전국 지방자치단체 공공 AX 현장 리포트

2년간 전국 현장 협의에서 확인한 생성형 AI 공통 과제, 선택 기준, 기관 유형별 실행전략

발행일: 2026. 10. 07

발행인: 스캐터엑스 대표이사 강명수

한눈에 보는 7대 핵심 제언

01
벤치마크 1위 ≠ 행정업무 최적
코딩·수학 아닌 지식 정확성·지시 이행·한국어 처리가 기준

제3장

02
공개정보는 내부에 다시 쌓지 마세요
법령·통계 등 실시간 검색으로 전달 — 구축비용↓ 현행성↑

제4장

03
광역은 인프라, 기초는 서비스
광역이 기초의 업무화면까지 일괄 구축하면 현장 적합성 하락

제5장

04
하나의 AI로 모든 업무 처리 금지
보안등급·업무 목적별 혼합형 LLM 운영체계 필요

제5장

한눈에 보는 7대 핵심 제언

② 제언 5~7

05
성과는 플랫폼이 아닌 에이전트 이용률

GPU 수·문서 적재 수가 아닌 공직자 실사용률로 측정

제6장

06
광역·기초·사업자·현업 역할 명확히

주체별 역할을 처음부터 구분해야 사업이 작동

제5·7장

07
지금 당장 시작할 수 있습니다

공개정보 기반 에이전트로 즉시 파일럿 가능

제7장

광역자치단체	기초자치단체	전문사업자	현업 공무원
인프라	서비스	구축	실사용

목차

Table of Contents

목차

1 제1장 **2년간 전국 현장에서 얻은 것**
현장 협의 기록·범위·읽는 방법

2 제2장 **전국 현장에서 확인한 공통 과제**
7가지 반복 과제

3 제3장 **생성형 AI, 무엇을 선택할 것인가**
LLM 선택 6대 기준

4 제4장 **공공데이터 3단계 활용체계**
찾아주기·만들어주기·설명하기

5 제5장 **광역·기초 공동 AI 운영모델**
역할 분담·혼합형 LLM

6 제6장 **플랫폼에서 업무 에이전트로**
에이전트 유형·로드맵·성과지표

7 제7장 **기관 유형별 실행 가이드**
광역·기초·현업·즉시 시작

제1장

2년간 전국 현장에서 얻은 것

2024년 11월 철원군청을 시작으로,
전국 광역·기초자치단체 현장 협의 2년의 기록

현장 협의 기록 | 리포트 범위 | 읽는 방법

제1장 | 현장 협의 기록과 리포트 작성 원칙

2024년 11월
강원도 철원군청 첫 방문

2024~2026
전국 광역·기초 수백 차례 협의
제주·울릉 제외 전국 방문

2026년 10월
리포트 발간

start — end



평가하지 않습니다
개별 기관·사업의 우열 판정 없음



통계가 아닌 현장 분석
정형 설문이 아닌 질적 분석



사전 검토 필요
보안 규정·법령 사전 확인 후 적용

제1장 | 이 리포트를 읽는 방법

담당 역할에 따라 먼저 읽어야 할 장이 다릅니다

1



광역자치단체 정보화·AI 담당자

제5장 + 제7장 7.2 먼저

공통 인프라 설계·광역-기초 역할 분담

2



기초자치단체 담당자

제7장 7.3 먼저

광역에 요청할 것·스스로 할 것·규모별 접근

3



현업 공무원

제7장 7.4 먼저

지금 당장 쓸 수 있는 에이전트·활용 원칙

4



의사결정자

핵심 제언 요약 + 제7장 먼저

사업 방향·기관 유형별 체크리스트

제2장

전국 현장에서 확인한 공통 과제

규모도, 예산도, 지역도 달랐지만 — 고민은
놀랍도록 비슷했습니다

도입목적 | 모델종속 | 데이터 | 외부정보 | 운영인력 | 역할분담 | 성과측정

01 도입 목적이 업무 단위로 내려오지 않는다

AI 도입 방침은 있으나 → 어떤 업무에? 어떤 데이터로?
가 정리 안 된 채 예산부터 편성

**구축됐지만
아무도 안쓰는 플랫폼**

02 특정 모델·공급자 종속에 대한 우려

전용 임베딩·벡터DB 일체형 구축
→ 모델 교체 시 처음부터 재구축

**계약 주기가
기술 발전 속도를
따라가지 못함**

03 내부 데이터 중복 적재와 최신성 문제

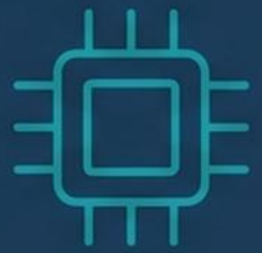
이미 공개된 법령·통계를 내부에 재수집
→ 현행성 관리 부담 + 구축비용 ↑

**개정 법령 반영 못하면
AI가 오래된 정보로 답변**

제2장 | 공통 과제 ④ ⑤ ⑥ ⑦

04	외부 공개정보 활용이 막혀 있다	보안 이유로 외부 검색 차단 → 법령·통계 제공 못 하는 AI → 공직자가 기존 방식으로 복귀	도입했지만 실제로는 안 쓰는 가장 흔한 이유
05	구축 후 운영·고도화 인력이 없다	구축 예산 ○ / 운영 인력 계획 × → 시간이 지나면 사용 급감	AI는 멈추지만 업무 수요는 계속 변한다
06	광역과 기초의 역할이 정리되지 않았다	광역이 기초 업무화면까지 일괄 구축 → 범위 과대 + 현장 불만	"우리 기관 업무와 맞지 않는다"
07	성과를 어떻게 측정할지 모른다	접속자 수·문서 적재 수 → 실제 업무 개선 여부와 무관	업무 처리시간 단축률·재사용률·만족도가 진짜 지표

제3장



생성형 AI, 무엇을 선택할 것인가

뉴스에서는 매달 새 모델이 나옵니다.
*공공부문 담당자가 **실제로 쓸 수 있는***
***선택 기준**을 제시합니다.*

벤치마크의 함정 | 3대 핵심 능력 | 내부망 운영 | 6대 선택 기준

2026년 10월 현재 BenchAlign 1위: GPT-6 Astra (88.75점)

글로벌 벤치마크 평가 항목

- · 에이전트 작업
- · 코딩 능력
- · 수학 난제 해결
- · PhD급 과학 추론
- · 멀티모달 처리

Artificial Analysis Intelligence Index v4.3.2 기준

≠

공직자 실제 행정업무

- · 법령·조례 검색 및 요약
- · 공문·보도자료 초안 작성
- · 회의록·보고서 요약
- · 민원 답변 초안 생성
- · 정책 자료 조사

PhD급 과학 추론이 필요한 업무는 없습니다

벤치마크 1위 ≠ 행정업무 최적 — 이 단순한 사실이 가장 자주 간과됩니다

제3장 | 공직자 업무에 실제로 필요한 AI 능력

Artificial Analysis는 이 세 가지를 별도 항목으로 측정합니다.
전체 1위 모델이 이 세 항목에서도 반드시 1위가 아닙니다.

01



지식 정확성

Knowledge

법령 조문을 정확하게 인용하고
통계 수치를 틀리지 않으며
사실에 기반한 답변을 생성하는 능력

오류 하나가 민원 처리 전체를 흔든다

02



지시 이행 능력

Instruction Following

"공문 형식으로 작성해 달라"
"세 줄로 요약해 달라"
"존댓말로 바꿔 달라"
— 지시를 정확하게 따르는 능력

형식이 맞지 않으면 그대로 쓸 수 없다

03



한국어 행정문서 처리

Multilingual (Korean)

한국어 법령 문체
행정 용어·공문 형식
— 정확하게 이해하고 생성하는 능력

글로벌 벤치마크는 대부분 영어 기준



Claude Sonnet 5.5는 최상위 모델 대비 **94%** 성능 유지 + 출력 비용 **50%** 절감
나머지 6% 성능 차이가 행정업무에서 실질적으로 필요한지 검토하십시오

제3장 | 내부망 운영 가능성이 선택 기준이 되어야 한다

GPT-6 Astra, Claude, Gemini 등 최상위 글로벌 모델은 모두 클라우드 API 방식
— 민감한 행정 데이터에 그대로 사용하는 것은 보안 측면에서 적절하지 않습니다

GPU 기준

기준 장비

NVIDIA H200

단일 H200에서 Falcon-180B 수준
운영 가능 (양자화 적용 시)

양자화 효과

INT4 AWQ 양자화

메모리 최대 4배 ↓

TensorRT-LLM 활용 /
정확도 95% 이상 유지

처리 성능

Llama 2 70B 기준

3,000 토큰/초 ↑

단일 H200에서 달성 가능한 처리량

모델의 크기보다 **데이터의 품질**이 실제 업무 성과를 결정합니다
LLM 추가 학습보다 — 검증된 모델에 정확한 데이터를 전달하는 것이 효율적

제3장 | 공공부문 LLM 선택 7대 기준

① 행정업무 적합 벤치마크
코딩·수학이 아닌
지식 정확성·지시 이행·한국어 처리 항목 기준으로 평가

② 독립 기관 검증 순수치 활용
Artificial Analysis, benchlm.ai 등
공급사와 무관한 중립 벤치마크 기관 수치 기준

③ 내부망 운영 가능성
H200 기준 실제 운영 가능한 GPU 요구량 확인
양자화 적용 시 운영 가능 모델 범위 검토

④ 한국어 행정문서 처리 품질
글로벌 벤치마크와 별도로
실제 한국어 법령·공문·보고서 처리 품질 직접 검증

⑤ 충분한 성능과 지속 운영 가능성
최고 성능보다 행정업무에 충분한 성능
교체와 운영이 지속 가능한 모델 우선

⑥ 벤더·기술 스택 종속 금지
특정 LLM 전용 임베딩 모델·벡터DB에 종속되지 말 것
→ 모델 교체 시 처음부터 재구축하는 상황을 피해야 합니다

영구 라이선스 계약 금지 — 연간 라이선스 또는 오픈소스 모델 사용

생성형 AI 기술은 2~3개월 단위로 성능이 크게 달라집니다
계약 시점 모델 버전을 수년간 사용하는 구조는 반드시 피해야 합니다

⑦ 모델 교체 유연성
특정 모델에 종속되지 않고
2~3개월 주기로 더 나은 모델로
교체할 수 있는 구조 유지

STEP 1
①②로 후보 추리기

STEP 2
③④로 범위 좁히기

STEP 3
⑤, ⑥ ⑦로 최종 결정

모든 업무에 동일한 AI를 사용할 필요는 없습니다

🌐 글로벌 LLM

- ✔ · 법령·조례·통계 검색 및 요약
- ✔ · 보도자료 초안 작성
- ✔ · 정책자료 조사
- ✔ · 외부 공개정보 검색·요약
- ✔ · 타 기관 사례 분석

민감한 내부 데이터를 다루지 않으므로
보안상 문제 없음 + 최신 성능 활용 가능

🔒 내부망 LLM

- 🔒 · 민원인 정보 포함 문서 처리
- 🔒 · 내부 보고서 작성
- 🔒 · 인사·예산·계약 데이터 분석
- 🔒 · 보안등급 높은 내부 문서 처리

데이터가 외부로 나가지 않음
— 가장 높은 보안 수준

✅ 보안↑ + 비용↓ — 업무 성격에 따라 AI를 구분하면 두 가지를 동시에 얻습니다

제4장

공공데이터 3단계 활용체계

AI 성능보다 데이터 전달 방식이 결과를 결정합니다

— 찾아주기 · 만들어주기 · 설명하기

1단계 찾아주기 → 2단계 만들어주기 → 3단계 설명하기

60~70% 공직자 업무에 필요한 데이터는 이미 외부에 공개되어 있습니다

법령·조례·판례·행정규칙·정부통계·연구보고서·언론자료·타기관 사례

✕ 내부 재수집·임베딩 방식의 문제

- ✕ 법령 개정 시 내부 데이터 재처리 필요
- ✕ 현행성 관리 부담 지속 발생
- ✕ 구축·저장 비용 불필요하게 상승
- ✕ AI가 오래된 정보로 답변하는 오류 발생

✓ 실시간 검색 방식

- ✓ 질문 시점에 최신 법령 직접 검색
- ✓ 항상 현행 데이터 기반 답변
- ✓ 내부 저장·관리 데이터 없음
- ✓ 구축 비용·운영 부담 동시 절감

국가법령정보센터 | 공공데이터포털 | 통 계 청 | 국회도서관 | 나라장터

제4장 | 2단계: AI가 읽을 수 있게 만들어주기 (내부 데이터)

20~30% 기관 내부에만 존재하는 데이터 — 외부에서 검색할 수 없습니다

민원 데이터·예산·결산·계약·내부 보고서·회의록·업무 매뉴얼

1 **데이터 추출**
행정시스템에서 필요한 데이터 가져오기

2 **비식별화**
이름·주민번호·연락처 등 개인정보 제거·마스킹
가장 중요한 단계

3 **구조화**
비정형 문서에서 핵심 정보 추출 AI가 이해할 수 있는 형태로 정리

4 **데이터 리니지**
출처·생성일·처리 이력 추적 가능하도록 기록

5 **데이터 카탈로그**
어떤 데이터가 어디에 누가 접근 가능한지 목록화

업무 사례: *"올해 1분기 민원 처리 현황을 요약해 달라"*
→ AI가 비식별화된 민원 데이터를 기반으로 건수·유형별 분류·처리 기간·전년 대비 변화 정리
→ 권한 기반 접근통제 적용 — 접근 권한 있는 데이터만 AI에 전달

제4장 | 3단계: AI에게 맥락을 설명해주기 (센서·시설 데이터)

10~20% **수치 자체만으로는 의미를 알 수 없는 데이터 — 맥락이 없으면 AI도 판단할 수 없습니다**

상하수도 압력·소방 온도·도로 진동·대기질 센서·재난 감지 신호

❌ 수치만 전달하면

상수도 관로 압력값: 2.3bar
→ AI가 판단할 수 없는 것들:
- 이 수치가 정상인지 이상인지
- 어느 지점의 관로인지
- 계절별 정상 범위가 얼마인지
- 과거에 어떤 문제가 있었는지

시설 기본 정보 | 정상 범위 | 이상 기준

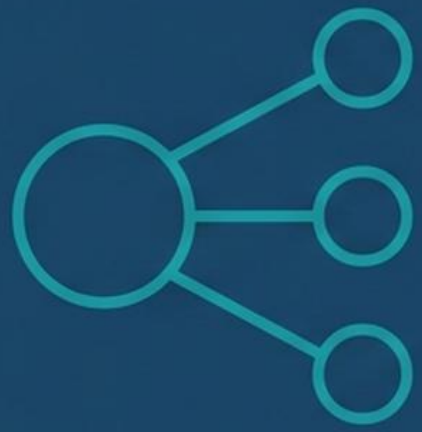
✅ 맥락과 함께 전달하면

AI가 판단할 수 있는 것:
"이 수치는 정상 범위를 벗어났으며 즉시 점검이 필요합니다"

점검 매뉴얼 | 과거 이력 | 계절·시간대 패턴

⚠️ 3단계는 1·2단계보다 준비 기간이 길고 전문성이 더 요구됩니다
단계적 추진에서 가장 나중에 시작하는 것이 현실적입니다

제5장



광역·기초 공동 AI 운영모델

소규모 군 단위 기관이 AI 인프라를 독자 구축하는 것은 현실적으로 **불가능에 가깝습니다.**
광역이 인프라를, 기초가 서비스를 담당합니다.

광역의 역할 | 기초의 역할 | 혼합형 LLM | 공동비용 분담

*무엇을 해야 하는지만큼 — **무엇을 하지 말아야 하는**지도 중요합니다*

✅ 광역이 해야 할 것

공통 AI 엔드포인트 구축
기초자치단체가 복잡한 기술 설정 없이
AI를 쓸 수 있는 표준 접속 환경 제공
→ 내부망 LLM·광역 공동 LLM·글로벌 LLM 자동 라우팅 포함

보안정책·데이터 처리 기준 수립
어떤 데이터를 어떤 AI에 입력할 수 있는지
광역이 먼저 기준을 정해 산하 기관에 제공

글로벌 LLM·범정부 AI 공동계약 주도
광역 일괄 계약 → 비용 절감 + 계약 조건 통일
개별 기초자치단체 각자 계약 대비 협상력 ↑

에이전트 빌더 환경 제공
기초자치단체 담당자가 직접 에이전트를
기획·테스트할 수 있는 도구 공유

❌ 광역이 하지 말아야 할 것

기초자치단체 업무화면·에이전트 일괄 개발
기초마다 업무 방식·데이터 구조·현장 수요가 다름
→ 일괄 개발 시 사업 범위 과대 + 현장 불만 발생

"우리 기관 업무와 맞지 않는다"

기초자치단체 내부 데이터 중앙 통합 관리
민원·예산·인사 데이터는 해당 기관이 직접 관리
광역이 통합 관리 시 데이터 주권·개인정보 보호 문제

제5장 | 기초자치단체: 해야할 것

광역에 요청할 것

→ 공통 AI 엔드포인트 접속 권한·사용 한도 확인
→ 보안정책 기준 제공
→ 글로벌 LLM 비용 분담 기준 안내
→ 에이전트 빌더 사용 교육
→ 광역 개발 공통 에이전트 목록·활용 방법

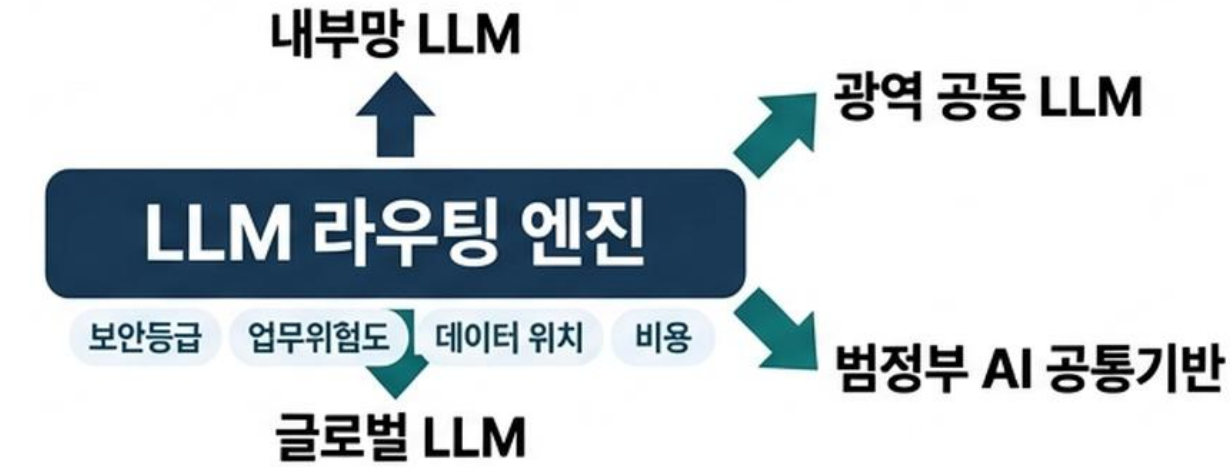
스스로 해야 할 것

1. **기관 내부 데이터 정비·권한 분류**
   — 어떤 데이터가 어디에, 누가 접근 가능한지 먼저 정리
2. **현업 담당자와 에이전트 3~5개 선정**
   정보화 부서 단독 결정 ✕ / 현업 참여 필수
3. **사용 피드백 수집 후 광역과 공유**
   — 좋은 점·부정확한 답변·추가 필요 기능 체계적 수집

SCALE-BASED APPROACH TABLE

기관 규모	권장 접근 방식
대형 시 (인구 30만 이상)	광역 인프라 기반으로 자체 에이전트 직접 개발·운영 전담 인력 또는 전문 사업자와 장기 운영 계약 권장
중형 시·군 (인구 10만~30만)	광역 공통 에이전트 우선 활용 기관 특성에 맞는 부분만 커스터마이징 에이전트 빌더로 일부 직접 개발 가능
소형 군·구 (인구 10만 미만)	광역 공통 에이전트 그대로 활용 독자 인프라 구축보다 광역 공동 활용 체계에 집중

모든 업무에 동일한 AI를 사용하는 것은 비효율적이거나 위험합니다



내부망 LLM

적용

민원인 정보 포함 문서
인사·예산·계약 데이터
보안등급 높은 내부 보고서

데이터 외부 유출 없음 — 최고 보안

광역 공동 LLM

적용

광역 공통 행정업무
개인정보 미포함 일반 행정

기초자치단체 공동 활용
광역 구축·산하 기관 사용

범정부 AI 공통기반

적용

중앙-지방 연계 업무
정부24 등 범정부 시스템 연계

NIA 66억 2차 구현 사업
GPU 별도 구매 없이 활용 가능

글로벌 LLM

적용

법령 검색·보도자료 작성
정책자료 조사·외부 공개정보 요약

최신 성능 활용 + 운영 비용↓
민감 데이터 미포함 업무 한정

공동 운영모델의 가장 실질적인 장점은 비용 절감입니다



중앙정부 → 지자체 공동 활용

중앙정부 구축 인프라를
지방자치단체가 공동 활용
→ 별도 인프라 구축 비용 없음

GPU 서버 구매·임차 불필요



광역 구축 → 기초 사용량 비례 분담

광역이 공통 인프라 구축
기초자치단체는 사용량·기관 규모에
비례한 분담금 납부

사용하지 않는 비용 부담 없음



광역 일괄 계약 → 단가 절감

개별 기초자치단체 각자 계약
→ 협상력 낮음·계약 조건 불리
광역 일괄 계약
→ 사용량 규모 단가 인하 확보

실제 사용량 비례 분담

기초자치단체 입장: 독자 구축 대비 훨씬 적은 비용으로
더 높은 수준의 AI를 활용할 수 있습니다

제6장



플랫폼에서 업무 에이전트로

~~GPU 수·문서 적재 수·플랫폼 구축 완료~~ 는
AI 사업의 성과가 아닙니다.
공직자가 실제 업무에서 쓰는 **에이전트 이용률**
이 성과입니다.

즉시 활용 에이전트 | 개발·운영·폐기 원칙 | 4단계 로드맵 | 성과관리 지표

제6장 | 성과 단위를 바꿔야 합니다 — 지금 쓸 수 있는 에이전트 13종

❌ 지금까지의 성과 지표

- GPU 수량
- 문서 적재 수
- 플랫폼 구축 완료
- 접속자 수

✅ 바꿔야 할 성과 지표

- 월간 활성 사용자·재사용률
- 업무 처리시간 단축률
- 답변 정확도·근거 제시율
- 사용자 만족도

기본 5종 — 내부망 구축 없이 지금 당장 시작 가능

- 법령·조례 검색 및 요약
- 보도자료·공문 초안 작성
- 회의록·보고서 요약
- 민원 분류·답변 초안
- 정책자료·연구보고서 조사

전국 현장 협의에서 추가 확인된 8종

현장 확인 8종 — 기관 특성·업무 수요에 따라 선택 적용

- 민원 분석
- 법률 검토
- 정책 기획·아이디어 도출
- 회의록 작성 (녹음→전문+팩트체크)
- 실시간 관광 정책
- 정책 브리핑
- 시민소통 뉴스레터
- 의회 자료 검색

업무 매뉴얼 기반 에이전트 (추가)

기관 보유 업무 매뉴얼·당직 매뉴얼을 AI 검색 구조로 구성
매뉴얼별 독립 에이전트 → 단기간에 다수의 에이전트를 현장에 제공 가능

처음부터 많은 에이전트를 만들려 하지 마십시오
소수의 고효율 에이전트를 먼저 실증하고
효과가 확인된 에이전트를 중심으로 확대합니다



공직자가 여러 에이전트를 쉽게 찾고 쓸 수 있도록
에이전트를 선택하는 구조가 필요합니다 — **하나의 화면**에서 업무 유형에 따라
에이전트마다 별도 접속 방법이 있으면 사용 빈도가 낮아집니다

제6장 | 4단계 추진 로드맵 & 성과관리 지표

즉시 시작 가능

1

공개정보 기반 에이전트 실증

법령 검색·보도자료·정책조사
내부망 구축 없이 바로 시작
현업 5~10명 × 4주 파일럿

2

내부 데이터 연계 에이전트

1단계 효과 확인 업무 중심
민원 분류·예산 현황·내부 보고서
데이터 정비·비식별화·권한 설계 병행

3

기관 간 공동 활용

검증된 에이전트 → 광역 공통 전환
기초자치단체 공동 활용
에이전트 빌더로 기관별 커스터마이징

맥락 정보 구축 선행 필요

4

재난·시설·센서 분야 확장

상하수도·소방·도로·재난 감지
시설·센서 데이터 활용 에이전트
가장 나중에 시작

월간 활성 사용자·재사용률

에이전트가 실제로 쓰이는지
재사용률 = 공직자 신뢰도

업무 처리시간 단축률

보도자료 초안·법령 검토 시간
도입 전후 비교 측정

답변 정확도·근거 제시율

출처 없는 답변은 신뢰 불가
법령 조문·출처 함께 제시 비율

사용자 만족도

현업 공직자 정기 조사
만족도 낮으면 개선·폐기 검토

제7장



기관 유형별 실행 가이드

광역이냐, 기초냐, 현업 공무원이냐에 따라
지금 당장 해야 할 일이 다릅니다.

본인의 역할에 해당하는 절을 먼저 읽으십시오.

광역 담당자 → 7.2
기초 담당자 → 7.3
현업 공무원 → 7.4
즉시 시작 → 7.5

모든 것을 직접 개발하는 것이 아니라 기초자치단체가 스스로 할 수 있도록 기반을 제공하는 것이 핵심입니다

단계	해야 할 것	결정해야 할 것	준비해야 할 것
1단계 착수	공통 AI 엔드포인트 구축 착수	LLM 선택 기준 및 라우팅 정책 수립	보안정책 초안 작성 기초 담당자 협의 채널 개설
2단계 구축	에이전트 빌더 환경 제공 공통 에이전트 1차 배포	기초자치단체 분담금 산정 기준	글로벌 LLM 공동계약 체결 운영 전담 인력 또는 사업자 확정
3단계 확산	성과 측정 체계 운영 공통 에이전트 고도화	추가 에이전트 개발 우선순위	고도화 예산 편성 기초 피드백 수집 체계 가동

✕ 광역이 하지 말아야 할 것

✕ **기초자치단체 업무화면·에이전트 일괄 개발** → 사업 범위 과대 + 현장에서 *"우리 기관 업무와 맞지 않는다"* 불만 발생

✕ **기초자치단체 내부 데이터 광역 중앙 통합 관리** → 데이터 주권·개인정보 보호 측면에서 적절하지 않음

제7장 | 7.3 기초자치단체: 현장 서비스 운영자

광역에 요청할 것	스스로 해야 할 것
→ 공통 AI 엔드포인트 접속 권한·사용 한도 확인 → 보안정책 기준 제공 → 글로벌 LLM 비용 분담 기준 안내 → 에이전트 빌더 사용 교육 → 광역 공통 에이전트 목록·활용 방법	1 기관 내부 데이터 위치·형태 파악 부서별 접근 권한 분류 2 현업 담당자와 에이전트 3~5개 선정 정보화 부서 단독 결정 ✕ 독자 인프라 구축보다 광역 공동 활용 체계에 집중 3 사용 피드백 수집 → 광역과 공유 좋은 점·부정확한 답변·추가 필요 기능

기관 규모	권장 접근 방식	
대형 시 (인구 30만 이상)	광역 인프라 기반 자체 에이전트 직접 개발·운영	전담 인력 또는 전문사업자와 장기 운영 계약 권장
중형 시·군 (인구 10만~30만)	광역 공통 에이전트 우선 활용	기관 특성에 맞는 부분만 커스터마이징 에이전트 빌더로 일부 직접 개발 가능
소형 군·구 (인구 10만 미만)	광역 공통 에이전트 그대로 활용	독자 인프라 구축보다 광역 공동 활용 체계에 집중

단계	해야 할 것		결정해야 할 것	준비해야 할 것
1단계 착수	광역 엔드포인트 접속 권한 확보	공통 에이전트 파일럿 시작	우선 도입할 에이전트 3~5개 선정 현업 파일럿 참여자 5~10명 선정	피드백 수집 양식 준비
2단계 확장	내부 데이터 정비·권한 분류 착수	추가 에이전트 기획	내부 데이터 연계 범위·보안 기준 데이터 비식별화 절차 수립	운영 담당자 지정

제7장 | 7.4 현업 공무원: 에이전트 실사용자

기술을 깊이 알지 못해도 괜찮습니다.
자신의 업무에서 반복되는 불편함을 찾고
AI 에이전트에 맡겨볼 수 있는지 시도하는 것이 시작입니다.

광양시 — AI 도구 직접 개발 | 성남시 — 30분→3분 단축 | 광진구 — 오픈소스 도구 공직사회 공유

✓ AI가 잘하는 업무	✕ 아직 맡기면 안 되는 업무
✓ 법령·조례 검색 및 핵심 내용 요약	✕ 법령 해석에 대한 최종 판단
✓ 보도자료·공문 초안 작성	✕ 민원인에게 발송할 최종 공문
✓ 긴 문서의 핵심 내용 요약	✕ 예산 집행 결정 및 계약 승인
✓ 유사 사례 및 타 기관 정책 조사	✕ 개인정보가 포함된 민원 처리
✓ 반복되는 양식 문서 초안 생성	✕ 재난·안전 상황의 최종 대응 결정
✓ 민원 접수 쟁점·관련 법령·협력 부서 안내	✕ 민원 처리 결과에 대한 최종 행정 결정
✓ 사업 추진 전 관련 법령·조례 검토 및 미비사항 확인	✕ 법률 검토 의견의 최종 행정 해석
✓ 공약·사업 기획 아이디어 확장 및 구조화	✕ 사업 추진 여부에 대한 최종 결정
✓ 회의 녹음 파일 기반 전문·요약·팩트체크	✕ 회의 결정사항에 대한 공식 확정
✓ 관광 이슈 파악 및 정책·홍보 방안 제안	✕ 관광 예산 배분 및 사업 승인
✓ 중앙정부 정책·언론보도·시민 관심사 검색·분석	✕ 기관 공식 입장 발표 및 대외 발언
✓ 시민 대상 뉴스레터 초안 작성	✕ 시민에게 발송할 최종 공식 문서
✓ 의회 발언·정책 논의 검색 및 정리	✕ 의회 답변에 대한 최종 공식 입장

출처 확인
AI 법령 인용 → 국가법령정보센터에서 직접 확인
존재하지 않는 조문을 생성하는 경우 있음

담당자 최종 판단
AI 결과물은 참고자료
행정적 판단·책임은 담당 공무원에게 있음

민감정보 입력 금지
이름·주민번호·연락처 → 글로벌 LLM 입력 금지
개인정보 포함 업무 → 내부망 AI 또는 제거 후 처리

광역이든 기초든, 인프라가 완성되지 않았든 상관없습니다
지금 당장 시작할 수 있는 것이 있습니다

지금 당장 가능

첫째

공개정보 기반 에이전트로 파일럿을 시작하십시오

내부망 구축·보안 검토·대규모 예산이 없어도 됩니다
법령 검색·보도자료 초안 작성·정책자료 조사처럼
공개정보만으로 처리할 수 있는 업무에서 에이전트를

완벽한 시스템을 기다리다가 아무것도 시작하지 못하는 것보다
작은 에이전트 하나로 현장 경험을 쌓는 것이 훨씬 가치 있습니다

둘째

현업 담당자 5~10명과 4주간 실증을 운영하십시오

5~10명 × 4주

파일럿은 정보화 부서 혼자 하는 것이 아닙니다
민원·복지·도시·안전 등 실제 업무를 담당하는
현업 공무원이 직접 써봐야 합니다

→ 어떤 점이 도움이 됐는지
→ 어떤 답변이 틀렸는지
→ 어떤 기능이 더 필요한지

셋째

이용률과 불편사항을 수집해 즉시 개선하십시오

4주 파일럿이 끝나면 피드백을 바탕으로 에이전트를 개선합니다
잘 쓰이지 않는 에이전트는 원인을 분석하고
현업에서 반복 요청하는 기능은 우선적으로 추가합니다

이 개선과 실증의 반복이 공공 AX의 실제 작동 방식입니다

완벽한 계획보다 빠른 실증과 반복 개선이 공직자가 실제로 쓰는 AI를 만드는 가장 확실한 방법입니다

이 리포트가 전하는 7가지 메시지

① 벤치마크 1위 ≠ 행정 최적
② 공개정보는 실시간 검색으로
③ 광역은 인프라·기초는 서비스
④ 혼합형 LLM 운영
⑤ 성과는 에이전트 이용률
⑥ 역할을 명확히 나누어야
⑦ 지금 당장 시작할 수 있습니다

본 리포트는 전국 지방자치단체 담당자들과의
2년간의 현장 협의에서 나온 내용입니다.
현장의 고민이 이 리포트를 만들었습니다.

발행일: 2026. 10. 07
발행인: 스캐터엑스 대표이사 강명수