

VOL .6

공공부문 AI 에이전트 도입 및 활용 현황과 시사점

최병윤 부연구위원(행정혁신연구실)

황한찬 교수(순천대학교 행정학과)

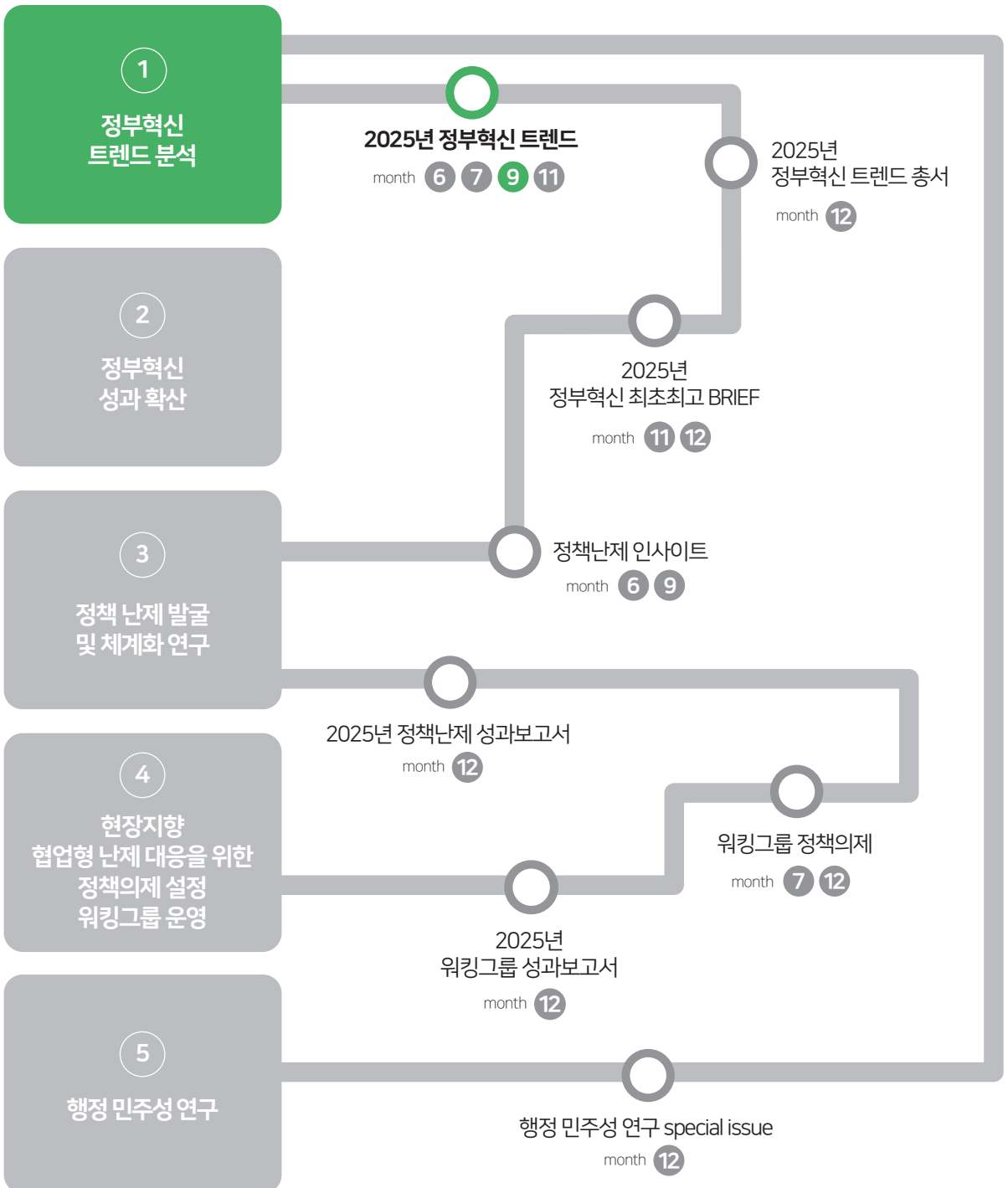
정부혁신 트렌드




KIPA
KOREA
INSTITUTE OF
PUBLIC
ADMINISTRATION

행정선진화 사업 발간물 로드맵

행정혁신연구실에서는 정부 조직, 재정, 혁신, 디지털 정부 등의 분야를 중점적으로 연구하고 있습니다. 저희 실에서 진행하는 '행정선진화 사업'은 1) 정부혁신 트렌드 분석, 2) 정부혁신 성과 확산 지원, 3) 정책 난제 발굴 및 체계화 연구, 4) 현장지향 협업형 난제 대응을 위한 정책의제 설정 워킹그룹 운영 등의 세부사업으로 구성되어 있습니다. 본 로드맵은 각 세부사업의 발간물 게시 일정을 보여줍니다.



공공부문 AI 에이전트 도입 및 활용 현황과 시사점

 **최병윤** 부연구위원(행정혁신연구실) | **황한찬** 교수(순천대학교 행정학과)



공공부문 AI 에이전트 도입 및 활용 현황과 시사점

서론



1. 서론

1) 연구의 배경과 목적

● 연구의 배경

● AI 에이전트 전략의 부재와 탐색적 연구로서의 의의

- 국내에서는 아직 공공부문 AI 에이전트 전략 보고서나 실증 사례가 충분히 축적되지 않았고, AI 에이전트의 구성 요소와 개념적 경계에 대한 합의 역시 초기 단계에 머물러 있음
- 이에 본 연구는 탐색적(exploratory) 연구로서, 국제 동향과 초기 사례를 작동적 정의에 따라 체계적으로 정리·비교하고, 한국형 AI 에이전트 정책 설계를 위한 기준선(baseline)을 마련하는 것을 목적으로 함
 - 특히, 법·제도 및 윤리, 인프라, 인적 역량, 거버넌스 등 핵심 전제조건을 토대로 정책연구의 주춧돌을 놓는 데 본 브리프의 의의가 있음

● 글로벌 AI 에이전트 도입 가속화

- 2025년 영국 정부는 AI 도우미 ‘험프리(Humphrey)’를 도입해 정책 제안에 대한 시민 응답을 자동 분류하기 시작했고(DSIT, 2025), 미국 식품의약국(FDA)도 GPT-4 기반의 AI 에이전트 ‘엘사(Elsa)’를 통해 임상 검토 기간을 단축하고 과학적 평가를 가속화하고 있음(Reuters, 2025)
 - 이러한 사례들은 공공부문에 AI 에이전트가 본격적으로 도입되고 있음을 보여주고 있는데, AI 에이전트란 단순히 콘텐츠를 생성하는 것을 넘어, 주어진 목표를 달성하기 위해 스스로 계획을 수립하고 실행하며 환경과 상호작용하는 자율적인 AI 시스템을 의미(Russell & Norvig, 2020; GAO, 2025)
 - AI 에이전트는 데이터를 인지하고 처리하여 의사결정을 내린 뒤 그 결과에 따라 행동을 취하는 일련의 과정을 인간 개입을 최소화하여 수행(GAO, 2025)
 - 이러한 능력은 기존의 챗봇이나 자동화 도구를 넘어서는 것으로, AI 에이전트는 상황 변화에 맞춰 즉각 판단하고 복잡한 작업도 연속적으로 처리할 수 있음
- 2025년 현재, AI 에이전트 도입은 전세계적으로 급속히 확산되고 있음
 - 글로벌 AI 에이전트 시장은 2023년 37억 달러에서 2025년 73.8억 달러로 거의 두 배 증가했으며, 2032년까지 1,036억 달러에 달할 것으로 전망(Index, 2025)
 - 기업의 85%가 2025년까지 AI 에이전트를 도입할 예정이며(Deloitte, 2025), 이는 단순한 디지털화를 넘어 지능형 자동화로의 근본적 전환을 의미

● AI 에이전트 도입의 사회적·정책적 배경

- 첨단 기술 발전에 발맞춰 행정환경도 변화되고 있으며, 기술적으로 대규모 언어모델(LLM) 등의 생성형 AI 기술의 발전과 에이전틱 AI(agent AI) 아키텍처 성숙이 AI 에이전트 구현과 발전을 가능케 했으며, 국민들이 행정 서비스의 신속성·맞춤성을 강하게 요구하는 디지털 시대의 사회적 분위기 속에 정부 역시 생산성 향상과 행정 효율성 압박에 직면

- 2024년 미국 연방정부의 서비스 만족도는 민간 평균보다 낮은 수준에 머물렀고(FedScoop, 2024), 시민의 98%는 지방 정부에 현대적인 온라인 서비스를 기대하지만 정작 24시간 이용 가능한 공공서비스는 30% 미만에 불과한 것으로 나타났는데, 서비스 품질 격차와 공공부문 생산성 정체 문제가 대두되면서, 행정 분야에서도 지능형 자동화를 통한 혁신 필요성이 커짐

• AI 에이전트 도입은 초거대 AI 기술 발전으로 인한 도구적 가능성과 효율적이고 투명한 행정을 바라는 사회적 수요가 맞물려 등장한 디지털 정부 혁신의 새로운 물결

- 정부들은 AI 에이전트를 통해 업무를 자동화하고 국민에게 개인화된 상시 서비스를 제공함으로써 행정의 신뢰성과 대응성을 제고하고자 함

- 이러한 배경하에 영국, 미국 등 주요국이 AI 전략과 투자 계획을 앞다투어 발표하며 공공부문 AI 에이전트 활용을 본격화하고 있으며, 대한민국도 디지털 정부 추진과 연계하여 AI 에이전트 활용 전략을 마련할 시점

● 디지털 정부 혁신과 AI 에이전트의 필연적 결합

• 공공부문의 AI 에이전트 도입은 디지털 정부 혁신의 자연스러운 진화 과정(Deloitte, 2025)

- 기본적인 양식의 디지털화에서 통합 디지털 서비스, 그리고 AI 증강 의사결정을 거쳐 자율적 AI 에이전트가 복합적인 시민 상호작용을 관리하는 단계로 발전하고 있음

- 이 과정에서 클라우드 인프라, 데이터 통합, API 우선 아키텍처, 디지털 신원 시스템이 필수 기반이 되고 있음

• 신공공관리(NPM)에서 디지털 거버넌스로의 진화에서 AI 에이전트가 촉매 역할

- 무어(Moore)의 공공가치 이론을 AI에 적용하면 (1) 운영 역량 - AI가 행정 효율성과 서비스 제공을 향상, (2) 정치적 승인 - 민주적 정당성과 시민의 AI 시스템 수용, (3) 공공 성과 - 향상된 거버넌스, 투명성, 시민 만족도의 세 가지 차원이 나타남

• OECD 디지털 정부 지수 연구는 디지털 정부 전환이 “AI와 같은 새로운 기술을 관리하기 위한 적응 가능한 거버넌스 체계, 신뢰할 수 있는 디지털 인프라, 전향적 접근법을 포함한 견고한 기반”을 필요로 한다고 밝힘(OECD, 2024)

• 기술-조직-환경(TOE) 프레임워크는 정부의 AI 도입을 설명하는 이론적 기반을 제공

- 기술적 요인으로는 AI 성숙도와 기존 시스템과의 호환성, 조직적 요인으로는 디지털 기술, 리더십 지원, 문화적 준비성, 환경적 요인으로는 규제 요구, 시민 기대, 기관 간 협력 필요

● 공공서비스 혁신 요구와 기술 발전의 교차점

• 현재 공공부문은 전례 없는 서비스 혁신 압박에 직면해 있음

- 전세계적으로 공공부문이 공식 고용의 38%, GDP의 25%를 차지하는 상황에서(World Bank, 2021), 민간부문의 연간 2% 생산성 증가에 비해 공공부문 생산성 향상은 크게 뒤처져 최저 수준 대비 15-25%의 생산성 격차가 발생하고 있음

• 시민들의 서비스 품질 기대수준은 지속적으로 상승 중

- 2024년 미국 고객만족도 지수에서 연방정부 서비스는 69.7점으로 민간부문 평균 73.2점보다 낮았으며(FedScoop, 2024), 시민의 98%가 지방정부로부터 현대적인 온라인 기술을 원한다고 응답(Springbrook, 2025)

- 특히 24시간 서비스 이용 가능성, 개인화된 서비스, 일관된 서비스 품질에 대한 요구가 급증하고 있으나, 현재 정부 서비스의 30% 미만만이 연중무휴 접근이 가능한 상태

- 동시에 인구 고령화가 서비스 수요 증가를 가속화하고 있음
 - 2050년까지 고소득 국가의 부양비는 연간 0.4%씩 감소할 예정이며, 65세 이상 인구의 의료비 지출은 300% 증가하고, 사회보장 및 의료보험 지출이 연방예산의 36%를 차지하게 될 전망
- 영국의 중앙정부 주도의 체계적 전략에 기반한 AI 배치에서는 USAJobs의 기술 식별 기반 AI 구동 채용 추천, HHS AI 챗봇의 보안 질문 자동 응답, 예측 기능을 사용한 기상 위험 분석 시스템, AI 도구를 사용한 연방 웹사이트 콘텐츠 최적화 등의 사례가 나타남

● 연구의 목적과 의의

● 정책 지향적 연구의 목표

- 현재 공공부문에 존재하는 AI 기술 적용 수준과 범위의 부차·분야 간 격차, 그리고 이를 심화시키는 예산 배분 불균형과 기술 데이터 표준화 미비 등의 문제 인식을 바탕으로, 본 연구는 한국 공공부문의 AI 에이전트 도입 현황과 쟁점을 체계적으로 분석하여 정책적 시사점을 도출하고자 함
 - 첫째, 국내외 공공부문 AI 에이전트 도입 현황에 대한 종합적 현황 분석
 - 둘째, 주요국의 AI 거버넌스 접근법에 대한 비교 연구를 통한 한국형 AI 거버넌스 모델 탐색
 - 셋째, AI 에이전트 도입 과정에서 발생하는 정책적·기술적·윤리적 쟁점의 체계적 분석과 해결방안 모색

● 국내외 AI 에이전트 도입 사례 비교분석을 통한 시사점 도출

- 본 연구는 국내외 AI 에이전트 도입 사례를 분석하여, 한국의 공공부문 AI 에이전트 전략 수립에 필요한 실증적 근거를 제공하는 것을 목표
 - 분석 결과를 토대로 기술적, 조직적, 정책적 측면에서의 시사점을 도출하고, 한국 공공부문이 글로벌 AI 혁신 트렌드에 효과적으로 대응할 수 있는 전략적 방향 제시

● 공공부문 AI 에이전트 활용 모범사례 발굴

- 분야별 특성을 고려한 AI 에이전트 활용 모범사례를 발굴하고, 이를 확산 가능한 모델로 체계화
 - 국내 성공사례와 함께, 해외의 혁신적 접근법을 분석하여 벤치마킹 가능한 모델을 제시하고 단계적 도입 전략, 데이터 품질 관리, 조직 역량 강화, 시민 수용성 제고 등 성공 요인별 세부 실행 방안을 구체적으로 제시

● 국내 공공부문 AI 에이전트 도입 전략 수립을 위한 정책 함의 제시

- 한국 공공부문의 AI 에이전트 도입을 체계적으로 추진하기 위한 정책 방향과 구체적 실행 방안 제시
 - 인력 양성, 인프라 구축, 제도 개선, 거버넌스 강화 등 4개 핵심 영역별로 단계적 로드맵을 제공하며, 예산 배분 우선순위와 투자 전략을 구체화
 - AI 윤리, 개인정보보호, 사이버보안 등 새로운 위험 요소에 대한 관리 체계와 시민 수용성 제고 방안을 포함하여, 지속가능하고 책임감 있는 AI 도입이 이루어질 수 있도록 정책적 가이드라인 제시

2) GenAI vs AI Agent vs Agentic AI: 개념적 구분과 정책적 의미¹⁾

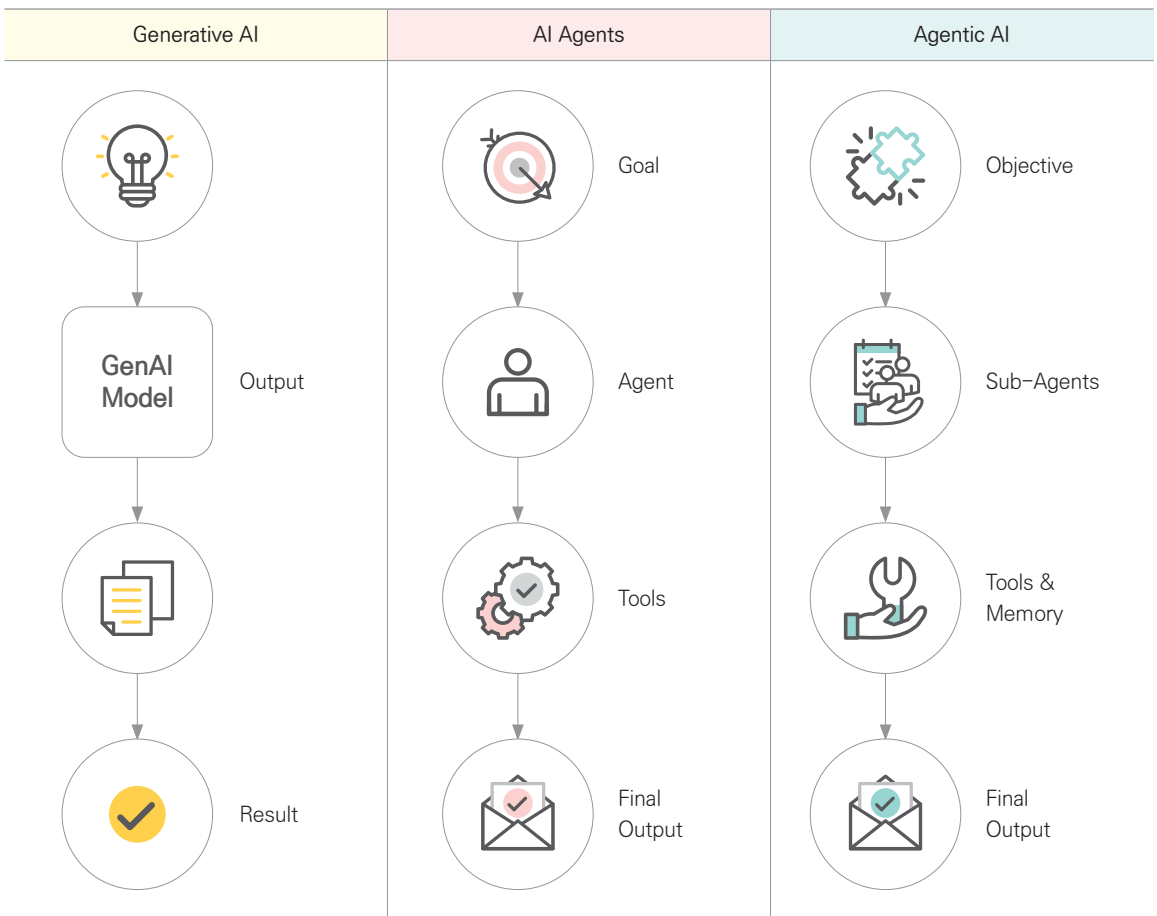
- **GenAI(생성형 AI)는 텍스트, 이미지 및 코드를 생성하는 기술로, 대규모 언어 모델(LLM)을 기반**
 - 대표적으로 ChatGPT, Claude, Bard 등이 이에 해당하며 GenAI는 사용자 입력 프롬프트에 응답하고 즉시 출력을 생성하지만 자율적으로 목표를 설정하거나 행동을 위한 계획 및 전략을 만들지는 않음
 - 행정실무에서는 보고서 작성, 질문에 대한 답변, 텍스트 번역, 데이터 요약 및 번역을 지원하는 단일 작업 지원 도구로 사용
- **AI Agent는 GenAI의 기능을 기반으로 하여 자율적으로 일련의 프로세스를 구조화, 실행 및 조정하여 미리 정해진 특정 목표나 목적을 달성하는 시스템**
 - 예컨대 에이전트는 정책 자료와 관련된 일련의 작업의 전체 워크플로를 독립적이고 자율적으로 수행할 수 있으며, 이는 수집, 요약, 분석, 초안 작성 및 초기 버전 배포 등이 해당
 - 행정실무에서는 현재 논의 중인 자동화된 에이전트가 문서 처리, 업무 분석 및 표준화된 질문 세트에 대한 반복적인 질적 및 양적 데이터 분석에 사용되며, 이 수준의 AI와 가장 유사
- **Agentic AI는 AI 에이전트의 기능을 강화하고 확장하는 보다 발전된 시스템으로, 독립적으로 목표를 설정하고 수정하고, 변화하는 상황에 대응하기 위해 새로운 정보를 적극적으로 탐색하고 통합하며, AI 에이전트, 인간 및 여러 이해관계자와 협력하여 복잡한 문제를 해결할 수 있음**
 - 예를 들어, AI는 지역 세금 체납 수집률 개선과 같은 광범위한 목표를 스스로 설정하고, 관련 데이터를 분석하여 정책 대안을 개발하거나, 옵션을 발전시키는 데 필요한 부서 간 협력 프로세스를 적극적으로 촉진할 수 있음
 - 이러한 의미에서 AI는 준자율적 정책 행위자 역할을 공유하는데, 초기 개발 중인 에이전틱 AI 시스템의 예로는 미국 GSAi (DOGE), 영국 Gov.uk Chatbot 2.0 및 싱가포르 Ask Jamie+가 있음

1) 두 용어는 영어 어법상 거의 유사하게 사용되지만, 기술 산업에서는 'AI agent'를 기능적 시스템 수준에서, 'Agentic AI'를 자율적·목표 지향적 행동을 수행하는 상위 개념으로 구분하는 경향이 있음(Mittal, 2024; Russell & Norvig, 2020; IBM, 2025). 따라서 정책 문맥에서는 'AI agent'를 행정 실무 단위의 구현 기술로, 'Agentic AI'를 자율성과 협업능력을 강화한 차세대 야기텍처로 이해

그림 11 AI 에이전트 개념 비교 참고 자료

Model Type	Architecture	Memory	Autonomy	Use Case
GenAI	Input → Model → Output	Stateless	None	Text/code generation, summarization
AI Agents	Goal → Agent → Tools → Output	Episodic	Limited	Task execution via APIs, basic RAG
Agentic AI	Objective → Sub-Agents → Tools & Memory → Output	Persistent	Full	Autonomous workflows, contextual decision-making, adaptive defense

출처: <https://www.samsungsds.com/kr/insights/agentic-ai-the-autonomous-era-of-artificial-intelligence.html> 저자 재구성



출처: <https://www.linkedin.com/pulse/generative-ai-vs-agents-agentic-ai-generative-bhuwan-mittal-xskqc/> 저자 재구성

공공부문 AI 에이전트 도입 및 활용 현황과 시사점

AI 에이전트의 개념과 기술 활용의 효과



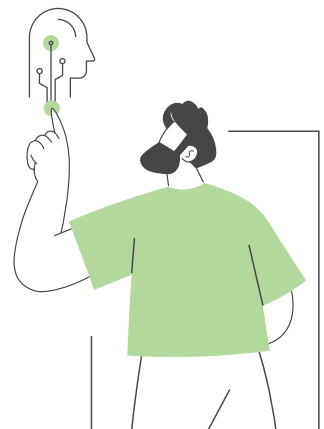
2. AI 에이전트의 개념과 기술 활용의 효과

1) AI 에이전트의 개념과 구성 요소

● AI 에이전트 개념의 발전과 주요 정리

● AI 에이전트의 개념 정의

- AI 에이전트는 최소한의 인간 개입으로 독립적인 의사결정을 수행하는 자율성(Autonomy), 명확한 목적을 설정하고 이를 달성하기 위한 계획을 수립·실행하는 목표 지향성(Goal-Oriented), 그리고 외부 도구, 시스템, 데이터베이스와 동적으로 상호작용하며 환경 변화에 반응하는 환경 상호작용(Environmental Interaction)이라는 3대 핵심 기준을 충족하는 지능형 소프트웨어 시스템¹⁾
 - 이러한 AI 에이전트는 단순히 사용자의 프롬프트에 반응하는 챗봇이나 사전 정의된 규칙만을 실행하는 RPA(로봇 프로세스 자동화)와 달리, “사용자를 대신하여 복잡한 프로젝트를 처음부터 끝까지 독립적으로 처리할 수 있는 가상의 협력자”로서 기능
 - 특히 공공부문에서는 “자율성을 바탕으로 스스로 업무를 수행하되, 투명성과 설명가능성, 그리고 인간의 최종 검토를 통한 책임성을 보장하는 신뢰할 수 있는 인공지능 시스템”으로 정의
- AI 에이전트 개념은 1995년 Russell & Norvig의 고전적 정의에서 시작하여 2025년 현재 대화형 AI와 대규모 언어모델의 발전과 함께 급속히 진화 중(IBM, 2025)
 - 환경을 인식하고 행동하는 모든 존재라는 초기 개념(Russell & Norvig, 1995)에서 자율적으로 목표를 추구하며 복잡한 업무를 완수하는 시스템으로 발전



1) 실용성과 단기 정책 적용을 우선한다면 3대 기준(자율성, 목표 지향성, 환경 상호작용)이 적절하지만, 미래 비전과 선도적 프레임 제시를 목표로 한다면 '학습 루프'가 더해진 4대 기준이 적절. 본 브리프에서는 현실적으로 3대 기준을 기본 정의로 두고 4대 기준을 장기적 발전 요소로 제시

표 1 | 기존 AI 시스템과의 구분

구분기준	AI 에이전트	챗봇	RPA	가상비서
자율성	높음: 독립적 의사결정 및 다단계 업무 수행	낮음: 사용자 프롬프트에 반응	없음: 사전 정의된 규칙만 실행	중간: 명령 실행 중심, 제한적 자율성
목표 지향성	높음: 추상적 목표를 구체적 계획으로 전환	낮음: 대화 유지가 주 목적	없음: 목표 설정 불가, 프로세스만 실행	중간: 특정 요청에 대한 직접적 응답
환경 상호작용	높음: 다중 시스템 통합, 동적 도구 활용	낮음: 제한적 외부 연동	중간: UI 기반 단순 상호작용	중간: 서비스별 API 연동
학습능력	높음: 경험 기반 적응과 개선	낮음: 사전 훈련된 응답 패턴	없음: 학습 능력 부재	낮음: 사용 패턴 기반 제한적 개인화
복잡도 처리	높음: 다단계 추론과 계획	낮음: 단순 질의응답	낮음: 단순 반복 작업	중간: 구조화된 작업 실행
대표사례	Amazon Q Developer, Genentech 바이오마커 연구 시스템	고객 상담 챗봇, FAQ 응답 시스템	송장 처리 자동화, 데이터 입력 봇	Siri, Alexa (기본 기능 사용 시)

출처: 저자작성

● 기술 기업들의 정의

- Google은 “사용자를 대신해 목표를 추구하고 업무를 완수하는 소프트웨어 시스템으로, 추론, 계획, 기억 능력을 보유하며 의사결정, 학습, 적응을 위한 일정 수준의 자율성을 갖춘 시스템”으로 정의(Google Cloud, 2025)
 - 특히 ReAct 프레임워크(추론-행동-계획-기억-학습)를 통해 단순 반응형 시스템과 구분
- Microsoft는 “특정 프로세스를 처리하거나 비즈니스 과제를 해결하기 위해 구축된 전문화된 AI 도구”로 규정 (Microsoft News, 2024)
 - 두 가지 아키텍처로 구분하는데 1) 선언적 에이전트(기존 인프라 활용), 2) 맞춤형 엔진 에이전트(독립적 오케스트레이션)
- OpenAI는 “사용자를 대신해 독립적으로 업무를 완수하는 시스템”으로 정의(OpenAI, 2025)
 - 5단계 AGI 경로에서 3단계 ‘에이전트’를 업무 지원과 자율 운영이 가능한 수준으로 설정(OpenAI, 2025)
- Anthropic은 “LLM이 동적으로 자체 프로세스와 도구 사용을 지시하며 업무 수행 방식을 제어하는 시스템”이라고 정의(Anthropic, 2024)
 - 워크플로우와 에이전트를 명확히 구분하는데 워크플로우는 “사전 정의된 코드 경로를 통해 LLM과 도구가 조율되는 시스템” (Anthropic, 2024)

● **학술적 정의와 이론적 근거**

- Russell & Norvig의 고전적 프레임워크는 여전히 학술 연구의 기초가 되고 있음
 - 최신 연구에서 Wang et al.(2023)은 LLM 기반 자율 에이전트를 “자연어로 완전한 경험 기록을 저장하고, 기억을 상위 수준의 성찰로 합성하며, 행동 계획을 위해 정보를 동적으로 검색하는 시스템”으로 정의
- MIT CSAIL은 대중 에이전트 환경과 구현된 AI 연구를 통해 에이전트의 환경 상호작용 능력을 강조
 - Stanford AI Lab의 생성 에이전트 연구는 인간 행동의 믿을 만한 시뮬레이션을 통해 에이전트의 사회적 지능을 탐구
 - Carnegie Mellon의 다중 에이전트 시스템 연구는 협력적 조정 행동을 강조
- BDI(Belief-Desire-Intention) 모델은 여전히 에이전트 아키텍처의 핵심 이론적 토대로 작용(Jennings, 1998; IBM, 2025)하며, 최근 연구는 이를 대규모 언어모델과 통합하는 방향으로 발전 중

● **국제기구 및 정부의 접근법**

- OECD는 AI 시스템을 “주어진 목적을 위해 예측, 권장사항, 또는 실제/가상 환경에 영향을 미치는 결정과 같은 결과를 생성할 수 있는 공학적 또는 기계 기반 시스템으로, 다양한 수준의 자율성으로 작동하도록 설계된 시스템”으로 정의(OECD, 2024)
 - NIST AI 위험관리 프레임워크는 신뢰할 수 있는 AI의 7가지 특성(유효하고 신뢰할 수 있음, 안전함, 보안성과 복원력, 책임감, 투명성, 설명 가능하고 해석 가능함, 개인정보 보호 강화, 공정함)을 제시하며 위험 기반 접근법을 채택(NIST, 2025)
- EU AI 법은 위험 기반 분류를 통해 AI 시스템을 규제하며, 특히 상당한 자율성으로 운영되는 시스템에 대한 특별한 주의를 기울임
 - 일반목적 AI 모델에 대한 별도 규정을 두어 다양한 용도로 활용될 수 있는 기초 모델을 관리(EU AI Act, 2025)

표 2 | AI 에이전트의 개념 정의

출처/조직	정의	핵심 구성요소/특성	연도
Google	“시를 활용하여 사용자를 대신하여 목표를 추구하고 작업을 완료하는 소프트웨어 시스템입니다. 추론, 계획, 기억 능력을 갖추고 있으며, 의사 결정, 학습 및 적응을 위한 일정 수준의 자율성을 갖추고 있습니다.”	ReAct 프레임워크(추론 + 행동), 다중 모드 처리, 자율성, 학습/적응, 도구 통합, 다중 에이전트 조정	2024년
Microsoft	“에이전트는 결정을 내리고, 도구를 활용하고, 워크플로에 참여합니다. 때로는 독립적으로, 때로는 다른 에이전트나 사람과 협력하여 참여합니다. 에이전트를 어시스턴트와 차별화하는 것은 자율성입니다. 어시스턴트는 사람을 지원하고, 에이전트는 목표를 달성합니다.”	3가지 핵심 구성 요소(모델, 지침, 도구), 자율성 대 지원 구분, 메모리/자격/도구, 구성 가능한 단위, 프로덕션 준비 보안	2024년
OpenAI	“에이전트는 사용자를 대신하여 독립적으로 작업을 수행하는 시스템입니다. 에이전트는 LLM을 사용하여 지시를 실행하고 결정을 내립니다.”	4가지 기본 요소(에이전트, 핸드오프, 가이드 레일, 세션), LLM + 지침 + 도구 아키텍처, 자율적 작업 실행, 경량 프레임워크	2024년
Anthropic	“AI 에이전트는 목표가 주어지면 자율적으로 작업을 수행합니다. 에이전트를 복잡한 프로젝트를 처음부터 끝까지 독립적으로 처리할 수 있는 가상의 협력자라고 생각해 보세요.”	자율적인 작업 추구, 가상 협업, 프로세스 제어, 최소한의 인적 투입, 안전/신뢰성 원칙	2024년
MIT CSAIL	“최소한의 인간의 감독으로 목표 지향적 행동을 취하고, 상황에 맞는 결정을 내리고, 변화하는 상황에 따라 계획을 조정함으로써 생산성을 높이는” “자율성과 예측 가능성을 갖추고 복잡한 목표를 추구하도록 설계되었습니다.”	행동 지향적(생성적 대비), 자율적 운영, 독립적인 작업 실행, 동적 반응성, 인간-AI 협업 패턴	2024-2025년
Stanford University	“자연어로 된 완전한 경험 기록”을 통해 “과거 경험을 기반으로 행동 계획”을 가능하게 하는 “믿을 수 있는 인간 행동을 시뮬레이션하는 계산 소프트웨어 에이전트”입니다.	기억 합성, 동적 검색, 3가지 핵심 구성 요소(관찰, 계획, 반성), 다중 모드 지각, 체화된 지능	2023-2025년
Carnegie Mellon	“여러 단계로 이루어진 작업 완료”를 필요로 하는 “탐색, 사회적 상호 작용, 상식적 추론”이 필요한 “복잡한 사무 작업을 자율적으로 수행하는” 시스템입니다.	환경 탐색, 사회적 상호 작용 가능, 상식적 추론, 단단계 작업 완료, 실제 성능 벤치 마킹	2024-2025년
OECD	“명시적 또는 암묵적 목적에 따라 입력된 데이터를 바탕으로 예측, 콘텐츠, 추천, 의사 결정 등의 출력을 생성하는 방법을 추론하여 물리적 또는 가상 환경에 영향을 미칠 수 있는 기계 기반 시스템입니다.”	추론 기능, 명시적/암시적 목표, 다양한 자율성 수준, 물리적/가상 환경 영향, 신뢰할 수 있는 AI 원칙	2019-2024년
IEEE	상황에 따른 정의: “사이버 물리 시스템에서 저수준 자동화 장치와 지능형 소프트웨어 에이전트 통합”(IEEE 2660.1-2020)	인터페이스 솔루션, 재사용/일관성/투명성, 산업 자동화 중심, 상호 운용성 표준	2017-2020년
ISO/IEC	“신뢰성, 안정성, 윤리적 고려 사항”을 강조하는 위험 기반 접근 방식을 적용한 포괄적인 AI 용어 프레임워크의 일부입니다.	위험 관리 중심, 수명 주기 관리, 계획-실행-확인-조치 방법론, 기존 프레임워크와의 통합	2022-2023년
NIST	“주어진 일련의 목표에 대해 예측, 권장 사항, 실제 또는 가상 환경에 영향을 미치는 의사 결정 등의 결과를 생성할 수 있는 엔지니어링 또는 기계 기반 시스템으로... 다양한 수준의 자율성을 가지고 작동하도록 설계되었습니다.”	다양한 자율성 수준, 신뢰성 특성(타당성, 안전성, 보안성, 책임성, 설명 가능성, 개인 정보 보호, 공정성), 4가지 기능 접근 방식	2023-2025년
EU AI Act	“다양한 수준의 자율성으로 작동하도록 설계된 기계 기반 시스템으로, 배포 후 적응성을 보일 수 있으며... 물리적 또는 가상 환경에 영향을 줄 수 있는 출력을 생성하는 방법을 추론합니다.”	다양한 자율성 수준, 배치 후 적응성, 위험 기반 분류 시스템, 기본권 보호	2024년

출처: 저자작성

● AI 에이전트의 구성 요소

● 권위 있는 모든 출처의 정의들 간에 다섯 가지 보편적 특성이 공통 요소로 도출

- 경계가 있는 자율성은 가장 일관된 주제
 - 모든 정의는 시스템이 독립적으로 작동할 수 있으면서도 적절한 감독 메커니즘이 필요함을 강조하는데 이는 구글의 “의사결정 자율성 수준”부터 NIST의 “다양한 수준의 자율성”, EU의 위험 기반 자율성 분류까지 다양하게 나타남
- 목표 지향적 행동은 모든 프레임워크에서 나타남
 - “목표 추구”(구글, Anthropic), “과제 완수”(OpenAI, CMU), “명시적/암묵적 목표”(OECD, EU)로 묘사되며, 이는 단순 반응형 시스템과 달리, 정의된 목표를 향한 능동적인 실행이 강조
- 환경과의 상호작용은 보편적으로 인정되는 시스템의 필수 요소
 - OECD, EU, NIST는 “물리적 및 가상 환경”을 포함하는 시스템을 설명했으며, 학술 문헌에서는 “환경 인식 및 행동”에 초점을 맞추어 경계가 없는 AI 모델과 에이전트를 구별하는 기준으로 삼음
- 의사결정 능력은 기술적 및 규제적 출처 전반에 걸쳐 지속적으로 강조됨
 - Microsoft는 “에이전트가 독립적으로 목표를 달성한다”고 강조하며, 정부 프레임워크는 “예측, 추천 또는 의사결정”을 요구 (NIST, EU)
- 도구와 시스템의 통합은 산업(구글의 통합 도구, Microsoft의 모듈형 유닛)과 학계 (CMU의 다중 도구 사용 요구 사항) 전반에 걸쳐 실용적인 필요조건으로 강조
 - 이는 에이전트가 실제 환경에서 복잡한 다단계 작업을 수행할 수 있게 함

● AI 에이전트의 3가지 핵심 판별 기준과 학술적 근거

● 기존 연구에서 제시된 5가지 요소를 정책적 실용성을 고려하여 자율성, 목표 지향성, 환경 상호작용이라는 3가지 핵심 기준으로 통합(arXiv, 2025)

- 이는 Woodridge & Jennings(1995)의 반응성, 능동성, 사회적 능력과 일치하며, 현대 LLM 기반 시스템의 특성을 반영
 - 이를 토대로 자율성, 목표 지향성, 환경 상호작용이라는 3가지 핵심 기준이 기존 AI 시스템과의 명확한 구분점을 형성한다는 것은 핵심 발견사항(Nb-Data, 2025; TechTarget, 2025)
 - 이러한 기준은 Russell & Norvig의 고전적 5요소 프레임워크를 현대적으로 통합한 것으로, 정책 입안자들이 AI 에이전트의 규제와 거버넌스에 활용할 수 있는 실무적 정의를 제공

● 자율성(Autonomy)

- 최소한의 인간 개입으로 독립적인 의사결정과 행동이 가능한 능력(Google Cloud, 2025)
 - Russell & Norvig의 “지속적인 시간동안 자율적으로 운영”하는 특성과 일치하며(IBM, 2025), Jennings et al.(1998)의 “외부 통제 없이 자체 행동과 내부 상태를 제어”하는 정의를 포함(Jennings, 1998)
 - 기술적 지표: 사람의 단계별 지시 없이 다단계 업무 수행, 예상치 못한 상황에서의 적응적 대응, 오류 발생 시 독립적인 복구 시도, 업무 우선순위의 동적 조정

● **목표 지향성(Goal-Orientation)**

- 명확한 목적을 설정하고 이를 달성하기 위한 계획 수립 및 실행 능력(IBM, 2025)
 - BDI 모델의 Desire(욕구)와 Intention(의도) 구성요소와 직접 연관되며, 목표 기반 에이전트와 효용 기반 에이전트의 핵심 특성(IBM, 2025)
 - 기술적 지표: 추상적 목표를 구체적 하위 업무로 분해, 목표 달성을 위한 동적 계획 수정, 다중 목표 간 우선순위 조정, 목표 달성도 자체 평가 및 보고

● **환경 상호작용(Environmental Interaction)**

- 외부 도구, 시스템, 데이터베이스와 동적으로 상호작용하며 환경 변화에 반응하는 능력(n8n, 2025)
 - Russell & Norvig의 “센서를 통한 환경 인식과 액추에이터를 통한 행동”이라는 기본 정의의 현대적 해석이며, Cheng et al.(2024)의 “인식-의사결정-행동” 주기와 일치(n8n, 2025)
 - 기술적 지표: 외부 API, 데이터베이스, 웹 서비스 활용, 실시간 정보 수집과 처리, 다중 시스템 간 데이터 통합, 환경 변화에 따른 전략 수정

● **이론적 통합: 5요소에서 3기준으로의 발전**

• **이론적 간결성(Theoretical Parsimony)**

- 더 단순한 프레임워크가 구현과 평가를 용이하게 한다는 인지과학적 원리에 기반
- Occam’s razor의 원리에 따라 “필요 이상으로 복잡하지 않은 단순한 설명”이 선호되며, 인지과학 연구는 단순성이 이해와 학습을 촉진한다는 증거를 제시(ScienceDirect, 2024)

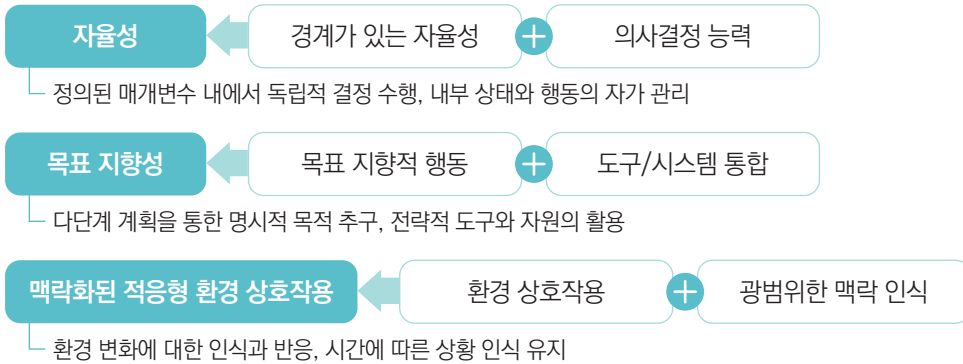
• **실용적 구현**

- 3기준이 시스템 아키텍처에 더 직접적으로 매핑된다는 MIT와 Stanford의 공동 연구 결과 반영
- 개발과 유지보수가 간소화되며, Microsoft Azure의 AI 에이전트 연구에 따르면 “모듈성은 개별 구성요소를 독립적으로 개발, 테스트, 유지 관리할 수 있게 함(Microsoft Learn, 2025)

• **평가 명확성**

- 적은 차원이 더 집중된 평가를 가능하게 한다는 측정학적 원리를 적용
- AI 에이전트의 성능 평가를 보다 체계적으로 수행할 수 있게 함(IBM, 2025)

● 통합 매핑 구조



2) 기술 활용의 효과

● AI 에이전트의 효과는 먼저 민간 부문에서 업무의 자동화, 서비스의 혁신, 생산성의 개선과 같은 구체적인 성과를 통해 증명되고 있음

- 반면 공공 부문은 아직 본격적인 실증연구와 적용 사례가 쌓이지 않은 상태이나 AI 기술의 행정 적용 사례와 민간의 적용 성과를 종합적으로 고려할 때 행정의 서비스 수급과 서비스의 질 향상 측면에서 상당한 잠재적 효과를 미칠 것으로 예상
 - 이는 단순한 기술 수용 단계를 넘어서 행정 시스템 혁신의 측면으로 AI 에이전트 도입의 논의가 진행되어야 한다는 점을 시사

● 행정 생산성 및 업무 효율성 향상

● AI 에이전트는 반복적인 행정업무를 자동화하고 대국민 서비스를 개인화하는 도구로서, 공공부문의 효율성을 크게 향상

- 공무원은 더 복잡한 업무나 정책기획에 집중할 수 있어 행정의 생산성이 높아지고, 국민은 언제 어디서든 편리한 서비스를 받을 수 있음
 - 또한 AI 에이전트가 축적된 데이터를 분석해 정책의 사전예측과 평가를 지원하면, 정부는 객관적인 근거를 바탕으로 의사 결정을 내릴 수 있음
- 측정 가능한 성과 지표로는 영국 정부 시험에서 정부 직원 1인당 일일 26분 절약(연간 1,130명에게 1년을 돌려 주는 것과 동등), 교사들의 수업 계획에 소요되는 일일 100분 중 최대 75% 절약, 공무원들의 이메일 처리 시간 70% 이상 단축(DSIT, 2025)
 - 운영 우수성 사례: (1) 런던 온타리오의 만성 노숙 예측 AI 모델 93% 성공률, (2) 자동화된 특허 출원 처리로 심사관 효율성 증가, (3) AI 기반 사기 탐지가 인간 감사관이 놓치는 패턴 식별, (4) 예측 유지보수를 통한 정부 시설 운영 비용 절감(LGIU, 2024)

- 시민 만족도 개선은 24시간 AI 챗봇 지원, 싱가포르의 70개 이상 정부 웹사이트에 배치된 “Ask Jamie” 가상 비서, 대기 시간 단축과 응답 정확도 향상, 장애인이나 언어 장벽이 있는 시민을 위한 접근성 개선 등으로 나타남 (Jackson, 2024)
- **연구에 따르면 영국 공공부문 업무의 56%가 AI로 보완 가능하며, 이를 통해 행정 처리량을 획기적으로 늘릴 수 있다고 함(김은영, 2024)**
- 실제 시뮬레이션 결과 AI 도입 시 주당 약 370만 건의 추가 의료진료를 소화하고 교사 1인당 학생비율을 16% 개선하는 등의 공공서비스 증대 효과를 기대(김은영, 2024)
 - 이렇듯 진보한 형태의 AI 에이전트 기술은 동일 시간에 더 많은 업무를 처리하고 의사결정을 지원함으로써 전반적인 행정 생산성 혁신을 가능케 할 것

● 도입의 경제적 효과

- **영국 정부는 2025년 1월에 발표한 ‘AI 기회 행동 계획’에서 AI 도입으로 2030년까지 영국 경제에 약 4,000억 파운드 규모의 추가 시장 잠재력이 있다고 평가**
- 투자 수익률(ROI) 분석에서는 Q4 Inc.의 3년간 212% ROI, Microsoft Azure Arc의 3년간 304% ROI, Dell Technologies의 투자 달러당 2-3배 수익 추정, Tony Blair Institute의 영국 공공부문 AI 구현에 대한 9:1 편익-비용 비율 등 200% 이상의 ROI가 다수 문서화
 - 이는 특정 생성형 AI 도구 도입 사례의 결과로, AI 에이전트 도입 시 기대할 수 있는 잠재적 효과를 보여주는 참고 지표이며, AI 에이전트 도입에 따른 생산성과 혁신 증대 및 공공서비스 개선을 통한 경제 전체에의 기여를 시사
- **업무 자동화와 최적화된 자원배분을 통해 대규모 비용 절감 효과가 입증되고 있음**
- 미국에서는 Deloitte 분석에 따라 정부 자동화가 연간 9,670만 시간에서 12억 시간의 노동 시간을 절약하여 연간 33억-411억 달러의 잠재적 절약을 가능(Deloitte, 2025)
 - 미국 뉴저지주 등에서도 생성형 AI 활용을 통해 주 정부 예산을 수억 달러 규모로 절약한 사례가 보고(ZDNET, 2025)
- 영국에서는 Tony Blair Institute 분석이 공공부문에서 AI를 완전히 활용할 경우 2030년까지 연간 약 380억 파운드 (GDP의 1.3%)의 공공 비용 순 절약을 예상(김은영, 2024)
 - 이는 동일 인력으로 더 많은 서비스를 제공하거나 절감된 재원을 다른 필수 분야에 투자하는 선순환으로 이어짐
- **행정 처리, 사기 탐지, 문서 관리, 인프라 비용 등에서 구체적인 시간과 비용 감소 사례 입증(Link, 2023; Krishnamurthy, 2025)**
- **행정 처리:** 일상 업무 처리 시간 75% 단축, 사기 탐지: 캐나다 AI 시스템이 첫해에 5억 파운드의 미납 세금 회수
- **문서 관리:** 자동 분류 및 라우팅으로 처리 시간 50% 이상 단축, 인프라 비용: AI 최적화 시스템을 통해 공공부문 조직의 데이터 저장 비용 75% 감소

● **미국 연방정부는 AI 권리장전과 NIST AI 위험관리 프레임워크를 통해 각 부처의 AI 활용을 장려하고 있음**

- 비자 심사·세금 징수·재난 대응 등에서 AI 에이전트를 시범 적용하여 업무 효율성과 응답 속도 향상을 경험하고 있음
 - 고도화된 AI 에이전트 기술은 시장 잠재력과 투가 가치를 기반으로 최적화된 자원배분, 더 큰 가치 창출, 대규모 비용 절감 등의 효과로 세계 각국 정부의 활용이 증대될 것으로 예상

● **공공서비스 혁신 및 품질 향상**

● **영국의 계획은 교육·보건·인프라 등 12개 핵심 분야에 AI를 적용해 공공서비스 품질을 높이고 디지털 일자리를 창출하려는 목표를 포함**

- 미국도 AI를 통해 민원 상담, 음성봇, 문서 분석 등 반복 업무를 자동화해 공무원이 고부가가치 업무에 집중할 수 있도록 하고 있음

● **AI 에이전트는 대민 서비스의 접근성과 반응속도를 높임**

- 예를 들어 미국 미시간주 앤아버시는 71개 언어로 24시간 응대하는 AI 챗봇 ‘Ask Ann’을 도입하여 시민들이 시간이나 언어 장벽 없이 시정부 서비스를 이용할 수 있게 함으로써 서비스 편의성과 만족도를 제고(김은영, 2024)
 - 이처럼 AI 에이전트는 상시 맞춤형 서비스 제공을 통해 국민 경험을 향상시키는 효과

● **행정 조직과 거버넌스 구조 변화**

● **알고리즘에 의한 의사결정이 늘어남에 따라 투명성과 책임성 확보를 위한 새로운 제도와 절차가 도입되고 있음**

- 영국 정부는 알고리즘 투명성 기록(Standard) 준수를 중앙부처에 의무화하고 고위험 AI 시스템을 검증할 AI 안전 연구소를 설립하는 등 거버넌스 측면의 대응에 나섬
 - 공공조직 내부에서는 AI 활용을 관리·감독하기 위한 위원회 신설, 내부 감사 강화, 결과 설명 절차 마련 등이 추진되고 있으며, 의사결정 과정에서 인간의 검증 역할을 재정립하는 노력이 진행 중
 - AI 에이전트의 도입 이전에 행정의 거버넌스와 책임 및 검증 시스템에 대한 논의는 이미 진행 중이었으나 AI 기술의 채택 및 구현이 증가함에 따라 자율적인 의사결정 및 실행 능력을 갖춘 시스템을 고려한 새로운 거버넌스 원칙 및 검증 프로세스의 필요성이 점점 더 분명해지고 있음
 - 요컨대, 기존의 AI 거버넌스가 ‘윤리·책임·투명성’ 중심의 관리 체계였다면, AI 에이전트 거버넌스는 자율적 의사 결정과 상호 작용 능력을 전제로 한 ‘행위 책임(accountability of action)’과 ‘결과 검증(result traceability)’을 추가로 다룰 필요가 있음
 - 다만 이러한 별도 거버넌스의 필요성에 대해서는 기술 성숙도와 행정 통제 가능성을 고려한 점진적 접근이 요구

공공부문 AI 에이전트 도입 및 활용 현황과 시사점

국내 AI 에이전트 도입 전략 및 사례



3. 국내 AI 에이전트 도입 전략 및 사례

1) 국내 공공부문 AI 정책의 에이전트 관련성 분석

● 국내 공공부문 AI 정책·전략 동향

● AI 에이전트 전략적 진화 과정

- 국내 공공부문 AI 도입 전략은 2022년 에이전트 개념이 없는 지능형 시스템 수준에서 출발하여, 2023년 AI 시스템의 자율성 강화 방향을 제시하고, 2025년 AI 에이전트를 명확히 정의하며 실제 구현 사례를 추진하는 단계로 발전
- **1단계:** 에이전트 개념 부재 단계 (2022-2023년 초)
 - AI 에이전트 개념 직접 언급 없이 '지능형 시스템' 수준의 접근
 - 자동관제, 지능형 망 관리 등 초기 형태의 자율 시스템 구상
- **2단계:** 자율성 강화 방향 제시 단계 (2023년)
 - AI 시스템의 자율적 운영 기반 마련 방향 최초 제시
 - 논리적 사고, 실시간 정보 반영 등 에이전트 핵심 기능 개발 착수
 - 실시간 모니터링·예측 기반 자율적 대응 시스템 구축 계획
- **3단계:** 명시적 정의 및 실증 단계 (2025년)
 - 국내 공공부문 최초로 AI 에이전트 명확한 개념 정의
 - 에이전틱 시를 차세대 AI 발전 방향으로 명시
 - 실제 AI 에이전트 구현 사례 추진

표 3 | 국내 공공부문 AI 정책문서별 에이전트 관련성 분석

정책명	발표시기	주무부처	핵심목표	AI 에이전트 관련성
대한민국 디지털 전략	2022.09	관계부처합동	세계 최고 디지털 역량 확보	낮음: 간접적 연관 (지능형 망 관리)
인공지능 일상화 계획	2023.01	과기정통부	전국민 AI 일상화, 공공·산업 융합	보통: AI 시스템 자율성 강화 방향
초거대 AI 경쟁력 강화 방안	2023.04	관계부처합동	초거대 AI 기술·산업 인프라 확충	보통: 논리적 사고, 실시간 정보 반영
전국민 AI 일상화 실행계획	2023.09	관계부처합동	AI 활용 양적·질적 확대	보통: 실시간 모니터링·예측 자율 대응
국가 AI 전략 정책방향	2024.09	관계부처합동	AI G3 도약, 국가 AX 전면화	낮음: 간접적 연관 (지능형 업무관리)
공공부문 초거대 AI 가이드라인 2.0	2025.04	디플정위	공공 AI 3대 전략목표	높음: AI 에이전트 개념 정의 및 적용
초거대 인공지능 서비스 개발 지원 사업	2025	디플정위	행정 효율화, 대국민 서비스 혁신	높음: AI 에이전트 실증 서비스 구현

출처: 저자작성

● **대한민국 디지털 전략 (관계부처합동, 2022.09)**

- 디지털 역량 강화를 위한 범정부 차원의 기본 전략으로 AI를 6대 디지털 혁신 기술 중 핵심 요소로 위치
- 5대 전략 방향으로 디지털 역량, 디지털 경제, 디지털 사회, 디지털 플랫폼 정부, 디지털 문화 혁신 제시
 - AI-데이터 기술 기반 정부 업무 방식 혁신과 AI 융합 시대 본격화를 핵심 추진과제로 설정
 - 전 분야 AI 학습용 데이터 구축 확대 및 데이터 문제해결은행 구축을 통한 중소기업 지원 강화
- **AI 에이전트 관련성:** 에이전트 개념의 직접적 논의는 부재하나, 향후 AI 에이전트 사업으로 발전하는 물관리·스마트 복지 등 지능형 시스템 구축 계획 포함
 - AI 등 신기술 기반 지능형 망 관리 시스템 도입('23년~)을 통한 자동관제 체계 구축

AI 에이전트 정책과의 연계 가능성

해당 전략은 AI 에이전트의 개념이나 정책적 추진전략을 명시적으로 다루고 있지 않으며, 주로 AI-데이터 인프라 확충과 디지털 정부 전환의 기본 방향 설정에 초점이 맞춰져 있음

- 다만, 물관리·스마트복지 등 지능형 시스템 구축 과제를 통해 향후 AI 에이전트 적용 분야가 확장될 수 있는 잠재적 기반을 제공하고 있다는 점에서 간접적 관련성을 가짐

● **인공지능 일상화 및 산업 고도화 계획(안) (과기정통부, 2023.01)**

- 단일지능을 넘어 복합·초거대 인공지능으로의 기술 발전 패러다임 전환 및 신뢰성·윤리 기반 산업 성장 견인
- 신뢰성·윤리 기반의 산업 성장 견인을 위한 2대 핵심 프로젝트 추진
- **전 국민 AI 일상화 프로젝트:** 국가 대표 AI를 통한 사회적 약자 보살핌 및 민생문제 해결
- **공공·산업 AI 전면 융합 프로젝트:** 제조·물류·반도체 등 기존산업의 AI-데이터 기반 전환
- **AI 에이전트 관련성:** 기존 로드맵 과제의 자율성 강화 방향 제시
 - 사전 연구 수준의 과제 9건¹⁾에 대한 법·제도 개선방안 도출 및 정비 방향 구체화로 AI 시스템의 자율적 운영 기반 마련

AI 시스템 자율성 강화를 통한 에이전트 전환 기반 마련

본 계획(안)은 단순한 AI 활용 확산이 아니라, 단일 지능에서 복합·초거대 지능으로의 기술 패러다임 전환을 전제로 하며, AI 시스템의 자율적 운영을 위한 법·제도적 기반을 구체화하기 시작했다라는 점에서 AI 에이전트 발전과 밀접한 관련성

- 사전 연구 단계의 9개 과제에 대해 법·제도 개선 방향을 정비함으로써, 향후 다양한 산업 및 공공 서비스 분야에서 AI 시스템이 인간의 직접적 개입 없이 자율적으로 판단·운영될 수 있는 제도적 환경을 마련하려는 시도가 이루어짐
- “전 국민 AI 일상화 프로젝트”와 “공공·산업 AI 전면 융합 프로젝트”는 각각 사회적 약자 돌봄·민생문제 해결, 그리고 산업 전반의 지능형 전환을 목표로 하고 있어, 중장기적으로는 실시간 환경 인식-판단-행동이 가능한 에이전트형 서비스로 발전할 여지가 큰 영역
- 이러한 맥락에서 본 계획은 기술·제도·사회적 인프라 측면에서 AI 에이전트 확산을 위한 토대를 마련한 초기 전략으로 해석 가능

1) 인공지능 행위에 대한 손해배상 방안 마련, 인공지능 개입 범죄에 대한 제재방안 다양화, 인공지능 법인격 부여 방안 정립, 인공지능 사고 처리를 위한 보험제도 개편 방안 마련, 인공지능 기술기준 마련, 인공지능 성년후견제도 도입, 인공지능 보호·안전 관리체계 구축, 인공지능 투명성·윤리 확보 방안, 인공지능 국제협력 체계 마련

● **초거대 AI 경쟁력 강화 방안 (관계부처합동, 2023.04)**

- 초거대 AI 경쟁력 강화를 위해 인프라 확충, 혁신 생태계 조성, 제도·문화 정착을 3대 전략으로 제시
 - 전략 1: 초거대 AI 기술·산업 인프라 확충 (학습용 데이터, 국가 지식 개방, K-클라우드 개발)
 - 전략 2: 초거대 AI 혁신 생태계 조성 (민간 5대 전문영역 플래그십 프로젝트, 공공·행정 서비스 개발)
 - 전략 3: 범국가 AI 혁신 제도·문화 정착 (규제 개선, 신뢰성 평가, 사회적 수용력 향상)
- 초거대 AI 한계 돌파를 위한 차세대 기술 개발 착수('24년~)
 - 인과관계 이해, 실시간 정보 반영, 편향성 필터링 등 AI 에이전트의 자율적 추론·판단 능력 강화 기술 개발
- **AI 에이전트 관련성:** 초거대 AI 한계 돌파를 위한 차세대 기술 개발이 AI 에이전트 핵심 기능과 직결
 - 에이전트 핵심 기능과 직결되는 자율적 추론·판단 능력 강화 기술 개발 및 디지털플랫폼정부 테스트베드에 소스코드 자동 생성 기능 제공

AI 에이전트 기술 발전과의 직접적 연계성

- 본 강화 방안은 초거대 AI 모델 자체의 기술적 성숙을 목표로 하지만, 차세대 기술 개발 과제로 제시된 인과관계 이해, 실시간 정보 반영, 편향성 필터링 기술은 AI 에이전트의 자율적 추론·판단 기능을 구현하는 핵심 요소와 맞닿아 있음
 - 디지털플랫폼 정부의 테스트 베드에 자동 소스 코드 생성 기능을 제공하는 계획은 미래의 행정 및 정책 AI 에이전트가 자율적으로 기능 모듈을 생성하고 수정하며 자기 구성 및 도구 사용 기능으로 작업을 수행할 수 있는 기초를 마련
 - 따라서 이 전략은 AI 인프라의 단순한 확장을 넘어 AI 에이전트의 정교함과 국가 수준 시연 환경 통합을 위한 필요한 기술 진보를 준비하며, 이는 핵심 전략적 발전에 대한 초기 노력과 직접적으로 연관

● **전국민 AI 일상화 실행계획 (관계부처합동, 2023.09)**

- 국민의 삶의 모든 분야에서 인공지능의 활용을 질적·양적으로 확대하기 위한 구체적 방안 제시
 - 일상, 생산, 공공 3대 분야의 AI 내재화 속도 가속을 위한 변화 촉진
- 사회적 약자의 돌봄을 위한 AI 기반 복지 서비스의 고도화
 - 복지 위기가구 AI 분석을 통한 선제적 발굴 시스템 구축 및 AI 초기상담 서비스 시스템 구축(복지부, '24년) AI의 복지 상담 서비스 구축(복지부, '24년)
- **AI 에이전트 관련성:** 실시간 모니터링, 예측 및 자가 반응 행동 시스템을 기반으로 한 자율 시스템 개발
 - 환경 변수 모니터링 및 대응을 위한 자가 대응 능력이 활성화된 전국 AI 홍수 예측 시스템의 구축 및 운영(환경부)
 - 스마트 도시 홍수 시스템의 시연을 통해 공간-시간 데이터에 대한 실시간 AI 분석을 기반으로 한 홍수 위험 모니터링 및 예측의 자동화(과학기술정보통신부, 2024년 예상)

실시간 모니터링 및 예측에 기반한 자율 시스템과의 통합

- 이 계획은 사회 복지 및 환경 분야에서 공공 서비스 제공에 AI를 통합하는 데 중점을 둠
 - 특히 사회 복지 서비스 분야에서는 자동화된 위험 유발 분석 및 대응 시스템을 위한 AI 에이전트의 적용이 주목받고 있음
 - 이때 '사회 복지 서비스와 주변 환경 시스템'은 에이전트가 위험을 감지하고 효과적인 대응 방안을 수립할 때의 주요 분석 대상이 되며, 행정 서비스 제공 과정과 환경 데이터를 모두 포함
- '센서/데이터 수집 → AI 분석 및 예측 → 정책/행정 대응의 자동화 실행'이 포함된 프로세스의 구축은 자급 자족적인 행동 구조와 자동화된 위험 분석 대응 시스템을 예시하며, 이는 실제로 핵심 AI 에이전트 기술 구조(감지-결정-행동 루프)를 구현
 - 따라서 이 계획은 공공 부문 AI 에이전트가 환경에 대한 실시간 '감지' 변화에 적용할 수 있도록 하여, 환경에 제시된 자극에 대해 자가 결정 및 자가 행동할 수 있는 전략으로 구상

● 국가 AI 전략 정책방향(안) (관계부처합동, 2024.09)

- 'AI G3' 도약을 위한 범국가적 AI 혁신 비전 제시
- 4대 프로젝트를 통한 정책 패러다임 전환 추진: AI컴퓨팅 인프라 확충, 민간투자 확대(65조원), 국가 AX 전면화, AI 안전·안보 확보
- **AI 에이전트 관련성:** 직접 명시된 내용은 없으나, 공공부문 계획 중 AI 에이전트 서비스와 연장선상에 있는 사업들이 포함
 - 지능형 업무관리 플랫폼 구축 계획이 향후 에이전트형 서비스로 발전 가능성 시사

공공부문 지능형 플랫폼 구축과 AI 에이전트 전환 가능성

- 본 정책방향(안)은 AI 컴퓨팅 인프라 확충, 민간 투자 확대, 국가 전 부문의 AX(Agile Transformation) 추진 등을 통해 행정 전반의 디지털 전환과 업무 혁신을 촉진하는 것을 목표로 함
 - 특히 주목할 점은 공공부문 지능형 업무관리 플랫폼 구축 계획으로 단순한 정보화 시스템을 넘어, 행정업무의 자동화·지능화를 단계적으로 추진하기 위한 것
 - 특히, 국가 차원의 컴퓨팅 자원 확충 및 표준화된 AI 서비스 인프라 구축 계획은 AI 에이전트의 대규모 운영·확산을 위한 기반(Backend Ecosystem)으로 작용하며, 범정부 차원의 AX 전면화 정책과 결합되어, 부처별·기관별로 분절적으로 존재하던 디지털 시스템을 자율적 의사결정 및 환경 적응이 가능한 지능형 에이전트 네트워크로 확장할 수 있는 여지 제공
 - 따라서 이 정책방향은 AI 에이전트를 직접 언급하지 않더라도, 공공부문에서의 AI 에이전트형 행정 서비스로의 전환 가능성을 전략적으로 내포

● **공공부문 초거대 AI 도입·활용 가이드라인 2.0 (디플정위, 2025.04)**

- AI 에이전트에 대한 명확한 개념 정의 제시
 - “사람이 직접 프롬프트를 입력하는 LLM 기반 AI 챗봇과 달리 자율성을 바탕으로 스스로 업무를 수행하는 인공지능”으로 규정
- 공공 AI 3대 전략목표를 통한 체계적 도입 방향 설정
 - 대국민 서비스 혁신: AI로 실현하는 초개인화·無장애 대국민 서비스 제공 확대
 - 사회문제 해결: 복잡한 사회 문제 해결을 위해 데이터 분석 및 실시간 대응 역량을 활용한 AI 활용
 - 작업 효율성 향상: 일상 업무의 간소화와 AI 협력을 통한 정책 정보의 향상을 통한 최적의 의사결정 지원
- **AI 에이전트 관련성:** 차세대 AI 개발을 위한 경로로서 AI 에이전트의 최근 기술 발전을 구체화
 - 사용자 명령에 따라 목표 달성을 위한 단계를 자율적으로 구상하고 실행할 수 있는 가상 인력의 잠재 가능성 제시
 - 복수의 교통수단이 결합된 티켓 발급의 복합업무를 실행하는 서비스 모델은 LLM 에이전트 상호 연계 서비스를 통해 구체화

AI 에이전트 개념의 제도적 정립과 공공서비스 적용 모델 구체화

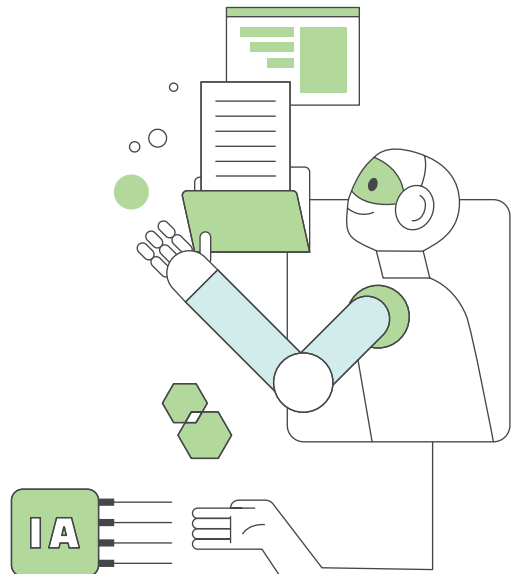
- 본 가이드라인은 공공부문 정책 문서 중 처음으로 AI 에이전트의 개념을 명확히 정의하고, 이를 공공서비스 혁신의 핵심 기술로 위치시켰다는 점에서 중요한 전환점
 - 단순한 질의 응답자가 아니라, AI 에이전트는 다양한 행정, 공공 데이터 및 시스템 아키텍처를 통합하고 자율적으로 복잡한 작업 흐름을 수행
 - 이러한 유형의 서비스는 기본 기능 루프로서 에이전트 관리에서 전형적인 지각-결정-행동 프로세스를 통합
 - 사용자 요청을 통한 목표 설정, 이질적인 자원의 탐색 및 통합, 계획 수립 및 행동 실행. 따라서 이러한 점에서 가이드라인은 AI 에이전트를 근본적인 정책 및 제도적 개념이 아닌 기술적 개념으로 위치 지으며, 그 결과 공공 서비스 디자인을 명시
 - 따라서 이는 공공 부문에서 AI 에이전트와 관련된 향후 프로젝트의 기준선 및 이정표로 역할

● **초거대 인공지능 서비스 개발 지원 사업 (디플정위, 2025)**

- 공공분야 초거대 인공지능 도입·확산을 통한 행정 효율화, 대국민 서비스 혁신, 사회 현안 해결 목표
 - 국민이 체감할 수 있고 실질적 변화를 가져올 수 있는 과제 중점 발굴
- 중앙 부처·지자체·공공기관 대상 공모를 통한 총 5개 과제 선정
 - 국민권익위원회의 생성형 AI 기술을 활용한 소통·민원 분석 체계 구축 과제, 국세청의 시나리오 기반 상담 시스템을 생성형 AI 기반으로 고도화하기 위한 적용 과제가 포함
- **AI 에이전트 관련성:** AI 에이전트 기술 적용 사례의 구체적 실현
 - 산업통상자원부: AI 에이전트 기반 통합 인증상담 서비스, 멀티채널 연동 및 All-in-One 서비스 포털 제공
 - 국민건강보험공단: 에이전틱 AI 구축을 통한 업무별 전문화된 상담 지원 및 AI 에이전트 간 연계 프로토콜 설계

공공기관 실증 과제를 통한 AI 에이전트 기술의 초기 적용 확산

- 본 사업은 공공부문에서 AI 에이전트 기술을 실제 행정 서비스와 업무 프로세스에 접목하는 구체적 과제들이 등장했다는 점에서 주목
 - 산업통상자원부는 멀티 채널 커뮤니케이션을 통합하고 클라이언트를 위한 올인원 포털을 완성하기 위해 AI 에이전트 구현을 시도하고 있는데 서로 다른 정보 출처와 절차적 관리 사일로를 상호 연결된 단일 관리 인터페이스로 융합하여 계획-정보 검색-행동의 목표를 달성하고 행정 시의 차별화된 서비스를 실행하기 위한 시도
 - 국민건강보험공단의 사례는 다수의 에이전트가 협업하여 통합된 컨설팅 및 행정 서비스 세트를 점진적으로 제공할 수 있도록 지정된 작업에 맞게 구성된 전문 AI 에이전트를 위한 상호 에이전트 조정 프로토콜을 설계하는 것
 - 이러한 발전은 기술적 관점에서 중요하며 시는 단일 에이전트 환경에서 제공될 의도를 가지고 있으며 제안된 상호 연결된 전문 시스템은 지능형 서비스 제공을 자동화하고 향상시킬 가능성이 높음
 - 이는 정책/행정 영역에 중점을 둔 실질적인 AI 기반 행정 기술을 제공하고 정부 부처, 지방 정부 및 공공 기관의 조정된 노력은 공공부문에서 제안된 시스템을 대규모로 구현하기 위한 기초 역할
 - 지역 간에 분산되어 있던 운영 및 기술 체계를 통합하여, 공통의 테스트베드 환경을 조성하고 보다 엄격하고 정밀한 실증 검증이 가능하도록 함



2) 국내 공공부문 AI 에이전트 도입 시도 사례

● 국내 AI 에이전트 주요 활용 사례

● AI 에이전트 활용 현황

- 국내 공공부문에서도 AI 에이전트의 자율성, 목표 지향성, 환경 상호작용 특성을 활용한 다양한 서비스가 구현되고 있음
- 본 분석에서는 3가지 기준을 모두 충족하고 실제 운영 중인 7개 대표사례를 선정
- 앞서 1장에서 제시한 AI 에이전트의 3가지 핵심 기준에 따라 국내 주요 사례를 분석
 - 자율성(Autonomy): 사람의 직접적 개입 없이 스스로 판단하고 행동을 수행하는 능력
 - 목표 지향성(Goal-oriented): 특정 목표나 과업을 달성하기 위해 계획을 수립하고 실행하는 능력
 - 환경 상호작용(Environmental Interaction): 주변 환경을 인식하고 환경 변화에 따라 적응적으로 대응하는 능력
- 표 4는 국내 AI 에이전트 도입 시도 사례를 보여주며, 각 사례의 자율성·목표지향성·상호작용 보유 여부를 평가 하되, 대부분 일부 요소만 갖춘 초기 단계 시스템임을 부각 (예: 지능형 법령검색 시스템의 경우 사용자 질의에 반응하는 수준으로, 완전한 자율성은 갖추지 못한 사례임을 명시)

표 4 | 국내 공공부문 AI 에이전트 도입 시도 사례 분석

사례명	운영주체	주요 기능	자율성	목표 지향성	환경 상호작용	종합평가	근거
지능형 법령검색 시스템	법제처	일상용어로 법령 검색, AI 인지검색 기술 적용	○	○	○	에이전트 지향 시스템	법제처 공식 보도자료
KONI (한국형 과학기술 특화 LLM)	KISTI	과학기술정보 기반 질의 응답, RAG 기술 적용	○	○	○	에이전트 지향 시스템	KISTI 공식발표
ScienceON AI (SAI)	KISTI	과학기술정보 기반 대화형 검색	○	○	○	에이전트 지향 시스템	KISTI 공식발표
KISTEP-GEN (업무 AI 어시스턴트)	KISTEP	과학기술 현안 검색, 문서 요약, 보고서 작성	○	○	○	에이전트 지향 시스템	KISTEP 공식발표
AI 스마트 스크린 보행안전시스템	화성시	AI CCTV + 안전차단기, 실시간 모니터링	○	○	○	실제 에이전트	화성시 공식발표
360° AI+ 돌봄 서비스	경기도사회 서비스원	어르신 안부 통화→정서 지원· 위기 징후 감지→복지연계	○	○	○	실제 에이전트	언론보도
AI 노동법 상담 및 근로 감독관 AI 지원 시스템	고용노동부	법령 검색·요약·제안→ 근로감독관 업무 자동 수행	○	○	○	실제 에이전트	공식 발표자료

출처: 저자작성

● **지능형 법령검색 시스템 (법제처)**

- 620만 건의 방대한 법령정보 데이터와 연계한 생성형 AI 검색 모델을 통한 국민 친화적 법령서비스 구현
 - 일반 국민이 생활 속 용어로 질문하면 시가 의도를 분석하여 관련 법조문, 입법 취지, 판례, 해석례 등을 통합 제공
- 법률 분야 특화 기술 개발을 통한 서비스 고도화
 - 법률 분야 특화 sLLM 개발로 환각 현상 최소화 및 법령정보 한정 학습을 통한 답변 정확도·신뢰성 확보
 - 멀티모달 기술로 텍스트를 넘어 이미지, 음성, 영상 학습을 통한 그림, 도표, 웹툰 형태의 시각적 법령정보 제공
- 3단계 발전 계획을 통한 체계적 고도화 추진
 - 1단계('23): 지능형 법령 검색 기본 구축, 2단계('24): 질의답변 데이터셋 2만 건 구축, 3단계('25): 카테고리화 능동적 검색 결과 제공
- AI 에이전트 판별기준별 분석
 - 자율성: 사용자 검색 패턴을 자동 분석하여 맞춤형 법령 분야 추출 및 개정 법령 알림 서비스 제공
 - 목표 지향성: 입력된 일상용어를 카테고리화 하여 사용자가 스스로 법조문을 찾아가는 능동적인 검색 결과 제공 (ex. "강아지" → '동물' → '동물관리' → '등록' → '동물보호법 제15조 제1항' 과 같은 체계적 범주 분류)
 - 환경 상호작용: 법령정보의 변경, 사용자 검색어에 대한 No-Hit 및 신조어의 지속적인 추적·관리를 현행화
 - 자율성이 제한되어 사용자 질의에 반응하는 수준에 그치기 때문에 에이전트라기 보다는 에이전트를 지향하는 고도화된 검색 시스템에 가까움

● **KONI (한국형 과학기술 특화 LLM, KISTI)**

- Meta-Llama-3-8b-Instruct 기반 과학기술정보 특화 한국어 대규모 언어모델 개발
 - 2배 이상의 과학기술정보 데이터 학습을 통한 추론, 글쓰기, 이해 등 다양한 작업 성능 대폭 향상
- 과학기술정보 제공을 위해 신뢰성 확보 및 개발 환경 구축
 - 검색 증강 생성(RAG) 기술 연계를 통한 환각 현상 최소화 및 과학기술 관련 법령, 규정, 지침 대상 성능 검증 및 연구원 피드백 기반 추가 학습
 - LLMOps 플랫폼 구축으로 웹 기반 모델 제작·훈련 지원(PT, FT, DPO) 및 다양한 평가 툴·메트릭 지원을 통한 체계적 개발 환경 조성
- AI 에이전트 판별기준별 분석
 - 자율성: RAG 기술 연계를 질의에 적합한 정보를 자동 검색하고 과학기술 맥락에 맞는 답변 생성
 - 목표 지향성: 연구자의 과학적 발견 지원이라는 목표로 질의응답→정보분석→연구지원 과정 체계적 수행
 - 환경 상호작용: 과학기술 관련 법령, 규정, 지침을 대상으로 연구원 피드백을 통한 추가 학습 시스템
 - 정보 검색과 답변 생성으로 자율성이 높다고 보기 어렵기 때문에, 현재는 LLM 모델로 보이며 향후 에이전트를 지향하는 특성을 보임

● ScienceON AI (SAI, KISTI)

- KONI 기반 1억8천만 건 과학기술정보를 활용한 대화형 검색 및 논문 분석 서비스
 - 과학기술 지식인프라 ScienceON과 연계하여 논문 요약, 번역, 용어 설명 등 AI-Helper 서비스 제공
 - AI-Reviewer 서비스를 통한 다중 논문 비교분석 및 질의응답 형식 대화를 통한 추가 인사이트 도출 기능 제공
- 신뢰성 확보를 통해 논문 분석 서비스 고도화
 - 검색 증강 생성(RAG) 기법 활용을 통한 환각 현상 최소화 및 신뢰도 있는 분석 결과 제공
- AI 에이전트 판별기준별 분석
 - 자율성: 최대 10개 논문을 11개 관점에서 자동 비교분석하고 질의응답을 통한 추가 인사이트 도출
 - 목표 지향성: 11개 관점 논문 비교분석→AI 추천 질문 제시→추가 인사이트 도출→AI-Helper 연계 검토의 단단계 계획을 통한 연구 효율성 향상 추구
 - 환경 상호작용: 연구자 업로드 논문과 기존 DB를 통합 분석하는 다중 데이터 소스 통합
 - 주어진 데이터 내에서만 분석 및 요약이 가능하고 환경 변화에 반응이 어려워 자율성을 만족한다고 보기 어려움

● KISTEP-GEN (업무 AI 어시스턴트, KISTEP)

- 과학기술혁신 정책 분야 특화 SLM 'SPARK' 기반 업무지원 생성형 AI 서비스
 - 오픈소스 파운데이션 모델(Gemma 27B, Mistral 12B)에 정부 위원회 안건, 정책 연구자료, KISTEP 발간 보고서 등 추가 학습
 - 허깅페이스 공개를 통한 유관 부처·기관 무료 활용지원을 통해 기관 간 정보 및 시스템 연계 협력을 통한 공공기관 AI 전환 선도
- 정책업무 지원을 위한 3대 핵심 서비스
 - 검색 증강 생성(RAG) 기반 정확하고 신뢰성 있는 질의응답, 문서 요약 서비스, 보고서 목차 및 초안 생성 서비스
- AI 에이전트 판별기준별 분석
 - 자율성: 정책 현안 질의 시 관련 문서 자동 검색하여 요약·분석하고 보고서 목차 및 초안 자동 생성을 통한 업무 자동화
 - 목표 지향성: 보고서 유형(연구보고서, 안건, 기타 보고자료, 브리프) 선택→템플릿 적용→목차 생성→초안 작성으로 이어지는 체계적 문서작성 지원
 - 환경 상호작용: 운용 과정에서 수집한 피드백 데이터와 타 기관의 활용 경험을 바탕으로 SPARK를 지속적으로 고도화
 - 자동 검색, 분석이라는 자율성, 연구지원이라는 목표지향성은 만족하나 피드백이 약해 환경 상호작용이 부족하고 피드백 기반 시스템이 있으나 실시간성이 약해 환경변화나 동적 데이터 처리 반응이 부족하여 제한적

● AI 스마트 스쿨존 보행안전시스템 (화성시)

- 전국 최초 AI CCTV와 안전차단기 결합을 통한 실시간 모니터링 및 원격관리 시스템
 - 초등학교 앞 횡단보도에서 어린이들의 무단횡단과 돌발행동을 시가 자동 감지 및 차단
- AI 기반 실시간 안전 관리 시스템 운영
 - 지능형 상황 인식을 통한 잔여 보행자 감지 시 차단기 자동 개방 및 위험 상황 시 음성안내 등 자율적 현장 대응
 - 관제센터를 통한 실시간 현장 모니터링 및 중앙집중관리 기반 긴급상황 대응 체계 구축

• AI 에이전트 판별기준별 분석

- 자율성: 횡단보도 상황을 실시간 분석하여 잔여 보행자 감지 시 차단기 자동 개방, 위험 상황 시 음성안내 제공
- 목표 지향성: 어린이 교통사고 예방 목표로 상황 인식→위험 판단→안전조치→관제센터 알림으로 이어지는 단단계 안전 대응 체계
- 환경 상호작용: 교통정보센터에서 실시간 모니터링과 원격관리를 통한 중앙집중관리 시스템
- 제시된 사례들 중 가장 완전한 AI 에이전트로 평가되며 자율성(자동 분석 및 대응), 목표지향성(교통사고 예방), 환경 상호작용(실시간 모니터링 및 차단기 반응)까지 모두 만족

기술적 구현 측면의 심층 분석

- 화성시가 구축한 AI 스마트 스쿨존은 단순히 영상을 분석하는 수준의 시스템이 아니라 센서와 AI, 그리고 물리적으로 장치가 하나의 통합 구조로 작동하는 실시간 지능형 안전관리 체계
 - CCTV 영상, 차단기, 음성 안내 장치 등이 상호 연동되어 있으며, AI 모델이 보행자의 잔류 여부나 돌발 행동을 탐지
 - 필요한 경우 차단기의 작동이나 안내 방송 송출 등 물리적 제어를 자동으로 수행하고 관제센터로 상황이 즉시 전달
 - 이러한 구조는 “감지 → AI 분석 → 판단 → 실행”으로 이어지는 완결된 루프를 완성
 - 교통정보시스템(TMS)과 CCTV망과의 실시간 연계를 통해 주변 환경 변화에 즉각 반응할 수 있는 동적 상호작용성을 확보하고 있는데, 이는 실시간 자율적 의사결정 능력, 교통사고 예방이라는 명확한 목표 지향성, 환경과의 상호작용 측면에서 높은 수준의 AI 에이전트 특성을 구현하고 있음
- 화성시 AI 스마트 스쿨존은 공공영역에서 구현된 사례 중에서도 물리적 공간 제어와 AI 판단이 통합된 고도화된 ‘실시간 환경 연동형 AI 에이전트’로 평가할 수 있음

● 360° AI+ 돌봄 서비스 (경기도사회서비스원)

- 인간중심 AI 사업을 지향하는 7대 취약계층 돌봄 서비스 통합 제공
 - AI 노인말뚝, AI 어르신 든든지킴이, 불편한 AI 케어, 발달장애인 AI 돌봄 등 인간중심 AI 사업 추진
- AI 기반 다중 채널 모니터링 및 긴급 대응 체계
 - 네이버클라우드 '하이퍼클로바' 기반 주 1회 안부 전화로 건강, 식사, 수면 등 일상 모니터링 및 전화 3회 이상 미수신 시 담당자 직접 방문(AI 노인말뚝)
 - 긴급 메시지 음성 감지 시 112·노인보호전문기관 자동 호출(AI 어르신 든든지킴이)
 - 휴대폰 카메라 15초 터치로 혈류·심혈관 건강 상태 확인 및 AI 건강리포트 자동 작성(불편한 AI케어)
- AI 에이전트 판별기준별 분석
 - 자율성: AI 노인말뚝이 정기 안부 전화로 위기 징후 자동 감지, AI 스피커가 긴급 메시지 인식 시 112 자동 호출
 - 목표 지향성: 고독사 예방 및 학대 재발 방지 목표로 모니터링→위기 감지→즉시 대응→사후관리의 단단계 돌봄 체계

- 환경 상호작용: 24시간 동안 휴대폰이 작동하지 않으면 돌봄 매니저에게 SOS 긴급호출로 알림이 가는 실시간 모니터링 체계
- 자율성(자동 안부 전화 및 위기 감지), 목표 지향성(고독사 예방), 환경 상호작용(실시간 모니터링 및 긴급 호출)까지 모두 만족하는 완전한 AI 에이전트

기술적 구현 측면의 심층 분석

- 경기도의 360° AI+ 돌봄 서비스는 모바일 전화의 센서, 음성 인식, 대화 능력을 포함한 여러 입력 채널을 동시에 처리하는 고급 기술 구조를 갖추고 있으며, 실시간으로 취약 계층의 상황을 모니터링함
 - 본 시스템은 대화 중 비정상적인 침묵 패턴과 비상 단어의 발화를 감지하여 즉각적인 주의 조치를 위해서 돌봄매니저에게 긴급 전화를 자동으로 전송
 - 이는 실시간 환경 평가와 적응적 행동 실행을 중심으로 설계된 고급 AI 에이전트 시스템의 특성을 구현
 - 누적 데이터를 분석하고 종합하여 개인화된 건강 및 돌봄 보고서를 생성하는 능력은 목표 지향적 추론을 보여줌
- 전체 시스템은 "센서 → 인식기 → 평가자 → 실행자"로 이루어진 '지각-결정-행동 루프'를 통해 돌봄 서비스 시나리오를 실행하는 복합적 에이전트 구조를 구현
 - 경기도의 360° AI+ 돌봄 서비스의 이러한 기술 시스템은 단순한 작업 지원 및 원격 상담을 넘어서 실시간 적응형-상황인식형 공공 AI 에이전트 서비스로 발전할 수 있게 하는 대표적 사례

● AI 노동법 상담 및 근로감독관 AI 지원시스템 (고용노동부)

- 근로감독관 업무 지원과 국민 대상 24시간 개인화 답변 제공하는 이중 체계 구축
 - 근로감독관은 업무망 내에서, 국민은 웹사이트·SNS에서 신속하고 정확한 노동법 상담 서비스 이용
- 4대 핵심 서비스를 통한 종합적 노동법 지원
 - AI 노동법 지식 검색, 법령·판례·질의회시·매뉴얼 대화형 검색, 진정서 접수 지원, 문서와의 대화
- 한국공인노무사회와 협업을 통한 상담 정확도 향상
 - 공인노무사 173명 규모 'AI 노동법 상담 서비스 개선 지원단' 구성으로 AI 재학습 및 상담 품질 평가 참여
- AI 에이전트 판별기준별 분석
 - 자율성: 노동법 질의 시 관련 법령·판례·해석을 자동 검색·분석하여 개인화된 답변 생성 및 진정서 접수 지원
 - 목표 지향성: '사례법과 노동 기준 법규의 통합, 사례법 분석, 노동 행정 법률 상담 및 노동 감사 후 작업 흐름 제공'의 체계적인 상담 지원과 노동 감독자를 위한 업무 흐름의 간소화를 통해 노동 시장에서의 약자를 돕는 것이 주된 목표
 - 환경 상호작용: 변화하는 법률 조항과 실시간 동적 업데이트 정보 시스템을 통합하는 고용 감사 전문가와 협력하여 법적 지원을 요청하는 협력을 통해 이루어짐
 - 자율성(관련 법률 검색 및 분석), 목표 지향성(고용 감사원에 할당된 업무 지원), 환경 상호작용(실시간 통합을 위한 법률 개정)의 세 가지 지표가 충족

기술적 구현 측면의 심층 분석

- 본 시스템은 질문(자연어 입력) → 판단 프롬프트(의도 분석) → 관련 법령·판례·질의·매뉴얼 검색 → 법률 응답 생성 → 진정서 접수 및 행정 조치 연계의 다단계 구조로 설계되어 있음
 - 사용자의 질의는 의도에 대한 분석 과정을 거쳐 분류되며, 관련 법령·판례·행정 해석이 체계적으로 검색되고, 맥락에 부합하는 답변이 자동으로 생성되어 상담의 일관성을 유지하고 행정 효율성을 제고
 - 생성된 응답은 진정서 접수 및 행정 조치로 연계되어 질의→판단→행동으로 이어지는 행정 처리의 자동화 루프를 구현
 - 인증된 노동법 상담 전문가의 피드백과 법령 개정사항을 실시간 반영하여 동적 업데이트 루프를 지속 개선함으로써 규제 환경 변화에 대한 적응력과 상담 정확도를 향상
- 이러한 기술적 특성으로 인해서 본 시스템은 물리적 환경을 직접 제어하지 않지만 복합적인 검색-분석-판단-행동의 다단계 자동화 구조와 실시간 정보 갱신 능력을 갖춘 Operational AI Agent로 평가할 수 있음

- **7가지 사례 중 가장 고도화된 AI 에이전트는 화성시 AI 스마트 스쿨존, 경기도 돌봄 서비스로 실시간 의사결정이 가능할 뿐만 아니라 이를 통해 실제 물리적 업무를 수행**
 - 반면에 법체처나 KISTI, KISTEP, 고용노동부의 사례는 검색 결과를 화면에 보여주는 수준일 뿐 물리적 행동이 포함되어 있지 않음
 - 멀티모달 기능(텍스트·음성 등 복합 입력 처리)은 AI 에이전트 구현의 핵심 요소로, 현재 공공부문에서는 경기도 화성시와 고용노동부 사례에서만 확인
 - 두 기관은 음성·텍스트 기반 질의응답을 결합해 초기 단계의 에이전트형 행정서비스를 시도하고 있음
- **기술 구현 측면에서 화성시의 스쿨존 서비스와 경기도의 돌봄 서비스는 실시간 의사결정을 수행하고 물리적 행동을 실행하기 때문에 본 연구의 분석 대상 가운데 가장 높은 수준의 AI 에이전트 기술을 구현하고 있는 것으로 평가할 수 있음**
 - 고용노동부의 모델은 자동화된 복잡한 업무와 개인 서비스를 제공하기 때문에 중간 수준의 자동화 비즈니스 에이전트로 평가
 - 기타 사례(고용노동부의 지능형 법안 검색, KISTI KONI, KISTI ScienceON, KISTEP-GEN)는 지능형 검색과 기본 상호작용만 제공하는 에이전트 지향 시스템으로 구성되어 있음
- **국내 AI 시스템 중 화성시의 스마트 스쿨존과 경기도의 돌봄 서비스는 서비스 자동화를 넘어 실제로 변혁적인 서비스를 제공하는 AI 임베디드 에이전트 기술의 대표**
 - 반면, 대부분의 중앙 정부 부처와 연구 기관의 시스템은 RAG+LLM 기반의 지능형 검색 및 분석에 국한되어 있으며, 위에서 설명한 자율적인 환경 탐색, 실시간 계획 및 적응 제어를 사용하지 않음

공공부문 AI 에이전트 도입 및 활용 현황과 시사점

해외 주요국 AI 에이전트 도입 방향 및 사례



4. 해외 주요국 AI 에이전트 도입 방향 및 사례

1) 해외 주요국 공공부문 AI 에이전트 도입 방향

● 미국: AI 에이전트보다는 포괄적인 AI 개발·활용 전략

- **바이든 행정부에서는 AI의 안전, 보안, 책임성 문제를 중요하게 고려하는 행정명령을 발표하였음(행정명령 14110호)**
 - 바이든 행정부의 행정명령 14110호는 안전하고(safe) 보안이 보장되며(secure) 신뢰할 수 있는(trustworthy) 인공지능 개발과 활용을 지향함
 - 그래서 이 행정명령은 AI 시스템의 구조적인 결함과 취약성을 식별하는 것을 강조함
 - 이 조치는 AI 진흥 정책만큼 미국 정부와 기업이 AI를 활용하고 개발하는 데 책임을 강조하며 특히 위험성이 높은 AI 적용을 규제하려는 의도를 가짐(이해수 외, 2025)
 - 이와 같은 행정명령의 후속조치로 미 예산관리국(Office of Management and Budget: OMB)은 연방정부의 AI 활용 지침(M-24-10)을 발표함¹⁾
 - 이 지침은 미 연방정부의 AI 활용과 관련한 위험(예: 시민의 권리와 안전)을 관리할 수 있도록 각 연방기관이 CAIO를 지명하는 동시에 안전한 활용을 위한 전략과 실천방안을 수립할 것을 제시함
 - 정리하면, 바이든 행정부의 공공부문 AI 활용 전략은 기본적으로 공공부문의 AI 활용이 초래할 수 있는 위험을 관리하는 AI 거버넌스를 확립하는 데 있었음
- **그러나 트럼프 행정부(2기)에서는 미국의 AI 혁신 선도를 촉진하기 위해 기존 행정명령을 폐지하고 대체하는 행정명령을 발표함(행정명령 14179호)²⁾**
 - 트럼프 행정부(2기)의 행정명령 14179호는 AI의 안전, 보안, 책임성 문제에 초점을 맞추기보다는 이와 관련된 AI 규제를 축소하고 이를 통해 미국이 AI 산업을 선도국의 지위를 공고히 하는 것을 추구함

1) <https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf>

2) <https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/>

- **트럼프 행정부(2기)의 새로운 행정명령에 따라 미 예산관리국은 2025년 4월 3일에 미 연방정부의 효과적인 AI 활용을 촉진하는 두 건의 조치를 동시에 발표함³⁾**
 - 먼저 새롭게 발표한 ‘혁신, 거버넌스, 공공 신뢰를 통한 연방정부의 AI 활용 가속화’ 지침(M-25-21)은 바이든 행정부에서 발표한 기존 연방정부의 AI 활용 지침(M-24-10)을 대체하는 것임
 - 이 지침은 CAIO 지정 등 기존 지침에서 강조한 내용을 포함하는 동시에 AI 도입의 속도를 빠르게 하기 위해 연방정부 기관이 미 예산관리국의 템플릿을 활용해서 AI 활용 전략을 수립하는 한편 AI 관련 절차를 간소화하고 규제를 완화하는 것을 강조함
 - 또한 이 지침은 미국에서 개발되고 생산된 AI 활용을 강조하는 것이 특징적이며, 이는 최근에 논의되는 AI 소버린(AI sovereign)과 연관되어 있음
 - 다음으로 미 예산관리국은 ‘정부의 효과적인 AI 조달 활성화’ 지침(M-25-22)을 발표하여 연방정부가 AI를 개발하고 활용할 때 어떻게 조달할 것인지에 관한 세부적인 사항을 안내함
 - 이 지침은 조달의 생애주기(요구사항 식별 - 시장 조사 및 기획 - 공고 - 선정/계약 - 계약 관리 - 계약 종료) 전체에 걸쳐 연방기관이 따라야 하는 세부사항을 규정함
- **미 연방정부는 이어서 OpenAI와 협업하여 2025년 1월 28일에 미국의 연방·주·지방정부 공무원이 사용할 수 있는 ChatGPT Gov를 공개함(OpenAI, 2025)**
 - ChatGPT Gov는 정부 전용 버전으로 특히 공공부문의 보안 관련 규정을 자가 통제하기 때문에, 공무원이 보안에 대한 우려로부터 자유롭게 ChatGPT 서비스를 이용할 수 있는 이점을 가짐
 - ChatGPT Gov의 출시는 미국 공무원의 업무 수행 능력을 개선할 것으로 기대됨
- **그러나 ChatGPT Gov는 그 자체로 AI 에이전트가 아니라 공무원이 사용할 수 있는 대화형 생성형 AI 서비스이며, 미국 정부는 아직 AI 에이전트에 관한 구체적인 활용 계획이나 전략을 발표하지 않음**
 - 미국 정부는 트럼프 행정부(2기) 이후 공공부문에서 효과적인 AI 활용을 강조하기 위한 지침들을 발표하고 있으나, 구체적인 AI 에이전트 정의나 활용 전략은 제시하지 않고 있으며, 현재 사례로 소개된 DHSSChat 역시 생성형 AI 기반 업무 보조도구로, 독립적 자율성을 갖춘 AI 에이전트라기보다는 초기적인 활용 단계에 해당

3) <https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-21-Accelerating-Federal-Use-of-AI-through-Innovation-Governance-and-Public-Trust.pdf>
<https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-22-Driving-Efficient-Acquisition-of-Artificial-Intelligence-in-Government.pdf>

● 미국 AI 에이전트 도입 방향성 평가 및 시사점

- 미국의 사례에서 AI 에이전트에 대한 정의는 부족하지만, AI 기술의 위험 관리와 행정 효율성의 차원에서 정책 방향의 전환이 이루어지고 있는 정책 단계의 사례로 바이든 행정부는 AI의 안전성, 보안, 책임성을 강조하며 공공 분야에 AI 기술 활용에 대한 규제적 거버넌스의 기초를 마련
 - 이후 트럼프 행정부(2기)에서 AI 활용 촉진과 산업 경쟁력 강화를 주요 목표로 정책의 방향이 전환
 - 그러나 그 정책 전환의 맥락에서 AI 에이전트에 대한 구체적 정의와 활용 구상이 부족하며, 행정업무에 ChatGPT Gov와 같은 범용 생성형 AI 서비스를 도입하는 수준에 그침
- 이러한 미국의 태도는 AI 에이전트를 독립적 정책 의제로 선도적으로 다루기 보다는 광범위한 AI 기술 거버넌스와 행정 혁신의 하위 영역으로 다루고 있음을 뜻함
 - AI 정책의 미국적 사례는 종합적 일률적 로드맵 대신 기관별 자율전략 수립에 따른 템플릿, 지침, 틀을 통한 간접적 조율 방식 틀을 설계하는 분권적 기술 생태계와 중앙정부의 분권적 구조를 반영
 - 우리나라의 정책 설계는 미국 사례에서 AI 에이전트를 정의하지 않거나 전면에 내세우지 않더라도 AI 거버넌스 정비와 행정 디지털화를 통해 활용 기반을 점진적 조성을 시도할 수 있다는 점과 분권적 행정구조와 민간 기술 생태계를 통해 다양하고 폭 넓은 자율적 실험을 시도할 수 있다는 점에서 중요한 시사점을 제공

● 영국: 공공부문 AI 혁신의 핵심으로서 AI 에이전트 활용

- 영국 정부가 2025년 1월에 발표한 'AI 기회 행동 계획'은 영국의 AI 산업 육성과 공공·민간 활용에 관한 전략 계획의 성격을 가지며(Department for Science, Innovation and Technology, 2025), 정부가 AI의 주요 이용자인 동시에 민간 부문 채택을 촉진하는 역할을 해야 함을 강조함
 - 영국 정부의 'AI 기회 행동 계획'은 공공부문의 AI 도입 핵심 전략으로 '스캔(Scan)-파일럿(Pilot)-스케일(Scale)'을 제시함(Department for Science, Innovation and Technology, 2025)
 - '스캔' 단계는 정부가 AI 역량을 파악하고 어떤 미션을 위해 쓸지를 민간과 함께 매핑하는 단계임
 - '파일럿' 단계는 빠르게 활용 목적에 맞는 프로토타입을 만들고, 신속한 조달 절차를 통해 빠르게 시범사업을 하는 단계임
 - '스케일' 단계는 성공적인 공공서비스 제공을 한 시범사업을 파악하고 이를 대규모로 신속하게 확장하는 단계임
- 이처럼 'AI 기회 행동 계획'은 정부와 민간의 협업 속에서 공공부문 혁신을 위한 기본적인 공공부문에서 AI 채택·활용 전략을 구체화함

그림 21 영국 정부의 AI 행동 계획: AI 활용에서 공공부문의 역할



출처: Department for Science, Innovation and Technology(2025, p.16)

- **영국 정부는 이어서 2025년 2월에 공공부문에서의 AI 활용 안내서를 발표하였으며, 여기에서 AI 에이전트가 무엇인지 명시적으로 정의함**
 - 이 안내서에서 영국 정부는 AI 에이전트란 최소한의 인간 개입만으로 의사결정을 내리고 행동을 수행할 수 있는 자율 AI 시스템을 의미한다고 정의함
 - 이 안내서에 따르면, AI 에이전트는 자율적으로 환경과 상호작용할 수 있고 목표 달성을 위해 자신이 수행할 수 있는 도구와 기능을 식별하고 수행할 수 있음
 - 그러나 이 안내서는 AI 에이전트가 여러 잠재적 AI 활용 가능성의 예로 정의를 제시하는 데 그칠 뿐이며 AI 에이전트의 활용 전략이나 계획 구상을 포함한 것은 아님
- **2025년 8월 16일에 영국의 과학기술혁신부(Department for Science, Innovation and Technology)는 AI 에이전트가 공공서비스 제공에 시범 도입 후 2027년까지 전국적으로 확대 적용할 계획을 발표함**
 - 발표 내용에 따르면, AI 에이전트는 시민이 원하는 다양한 종류의 공공서비스 신청부터 처리까지 대신해 줄 수 있음
 - 예를 들어, 어떤 시민이 영국 내 한 지역에서 다른 지역으로 이사할 때 운전면허 주소 변경, 유권자 등록 변경, 새로운 주치의 등록 등의 파생되는 공공서비스를 AI 에이전트가 처리함
 - 영국의 과학기술혁신부는 민간 AI 기업과 전문가와 함께 6-12개월 동안 이와 같은 AI 에이전트의 프로토타입을 개발할 예정임
 - 이를 통해 공공서비스에 적용 가능한 AI 에이전트의 가능성과 한계를 실증적으로 탐색하고자 하며 특히, 초기 단계에서는 사용자 리서치, 데이터 가용성 조사, 잠재 서비스 영역 식별과 같은 탐색적 활동이 중심을 이루며, 이후에는 소규모 반복적 개발 및 평가를 통해 기능을 확장하는 방식이 구상
 - 이 계획은 험프리 톨세트의 지원 역할을 확대하여 에이전트 수준에서 자율적으로 기능하고 환경과 상호 작용하여 이중 협력을 보여주는 것을 목적으로 함
 - AI 정부 에이전트의 점진적인 채택은 기술적 위험 완화에 대한 보다 전략적인 접근 방식을 나타냄
 - 따라서 AI 정부 에이전트 전략적 접근 방식은 'AI 행동 계획과 함께하는 기회' 내에서 점진적인 '탐색-파일럿-확대' 방법과 일치

● 영국 AI 에이전트 도입 방향성 평가 및 시사점

- 미국은 연방정부가 전략을 제시하고 부처·지방정부·민간이 주도하는 분권적 실험 중심으로 AI 에이전트를 도입하는 반면, 영국은 중앙정부가 거버넌스 틀을 마련해 지방정부·민간과 협력하며 체계적으로 추진
 - 점진적인 전반적 통제를 허용하는 영국의 '스캔-파일럿-확대' 전략은 기술적 위험이 국가 시행 전에 실증적으로 평가되고 검증되도록 설계
 - 점진적인 공공 정책 접근 방식은 점진적인 정책 실험으로 설계되었으며 채택을 촉진할 가능성이 높음
 - 이는 영국 정부의 중앙 집중적인 생태계 조정, 민간 기업 및 전문가와의 협력, 그리고 시민이 직접 경험하는 서비스 내에서 영향력 있는 사용 사례의 명확한 설명을 통해 분명히 나타냄
 - 이는 영국 정부가 서비스 AI에 대한 시민의 수용을 증진하려는 의도를 나타냄
 - 중앙집중식 행정 시스템과 강력한 디지털 인프라를 갖춘 한국은 영국의 접근 방식과 유사한 정책의 혜택을 받을 가능성이 높음
 - 중앙 정부는 지정된 행정 서비스 파일럿 지역에서 명확한 개념 정의와 실행 전략을 가지고 효과적인 공공 서비스 제공을 강화할 것

2) 해외 주요국 AI 에이전트 도입 시도 사례

● 미국 정부의 AI 에이전트 도입 및 활용 가능성

- 미국의 국토안보부(Department of Homeland Security)는 2024년 12월 17일에 DHSChat이라는 AI 도구를 도입하여 부처 소속 공무원에게 제공한다고 발표함⁴⁾
 - 기존에는 공무원들이 상업용으로 출시된 생성형 AI 도구인 ChatGPT 등을 주로 사용했었으나, DHSChat은 안전한 환경에서 공무원들이 내부에 보안성 있는 민감정보를 활용할 수 있도록 도움
 - 미국의 국토안보부가 보도자료 발표한 내용에 따르면, DHSChat은 아래의 <표 5>와 같은 방식으로 활용되고 있음
 - 그러나 DHSChat은 현재 AI 챗봇으로 공무원들이 텍스트로 입력하는 것에 대해 자료를 탐색하고 그 결과물을 생성하는 AI 업무 보조도구에 가까운 것으로 보임
 - 다만, 향후에 DHSChat이 국토안보부의 문서관리·이메일·정보시스템과 연결되고 그에 기반하여 다양한 과업을 연결하고 목표 달성을 위해 자율적으로 실행될 수 있다면 AI 에이전트의 성격을 가지게 될 수도 있음

4) https://www.dhs.gov/archive/news/2024/12/17/dhss-responsible-use-generative-ai-tools?utm_source=chatgpt.com

표 5 | 미국 국토안보부의 DHSChat의 활용 사례

활용 사례	내용
보조금 프로그램 의사소통 개선	국토안보부 팀들은 이메일과 서면 의사소통을 간명하게 하도록 DHSChat을 활용함
사이버보안 용어 해설	사이버보안 전문가들은 사이버보안과 관계없는 공무원들에게 복잡한 사이버보안 관련 용어를 쉽게 요약해서 해설하기 위해 DHSChat을 사용함
교육훈련 프로그램 설계	국토안보부 공무원들은 변호사들을 대상으로 AI 법과 정책 관련 교육훈련을 위한 초안을 DHSChat을 사용하여 만들
학습자료 제작	국토안보부 공무원들은 미국 표준기술연구소(National Institute of Standards & Technology: NIST)의 위험관리 프레임워크를 학습하기 위한 시스템을 만들기 위해 DHSChat을 사용함

출처: 미국 국토안보부 홈페이지에 소개된 보도 내용에 기반하여 저자작성
 (https://www.dhs.gov/archive/news/2024/12/17/dhss-responsible-use-generative-ai-tools)

- 한편, 미국 연방총무청(General Services Administration)은 2025년 8월 14일에 미국의 모든 연방기관들이 사용할 수 있는 AI 평가 도구인 USAi를 출범하였음⁵⁾
 - USAi는 미국 연방기관들이 여러 AI 도구 개발과 사용을 검토하고, 기능과 효과, 장단점 등을 파악할 수 있도록 돕는 점에서 미국 연방기관들의 효과적인 AI 활용을 촉진하는 인프라 역할을 할 것으로 기대됨
- 따라서 향후 USAi를 통해서 미국 연방기관들이 적극적으로 AI 에이전트를 개발하고 시범적용하고, 활용하는 것을 검토할 가능성이 커질 것으로 기대됨
- DHSChat과 USAi의 등장은 미국 정부가 아직 AI 에이전트를 본격적으로 도입하지는 않았지만, AI 에이전트 생태계를 위한 기술적·제도적 기반을 점진적으로 마련하고 있음을 시사
 - DHSChat은 현재 단순한 챗봇 형태이지만, 향후 문서·정보시스템과의 연계를 통해 업무 자동화 및 자율적 과업 수행 기능을 확장할 경우 AI 에이전트로 진화할 잠재력을 지님
 - 동시에 USAi는 연방기관들이 다양한 AI 기술을 평가·선택·적용할 수 있는 공통의 테스트베드 및 조정 플랫폼 역할을 수행함으로써, 기관 간 이질적인 AI 도입을 조율하고 AI 에이전트의 시범사업 및 확산을 촉진하는 인프라로 기능할 수 있음
 - 결국 이러한 요소들은 미국이 명시적인 AI 에이전트 전략 없이도 '기술적 토대-도구 실험-적용 확산'의 경로를 통해 간접적으로 AI 에이전트 활용 기반을 형성하고 있음을 보여줌

5) https://www.gsa.gov/about-us/newsroom/news-releases/gsa-launches-usai-to-advance-white-house-americas-ai-action-plan-08142025?utm_source=chatgpt.com

● 영국 정부의 AI 에이전트 도입 및 활용 가능성: Humphrey

- 영국 정부는 2025년 1월 21일에 공공부문에서 공공서비스를 효과적으로 제공하기 위해 AI 도구 묶음인 **Humphrey**를 도입한다고 발표함
 - 아래의 <표 6>에 정리된 것처럼, Humphrey는 공공부문에서 일하는 정책결정자와 공무원의 업무 수행을 돕는 여러 AI 도구들을 포함하는 패키지임
 - 영국 정부는 ‘스캔-파일럿-스케일’ 전략에 따라서 Humphrey 시범 사업을 진행하고 있음
 - 예를 들어, Humphrey에 포함된 AI 도구 중 하나인 Consult는 스코틀랜드 정부에 의해 실제 공공의견수렴 과정에서 시범 적용되었으며, 비수술 미용 시술 규제와 관련한 2,000개가 넘는 응답을 분석해 핵심 테마를 정리했음
- 그러나 Humphrey에 포함된 AI 도구들은 대부분 AI 에이전트보다는 AI 업무 보조 도구의 성격을 가짐
 - 예를 들어, 앞서 살펴본 Humphrey에 포함된 Consult는 목표에 따라 자율성을 가지고 판단하고 작동하는 AI 에이전트라기보다는 공무원의 업무 수행 일부를 보조하는 AI 분석 도구라고 할 수 있음
 - 마찬가지로 Humphrey에 포함된 Minute는 영국의 25개 지방의회에서 시범적으로 활용되고 있으며 공무원의 회의록 작성을 대신하는 보조도구 역할을 할 것으로 기대됨

표 6 | 영국 정부의 Humphrey에 포함된 AI 도구

구분	개요	내용
Consult	공공의견수렴 분석 도구	공공의견수렴(government consultation)에서 온라인으로 접수되는 공중 응답을 분석하여 그 결과를 정책결정자와 공무원에게 제공함
Parlex	정책결정자의 의사결정 보조도구	정책결정자가 수십 년 간의 의회 토론을 탐색하고 분석하여 법안을 검토하고 관리하는 것을 도움
Minute	회의록 정리 도구	회의 내용을 보안에 유의하여 녹취하는 것을 도우며, 공무원이 요구하는 형태로 맞춤형의 요약물 생산함
Redbox	공무원의 일상적인 과업 보조도구	관련 정책을 요약정리하거나 브리핑 자료를 준비하는 것과 같이 공무원이 일상적인 과업 수행을 도움
Lex	법률 자문 도구	공무원이 일상 과업을 수행하는 데 있어 연관된 법률과 관련된 쟁점들을 분석하고 정리함

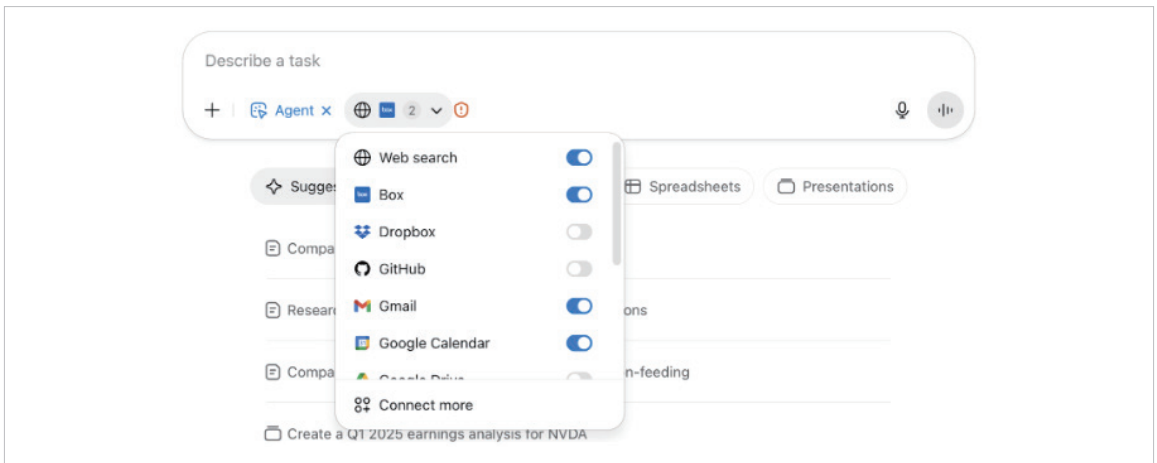
출처: Gov UK 홈페이지에 소개된 보도 내용에 기반하여 저자작성

(<https://www.gov.uk/government/news/shake-up-of-tech-and-ai-usage-across-nhs-and-other-public-services-to-deliver-plan-for-change>)

● 반면에, Humphrey에 포함된 일부 AI 도구는 단순한 AI 업무 보조도구를 넘어서는 AI 에이전트의 성격을 지님

- Humphrey라고 불리는 AI 도구 묶음은 아직 개발 완료된 것이 아니기 때문에, 이것들이 AI 업무 보조도구에 그칠 것인지 아니면 일부는 AI 에이전트로 발전할 것인지 명확하게 얘기하기는 어려움
- 그러나 앞서 살펴본 것처럼 영국 과학기술혁신부의 AI 에이전트 활용 계획을 고려할 때 일부 AI 도구들은 AI 에이전트로 발전할 가능성이 큼
 - 예를 들어, Redbox는 영국 정부의 보도자료에 따르면 공무원이 업무 수행시 관심 있는 정책을 요약하거나 브리핑 준비를 돕는 도구로 소개되었으며, 이것만 고려할 때는 AI 에이전트보다는 생성형 AI 기반 업무 보조도구에 가까움
 - 다만 Redbox는 이후에 공무원의 일정표, 문서관리시스템, 정보시스템 등 외부환경과 연동되어 공무원의 업무 일정별로 정책 요약, 브리핑 준비, 업무 회의 예약 등을 자율적으로 수행하는 형태로 진화할 가능성이 있으며, 이 경우 AI 에이전트의 성격을 가진다고 할 수 있음
- 이와 관련된 구체적인 가능성은 OpenAI에서 개발한 ChatGPT의 일반 채팅 모드와 에이전트 모드(agent mode)의 비교를 통해 쉽게 이해할 수 있음(그림 3)
 - ChatGPT의 일반 채팅 모드는 이용자가 텍스트를 입력할 때 그에 대해 산출된 답변을 제시하는 방식이라면, 에이전트 모드는 이용자가 설정한 목표를 달성하기 위해 여러 웹사이트를 탐색하고, 업로드된 파일을 분석하고, 외부 데이터 원천(이메일, 문서저장소 등)에 접근하고, 이용자가 원하는 형식과 파일로 분석된 내용을 정리하는 업무를 자율적으로 수행할 수 있음
 - 이와 같은 ChatGPT의 에이전트 모드는 공공부문에서 AI 활용이 단순하게 생성형 AI 기반 업무 보조도구 역할을 하는 것을 넘어서서 일정한 업무 과정을 외부환경과 상호작용하며 자율적으로 수행할 수 있는 가능성을 제시한 점에서 의미 있음

그림 3 | ChatGPT의 에이전트 모드



출처: <https://help.openai.com/en/articles/11752874-chatgpt-agent>

● **Humphrey 사례는 현재 공공부문에서의 AI 활용이 업무 보조 단계에서 AI 에이전트 단계로 이행하는 과도기적 양상임을 잘 보여줌**

- Consult나 Minute처럼 단일 기능을 수행하는 도구들은 정책 분석·행정 실무 지원에 집중하지만, Redbox와 같이 일정 관리·정보 접근·문서 분석 등 외부 환경과의 연속적 상호작용을 전제로 설계된 도구들은 향후 자율적 목표 수행 능력을 갖춘 AI 에이전트로 발전할 잠재력이 큼
- 특히, 영국 정부의 '스캔-파일럿-스케일' 전략에 따른 이러한 도구의 점진적인 실행 및 확산은 AI 에이전트의 완전한 구현에 앞서 공공 부문이 다양한 기능 AI 도구를 점진적으로 에이전트화하는 전략적 접근 방식을 보여줌
 - 이는 공공 부문 내에서 AI 에이전트 구현의 초기 단계와 공공 부문 내 AI 에이전트 생태계 개발을 의미



공공부문 AI 에이전트 도입 및 활용 현황과 시사점

결론



5. 결론

1) AI 에이전트 도입을 위한 정책적 전제조건

- AI 에이전트의 도입은 단순한 새로운 기술의 구현을 넘어서며, 법적 및 제도적 프레임워크, 데이터 및 디지털 인프라, 윤리, 인간 역량, 조직 문화 등 전체 공공 행정 시스템의 재구성을 포함
- AI 에이전트의 도입은 단순한 새로운 기술의 구현을 넘어서며, 법적 및 제도적 프레임워크, 데이터 및 디지털 인프라, 윤리, 인간 역량, 조직 문화 등 전체 공공 행정 시스템의 재구성을 포함
- 법·제도 및 책임성 정비
 - 공개적으로 책임을 갖는 AI 훈련 데이터의 관리 및 공공 데이터 세트의 보관, 품질, 통제 및 보안을 정의할 수 있는 법적 프레임워크를 구축하는 것이 필요
 - AI 에이전트의 자율성이 증가함에 따라 불투명한 의사 결정 및 편향의 위험이 커지므로 인간 중심 설계가 필수적
 - 데이터 라벨링 → 모델 학습 → 모델 평가의 3중 인간 개입 구조를 통해 위험성을 통제해야 함(홍승헌·황하, 2024)
 - 공공부문에서 AI 에이전트가 도입되기 위해서는 위험성 대응을 위한 인간의 3중 개입이 작동할 필요가 있음(홍승헌·황하, 2024). 첫 번째는 데이터 수집 시 라벨링 과정에서의 개입, 두 번째는 모델 학습 과정에서의 개입, 세 번째는 모델 평가 단계에서의 개입. 종국적으로는 AI 에이전트 기술의 고도화에 따라 인간 개입이 필요 없을 정도의 자율성에 대한 신뢰가 확보되는 것이 이상적인 청사진이지만, 이는 공공 영역의 특성을 세심히 고려하여 설계될 필요
 - 행정결정에서 인간의 최종 검토를 유지하고, 결정 과정과 근거 데이터를 기록·공개하는 책임성·투명성 확보 장치를 법령으로 마련해야 함
 - 개인정보보호법·행정절차법 등 기존 법률 역시 AI 시대에 맞춰 개정이 필요
 - 법·제도적 프레임워크 측면에서 AI 에이전트가 행정의사결정 및 업무수행 과정에서 일정 부분 역할을 담당할 경우 수반 되는 책임과 검증 그리고 규제 준비가 사전적으로 마련되어야 할 것

● 데이터 거버넌스와 디지털 인프라

- 고품질·표준화된 공공데이터와 안전한 인프라가 AI 에이전트의 기반
 - 영국은 국가 데이터 인프라 확충 및 공무원 AI 리터러시 교육을 강조하며, 미국은 NIST 프레임워크로 데이터 투명성을 관리
- 핵심 원칙은 데이터 스튜어드십, 계보 추적, 민간데이터 분류, 최소화 원칙 등(Mohan, 2025)
 - 데이터 거버넌스 프레임워크는 다음과 같은 핵심 원칙을 포함(Mohan, 2025). 데이터 스튜어드십(데이터 작업자 모두가 보안과 정확성에 책임을 지는 명확한 조직적 책임), 데이터 계보(AI 파이프라인 전반의 데이터 흐름과 변환 추적으로 투명성과 책임성 보장), 데이터 분류(민감한 데이터를 훈련 파이프라인에 입력하기 전 플래그하는 메타데이터 라벨링, 데이터 최소화 (AI 워크플로우에 특별히 설계된 접근 권한과 관행 배치))
- 안정적 클라우드, 엣지 컴퓨팅, 기관 간 데이터 연계, API 아키텍처 등 상호운용성 확보와 사이버보안 강화가 필수적(Stone, 2024; Krishnamurthy, 2025)
 - AI 에이전트가 안전하게 작동하려면 다음과 같은 인프라 사항이 요구(Stone, 2024; Krishnamurthy, 2025). AI 지원 인프라 (AI를 안전하고 접근 가능하며 책임감 있게 실행할 수 있는 컴퓨팅 성능과 복원력을 가진 시스템), 데이터 센터(적정 전력, 냉각 및 연결성을 확보한 안전하며 확장 가능한 시설), 엣지 컴퓨팅(실시간적 AI 응용을 위한 분산 컴퓨팅 능력), 통합 클라우드(보안, 비용, 확장성 간의 밸런스를 맞춘 하이브리드 통합 클라우드 전략 포함)
 - 데이터 및 디지털 인프라의 측면에서 AI 에이전트가 작동할 수 있는 기초가 되는 고품질의 행정데이터 및 시스템 간 상호 운용성이 갖추어진 디지털 인프라가 마련되어야만 에이전트가 자율적으로 정보를 탐색·분석·활용 할 수 있는 가능성이 있음

● 윤리 기준 및 책임성 확보

- UNESCO(2025), OECD, G7 등 국제 프레임워크는 인권, 공정성, 투명성, 설명가능성, 책임성을 핵심 원칙으로 제시
- 윤리영향평가(Ethics Impact Assessment) 절차와 AI 윤리위원회 설치를 통해 알고리즘 편향·차별을 모니터링하고, 오류 발생 시 시정 조치를 제도화
- 미국의 'AI 권리장전', 영국의 EU AI Act 기반 규제 샌드박스 등도 참고
 - 윤리 및 신뢰의 측면은 공공 행정에서 특히 중요한 구성 요소인 AI 에이전트가 행정서비스에 활용되는 경우 투명성·공정성·설명 가능성이 충분히 담보된 상태에서만 가능

● 공무원 교육과 인적 역량

- 공직사회 전반의 AI 수용성을 높이기 위해서는 AI 리터러시 정례 교육을 실시하고, 데이터 과학자·AI 엔지니어·윤리전문가 등 전문 인력을 확보하며, 민간 인력과의 교류를 확대할 필요
- 현재 부족한 AI 전문성을 보완하기 위해 민간과의 인사 교류 및 전문인력 채용을 적극 추진하고, 재직 공무원을 대상으로 AI 활용 방법, 데이터 해석 능력, 윤리 준칙 등을 포함한 체계적인 교육 프로그램을 운영
 - 이러한 노력을 통해 공직사회 전반의 AI 수용성을 제고하고, AI 에이전트 도입에 따른 업무 절차 변화를 효과적으로 관리 가능
 - 나아가 공무원의 기술적 활용 능력뿐 아니라 윤리적 판단 역량도 강화할 수 있을 것
 - 인간 역량의 강화도 중요한 조건인데, 공무원이 AI 기술을 이해하고 에이전트의 의사결정 과정을 점검·보완할 수 있는 역량이 없다면 행정 책임성과 통제가 약화될 것

● **조직문화와 거버넌스**

- AI 에이전트가 조직에 원활하게 통합되기 위해서는 조직이 협력적인 문화와 투명한 거버넌스 시스템을 가져야 함
- 이러한 협력적 거버넌스는 조직 내 AI 에이전트의 사용이 원하는 목표를 달성하고 기존 문제를 해결하도록 보장
 - 조직 내 거버넌스 효과화 및 AI 구현을 위해서는 협력적 인프라, 법률 및 윤리적 프레임워크가 필요
 - 마지막으로 공공 부문 내 조직 문화의 혁신으로 기술 통합을 제한하는 수직적이고 관료적인 문화에서 벗어나 부서 간 협업과 실험적 혁신의 문화로 나아가는 것을 포함

2) 국가 맥락에 따른 AI 에이전트 전략의 차별화

● AI 에이전트 도입 전략은 각국의 고유한 기술 역량, 제도적 프레임워크에 더해서 사회문화적인 패러다임에 기반하고 있음

- 영국과 미국 모두 첨단 기술과 함께 성숙한 행정 구조를 갖추고 있지만, 정책 실행에 대한 접근 방식은 차이

● **영국: 중앙집중형 조정과 친혁신적 규제의 병행**

- 영국은 중앙정부가 강력한 조정력을 바탕으로, 「AI 기회 행동 계획」을 통해 AI 도입을 국가 혁신 전략의 핵심에 배치하였으며(DSIT, 2024), 명확한 방향성을 제시하여 부처 간 전략을 통합하는 동시에 규제 샌드박스과 부문별 원칙 적용을 통해 혁신과 규제의 균형을 추구한다는 점이 특징
- 대표적 사례로는 NHS AI 진단기금(2,100만 파운드) 지원을 통한 보건의료 분야 AI 활용 촉진, 행정기관의 알고리즘 운용 투명성을 제고하기 위한 ATRS(Algorithmic Transparency Recording Standard) 도입, AI 안전 연구소 설립 등(GDS, 2025)
 - 이는 중앙정부가 명확한 국가 전략 목표를 제시하고 이를 제도와 예산자원으로 뒷받침하는 전형적인 중앙집중형 모델이라 할 수 있음

● 미국: 분권적 실험과 가이드라인 중심의 유연한 접근

- 미국은 포괄적 AI 법 제정보다는, 연방정부의 가이드라인과 부처별 자율성에 기반한 분권적 전략을 택하고 있으며, AI 권리장전과 NIST 위험관리 프레임워크를 중심으로 각 부처가 자율적으로 AI 전략을 수립·실행하며, 민간 부문과 학계의 참여를 통한 바텀업(bottom-up) 혁신 생태계를 조성하는 방식
 - 이러한 유연성 덕분에 AI 활용 사례와 예산 규모는 빠르게 증가했으며, 2024년 기준 연방기관의 AI 사용 사례는 2,133건으로 전년 대비 3배 증가했고, AI 관련 연방계약액도 2022~2023년 사이 1,200% 이상 급증 (Whitehouse, 2025; Jackson, 2024)
 - 그러나 부처별 추진 수준의 편차가 커 국가 차원의 통합적 전략 제시와 조정에는 한계가 있다는 점도 지적

● 국가별 전략 비교와 시사점

- 영국은 중앙정부가 중앙 차원에서 AI 전략 개발을 통할하고 있으며 혁신적인 모의실험 프로젝트를 규제 샌드박스 프로그램을 통해서 적극적으로 지원 중
 - NHS AI 보조 진단 프로그램과 ATRS, AI 안전 연구 기관 등 다양한 이해 관계자들의 참여로 정책 효과성이 극대화 되는 점이 특히 주목할 점
- 반면, 미국은 분산된 접근 방식으로 AI 권리 장전과 NIST 프레임워크에 기반하고, 통합된 법률 이 아니라 각 부서와 기관이 독립적으로 모의실험 프로젝트를 설계하고 시행
 - 미국은 민간의 연구 대학 및 기타 여러 기관들이 주도한다는 개방형 협력 접근 방식을 취하고 있다는 점이 특징
- 이들 두 나라의 사례를 참고하고 한국의 특수한 상황을 고려할 때, 중앙 정부가 전체적인 방향성과 기준 및 원칙을 정하고, 지자체와 개별 공공 기관들이 다양한 실증 프로젝트를 수행하여 현장에서의 성공과 실패 경험을 축적하여 체계적으로 분석하는 혼합 거버넌스 모델을 사용하는 것이 적절

● 한국에의 시사점: 중앙집권적 조정 + 현장 실험의 혼합 모델

- 한국은 중앙집권적 행정체계와 우수한 ICT 인프라를 모두 보유하고 있어, 영국과 미국의 전략을 절충한 혼합형 접근이 적합
 - 즉, 중앙정부가 표준과 법제를 정비하고 명확한 방향성을 제시하면서도, 지자체·일선기관에서 다양한 시범사업과 실증 프로젝트를 병행하여 현장 기반의 경험을 축적하고 이를 점진적으로 전국적으로 확산하는 방식
 - 이러한 혼합 모델은 한국의 신속한 정책 추진력과 높은 기술 수용성을 결합하여, AI 에이전트의 행정 적용을 효과적이고 신뢰성 있게 확산시킬 수 있는 전략적 선택이 됨
 - 한국의 경우 중앙집권적인 정부행정의 체계적 특성과 신속 추진을 특성으로 하는 정책 문화를 전략 설계 시에 반영 필요
 - 혼합적으로 영국처럼 중앙정부가 명확한 방향을 제시하고 이끌면서도, 미국처럼 현장 중심의 실험을 병행하는 전략이 바람직

- 한국 정부는 AI 에이전트에 관한 법적 프레임워크와 기준을 지속적으로 준비하고, 공공 데이터셋 구축을 위한 인프라를 더욱 개선해야 함
 - 또한, 국가 및 지방 현장 기관, 지방 정부를 위해 성공 및 실패 사례, 기회 및 위협 요인을 추적하고 분석할 수 있는 프레임워크 모델링 프로젝트를 개발해야 함
 - 중앙정부는 이를 조정하고 지원할 수 있으며, 검증된 모델을 전국적으로 확장 및 보급하기 위한 단계적 확장 전략을 취할 수 있음
- 영국 및 미국 모델을 활용함으로써, 한국은 AI 에이전트의 단계적 지역 배치를 통해 신뢰할 수 있고, 효율적이며, 효과적인 AI 에이전트 시스템을 구현할 수 있을 것
 - 한국의 행정 현실에 적합한 AI 에이전트 정책을 추진하기 위해 중앙 정부는 예산 배분 및 체계적 규제 프레임워크 조정의 임명된 후원자 역할을 수행해야 하며, 중앙 부처와 지방 정부는 혁신적인 실험 및 행정 지원을 수행해야 함



3) 정책 방향 및 단계별 로드맵

● 공공부문에 AI 에이전트를 도입하는 단계적 접근 전략 필요: 법·제도 정비 → 실증 확산 → 제도화 및 고도화

● 로드맵 수립 시 고위험 분야일수록 인간 개입 원칙과 설명가능성 확보 우선

- 중앙정부의 예산 배분 및 법제 정비·조정 역할, 부처 및 지자체의 실증 프로젝트 창의적 혁신 추진 역할 분담형 거버넌스 체계 필요
- 실행은 점진적 접근이 바람직하고, 이를 1단계(시범 추진); 2단계(모의 실험 확산, 제도화 및 표준화, 모니터링, 평가 체계 구축); 3단계(중간 위험도 확산, 조직 변화 관리, 공무원 재교육, 성과 확산); 4단계(높은 위험도 확산, AI 전문센터 설립(CoE), 국제 표준화 및 협력)로 추진
 - 우선 1단계(약 6개월)는 거버넌스 조직을 새로 신설하고, 법·제도적인 로드맵을 마련하며, 긴급성이 큰 분야를 중심으로 시범 사업을 착수하는 시기로 동시에 기본 교육과 기초 인프라 확충이 병행되어야 함. 이어서 2단계(7~18개월)에서는 모의실험을 폭넓게 확산하고, 표준화와 제도화 작업을 본격화하며, 모니터링 및 평가 체계를 구축. 3단계(19~36개월)에서는 위험도가 중간 수준인 분야로 활용 범위를 넓히고, 조직 변화 관리와 공무원 재교육을 통해 성과 확산을 뒷받침. 4단계 (37~48개월)에서는 고위험 분야까지 확대하면서 AI 전문센터(CoE)를 설립하고 전문가 조직을 운영하며 국제협력과 표준화 논의에 주도적으로 참여하는 단계가 될 것

● AI 에이전트 도입과 운영 원칙

- 행정 분야에서는 AI 에이전트를 효과적으로 사용하기 위해서는 개략적인 운영 원칙이 필요하며, 초기 배치 단계에서는 자가 의사 결정, 목표 지향적 행동 및 환경에 대한 적응력이 필요하며, 이는 필수적인 출발점
- 기술이 성숙해지고 생태계가 더 안정화되면, 지속적인 학습과 적응의 필요성을 포함하여 보다 광범위한 용어로 초점 영역을 다룰 수 있게 될 것
 - 원칙적으로, AI가 생성한 분석 및 정책 추천의 독단적인 수용에 대한 보호 장치를 갖춘 인간 중심의 의사 결정 모델이 현재로서는 반드시 필요하며, 공공의 이익과 제안의 합법성이 최종 검토의 주제가 되어야 함
 - 동시에 AI 사용에 대한 명확한 거버넌스 프레임워크가 필요하여 사회 중심의 감독 및 통제를 보장하고, 알고리즘의 운영 방식과 AI 사용이 충분히 투명해야 함

● 향후 과제 및 결론

- AI 에이전트는 기존의 챗봇이나 단순 자동화 시스템과 달리 자율적 목표 설정, 환경 상호작용, 학습 능력을 갖추고 있어 행정 효율성 제고, 정책 분석 고도화, 민원 서비스 품질 향상 등에서 혁신적 잠재력을 지님
 - 동시에 개인정보 보호, 알고리즘 편향, 책임소재 불명확 등 복합적인 법적 윤리 문제가 수반되므로 이를 체계적으로 대응할 전략 마련이 필수적
 - 향후 중점과제로 i) 데이터 관리 플랫폼의 자동화 품질·보안·API 통합 아키텍처 정비, ii) 국내외 전문가 유치, 교육 체계화, 인재 보상 강화를 통한 역량 기반 강화, iii) 실시간 평가 및 모니터링 시스템 구축과 준수 프레임워크 정교화, iv) AI 우수센터 설립 및 국제 협력 확대를 통한 지속가능한 거버넌스 체계 구축을 설정하여 추진
 - 기반이 갖춰진다면 AI 에이전트는 시민 중심의 행정 서비스 혁신, 공공조직의 생산성 제고, 중앙정부 정책결정 과정의 과학화를 촉진하는 공공부문 혁신의 전환점이 될 것
- 나아가 AI 에이전트의 도입은 단순한 기술 실험을 넘어 디지털 행정의 민주성 제고와 시민 신뢰 회복에도 기여하는 중대한 전환점이 될 것



참고문헌(미주)

- 경기도. (2024, May 30). 보도자료: 노인돌봄 정책에 인공지능 적극 도입. '경기노인 AI+돌봄 계획' 발표
- 경기도사회서비스원. (2025, June 17). 보도자료: 경기도사회서비스원, “아리아, 살려줘” 외침으로 AI가 지켜낸 어르신의 생명
- 고용노동부. (2024, July 15). 보도자료: 생성형 인공지능(AI)으로 노동약자 보호에 나선다
- 고용노동부. (2024, November 12). 보도자료: 어렵고 복잡한 노동법, 생성형 인공지능(AI)이 쉽게 알려드려요
- 과기정통부. (2023, January). 인공지능 일상화 및 산업 고도화 계획(안).
- 관계부처합동. (2022, September). 대한민국 디지털 전략.
- 관계부처합동. (2023, April). 초거대 AI 경쟁력 강화 방안.
- 관계부처합동. (2023, September). 전국민 AI 일상화 실행계획.
- 관계부처합동. (2024, September). 국가 AI 전략 정책방향(안).
- 김은영. (2024). AI 기반 공공서비스 혁신 방안. 한국정책학회보, 28(2), 123-145.
- 디지털플랫폼정부위원회. (2025). 공공부문 초거대 AI 도입·활용 가이드라인 2.0.
- 디지털플랫폼정부위원회. (2025). 초거대 인공지능 서비스 개발 지원 사업.
- 법제처. (2024, November 12). 보도자료: 법제처, 미래형 AI 법령정보 제공 서비스의 청사진을 제시하다!
- 법제처 (2025, June 9). 보도자료: 인공지능(AI)을 활용한 '지능형 법령검색 서비스' 고도화 착수.
- 법제처. (2025). 지능형 법령검색 시스템 구축(3차) 제안요청서.
- 솔트룩스. (2025). 국민비서 구배 기술 플랫폼 분석보고서. 솔트룩스.
- 유승익. (2024). 경찰청 범죄예측 시스템의 효과성 분석. 치안정책연구, 38(1), 67-89.
- 이건행. (2024). AI 기반 홍수예측 시스템 운영성과. 재난관리연구, 15(2), 45-63.
- 이혜수, 유재홍, 안미소, & 안성원. (2025). 2024년 국내외 인공지능 산업 동향 연구. 소프트웨어정책연구소.
- 이혜림, 김준호, & 박수진. (2025). 공공부문 챗봇 서비스 현황과 발전방향. 행정학연구, 43(1), 123-148.
- 정해경. (2023). 서울톡 챗봇 운영 성과와 개선방향. 지방정부연구, 27(3), 89-112.
- 한국지역정보개발원. (2024). AI를 활용한 노인돌봄 서비스 - 경기도 사례를 중심으로. 지역정보화, 144, 12
- 화성시. (2025, March 25) 보도자료: 화성특례시, 전국 최초 인공지능(AI) 스쿨존 보행안전·공영버스 안전운전시스템 구축... 기술과 시민 잇는 인공지능(AI) 선도도시로 부상.
- 홍승현, 황하. (2024). 누구를 위한 디지털 전환인가? 자동화된 복지행정의 위험성, 30(2), 61-84.
- Anthropic. (2024). AI agents: What are they and how do they work?.
- Cheng, L., Wang, X., & Zhang, Y. (2024). Perception-decision-action cycle in autonomous AI agents. arXiv preprint.
- Deloitte. (2025). Digital government evolution: From basic digitization to AI agents. Deloitte Insights.
- Department for Science, Innovation and Technology. (2025, January 13). AI opportunities action plan (Command Paper CP 1241). Government of the United Kingdom.
- Department for Science, Innovation and Technology. (2025, February). Artificial intelligence playbook for the UK government. <https://www.gov.uk/government/publications/ai-playbook-for-the-uk-government/artificial-intelligence-playbook-for-the-uk-government-html>
- Department for Science, Innovation and Technology. & The Rt Hon Peter Kyle MP. (2025, August 16). AI helpers could coach people into careers and help them move home. <https://www.gov.uk/government/news/ai-helpers-could-coach-people-into-careers-and-help-them-move-home>

- Department for Science, Innovation and Technology, & Ministry of Housing. (2025, May 23). AI experiments see “Humphrey” help townhalls cut costs and improve services. <https://www.gov.uk/government/news/ai-experiments-see-humphrey-help-townhalls-cut-costs-and-improve-services>
- Department for Science, Innovation and Technology, Government Digital Service. (2025, January 21). Shake up of tech and AI usage across NHS and other public services to deliver plan for change <https://www.gov.uk/government/news/shake-up-of-tech-and-ai-usage-across-nhs-and-other-public-services-to-deliver-plan-for-change>
- EU AI Act. (2025). Regulation on artificial intelligence (Artificial Intelligence Act). European Union.
- FedScoop. (2024, December). Federal customer satisfaction index 2024.
- GAO. (2025). Artificial intelligence: Federal government use and oversight challenges (GAO-25-123). U.S. Government Accountability Office.
- Google Cloud. (2025). What are AI agents? ReAct framework and capabilities. Google Cloud Documentation.
- Government Digital Service. (2025). Algorithm transparency recording standard hub. <https://www.gov.uk/government/collections/algorithmic-transparency-recording-standard-hub>
- IBM. (2025). AI agents: Definition, types, and applications. IBM Research.
- IBM (2025). “Dimensionality reduction techniques preserve essential features of complex data sets by reducing the number predictor variables.” What is Dimensionality Reduction.
- Index. (2025). Global AI agents market forecast 2023–2032. Market Research Report.
- Jackson, M. (2024). AI transformation in government: Current applications and future prospects. *Government Technology Review*, 15(3), 45–67.
- Jana, S., Kumar, R., & Patel, M. (2025). Ethical principles for AI governance: International perspectives. *AI Ethics Journal*, 7(2), 234–251.
- Jennings, N. R. (1998). Agent-oriented software engineering. *Artificial Intelligence*, 117(2), 277–296.
- Jennings, N. R., Sycara, K., & Wooldridge, M. (1998). A roadmap of agent research and development. *Autonomous Agents and Multi-Agent Systems*, 1(1), 7–38.
- KISTEP. (2025, February 25). 보도자료: 생성형 AI 서비스(KISTEP-GEN) 개발로 과학기술혁신 정책 분야 AI 전환 선도
- KISTEP. (2025, January). (AI 지능형 분석 시스템 및 KISTEP-GEN 유지관리 사업) 제안요청서
- KISTI. (2024, July 31). 보도자료: KISTI, 과학기술 특화 생성형 언어 모델 KONI 신규버전 공개
- KISTI. (2024, December 20). 보도자료: 논문 비교분석 통한 인사이트 도출까지! ScienceON AI-Reviewer 서비스 출시
- Krishnamurthy, S. (2025). Infrastructure requirements for government AI deployment. *Public Administration Technology*, 12(1), 78–95.
- LGIU. (2024). AI in local government: Performance metrics and citizen satisfaction. Local Government Information Unit.
- Link, D. (2023). Cost reduction through AI automation in government services. *Digital Government Quarterly*, 8(4), 234–256.
- Microsoft Learn (2025). “AI Agent Orchestration Patterns.” Azure Architecture Center.
- Microsoft News. (2024). Understanding AI agents: Specialized tools for business processes. Microsoft Corporation.
- Mohan, P. (2025). Data governance frameworks for AI systems. *Information Management Journal*, 39(2), 45–62.
- Moore, M. H. (1995). *Creating public value: Strategic management in government*. Harvard University Press.
- n8n. (2025). Environmental interaction capabilities of AI agents. n8n Documentation.
- Nb-Data. (2025). Core criteria for AI agent identification. Technical Report NB-2025-03.
- NIST. (2025). AI risk management framework (AI RMF 1.0). National Institute of Standards and Technology.

- OECD. (2024). AI principles and digital government index 2023. OECD Publishing.
- OpenAI. (2025). Introducing ChatGPT Gov
- OpenAI. (2025). AI agents: Independent task completion systems. OpenAI Documentation.
- OpenAI. (2025). ChatGPT agent: Learn about the features of ChatGPT agent mode and how to get started. OpenAI Documentation.
- Reuters. (2025, January). FDA introduces AI agent Elsa for clinical review acceleration.
- Russell, S., & Norvig, P. (1995). Artificial intelligence: A modern approach. Prentice Hall.
- ScienceDirect (2024). "The parsimony principle states that when there are multiple explanations for observed data, the simplest one is preferred." Machine Learning (Second Edition)
- Springbrook. (2025). Digital expectations in local government: 2025 citizen survey. Springbrook Software.
- Spri. (2023). AI 예산 배분 전략 분석. 소프트웨어정책연구소.
- Stanford HAI. (2024). Large language models and generative AI taxonomy. Stanford Human-Centered AI Institute.
- Stone, R. (2024). Digital infrastructure for AI-enabled government. Government Computing, 31(5), 112-128.
- TechTarget. (2025). AI agent classification and implementation guide. TechTarget Enterprise AI.
- UNESCO. (2025). AI ethics framework: Core values and principles. UNESCO Publishing.
- Wang, X., Li, Y., & Chen, Z. (2023). LLM-based autonomous agents with natural language experience storage. Proceedings of AAAI, 37, 15423-15431.
- Whitehouse. (2025, July). American AI action plan. The White House.
- Wooldridge, M., & Jennings, N. R. (1995). Intelligent agents: Theory and practice. Knowledge Engineering Review, 10(2), 115-152.
- World Bank. (2021). Public sector employment and productivity: Global trends. World Bank Publications.
- ZDNET. (2025, March). New Jersey saves millions through generative AI implementation. ZDNET Government Technology
- https://www.dhs.gov/archive/news/2024/12/17/dhss-responsible-use-generative-ai-tools?utm_source=chatgpt.com
- https://www.gsa.gov/about-us/newsroom/news-releases/gsa-launches-usai-to-advance-white-house-americas-ai-action-plan-08142025?utm_source=chatgpt.com
- <https://www.gov.uk/government/news/shake-up-of-tech-and-ai-usage-across-nhs-and-other-public-services-to-deliver-plan-for-change>
- <https://help.openai.com/en/articles/11752874-chatgpt-agent>
- <https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf>
- <https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/>
- <https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-21-Accelerating-Federal-Use-of-AI-through-Innovation-Governance-and-Public-Trust.pdf>
- <https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-22-Driving-Efficient-Acquisition-of-Artificial-Intelligence-in-Government.pdf>

KIPA KOREA INSTITUTE OF PUBLIC ADMINISTRATION

정부혁신 트렌드

- 본 KIPA 정부혁신 트렌드의 내용은 한국행정연구원의 공식 견해와 다를 수 있습니다.

발행처 한국행정연구원 | 발행인 권혁주 | 만든 사람들 최병윤, 황한찬

통권 제6호 | 발행일 2025년 9월 30일 | 주소 서울특별시 은평구 진흥로 235 한국행정연구원 | 전화 02-2007-0662

홈페이지 www.kipa.re.kr | 디자인·편집 디자인크레파스 02-2267-0663

