비스 에이전트를 이용한 기능 구현, 공통기반 제공 API(Agent, RAG 등) 및 공통 AI 서비스(연관 정보 검색, 문서 초안 작성, Chat 서비스) 이용 계획을 수립한다.
- 기관 내부의 정보화 추진 계획 또는 AI 서비스 운영 계획에 맞춰 공통기반 활용 시기(시범 사용, 계약, 서비스 구축 및 운영 등)를 설정한다.
- 각 기관은 설정된 활용 시기 전까지 단계별 필요 사항을 사전에 준비하여야 한다.
  - 시범 사용: 플랫폼 운영자와 협의하여 활용 가능 여부를 확인하며, 시범 사용 목적, 서비스(안), 사용 범위, RAG 구축용 데이터 등을 사전에 정의한다.
  - 계약 단계: 최종 결정된 공통기반 사용 방식에 따른 확정 예산과 관련 행정 서류를 준비한다.
  - 서비스 구축 및 운영: 실제 RAG 구축에 필요한 데이터, 확정된 서비스 아키텍처(모델, RAG 포함), 서비스 구성 요소(AI 모델, 프롬프트, 개발 도구 등)와 함께 개발·테스트·운영 인력을 확보한다.

3.3. 공통기반 이용 계약

○ **공통기반 사용 방식 결정**

- **사용 조건 선택**
  - 제공 중인 AI 플랫폼 중 기관에 적합한 환경과 기능을 갖춘 플랫폼을 선택한다. (다중 선택 가능)
- **활용 기능 확정**
  - 활용할 AI 모델, 공통 RAG, RAG 구축 도구 등을 확정하고 적합한 요금제를 선택한다.
- **사용 방식 확정**
  - RAG 구축 및 서비스 구현 활동을 공통기반 내에서 처리할지, 혹은 제공되는 API를 활용해 기관 내부 시스템에서 처리할지 결정한다.

○ **공통기반 이용 계약**

- 공통기반 이용 계약 절차를 확인하고 최종 계약을 체결한다.

3.4. 공통기반을 활용하여 AI 서비스 구현

○ 기관 AI 서비스 기획·설계

- **구성 요소 정의**
  - 서비스에 활용할 LLM을 선택하고, RAG 정의 및 세부 서비스별 RAG 맵핑을 수행한다.
  - 또한 프롬프트의 방향을 설정하고 필요한 도구(STT/TTS, 보안, 개인정보 처리 등)를 선택한다.
- **서비스 설계**
  - 서비스 및 데이터 아키텍처를 설계하고, 전체적인 서비스 플로우와 데이터 수집 방안을 정의한다.

○ 기관 특화 RAG용 데이터 수집 및 구축

- **데이터 수집**
  - 기획 단계에서 정의된 방안에 따라 내·외부 데이터를 수집한다.
- **RAG DB 구축**
  - 공통기반 기능을 활용해 데이터 전처리(파싱, 청킹, 임베딩 후 Vector DB 구축)를 수행한다.
  - 최상의 결과가 노출되도록 검색(리트리벌) 방식 설정 및 결과 우선순위 재조정(리랭킹) 작업을 진행한다.

○ AI 서비스 구현 및 배포

- AI 모델, 지시문, RAG, 플러그인 등 공통기반의 구성 요소를 조합하여 5단계 개발 과정(설계 〉 개발 〉 검증/보완 〉 평가 〉 배포)을 거쳐 서비스를 구현한다.
- 최초 배포 이후, 공통기반을 통해 수정 및 배포 이력에 대한 버전 관리를 수행한다.

3.5. AI 서비스 운영

○ 서비스 지속 최적화 및 모니터링

- RAG 데이터의 최신성을 유지하기 위해 지속적으로 데이터를 수집하고 전처리한다.
- 서비스 모니터링 및 사용자 피드백을 반영하여 프롬프트를 조정한다.

○ 공통기반 서비스 운영 및 과금 모니터링

- **서비스 운영 계획 수립**
  - 공통기반을 활용하여 운영 중인 AI 서비스의 운영 체계, 역할과 책임(R&R), 프로세스 등을 수립한다.
  - 개발시스템에 접근하여 서비스를 관리할 사용자를 등록한다.
- **서비스 운영 및 모니터링**
  - 공통기반 개발시스템을 통해 서비스 운영 현황과 과금 내역을 실시간으로 모니터링한다.

○ 장애 대응 및 서비스 개선

- 서비스 장애 발생 시 신속하게 대응하며, 정보 업데이트가 필요한 경우 추가 데이터를 수집하여 RAG DB를 재구축한다.
- RAG 추가, 프롬프트 변경, 서비스 개선 요구 사항에 따라 지속적인 기능 개선 활동을 수행한다.

4. 공통기반 활용 참고사항

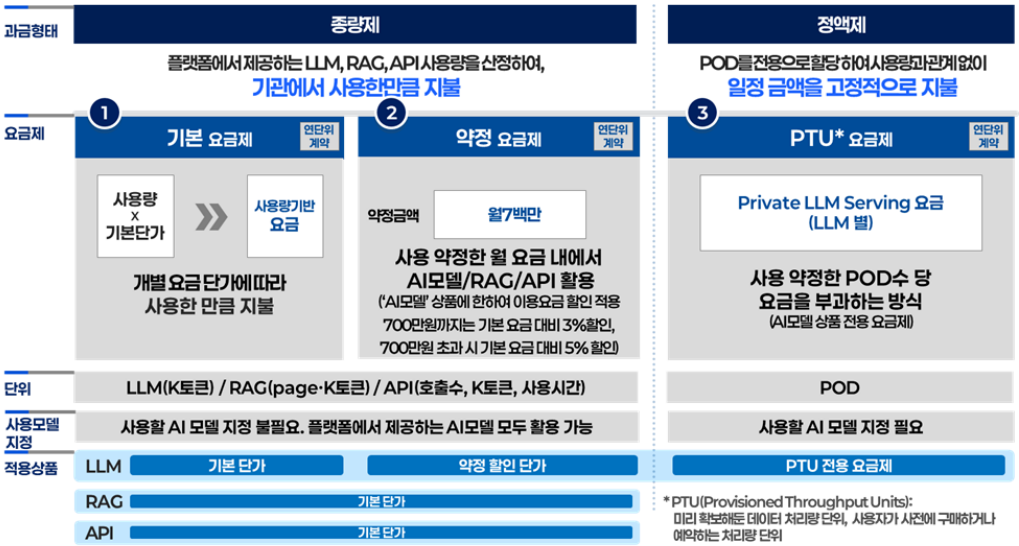
4.1. 공통기반 활용 소요 예산 산정

○ 과금 체계 및 요금제 선택

- 범정부 AI 공통기반의 과금 체계는 개별 기관의 사용 형태에 따라 결정된다.
- 사용량에 따라 비용을 지불하는 '종량제'와 기관의 수요에 따라 일정 금액을 고정적으로 지불하는 '정액제'로 구분되며, 각 기관의 운영 환경에 적합한 요금제를 선택할 수 있다.

○ 비용 포함 항목 및 기준

- 과금 체계에서 제시하는 모든 가격은 부가가치세(VAT)가 포함된 금액이다.
- 이용료에는 AI 플랫폼 활용에 필요한 인프라(GPU, CPU, 스토리지 등)와 솔루션(Kubernetes, Vector DB, 파서 등) 사용 비용이 모두 포함되어 있다.



[그림 14] 공통기반 과금 체계 유형

○ 과금 체계 유형

① 종량제

플랫폼에서 제공하는 LLM, RAG, 도구(Tools)의 실제 사용량을 산정하여 기관이 사용한 만큼 비용을 지불하는 요금제이다.

- **기본 요금제**
  - 사전 정의된 개별 요금 단가를 기초로 사용량만큼 지불한다.
  - 요금 산정 방식 = 사용량 x 기본 단가
- **약정 요금제**
  - 사용 약정한 월 요금 내에서 활용하는 요금제로, 약정 요금제 선택 시 AI 모델은 기본요금에서 할인된 가격으로 책정됨.
  - 약정 요금은 API 이용 금액을 선 차감 후, 700만 원까지는 기본요금 대비 3% 할인, 700만 원 초과 시 기본요금 대비 5% 할인됨
  - 요금 산정 방식 = 사용량 x 약정 단가

② **정액제**

기관별로 LLM 서빙 유닛(Serving Unit)을 분리하여 할당함으로써 안정적인 성능을 보장하는 방식이다. 전용 POD(LLM Serving Unit)를 할당받는 대신 고정 금액을 지불한다.

- 이용 기관은 사용할 LLM 모델을 사전에 선택하여야 한다.
- 선택한 모델 외에 다른 모델이나 API를 사용할 경우 기본 요금이 적용된다. (약정 체결 시 약정 요금 적용 가능)
  - 요금 산정 방식 = 1 POD 당 PTU*요금 x POD 수

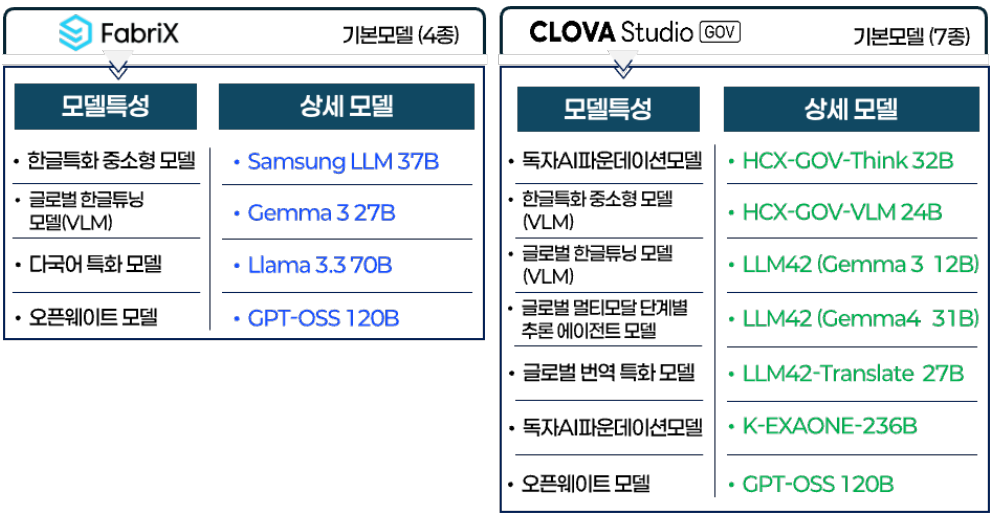
※ PTU(Provisioned Throughput Units): 사전에 확보한 데이터 처리량 단위로, 사용자가 미리 구매하거나 예약할 수 있는 처리 용량을 의미한다.

○ 과금 체계 상세

공통기반에 탑재된 LLM 모델은 모델별 토큰 사용량을 기준으로 과금하며, 도구(RAG 및 API)는 토큰 사용량 및 호출 횟수를 기준으로 비용을 산정한다.

- 현재 토큰 사용량별로 과금이 책정된 모델은 FabriX 4개, CLOVA Studio for GOV 7개 등 총 10개이다. 또한, 공통기반 서빙 모델 30종에 대한 PTU 단가가 별도로 책정되어 있다.
- AI 모델을 이용하고자 하는 기관은 플랫폼 운영자를 통해 상세한 PTU 단가를 확인한 후 이용하여야 한다.

* [부록1]'공통기반 서빙 모델' 참고



[그림 15] 플랫폼별 과금 산정 모델

- 도구(RAG 및 Tools)는 RAG 구축을 위한 도구와 생성형 AI 서비스를 사용하는 도구로 구분된다.

구분	형태	기능 옵션	단위	비고	
RAG 구축 API 요금	데이터 전처리	Parsing(a)	1페이지	FabriX 전용	RAG 초기 구축비용 / 운영·유지보수 기간 내 추가데이터에 대한 파싱·청킹·임베딩 등 데이터전처리 비용
		Parsing(b)	1페이지	FabriX / CLOVA Studio for GOV 사용	
		Chunking	K토큰		
		Parsing(c) + Chunking	1페이지 +K토큰	CLOVA Studio for GOV 전용	
		Embedding	K토큰		
서비스 API 요금	검색	Retrieval + Reranking	1호출		
		Reranking	1호출		
		RAG Chat	1호출	Retrieval + Reranking + Security Filter API + LLM 질의응답을 사용하는 기능	
		Chat	1호출	LLM에게 직접 질의하고 답변을 받는 기본적인 API	
		Code Interpreter	1호출	LLM이 파이썬 코드를 생성해주고 실행해주는 API	
	비식별화	PII Masking (Personal Indentifiable Information Masking)	1호출	데이터 전처리 시 개인식별정보(주민등록번호, 운전면허증번호, 여권번호, 휴대전화)를 비식별화하는 기능	
	음성변환	TTS(Text-To-Speech)	K토큰	텍스트를 음성으로 변환	
		STT(Speech-To-Text)	1초	음성을 텍스트로 변환	
	보안	Security Filter API	1호출	챗봇 대화 중 개인/민감정보 자동 식별 및 보호, 설정한 기밀정보 유출 여부 감지	

[표 12] RAG 구축 API / 서비스 API 요금 구성표

4.2. 공통기반 활용 계획 수립

4.2.1. 사용 환경 점검

○ AI 플랫폼은 국가정보자원관리원 대구센터의 삼성 클라우드를 통해 행정망에서 제공되므로, 사용 환경에 따른 사전 점검이 반드시 필요하다.

○ 활용 계획의 실현성을 검토하기 위해 기관의 네트워크 환경에 따라 아래 담당 부서에 방화벽 정책을 확인하여야 한다.

- **(공무원 행정망 대역 이용 기관인 경우)** 기관 전용 백본 네트워크 담당자에게 아웃바운드 정책을 확인한다. 단, 공통기반 Chat 서비스 접속을 위한 10번대 행정망 IP는 현재 개방된 상태이다.
  - 문의: 국가정보자원관리원 대전센터 방화벽 담당 (042-250-5747)
- **(공무원 행정망 대역 미이용 기관인 경우)** 기관 보안 담당자에게 방화벽 정책을 먼저 확인한 후, 인바운드 정책에 대해 문의한다.
  - 문의: 국가정보자원관리원 대구센터 PPP Zone 방화벽 담당 (053-669-6513)
- **(국가정보통신망 이용 기관 혹은 용역업체인 경우)** 중앙 HUB 백본 방화벽 운영자에게 정책을 문의한다.
  - 문의: 국가정보자원관리원 대전센터 방화벽 운영자 (042-250-5787)

○ TMS(Threat Management System) 연계

- 국가정보원「국가 클라우드 컴퓨팅 보안관제 가이드라인」준수를 위하여 AI 플랫폼 계약 시, 이용 기관은 SaaS 보안관제를 위한 솔루션(전용 네트워크 보안 시스템, 트래픽 복호화, VPN 등) 도입 및 운영에 필요한 제반 비용을 예산에 반영하여야 한다.

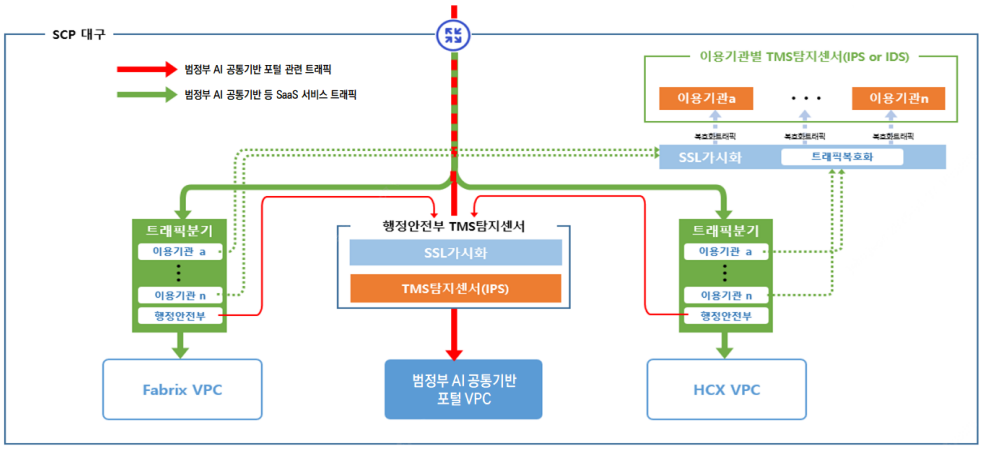
- 클라우드 영역에 구축된 보안 시스템은 이용 기관이 지정한 보안관제센터의 TMS와 연동하며, 최종적으로 국가사이버안보센터 TMS와 상시 연동되어야 한다.

- 각 기관은 SaaS 서비스 사용 전 자체적인 보안성 검토를 완료하고 직접 보안관제를 수행하는 것이 원칙이다.

  • SaaS 서비스 제공자는 보안관제를 위한 이용 기관별 트래픽 분리를 제공함

  • 분리된 트래픽은 이용 기관별 보안관제 솔루션을 이용하여 위협 탐지를 수행함

  • 다른 기관이 운영하는 보안관제 시스템을 활용하는 것이 더 효율적인 경우에는 관제 업무 위탁이 가능함 (또는 보안관제 전문업체의 인원 파견 가능)



[그림 16] 공통기반 보안관제 구성도 (예시)

4.2.2. RAG 구축 계획 수립

○ 공통기반이 제공하는 RAG 및 RAG 구축 도구의 활용 계획 수립

- 공통기반에서 제공하는 RAG 데이터의 필요성, 개발 및 유지보수의 용이성, 향후 확장성 등을 종합적으로 고려하여 공통 RAG의 활용 범위를 검토한다.

- 자체 RAG를 구축할 경우, 공통기반이 제공하는 API 형태의 RAG 구축 도구를 어느 정도 활용할 것인지 결정하여야 한다.

- 공통기반에서 제공하는 RAG 구축 관련 도구는 API 기능 모듈 형태로 제공하고 있으므로 (1) 개별 기관 자체 구축, (2) 일부 API 활용 구축, (3) 전체 공통기반 활용 구축 등 공통기반 활용 범위를 설정할 수 있다.

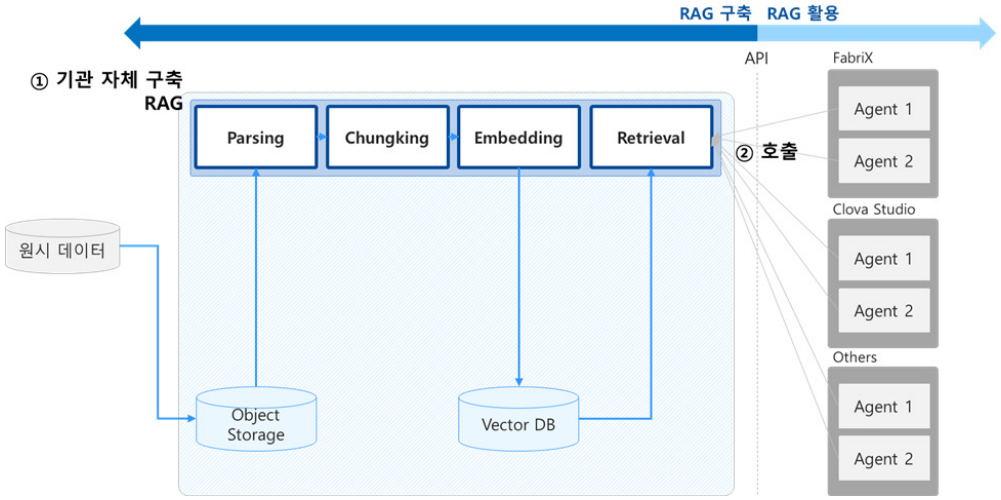
■ RAG 구축 Case 1 : 개별기관 자체 RAG 구축

○ 구축 방식

- 기관이 개별 인프라에 직접 설치하거나 자체 구축 도구를 활용하여 구현한 RAG 기능(파싱, 청킹, 임베딩, 리트리벌 등)을 공통기반 플랫폼에서 호출하여 활용하는 방식이다.

○ 장점 및 단점

- 장점: 기관별 특성에 맞춘 구축 자유도가 보장되며, 특정 플랫폼에 대한 의존성을 낮출 수 있다.
- 단점: 시스템 개발과 유지보수 전반을 기관이 직접 책임져야 하며, 최신 AI 기술을 지속적으로 적용하고 관리하는 데 상당한 자원과 노력이 소요된다.



[그림 17] 개별기관 자체 RAG 구축 예시

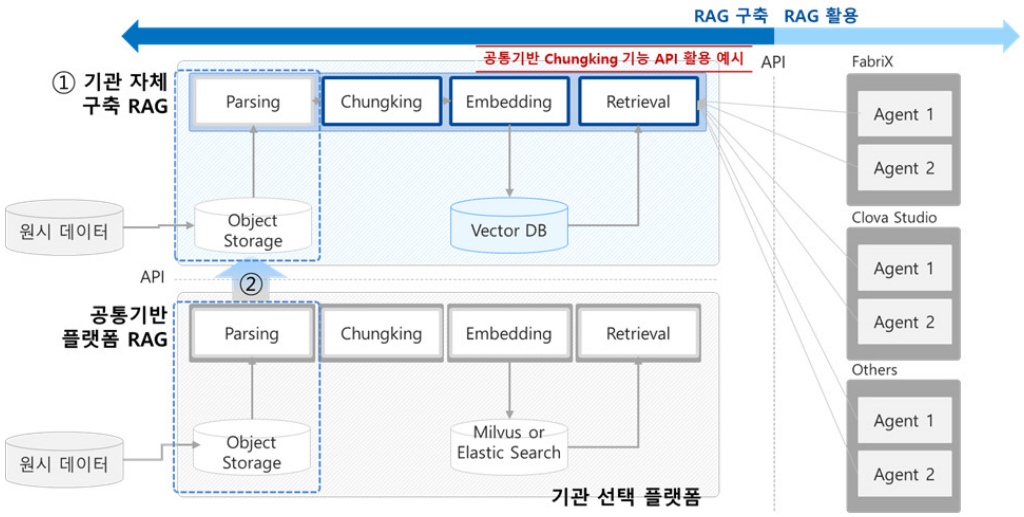
■ RAG 구축 Case 2 : 공통기반 API 일부 활용해 RAG 구축

○ 구축 방식

- 기관이 자체적으로 구현 가능한 RAG 기능은 직접 구축하고, 고도화된 기술이 필요한 일부 기능은 공통기반 플랫폼에서 제공하는 RAG API를 연계하여 구현하는 방식이다.

○ 장점 및 단점

- 장점: 기관의 자체 구축 역량과 공통기반 AI 플랫폼의 기술적 장점을 선택적으로 결합하여 유연하게 적용할 수 있다.
- 단점: 기관이 자체적으로 구축한 영역과 공통기반 플랫폼 API 활용 영역을 동시에 관리해야 하므로, 유지보수시 중복 투자가 발생하거나 관리 비용이 증가할 우려가 있다.



[그림 18] 공통기반 플랫폼 일부 기능 활용 RAG 구축 예시

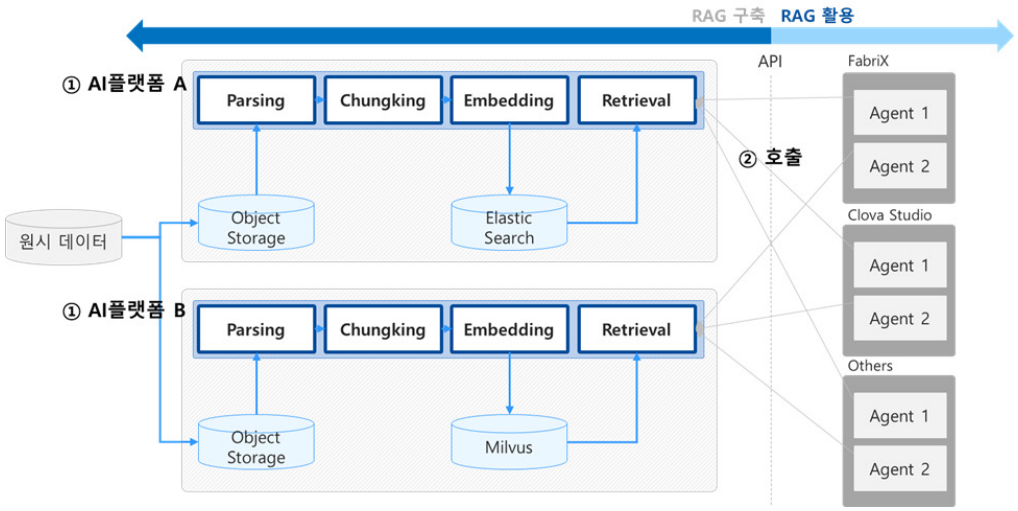
■ RAG 구축 Case 3 : 공통기반 전체 활용 RAG 구축

○ 구축 방식

- 사용 모델과 RAG용 데이터의 특성을 고려하여 적합한 공통기반 플랫폼을 선택한 후, 해당 플랫폼에서 제공하는 모든 RAG 기능을 활용하는 방식이다.

○ 장점 및 단점

- 장점: 기관별로 별도의 구축 비용을 들일 필요 없이 즉시 적용이 가능하며, 플랫폼에서 제공하는 최신 RAG 기술을 최적화된 품질로 이용할 수 있다.
- 단점: 공통기반 AI 플랫폼에서 제공하는 기능에 전적으로 의존하게 되므로, 특정 플랫폼에 대한 의존성이 발생할 수 있다.



[그림 19] 공통기반 플랫폼 활용 RAG 구축 예시

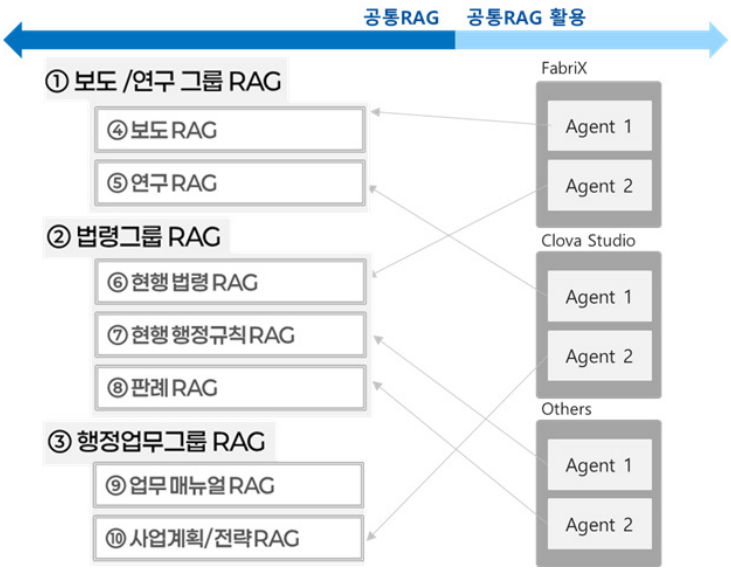
■ RAG 구축 Case 4 : 공통 RAG 활용

○ 구축 방식

- 범정부 AI 공통기반에 이미 구축되어 있는 공통 RAG를 활용하는 방식이다. 이용 기관은 플랫폼을 선택한 후, 해당 플랫폼 내에 마련된 공통 RAG를 지정하여 사용한다.

○ 장점 및 단점

- 장점: 별도의 데이터 수집이나 구축 과정 없이 즉시 행정 업무에 적용할 수 있으며, 공통으로 관리되는 고품질의 RAG 데이터를 경제적으로 이용할 수 있다.
- 단점: 공통기반 플랫폼의 서비스 환경에 귀속되므로 시스템 운영 및 기능 확장에 있어 플랫폼 의존성이 존재한다.



[그림 20] 공통 RAG 활용 예시

4.3. 공통기반 이용계약 체결 및 이용 신청

4.3.1. 이용 계약 체결

○ 「국가를 당사자로 하는 계약에 관한 법률」제7조(계약의 방법) 및 「행정기관 및 공공기관의 클라우드컴퓨팅 서비스 이용 기준 및 안전성 확보 등에 관한 고시」제5조(계약방법)에 따라 경쟁 입찰하여야 함

- 현재 제도상 한계로 중앙·지방정부가 공통기반을 이용 시 조달청 '카탈로그 계약'과 과기부 '디지털서비스 전문이용계약'불가능

※「국가를 당사자로 하는 계약에 관한 법률」시행령 제26조(수의계약에 의할 수 있는 경우) 제1항 제5호 아목 대상이 아님(과기부 미선정)

○ 중앙·지방정부가 국정자원 대구센터 PPP존 내 구축된 공통기반의 자원(AI 플랫폼, 모델, 서비스 등)이용 시 AI 정보화 사업(일반용역)에 공통기반 이용료(AI 구독료)를 포함하여 통합 발주하는 것이 효율적임

- 1) AP 구축 사업 또는 운영·유지 관리 사업에 포함하여 발주
- 2) AP 구축이 없거나 상용 SW 구매 시 MSP 사업(CSP 포함) 발주

※ 사업의 특성 상 필요한 경우 AP 구축/구매, MSP, CSP 각각 공개경쟁입찰 가능

4.3.2. AI 플랫폼 이용 신청

○ 공통기반 이용 계약이 완료된 후, 실제 플랫폼을 사용하기 위해 아래의 절차에 따라 신청을 진행하여야 한다.

- **시스템 접속:** 공통기반 개발시스템(dev.ai.go.kr)에 로그인한다.
- **정보 입력:** 이용하고자 하는 정보시스템의 명칭과 담당자 정보(소속, 성명, 연락처, 휴대전화, 이메일)를 정확히 입력한다.
- **증빙 서류 제출:** 공통기반 이용 계약 관련 자료(계약서 사본, 사업계획서 또는 제안요청서 등)를 첨부하여 이용 신청을 완료하여야 한다.

4.3.3. 테넌트 할당

- **(CLOVA Studio for GOV)** 이용 기관의 정보시스템 담당자가 개발시스템을 통해 AI 플랫폼 이용 신청을 완료하면, 공통기반 관리자가 확인 후 테넌트 코드를 부여한다.

- **(FabriX)** CLOVA Studio for GOV와 달리 테넌트 코드가 시스템상에서 자동 생성되지 않는다. 따라서 이용 기관의 정보시스템 담당자는 플랫폼사와 계약 시, FabriX 플랫폼 운영자에게 별도로 사전 신청하여 테넌트 코드를 발급받아야 한다.

  ** 정보시스템 담당자 → 공통기반 관리자(NIA 전달) OR 플랫폼 담당자 → NIA 전달

  ** 이용 신청 안내서 내 '정보시스템 이용 정보 등록' 절차 중, 이용 신청 정보 등록을 완료한 이후의 세부 단계를 반드시 확인

◈ 참고: AI 플랫폼의 테넌트

○ AI 플랫폼의 테넌트는 정보시스템과 **1:1로 매핑(Mapping)**되는 개념이다. 이는 각 테넌트마다 독립적인 데이터 환경을 구축하여 제공함을 의미한다.

○ 각 테넌트에 소속된 사용자는 권한이 없는 다른 테넌트의 설정이나 데이터에 절대 접근할 수 없다. 이러한 구조는 다중 사용자 및 다중 조직 환경에서 데이터 격리와 보안을 확보하는 핵심 메커니즘으로 작동한다.

○ 각 이용 기관이 사용할 AI 플랫폼은 하나의 정보시스템 단위로 관리된다. 따라서 해당 기관의 정보시스템 담당자가 AI 플랫폼의 테넌트 관리자 역할도 병행하여 수행하게 된다.

○ 테넌트를 생성하기 위해서는 테넌트 코드와 테넌트 관리자 계정 정보가 기본적으로 필요하다. 특히 테넌트 코드는 AI 플랫폼의 내부 운영 및 외부 시스템 연계 시 해당 테넌트를 고유하게 식별하기 위한 목적으로 사용된다.

4.3.4. 정보시스템 승인

- 공통기반 관리자는 이용 기관이 제출한 계약 관련 자료를 검토한 후, 개발시스템을 통해 접수된 내용을 바탕으로 플랫폼 약정 정보(테넌트 코드, 계약 정보 등)를 입력하고 관리한다.

- 이용 신청 정보 등록이 완료되면 해당 정보시스템 담당자에게 이메일로 승인 결과를 안내한다.

- 등록된 정보는 정보시스템 담당자의 요청에 따라 사후 변경이 가능하다.

4.3.5. 서비스 활용

- (공무원) GPKI 인증서 또는 모바일 공무원증으로 본인 인증을 거친 후 계정을 활성화하여 로그인한다. (최초 로그인 시에는(@korea.kr) 이메일을 통한 추가 인증이 필요하다.)

- (공공기관 직원*) 이메일과 이름을 입력한 후, 해당 이메일로 발송된 OTP 인증 번호를 입력하여 로그인한다.

- (민간 개발자**) 정보시스템 담당자가 서비스 개발을 위해 등록한 용역사 개발자를 의미한다. 등록된 개발자는 부여받은 권한으로 이메일을 이용하여 로그인한다.

  * 중앙·지방정부로부터 업무를 위임·위탁받은 정보시스템 관련 담당자가 이에 해당한다.

◈ 공통기반 사용자 구분

○ 범정부 AI 공통기반 사용자는 사용 권한에 따라 일반사용자(공무원 또는 공공기관 직원), 정보시스템 담당자, 정보시스템 개발자로 구분되며, 각 유형에 맞는 로그인 방식을 제공한다.

구분	개발시스템			AI플랫폼		
	사용자 관리	이용 현황/ 요금 조회	카탈로그/ 이용안내	Chat/ 공통서비스	개발도구	테넌트 관리
일반사용자* (공무원, 공공직원)			○ (조회)	○ (조회)		
정보시스템 담당자	○ (등록/수정)	○ (조회)	○ (조회)	○ (조회)	○ (생성/수정/삭제)	○ (생성/수정/삭제)
정보시스템 개발자			○ (조회)	○ (조회)	○ (생성/수정/삭제)	

[표 13] 사용 권한 별 공통기반 사용자 구분

○ 사용 권한과 관계없이 AI 플랫폼 사용 중 발생하는 모든 오류 및 문의 사항은 각 AI 플랫폼별로 마련된 전용 VOC(Voice of Customer) 채널을 활용하여 해결한다.

구분	1선	2선
FabriX	플랫폼 운영자 (플랫폼 내 Q&A게시판 활용)	fabrix.cs@samsung.com
CLOVA Studio for GOV	시스템 문의 창구는 없음 ('26년 내 추가 예정)	dl_clovastudio_gov@navercorp.com

[표 14] AI 플랫폼별 문의처

이용신청 절차	수행 주체	내용
AI 플랫폼 이용 신청	개발시스템 이용 신청기관	- 일반사용자로 로그인하여 진입한 공통기반 개발시스템의 카탈로그 관리 혹은 이용 안내 메뉴 하단에서 "AI 플랫폼 이용 신청" 버튼을 클릭 - 이용 방법을 확인한 뒤, 정보시스템 및 관리자 정보를 입력하고 공통기반 이용 계약 등 관련 서류를 첨부하여 저장
신청서류 확인	공통기반 관리자 (개발시스템 관리기관)	- 신청기관의 제출 서류를 확인하고 서비스 이용 충족 여부 확인 - 이용 신청이 반려*된 기관에 반려 사유를 신청기관 담당자 이메일로 전달 * 다음 각 호의 어느 하나에 해당하는 경우 신청을 반려할 수 있음 1. AI 플랫폼 이용 대상 또는 대상 업무에 해당하지 않는 경우 2. 계약 및 이용 신청 관련 첨부 등 필수 제출 서류가 누락된 경우 3. 중복·반복 신청으로 판단되는 경우
해당 정보시스템의 이용 정보 등록	공통기반 관리자 (개발시스템 관리기관)	- 정보시스템의 AI 플랫폼 이용 신청 정보(테넌트 코드, 계약 정보 등)의 등록 및 계정 동기화 - 이용 신청 정보 등록 완료 후 신청기관의 정보시스템 담당자 이메일로 전달
개발시스템 로그인	개발시스템 이용 신청기관	- (공무원) GPKI 인증서나 모바일 공무원증으로 인증 진행. 최초 로그인 시 이메일 추가 인증(최초 등록한 이메일과 일치 할 경우 로그인) - (공공직원 또는 민간 개발자) 등록된 이메일 계정으로 수신한 OTP 인증번호를 입력하여 인증 후 로그인
AI 서비스 활용	개발시스템 이용 신청기관	- AI 플랫폼 이동 및 담당 정보시스템의 AI 서비스 개발 진행 - 정보시스템 개발자 계정생성 및 권한 부여

[표 15] 정보시스템 이용 신청 절차

4.4. AI 서비스 개발/운영

4.4.1. AI 플랫폼 테넌트 관리

각 이용 기관의 정보시스템은 AI 플랫폼의 테넌트와 1:1로 매핑(Mapping)되어 운영되며, 이에 따라 이용 기관의 정보시스템 담당자는 해당 테넌트의 보안과 설정을 총괄하는 테넌트 관리자 역할을 함께 수행해야 한다.

플랫폼명	FabriX	CLOVA Studio for GOV
테넌트 관리 항목	- 공지사항 관리 - 권한 관리 - Open API 관리	- 공지사항 관리 - 권한 관리

[표 16] AI 플랫폼 별 테넌트 관리 항목

가) FabriX

FabriX의 경우 아래 URL로 제공되는 Admin 포털을 통해 테넌트 관리 기능을 제공하며, 상세한 기능 설명은 아래 경로에서 제공하는 관리자 매뉴얼을 참고한다.

○ FabriX Admin 포털: https://admin.fabrix.go.kr

○ FabriX 테넌트 관리 항목

- 공지사항 관리, 권한 관리(관리자/사용자), Open API 관리

나) CLOVA Studio for GOV

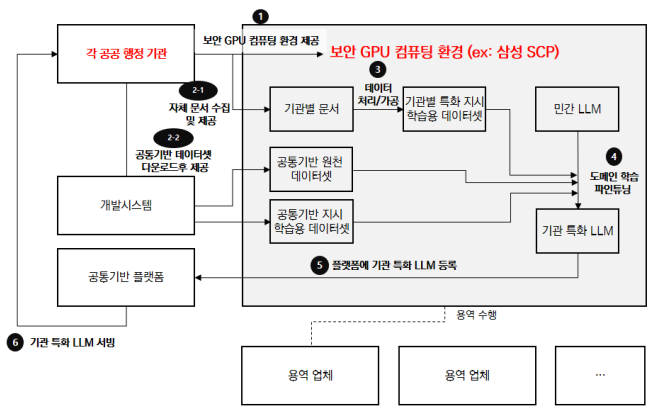
CLOVA Studio for GOV의 경우 개발시스템을 통해서 권한을 부여받고, AI 플랫폼으로 SSO 연동을 하여 접속한다. 정보시스템 담당자를 위한 별도 URL은 제공하지 않으며, 정보시스템 담당자와 개발자, 일반 사용자 각각의 경우에는 서로 다른 메뉴가 제공된다.

○ CLOVA Studio for GOV 테넌트 관리 항목

- 공지사항 관리, 권한 관리(정보시스템 담당자/정보시스템 개발자), Open API 관리

4.4.2. 학습데이터 활용 방법

○ 범정부 AI 공통기반에서는 일반적인 공공·행정업무를 잘 이해하고 수행할 수 있도록 학습시키기 위한 지시학습용 데이터를 제공하고 있다. 개별 기관이 AI 서비스 구축 시 기관 특화 학습데이터와 함께 공통기반에서 제공하는 지시학습용 데이터를 병행 학습할 경우 특화된 지식 뿐만 아니라 일반적인 공공·행정 지시도 잘 이해하는 서비스를 구현할 수 있다.



[그림 21] 원천/지시 학습용 데이터 활용 흐름도

No.	단계	내용
1	보안 GPU 컴퓨팅 환경 제공	각 기관이 특화 LLM 파인튜닝을 위한 보안 GPU 컴퓨팅 환경 준비
2	자체 문서 수집 및 제공	각 기관이 자체적으로 문서를 수집하여 보안 GPU 컴퓨팅 환경에 업로드
	공통기반 데이터셋 다운로드 후 제공	각 기관이 범정부 AI 공통기반 개발시스템에 접속, 데이터셋 다운로드 후 GPU 컴퓨팅 환경에 업로드
3	데이터 처리/가공	보안 GPU 컴퓨팅 환경 내 용역 업체 참여를 통한 기관별 수집 문서 처리 및 가공, 기관별 특화 지시 학습용 데이터셋 준비
4	도메인 학습 파인튜닝	기관별 특화 지시 학습용 데이터셋과 공통기반 지시 학습용 데이터셋을 통합하여 도메인 학습 파인튜닝 실시 또는 공통기반 원천 데이터셋과 기관별 문서 추출 말뭉치 데이터셋을 활용하여 연속적 사전 훈련(Continued Pre-training) 방식 적용
5	플랫폼에 기관 특화 LLM 등록	학습 완료된 기관 특화 LLM을 공통 플랫폼에 등록
6	기관 특화 LLM 서빙	기관 특화 LLM을 AI 플랫폼에서 서빙하여 각 기관 특화 서비스 개발 활용

[표 17] 원천/지시 학습용 데이터 활용 절차

4.4.3. FabriX 기반 개발/운영 방법

가) FabriX 기반 개발 방법

○ FabriX 개발도구인 AI Lab은 아래와 같이 구성된 구축 파이프라인을 제공하며 각 기능을 단계적으로 활용해 RAG 및 AI 서비스를 개발할 수 있다.

[그림 22] RAG 서비스 개발

[그림 23] Agent 서비스 개발

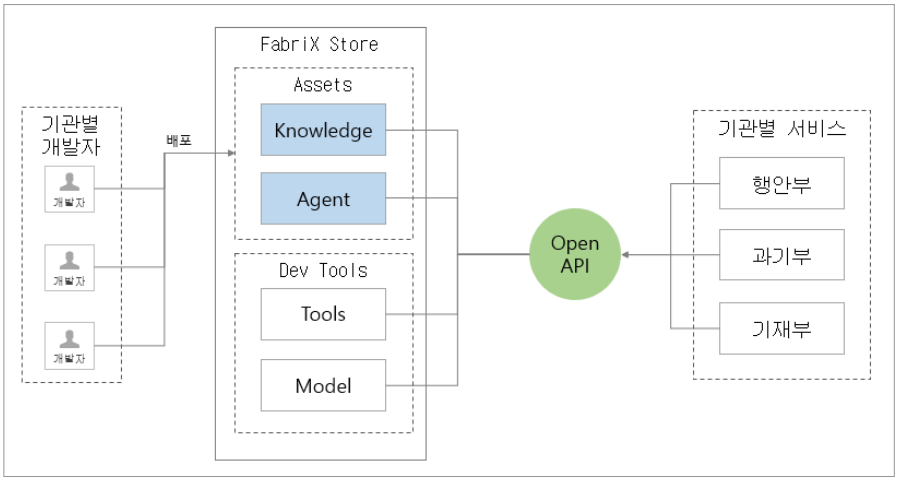
○ 부가적으로 제공되는 Evaluation은 생성된 Knowledge의 기본 성능을 평가하는 기능을, Prompt Template은 Chat에서 반복적으로 사용할 프롬프트를 저장해두는 기능을 제공한다.

○ AI Lab을 통해 개발된 RAG 및 AI 서비스는 Store를 통해 배포 가능하며 사용 권한에 따라 해당 서비스를 '사용 신청'하여 실행할 수 있다.

나) 기관 연계 및 운영 방안

○ Store 메뉴를 통해 사용자 전원 또는 일부에게 공유되는 단위를 Asset이라 부르며, 이는 개발도구를 통해 생성되는 Knowledge, Agent가 Asset에 해당한다.

○ Store를 통해 배포된 RAG 및 AI 서비스는 Open API로 외부 제공이 가능하다.



[그림 24] 기관별 AI 서비스 제공

◎ 사용자 매뉴얼의 7.2 Knowledge 사용하기 〉 Knowledge 배포하기 참고

◎ 사용자 매뉴얼의 8. Prompt Template 사용하기 〉 Prompt Template 배포하기 참고

◎ 사용자 매뉴얼의 10. Agent 사용하기 〉 Agent 배포하기 참고

◎ 사용자 매뉴얼의 11.2 Asset 사용하기 참고

다) Open API 사용 신청

○ Open API를 활용하고자 하는 사용자는 해당 AI 서비스에 대한 사용 신청을 진행하여야 하며, 기관별 관리자의 최종 승인 완료 후 사용이 가능하다.

○ FabriX의 Open API는 신청한 테넌트와 계정에 종속되어 권한이 부여된다. 따라서 권한을 부여받은 계정이 휴면 상태가 되거나 삭제되지 않도록 철저한 관리가 필요하다.

○ Open API 사용 권한은 활용 용도에 따라 개인 및 개발을 위한 Trial 권한과 운영을 위한 Prod 권한으로 구분된다. 특히 Prod 권한은 정보시스템 담당자의 승인이 필수적으로 요구된다. 각 권한별로 사용 기간 및 RPM/TPM*이 제한되며, 기본 설정된 수치를 변경하고자 할 경우 정보시스템 담당자의 검토를 거쳐 조정한 후 반영할 수 있다.

* RPM/TPM: API 사용 속도 및 사용량을 제한하기 위한 지표이다. 분당 호출 횟수와 토큰 사용량을 수치로 지정하며, 1분이 경과하면 초기화(Reset)된다.

◎ 사용자 매뉴얼의 11.3 Open API 사용하기 참고

4.4.4. CLOVA Studio for GOV 기반 개발/운영 방법

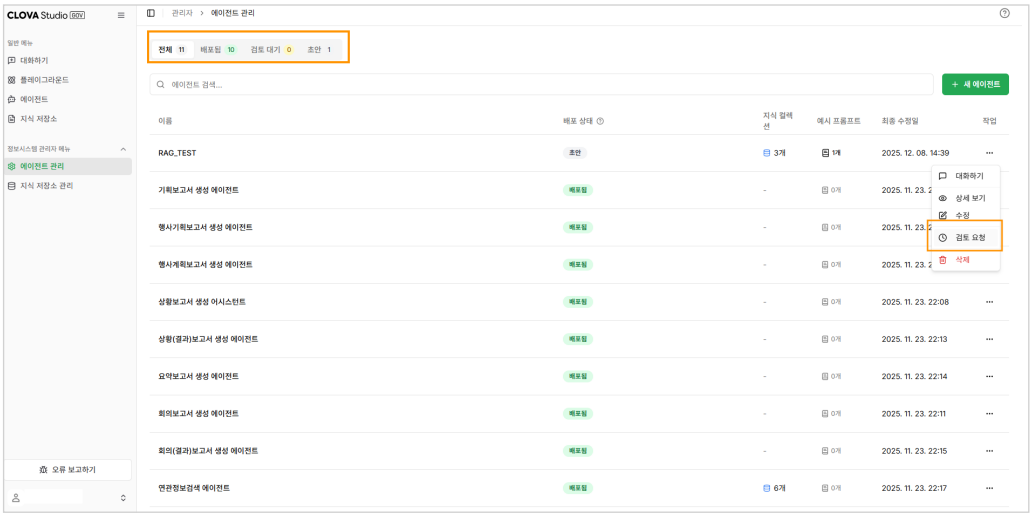
가) 개발 도구

① 에이전트 관리

○ CLOVA Studio for GOV는 문서 작성을 위한 에이전트와 RAG(Retrieval-Augmented Generation) 에이전트 제작 기능을 제공한다.

○ 에이전트 생성은 개발자 권한을 보유한 계정을 통해 가능하나, 생성된 에이전트를 해당 테넌트에 배포*하거나 삭제하기 위해서는 정보시스템 담당자의 최종 승인 절차를 거쳐야 한다.

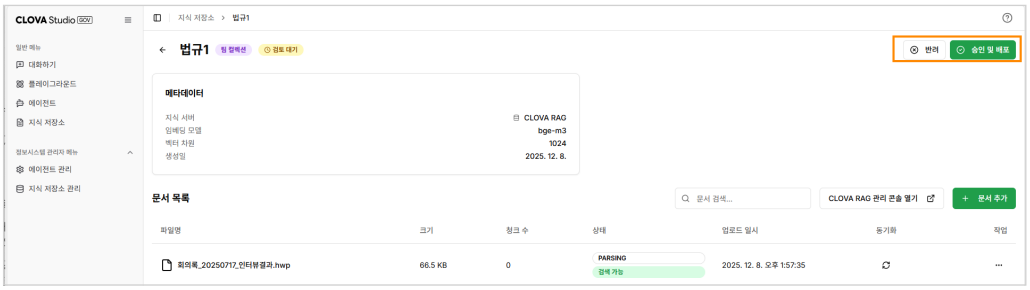
* 배포란 CLOVA Studio for GOV를 통하여 동일한 테넌트 내의 사용자들이 해당 에이전트를 사용할 수 있도록 권한을 활성화하는 기능을 의미한다.



[그림 25] CLOVA Studio for GOV 에이전트 관리 화면

② 지식저장소 관리

○ 지식저장소란, 같거나 비슷한 유형의 데이터를 생성형 AI를 활용하여 검색할 수 있도록 RAG(Retrieval Augement Generation)하는 것을 의미하며, 지식저장소 관리는 데이터를 전처리하는 과정을 확인 및 수정할 수 있도록 기능을 제공하고 있다.



[그림 26] CLOVA Studio for GOV 지식저장소 관리 화면

나) CLOVA RAG 관리 콘솔

① Dashboard

○ 지식저장소 내에서 처리된 데이터 정보(문서 수, 청크화된 수 등)를 한눈에 파악할 수 있도록 요약된 대시보드를 제공한다.

② Document

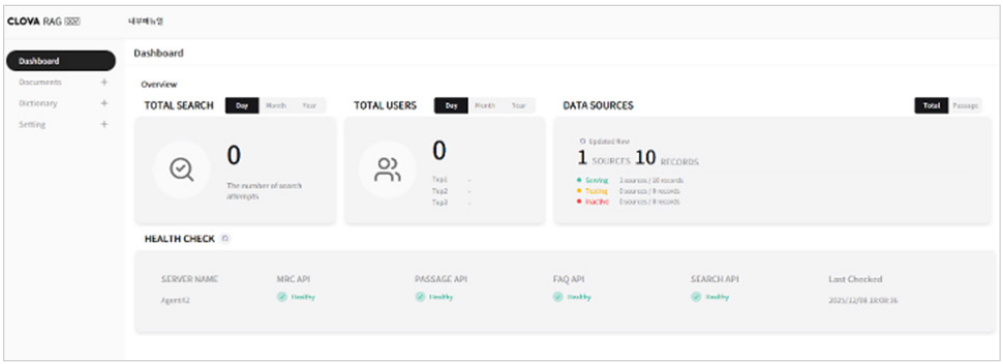
○ 지식저장소에서 처리된 데이터 정보를 확인하고 수정할 수 있는 기능을 제공한다. 또한, 전처리가 완료된 데이터의 원본 파일 상세 정보를 확인할 수 있다.

③ Dictionnary

○ RAG 검색 시 유사도가 높은 어휘의 검색 정확도를 향상하기 위하여, 특정 어휘를 용어사전으로 등록하는 기능을 제공한다.

④ Setting

○ RAG 검색 정확도 제고를 위해 사용자가 설정한 데이터 전처리 커스텀 항목을 시스템에 반영할 수 있는 기능을 제공한다.



[그림 27] CLOVA Studio for GOV RAG 관리 콘솔

FAQ 자주 묻는 질문

자주 묻는 질문은 범정부 AI 공통기반 사용과 관련하여 중앙·지방정부 담당자가 반복적으로 질의한 사항에 대한 답변을 정리한 것으로 사용자는 범정부 AI 공통기반을 활용하는 과정에서 주요 궁금증을 해소하고 업무의 효율성을 높일 수 있다.

Q1 **(계약)** 범정부 AI 공통기반 사용을 위한 계약은 어떻게 진행할 수 있나요?

A1 구축 또는 운영 사업에 포함하여 일반경쟁입찰 방식으로 공고 및 계약을 진행합니다. 향후 해당 서비스가 디지털서비스 전문계약 제품으로 선정될 경우, 관련 법령에 따라 보다 간소화된 구매(수의계약)가 가능합니다.

Q2 **(계약)** AI 플랫폼 계약 전, 플랫폼 선택을 위해 테스트를 해보고 싶은데 가능한가요?

A2 현재 정식 계약 절차에 앞서 각 AI 플랫폼별 무료 테스트를 공식적으로 지원하고 있지는 않습니다. 다만, AI 플랫폼 공급사와 사전에 협의하여 테스트 방안 등을 지원받으실 수 있습니다. 테스트 신청 및 관련 문의는 아래 플랫폼별 영업 담당자에게 연락하여 확인해 주시기 바랍니다.

- FabriX 영업 : 삼성SDS 공공사업 대표 이메일 publicbiz@samsung.com
- CLOVA Studio for GOV 영업 : 네이버클라우드플랫폼 대표 이메일 dl_gov_ai@navercorp.com

Q3 **(비용)** 공통기반 사용을 위한 요금 체계는 어디에서 확인할 수 있나요?

A3 공통기반 요금 체계는 '공통기반 개발시스템 〉 이용안내 〉 요금 시뮬레이션' 메뉴에서 [범정부 AI 공통기반_예산 시뮬레이터.xlsx] 파일을 다운로드하여 확인하실 수 있습니다.

Q4 **(비용)** AI 서비스 오픈 전 구현(개발) 단계에서도 비용을 지불해야 하나요?

A4 서비스 오픈 전 단계라 하더라도 구현 및 테스트를 위한 LLM 호출 비용, 데이터 처리(파싱/청킹/임베딩 등) 및 RAG 구축을 위한 API 사용료 등의 실비가 발생합니다. 단, 이용 기관의 사업 발주 방식에 따라 요금 청구 및 정산 방법은 상이할 수 있습니다.

Q5 **(비용)** 공통기반과 계약만 하면 AI 서비스 구현이 가능한 건가요?

A5 AI 공통기반은 서비스 구현에 필요한 자원과 기능을 갖춘 플랫폼(PaaS)을 제공합니다. 따라서 각 이용 기관은 기관별 요구사항에 맞춰 제공되는 자원을 활용해 AI 서비스를 직접 구현하셔야 합니다. 플랫폼 비용 산정 시 인건비 등은 포함되지 않으므로, 사업 발주 시 기관의 현황을 충분히 고려하여 공고해 주시기 바랍니다.

- **(범정부 AI 공통기반 역할)** RAG 구축·테스트 기능 제공, AI 서비스 구현·테스트·운영 기능 제공, Chat 서비스 및 공통서비스(연관정보검색, 문서초안작성) 제공, 공통기반 공통 RAG(10종) 제공, 공통활용 학습데이터 제공, AI 정보(AI 서비스·모델 등) 제공, 공통기반 관리기능(정보시스템 등록 및 관리, 사용량 및 사용 금액 조회, 오류 등록 및 이용 문의) 제공

- **(이용기관 역할)** 기관에서 필요로 하는 AI 서비스 기획 및 설계, 기관 특화 RAG용 데이터 정의 및 수집, 공통기반 자원·기능을 활용한 기관 특화 RAG 구축 및 AI 서비스 구현, 공통기반을 활용한 AI 서비스 운영, 과금 모니터링, AI 서비스 장애 및 민원 대응, 추가 기관 특화 RAG 구축, AI 서비스 고도화 등

Q6 (보안) 공통기반 사용 시 개인정보 등 데이터 보안은 안전한가요?

A6 네, 안전합니다. 공통기반은 보안성이 검증된 PPP(민관협력형 클라우드) 존 내에 위치하며, 기관이 자체적으로 보안 등급(C/S/O)을 설정하여 운영하실 수 있습니다. 또한, 개인정보는 PII Masking 및 비식별화 기능을 통해 철저히 보호됩니다.

Q7 (보안) 데이터의 보안등급 분류 주체는 누구인가요? 분류는 어떤 방법으로 할 수 있나요?

A7 데이터 보안 등급은 이용 기관에서 직접 분류하셔야 합니다. 국가 망 보안체계(N2SF) 가이드라인에 따라 각 기관은 업무상 취급하는 정보에 대해 C(기밀)/S(민감)/O(공개) 등급 분류 수행해야 합니다. 이때 기관 자체 업무기준(BRM) 및 N2SF 가이드라인 1.0에서 제시하는 기준을 참조하여 분류하실 수 있습니다.

*참고 C/S/O 등급 분류를 자동화할 수 있도록 지원하는 민간 솔루션들이 있으며 주요 키워드에 따른 분류, 문서의 내용과 문맥 기반 분류, 특정 저장소(NAS, 폴더 등)에 저장된 파일 기준 분류, 다운로드한 사이트 출처 기준에 따른 분류 등을 지원합니다.

Q8 (기능) 공통기반을 이용하는 경우, 공통기반의 지원 범위가 궁금합니다.

A8 공통기반은 RAG 구축 및 AI 서비스 개발 도구를 제공하는 'AI 플랫폼'과 사용자 관리 등을 위한 '개발시스템'으로 구성됩니다. 원활한 이용을 위해 두 시스템을 병행하여 사용하셔야 하며, 개발시스템만 단독으로 이용하실 경우 일부 기능(카탈로그 관리, 이용안내, 마이페이지)으로 사용 범위가 제한됩니다.

- **(AI 플랫폼 주요 기능)** LLM 서빙, RAG 구축·테스트, AI 서비스 구현·운영, Chat 및 공통 서비스(연관정보검색, 문서초안작성), 공통 RAG(10종) 등 (※ API 연계 방식으로 개별 기능 이용 가능)

  → 개별 이용 가능 기능: ① LLM 서빙 ② AI 서비스 구현·테스트·운영 기능+LLM ③ RAG 구축·테스트 기능(파싱/청킹/임베딩 각각 개별 이용 가능) ④ Chat 서비스 및 공통서비스(연관정보검색, 문서초안작성) ④ 공통기반 공통 RAG(10종) 제공

  * AI 플랫폼 개별 이용 기능은 API를 통해 제공됩니다.

- **(개발시스템 주요 기능)** 카탈로그 관리(학습데이터 제공, AI 서비스·모델 정보 제공 등) 기능 제공, 공통기반 관리 기능(대시보드, 정보시스템 등록 및 관리, 사용량 및 사용 금액 조회, 오류 등록 및 이용문의 등) 제공

  → 개별 이용 가능 기능: ① 카탈로그 관리 기능(ex.학습데이터 제공 등), ② 이용안내, ③ 마이페이지

  ※ 공통기반 내에서 GPU만을 단독으로 활용하는 것은 어렵습니다.(별도 GPU 필요 시 FAQ Q17.로 이동)

Q9 (데이터) 기관이 별도로 구축하지 않아도 공통기반에서 활용 가능한 RAG 데이터 범위는 어떻게 되나요?

A9 현재 약 33만 건의 공통 데이터를 제공하고 있습니다. 주요 범위는 법령자료(현행 법령·규칙, 판례), 보도자료(부처별 보도·설명자료), 연구자료(국정과제, 연구 보고서), 행정업무자료(업무 매뉴얼, 조직정보, 종합계획 등)입니다. 제공 데이터는 지속적으로 최신화하여 버전 관리를 수행할 예정입니다.

Q10 (AI 모델) 공통기반에서 제공하는 모델을 사용하는 것과 AI 모델을 개별적으로 사용하는 것의 차이는 무엇인가요?

A10 공통기반 제공 모델은 보안성이 검증된 안전한 PPP존 내에서 운영됩니다. 따라서 행정 업무 활용을 위한 별도의 추가 보안 평가 없이도 즉각적으로 업무에 도입하여 활용하실 수 있다는 강점이 있습니다.

Q11 **(AI 모델)** 공통기반에서 제공하는 모델의 범위는 어떻게 되나요?

A11 2026년 5월 기준으로 LLM, 멀티모달, 추론 모델 등 다양한 유형의 모델 30종이 탑재되어 있습니다. 향후 이용 기관의 수요에 따라 새로운 모델을 지속적으로 추가 도입할 계획입니다.

Q12 **(AI 모델)** 과학기술정보통신부 주관의 독자 AI 파운데이션이 범정부 AI 공통기반에서 탑재되어 서비스 계획이 있나요?

A12 네, 2026년 5월 기준 네이버클라우드의 HCX-Seed-Think가 탑재되어 서비스 중입니다. 또한, 2026년 상반기 중으로 독자 AI 파운데이션 모델을 추가 탑재하여 서비스하는 방안을 적극적으로 검토하고 있습니다.

Q13 **(AI 모델)** 기관에서 요청하는 특정 AI 모델을 공통기반에 추가할 수 있나요?

A13 기본 제공 모델 외에 특정 모델이 필요한 경우, 이용 기관에서 공통기반 관리자에게 AI 모델 추가 신청을 하실 수 있습니다. 신청된 모델은 공통기반 서빙 적합성 평가 및 등록 절차를 거쳐 도입 여부가 결정됩니다.

Q14 **(인프라)** AI 플랫폼 활용을 위한 방화벽 설정은 어떻게 할 수 있나요?

A14 출발지 및 도착지 IP에 대한 방화벽 허용 정책 등록과 각 IP의 로그 확인 가능 여부를 사전에 파악하셔야 합니다. 행정망 대역 이용 여부 및 이용 형태에 따라 정책이 상이하므로, 본 가이드라인의 방화벽 확인(30page) 섹션을 참조하여 설정해 주시기 바랍니다.

Q15 **(인프라)** AI 서비스 개발 시에 이용기관의 정보시스템과의 연계가 필요합니다. 저희 이용기관의 레거시 시스템은 국가정보자원관리원 대구센터 민관협력 클라우드존(PPP-대구)에 있지 않습니다. 대전 혹은 광주 정보센터 등에 위치하고 있습니다. 이럴 경우에도 범정부 AI 공통기반을 사용할 수 있나요?

A15 네, 가능합니다. 해당 시스템이 행정망에 구축되어 있다면 방화벽 허용 절차를 거쳐 공통기반과 연계하실 수 있습니다. 다만, 인터넷망이나 폐쇄망에 위치한 경우에는 「국가 정보보안 기본지침」 제2절에 의거하여 보안성 검토가 선행되어야 합니다.

Q16 **(인프라)** 행정망에 있지 않은 시스템은 공통기반과 어떻게 연계하나요?

A16 공통기반은 행정망 내에 위치하므로 보안 기준에 따라 기관 행정망을 통해서만 연계가 가능합니다. 행정망 외의 망(인터넷망 등)에서 연계가 필요한 경우, ① 기관 행정망 내 중계시스템 확보, ② 중계시스템을 경유한 연계(대상시스템-중계시스템-공통기반) 방식을 이용하셔야 합니다. 이때 망간 보안대책 준수는 필수사항입니다.

Q17 **(인프라)** 공통기반 외 하드웨어 임대 등 클라우드 서비스도 이용할 수도 있나요?

A17 인프라 하드웨어(G클라우드, PPP 등)는 국가정보자원관리원 소관이므로, 관리원의 현재 기관 인프라 담당 또는 대구센터 신기술기반과(PPP영역)에 문의하시면 됩니다.

Q18 **(교육 지원)** 공통기반 플랫폼 이용 및 활용을 위한 교육 지원이 가능한가요?

A18 각 AI 플랫폼별 담당 영업 채널을 통해 교육 지원을 요청하실 수 있습니다. 이후 담당 부서와 교육 일정 및 내용을 협의하여 맞춤형 지원을 받으실 수 있습니다.

부록 1

공통기반 서빙 모델

No	유형	모델명	특징 및 주요 기능
1	공공 특화 sLLM	Llama 3.1 8B	공공 도메인 특화 모델, 범용 질의응답, 요약, 번역에 적합
2	대형 LLM	Llama 4 MoE 109B	대형 모델로 고성능 Reasoning 및 Generation에 적합
3	대형 LLM	Llama 4 MoE 400B	초고성능 대형 모델로 복합 Reasoning, coding, 다국어 처리 탁월
4	한글 특화 중소형 LLM	Samsung LLM 37B	한글 성능 우수, 데이터 분석, QA 등 다양한 작업 지원
5	한글 특화 중소형 LLM	Llama 3.3 70B	한글 성능 우수, 고성능 모델로 범용 작업 지원
6	한글 특화 중소형 LLM	Gemma 3 27B	복합 Reasoning 및 Generation 등 대규모 작업에 적합, VLM 용도
7	초경량 sLLM	Gemma 3 1B	초경량 sLLM으로 모바일, 엣지 환경에 적합
8	오픈웨이트 모델	GPT-OSS 120B	범용 추론, 에이전트 작업에 최적화된 오픈형모델
9	경량 sLLM	Gemma 3 4B	성능과 속도의 균형, 범용 Assistant에 적합
10	초경량 sLLM	Llama 3.2 1B	초경량 sLLM으로 모바일, 엣지 환경에 적합
11	경량 sLLM	Llama 3.2 3B	성능과 속도의 균형, 범용 Assistant에 적합
12	멀티 모달(이미지 생성 전용)	Flux	빠른 이미지 인식 및 이미지 생성, 운영 비용 큼
13	멀티 모달(이미지 해석 전용)	Llama 3.2 vision 90B	이미지 해석 및 OCR 등을 결합한 질의응답 지원
14	Reasoning	Samsung LLM 37B Reasoning	Reasoning 중심 모델, 복합한 문제 해결에 적합
15	Reasoning	Nemotron	Reasoning 중심 모델, 추론 및 Tool 연동에 적합
16	임베딩 전용모델	Bge-M3	문서 검색, 벡터 임베딩, RAG 파이프라인 구축에 적합
17	한글 특화 중소형 LLM	민원도우미 서비스 학습데이터 생성 모델	민원도우미 자동 생성 모델 개발을 목적으로 Gemma 3 27B를 파인튜닝한 모델

[표 18] FabriX 언어모델 POOL (2026년 5월 기준)

No	유형	모델명	특징 및 주요 기능
1	한글 특화 추론형 LLM	HCX-GOV-THINK-V1-32B	네이버 독자 AI 파운데이션 모델(Text Only), Agent Tool Usage 기능 탁월
2	한글 특화 추론형 LLM	HCX-GOV-THINK-24B	한글 성능 우수, 고성능 Reasoning 및 Generation에 적합
3	글로벌 모델 튜닝모델	LLM42-12B	다국어 번역, 비전 언어 모델, TTA(한국정보통신기술협회)에서 보장하는 CAT(인공지능 신뢰성 인증) 획득
4	글로벌 모델 튜닝모델	LLM42-Gemma4-31B	비전 언어 모델, Gemma 4 모델에이전트 기능 우수
5	글로벌 비전번역모델	LLM42-Translate Gemma-27B	번역 전용 모델, Gemma 3 모델을 번역용으로 SFT, RL 학습한 모델
6	글로벌 오픈웨이트 모델	GPT-OSS-120B	범용 추론, 에이전트 작업에 최적화된 오픈형 모델
7	한글 특화 추론형 LLM	K-Exaone-236B*	LG 독자 AI파운데이션 모델
8	한글 특화 추론형 LLM	HCX-SEED-THINK - 14B	범용 추론, 에이전트 작업에 최적화된 오픈형 모델
9	한글 특화 비전 언어모델 (초경량)	HCX-Seed-Vision-INSTRUCT-3B	텍스트/이미지/비디오 입력, 텍스트로 출력
10	한글 특화 언어모델 (초경량)	HCX-Seed-Text-INSTRUCT-1.5B	저사용 GPU 지원용 모델
11	한글 특화 언어모델 (초경량)	HCX-Seed-Text-INSTRUCT-0.5B	엣지 다비스용 모델
12	임베딩 전용모델	Bge-M3-0.6B	문서 검색, 벡터 임베딩, RAG 파이프라인 구축에 적합
13	한글 특화 중소형 LLM	민원도우미 서비스 학습데이터 생성 모델	민원도우미 자동 생성 모델 개발을 목적으로 Gemma 3 27B를 파인튜닝한 모델

[표 19] CLOVA Studio for GOV 언어모델 POOL (2026년 5월 기준)

※ 독자 AI 파운데이션 모델 2종: HCX-GOV-THINK-24B, K-Exaone-236B

부록 2 상세요구사항 예시

요구사항 분류		기능 요구사항
요구사항 고유번호		SFR-01
요구사항 명칭		범정부 AI 공통기반 전면 활용
요구사항 상세설명	정의	범정부 AI 공통기반을 전면 활용하여 RAG DB 구축 및 LLM 이용
	세부내용	○ OO부 특화 AI 서비스에 필요한 데이터는 OO부 자체 보유 데이터를 활용하며, 내부 OOO시스템으로부터의 연계 및 수집 수행 및 필요 시 공통기반 제공 학습데이터 활용 방안 검토 및 제시 ○ 수집된 데이터를 RAG DB로 구축하기 위한 데이터 전처리(파싱, 청킹, 임베딩) 기능 및 RAG용 벡터DB(Vector DB)는 범정부 AI 공통기반 제공 기능을 전면 활용하여 구축 ○ 공통기반 전처리 기능 및 벡터DB는 공통기반 제공 API를 활용하여 연계하며, 데이터가 기관 외부로 반출되지 않도록 구성 ○ 공통기반 제공 기능을 활용하여 OO부 특화 AI 서비스에 필요한 RAG DB 구축 및 테스트 수행(공통기반 제공 API 활용) ○ OO부 특화 AI 서비스에 필요한 LLM은 범정부 공통기반 제공 LLM 활용을 원칙으로 하며, 필요 시 발주기관과 협의하여 외부 모델의 범정부 AI 공통기반 탑재(안) 제시 ○ 범정부 AI 공통기반 제공 공통서비스(AI 챗 화면인 Chat 서비스, 연관정보검색 서비스 등)를 OO부 업무 환경 내 직원 이용이 가능하도록 그대로 연계하여 제공
산출 정보		공통기반 활용계획서, RAG DB
관련 요구사항		
요구사항 출처		

요구사항 분류		기능 요구사항
요구사항 고유번호		SFR-02
요구사항 명칭		범정부 AI 공통기반 데이터 활용
요구사항 상세설명	정의	범정부 AI 공통기반 제공 기능을 활용하여 RAG DB 구축 및 연계 테스트 수행
	세부내용	○ OO부 특화 AI 서비스에 필요한 데이터 수집 및 관리는 기관 자체적으로 수행하며, 직원용 AI 챗 화면(UI) 또한 기관에서 직접 구축하여 맞춤형 환경으로 구성 ○ 수집된 데이터를 RAG DB로 구축하기 위한 데이터 전처리(파싱, 청킹, 임베딩) 기능 및 RAG용 벡터DB(Vector DB)는 범정부 AI 공통기반에서 제공하는 핵심 기능을 활용 ○ 공통기반 전처리 기능 및 벡터DB 연계 시 공통기반 제공 API를 활용하여, 기관 자체 데이터가 외부로 무단 반출되지 않도록 안전한 연동 구조 확보 ○ 공통기반 제공 기능을 활용하여 기관 데이터 기반의 RAG DB 구축 및 연계 테스트 수행(공통기반 제공 API 활용) ○ OO부 특화 AI 서비스에 필요한 LLM은 범정부 공통기반 제공 LLM 활용을 원칙으로 하며, 이를 자체 구축한 AI 챗 화면과 원활하게 연동할 수 있는 구조 제시 ○ 구현된 자체 화면과 공통기반 AI 모델을 연계하여 OO부 업무 환경 내 직원들의 안정적인 특화 AI 서비스 이용 환경 제공
산출 정보		공통기반 연계방안 정의서, RAG DB
관련 요구사항		
요구사항 출처		

요구사항 분류		기능 요구사항
요구사항 고유번호		SFR-03
요구사항 명칭		범정부 AI 공통기반 일부 기능을 활용한 RAG DB 구축
요구사항 상세설명	정의	범정부 AI 공통기반에서 제공하는 일부 기능을 활용한 RAG DB 구축 및 LLM 이용
	세부내용	○ OO부 특화 AI 서비스에 필요한 데이터 수집, RAG를 위한 임베딩 및 벡터DB 구축, 사용자 맞춤형 AI 챗 화면 구현 등 제반 인프라의 기관 직접 구축 및 관리 ○ 수집 데이터를 자체 RAG DB로 구축하기 위한 전처리 과정 중, 파싱 및 청킹 기능에 한하여 범정부 AI 공통기반 제공 API를 호출하여 활용 ○ 공통기반 파싱 및 청킹 API 활용 시, 데이터 처리 과정에서 기관 데이터가 외부로 유출되지 않도록 API 호출 구간의 보안성 확보 방안 마련 ○ 자체 구축한 임베딩 및 벡터DB 환경 기반의 RAG 체인 구성 및 기관 주도하의 구축·테스트 수행 ○ OO부 특화 AI 서비스에 필요한 LLM은 범정부 AI 공통기반 제공 LLM 활용을 원칙으로 하며, 자체 RAG 환경 및 AI 챗 화면과 API 형태로 연계하여 서비스 구현 ○ 최종 구현 시스템이 OO부 내부 업무 환경과 완벽히 통합되어 직원들에게 원활하게 제공될 수 있도록 연계 아키텍처 제시 ○ 향후 공통기반에서 제공 예정인 신규 서비스 및 기능에 대해 확장 가능한 연동 구조(Plug-in 또는 Open API 구조 등) 확보 ○ 공통기반과의 연계는 민관협력형 클라우드 또는 보안망 내 클라우드 통신 정책을 준수하며, 모든 통신 구간에 대한 암호화 및 접근 제어 정책 적용 ○ 연계된 공통기반 서비스의 호출 이력, 응답 결과, 장애 발생 이력 등은 통합 로그로 관리하고, 장애 발생 시 재처리 및 알림 기능 포함 ○ 공통기반 서비스의 SLA(가용성, 응답 시간 등) 기준에 따라 서비스 연계 시 처리 지연 감지 및 대체 처리 시나리오 포함
산출 정보		공통기반 활용계획서, 공통기반 연계방안 정의서, RAG DB
관련 요구사항		
요구사항 출처		

부록 3

AI 서비스 유형별 대표 사례

□ 범정부 AI 공통기반 활용 사례

○ 「지능형업무관리플랫폼」범정부 AI 공통기반 활용 사례

- 지능형 업무관리 플랫폼 내 'AI 행정지원 서비스'를 구현함에 있어 서비스 컨셉 정의, 기획, 기능·화면·DB 설계 등은 지능형 사업단에서 직접 수행하였다. 범정부 AI 공통기반은 고성능 생성형 AI 서비스 구현에 필수적인 AI 플랫폼 API 및 LLM API를 제공하여 안정적인 서비스 빌드를 지원하고 있다.

- 지능형 업무관리 플랫폼은 서비스의 특성에 따라 공통기반의 기능을 다음과 같이 활용하였다.
  - 업무공간 RAG 서비스: 범정부 AI 공통기반에서 제공하는 파서(Parser), 청킹(Chunking), 임베딩(Embedding), 데이터(Data), Chat API 등을 종합적으로 활용하여 지식 베이스를 구축하였다.
  - 법령정보 RAG 서비스: 신뢰도 높은 법령 정보 제공을 위해 범정부 AI 공통기반의 법령 RAG API를 연계하여 활용하였다.

- 지능형 업무관리 플랫폼은 범정부 AI 공통기반을 활용함에 있어 보안과 독립성을 최우선으로 고려하여 다음과 같은 요건을 준수하였다.
  - 각 부처별로 테넌트를 엄격히 분리하여 업무공간 자료 및 RAG에 대한 부처별 권한 관리와 접근 제어를 독립적으로 수행하였다.
  - 보안 강화를 위해 업무공간의 비정형 자료 및 RAG DB를 범정부 AI 공통기반의 스토리지나 테넌트에 별도로 저장·보관하지 않도록 구현하였다.
  - 자체적인 프론트엔드 개발 환경과 공통기반 API를 결합하여 AI 백엔드 서비스 및 에이전트(Agent) 구현을 위한 독자적인 개발 체계와 운영 환경을 구축하였다.

① 서비스 컨셉 정의

프로세스 No	공통기반 활용 프로세스	활용 사례
1-1	도입예정 AI 서비스 정의	• 고객 요건을 구체화하고 실사용자 니즈를 파악하기 위해 행정안전부 직원 총 22명 대상 사용자 인터뷰 수행(실/국장, 과장, 실무그룹) • 제안요청서(RFP) 내용과 사용자 인터뷰 내용을 종합해 AI 행정지원 서비스 16개 확정
1-2	서비스 구현 필요데이터 정의	• AI 행정지원 서비스에서 사용자 선택 범위에 따라 업무공간 데이터, SaaS 드라이브 파일, 법령정보를 활용한 답변을 제공하도록 설계 • 업무공간 데이터 정제/벡터화를 위해 범정부 공통기반의 전처리 모듈, 임베딩 모델을 API 호출해 활용
1-3	공통기반을 활용한 서비스 구현방안 검토	• 범정부 AI 공통기반에서 제공하는 언어모델, 전처리 포함 기능모듈, 인프라/GPU 자원을 활용하여 서비스 구현 • 사용자 대화창, 업무공간 데이터에 대한 RAG 구현은 SI 과업으로 수행하고, AI Agent는 민간 솔루션 도입/적용

[표 20] 공통기반 활용 프로세스 – ① 서비스 컨셉 정의

② 공통기반 활용 기관 특화 RAG/AI 서비스 구현

No	공통기반 활용 프로세스	활용 사례
4-1	공통기반을 활용한 AI 서비스 기획·설계	• 범정부 AI 공통기반 플랫폼에서 제공하는 복수의 LLM 모델 및 API를 지능형 연계 및 활용 • 민간의 솔루션을 활용해 사용자 질의 의도를 파악하고, 서비스 수행에 적합한 Tool을 호출할 수 있는 구조의 Agent 구현 및 서비스 구성
4-2	RAG용 데이터 수집 및 기관 특화 RAG 구축	• 범정부 AI 공통기반에서 제공하는 파서, 청킹, 임베딩 모델을 활용해 생성형 AI 구현을 위한 데이터 파이프라인 구축 • 지능형 업무공간 페이지 작성 콘텐츠, 업로드 파일과 사용자가 대화창에 업로드한 파일에 대해 범정부 공통기반 제공 API 활용한 데이터 전처리 수행 • 정제된 업무공간 데이터는 고객 요건에 따라 범정부 공통기반이 아닌 지능형 내 기관별 벡터DB에 적재 • 범정부 AI 공통기반의 법령 RAG를 이용하기 위해 제공된 API를 호출해, 법령 검색 서비스에 필요한 데이터 연계
4-3	공통기반을 활용한 AI 서비스 구현	• 행정망 업무 데이터를 활용한 AI 자료검색, 문서작성, AI 프롬프트 템플릿, 문서 요약/번역 서비스를 구현하여 AI 행정지원 서비스 제공

[표 21] 공통기반 활용 프로세스 – ④ 공통기반 활용 기관 특화 RAG/AI 서비스 구현

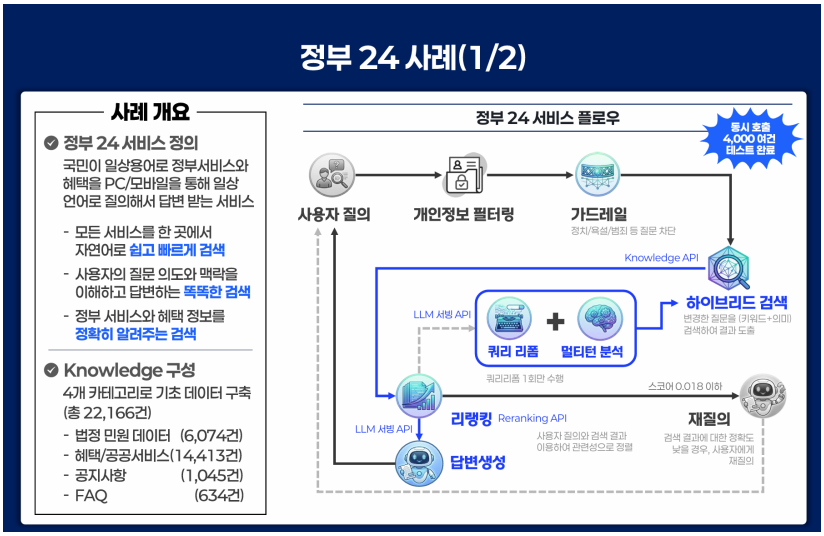
③ AI 서비스 운영

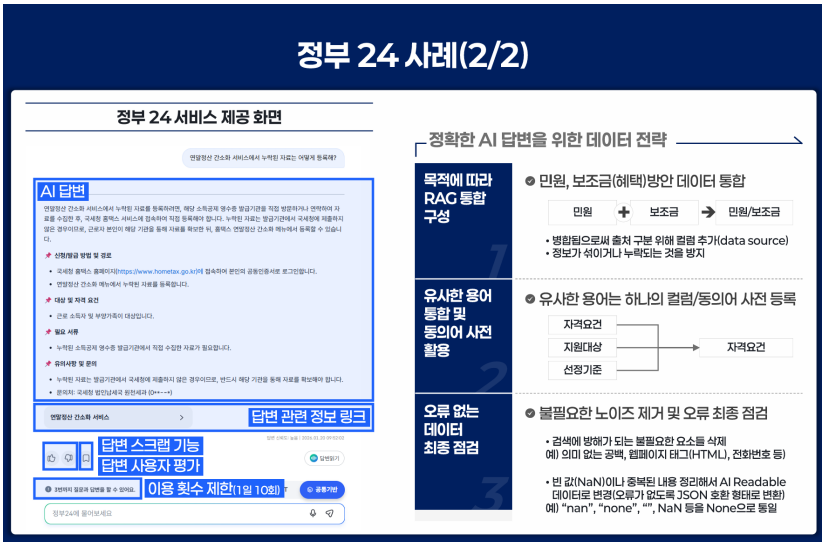
No	공통기반 활용 프로세스	활용 사례
5-1	공통기반 서비스 운영 및 과금 모니터링	• 해당 사항 없음. 지능형 업무관리 플랫폼 사용 부처별 사용량과 과금 및 모니터링 가능한 방안 필요
5-2	공통기반을 활용한서비스 장애 및 민원대응	• 해당 사항 없음. 장애 발생 시 문의 할 대응 창구 필요
5-3	AI 서비스 개선을 위한 추가 RAG 구축/AI 서비스 고도화	• 범정부 테넌트를 활용한 서비스의 경우, AI 서비스 제공에 필요한 정보가 업데이트 될 시 범정부 AI 공통기반에서 데이터 추가 수집 및 전처리를 수행한 RAG DB 필요 • RAG가 추가되거나 서비스 확대 요구 등이 있는 경우, 범정부 AI 공통기반으로부터 추가된 RAG를 이용하기 위한 API를 제공받아 지능형 AI 서비스 기획/고도화 가능

[표 22] 공통기반 활용 프로세스 – ⑤ AI 서비스 운영

○ 「정부24 AI」 범정부 AI 공통기반 활용사례

- 정부24 AI는 정부서비스와 혜택을 PC/모바일을 통해 사용자가 일상언어로 질의하여 답변을 받는 행정안전부의 대국민 서비스로 공통기반을 활용하여 구현하고 있다.
- 범정부 AI 공통기반 활용 범위
  - RAG 구축(약 2.2만 데이터)
  - Knowledge API, LLM Serving API만 사용





□ (참고) 공공부문 AI 활용 우수 사례 안내

○ 공공부문 AI 서비스 구축 및 활용에 대한 우수사례집은 아래 기관 홈페이지에서 확인 가능

- 행정안전부 홈페이지 〉 정책자료 〉 참고자료
  (URL:https://www.mois.go.kr/frt/bbs/type001/commonSelectBoardArticle.do?bbsId=BBSMSTR_000000000015&nttId=124688)
- 한국지능정보사회진흥원(NIA) 홈페이지 〉 지식 정보 〉 간행물 〉 Ai.gov
  (URL:https://nia.or.kr/site/nia_kor/ex/bbs/View.do?cbIdx=37989&bcIdx=29161&parentSeq=29161)

부록 4

공공 AI 서비스 개발 가이드북

- 공무원을 위한 실무 안내서 -

인공지능전략위원회 공공AX분과

이 문서는 "AI를 도입하자"는 제안서를 쓰기 위한 문서가 아니다.
이 문서는 이 서비스가 **계속할 가치가 있는지**, 아니면 **지금 접는 것이 맞는지**를 초기에 판단하기 위한 **안전장치**다.
AI는 도구일 뿐이며, 공공 서비스의 기준은 언제나 "문제가 실제로 줄어들었는가"다.

1 문제부터 써야 하는 이유

공공 AI 사업이 실패하는 가장 흔한 이유는 기술이 부족해서가 아니다.
"문제가 무엇인지 애초에 명확하지 않았기 때문"이다.
많은 사업이 'AI로 자동화한다'거나 'AI로 혁신한다'는 구호에서 시작하지만, 그 순간 이미 방향을 잘못 잡은 것이다.
이 가이드북이 반드시 문제 정의부터 작성하도록 요구하는 이유가 여기에 있다.

2 문제 정의는 이렇게 쓴다

문제 정의란 "우리가 만들고 싶은 것"이 아니라 "지금 겪고 있는 불편"을 쓰는 것이다.
좋은 문제 정의는 다음 질문에 답한다.
"지금 무엇이 불편한가? 누가 그 불편을 겪는가? 얼마나 자주 발생하는가? 이 문제가 지속될 때 어떤 비용이 발생하는가?"
문제 정의는 아래 문장을 완성하는 방식으로 쓰면 된다.
"현재 ○○ 업무에서 △△로 인해 ○○한 문제가 반복적으로 발생하고 있으며, 이로 인해 (시민/공무원)이 ○○한 불편과 비용을 겪고 있다."
여기서 중요한 점은 **AI, 시스템, 기술이라는 단어를 의도적으로 쓰지 않는 것**이다.

3 AI가 꼭 필요한지 먼저 점검한다

모든 문제에 AI가 필요한 것은 아니다. 다음 질문에 "예"가 많을수록 AI 적용이 의미를 가진다.
같은 판단을 하루에도 여러 번 반복하는가? 사람이 처리하기에는 업무량이 과도한가?
문서, 텍스트, 규정이 너무 많은가? 완전 자동이 아니라 "보조"만 해도 효과가 있는가?
여기서 "아니오"가 대부분이라면 AI를 쓰지 않는 것이 더 좋은 판단일 수 있다.

4 데이터는 목록으로 적는다

AI 서비스의 성패는 모델보다 데이터가 좌우한다. 따라서 데이터는 막연히 '있다/없다'가 아니라 상세한 목록으로 관리해야 한다.
각 데이터에 대해 다음을 적는다.
데이터 이름, 어느 부처가 가지고 있는지, 얼마나 자주 만들어지는지, 개인정보가 포함되는지, 최신성이 얼마나 중요한지, 이 과정을 거치면"이 서비스는 현실적으로 가능한가?"가 초기에 드러난다.

5 부처 협업은 역할로 정리한다

공공 AI 서비스는 단일 부처의 노력만으로 완성될 수 없다. 중요한 것은 "누가 참여하느냐"가 아니라 "누가 무엇을 책임지느냐"다.
각 부처의 역할을 한 줄로 명확히 적는다. 예를 들면 다음과 같다.

- 주관 부처 : 서비스 전체 책임
- 데이터 부처 : 데이터 제공 및 변경 통보
- 정책 부처 : 법·제도 해석
- 운영 부서 : 시스템 운영과 장애 대응

6 As-Is / To-Be는 비교표로 쓴다

AI 도입 효과를 설명할 때 가장 설득력이 있는 방법은 "전과 후 비교"다.
As-Is에는 지금의 업무 단계, 처리 시간, 사람이 개입하는 지점을 적는다.
To-Be에는 AI 도입 후 단계, 자동 / AI 보조 / 사람 판단을 구분해 적는다.
장황한 글보다 명확한 단계별 나열이 훨씬 효과적이다.

7 이 문제가 진짜 문제인지 확인한다

"우리가 보는 문제"와 "현장에서 느끼는 문제"는 다를 수 있다. 그래서 반드시 사전 검증을 한다.
다음 중 무엇을 했는지 적는다.
현업 공무원 인터뷰, 시민 민원·문의 분석, 실제 업무 관찰, 숫자는 많을 필요 없다.
하지만 "아무도 안 물어봤다"는 치명적인 결함이다.

8 성공 기준을 미리 정한다

성공 기준이 없으면 이 서비스는 절대 끝나지 않는다.
반드시 다음을 포함한다.
사용률(정말 쓰는가?), 처리 시간(빨라졌는가?), 오류율(품질은 유지되는가?), 사람 개입 비율(부담이 줄었는가?)
중요한 것은 "얼마나 좋아지면 성공인가"를 숫자로 적는 것이다.

9 로그는 운영을 위한 최소 장치다

로그는 감시를 위한 도구가 아니라, 운영을 가능하게 하는 기록이다.
최소한 다음은 자동으로 남아야 한다.
언제, 누가, 무엇을 요청했는지, 결과가 어땠는지, 얼마나 걸렸는지.
AI가 포함되면 AI 응답 성공/실패, 사람 개입 여부도 반드시 기록해야 한다.

10 알림과 보고서는 역할이 다르다

실시간 알림은 "지금 당장 개입해야 할 문제"를 알려준다.
주간 보고서는 "다음 주에 어떤 결정을 할지"를 알려준다.
알림에는 장애, 오류 급증, 데이터 갱신 실패를, 보고서에는 사용 추세, 성능 변화, 품질 변화를.
이 둘을 섞지 않는 것이 중요하다.

11 중단 기준을 미리 적는다

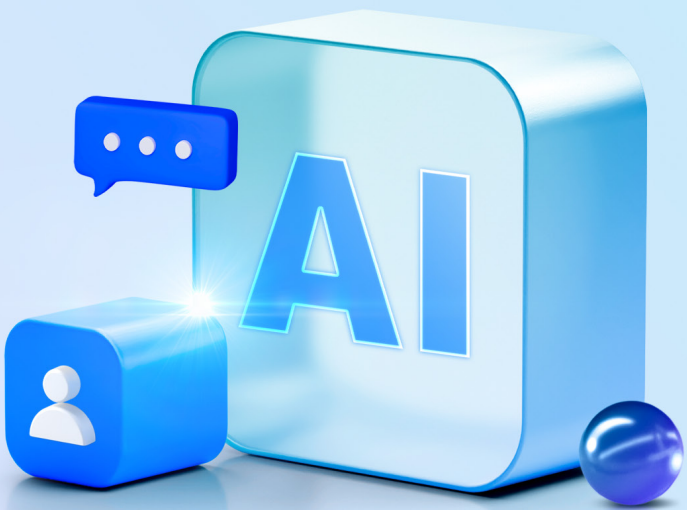
때로는 잘 만든 서비스보다 제때 잘 접은 서비스가 더 책임 있는 행정일 수 있다.
그래서 마일스톤마다 Pass 기준과 Stop 기준을 적는다.
예를 들면, 사용률이 일정 수준 이하로 유지되면 중단, 사람 개입 비용이 효과를 넘으면 중단.
이 기준은 나중에 바꾸기 어렵기 때문에 처음에 써야 한다.

12 모델은 언제든 바꿀 수 있어야 한다

공공 서비스는 특정 AI 모델보다 오래 가야 한다. 그래서 반드시 확인한다.
모델을 바꿔도 서비스가 유지되는가? 데이터와 모델이 분리돼 있는가? 복수 모델 비교가 가능한가?
이 질문에 답할 수 없다면 그 서비스는 위험하다.

공공부문
AI 도입·활용 가이드

- **발 행 처** | 행정안전부, 한국지능정보사회진흥원
- **문 의 처** |
  행정안전부 인공지능정부실 인공지능정부정책국 공공인공지능혁신과,
  한국지능정보사회진흥원 인공지능정부본부 AI정부서비스팀
- **발 행 일** | 2026년 5월



행정안전부
NIA 한국지능정보사회진흥원
NATIONAL INFORMATION SOCIETY AGENCY