「너무 이른 온톨로지. 지금은 메타데이터 RAG」

공공부문 AI 신뢰 인프라 구축을 위한 현실적 접근:

온톨로지 과잉 설계를 넘어, 메타데이터 RAG 기반 실시간 현행화·출처 통제·정책 인텔리전스로

---

1. 공공 AI 도입의 현주소와 기술적 과제

공공부문의 인공지능(AI) 도입은 행정 효율화와 대민 서비스 향상을 위한 중요한 과제로 자리 잡고 있습니다. 최근 이를 위해 지식그래프와 온톨로지(Ontology)를 선제적으로 구축하려는 시도가 늘고 있습니다. 그러나 실무 현장에서는 데이터가 정형화된 형태보다는 공직자들의 PC 에 비정형 문서로 산재해 있어, 이를 즉시 구조화하는 데 어려움을 겪고 있습니다. 데이터가 충분히 정비되지 않은 상태에서의 하향식(Top-down) 온톨로지 설계는 예상보다 많은 시간과 예산을 요구하며, 실제 행정 업무에 적용되기까지 상당한 시일이 소요되는 한계가 존재합니다.

또한, 복잡한 행정 문맥을 이해하고 정책을 분석하는 'AI 에이전트'를 원활히 구동하기 위해서는 거대언어모델(LLM)의 높은 추론 능력이 필수적입니다. 최근 발표된 객관적 지표(Artificial Analysis Intelligence Index 등)를 살펴보면, 현재 공공에서 주로 검토되고 있는 국내 LLM 이나 범정부 공통 기반 모델들은 글로벌 최상위 모델 대비 추론 성능에서 다소 아쉬운 지표를 보이고 있습니다. 실질적인 행정 업무 지원을 위해서는 고성능 글로벌 LLM 의 활용 방안을 진지하게 모색해야 할 시점입니다.

2. 현실적 대안으로서의 메타데이터 RAG 도입

스캐터엑스(ScatterX)는 이러한 공공 현장의 고민을 덜어드리기 위해 '메타데이터 RAG(Metadata-RAG)' 아키텍처를 제안합니다. 공공 행정에서는 AI 가 '얼마나 많은 지식을 생성하는가'보다 '어떤 근거를 바탕으로 답변하는가'가 훨씬 중요합니다. 메타데이터 RAG 는 단순한 텍스트 유사도 검색을 넘어 기관, 부서, 작성일, 문서 유형, 시행일 등의 메타데이터를 기준으로 출처를 엄격히 필터링하여 환각(Hallucination) 현상을 최소화합니다.

새로운 데이터 구조를 처음부터 구축하는 대신, 이미 체계적으로 분류되어 있는 정부 및 지자체 홈페이지, 법령 정보 사이트의 기존 구조를 십분 활용합니다. 여기에 조례 개정이나 지침 변경 사항을 자동으로 감지하는 '변경 감시 에이전트'를 결합하여, 과거 지침 인용으로 인한 행정 오류를 방지하고 체계적인 감사 대응 환경을 지원합니다.

3. 국정원 보안 가이드 2.0 준수 및 온프레미스 전처리 아키텍처

고성능 글로벌 LLM 을 공공 업무에 도입할 때 가장 면밀히 검토해야 할 부분은 데이터 보안입니다. 스캐터엑스는 2025 년 12 월 10 일 새롭게 발표된 **'국정원 AI 보안 가이드 2.0'**을 충실히 준수하는 하이브리드 보안 아키텍처를 마련했습니다.

대규모 예산이 수반되는 고가의 GPU 랙 서버 대신, 합리적인 수준의 엔비디아(NVIDIA) RTX 4090 GPU 가 탑재된 워크스테이션을 기관 내부망(폐쇄망)에 구축합니다. 인허가 도면, 민원 첨부 이미지, 회의 녹음 파일 등 민감한 비정형 데이터는 내부망의 AI OCR 및 STT 엔진을 통해 우선 텍스트로 변환됩니다. 이후 내부망의 AI DLP(데이터 유출 방지) 시스템이 개인정보와 민감 정보를 안전하게 비식별화(마스킹) 처리합니다. 이처럼 내부망에서 철저한 보안 처리를 거친 데이터만을 글로벌 LLM API 와 연동함으로써, 국가 보안 규정을 준수하면서도 우수한 AI 성능을 안전하게 활용할 수 있습니다.

4. 실무를 지원하는 12 종의 공공 특화 AI 에이전트

메타데이터 RAG 와 안전한 LLM 환경의 결합은 단순한 검색 보조를 넘어, 실질적인 정책 지원 플랫폼으로 나아가는 튼튼한 기반이 됩니다. 스캐터엑스는 공직자의 실무 부담을 덜어드릴 수 있는 12 종의 특화 에이전트를 준비했습니다.

국민신문고에 접수된 민원의 텍스트와 첨부 파일을 분석해 관련 법령과 처리 방향을 안내하는 **[민원 사전 분석 에이전트]**, 부서 회의를 텍스트화하고 주요 사항을 정리하는 **[회의 자동 비서 에이전트]**, 통신사 및 카드사 데이터를 바탕으로 지역 상권 활성화 입지를 제안하는 **[생활인구 기반 정책 설계 에이전트]**, 포트홀 발생 데이터를 분석해 합리적인 유지보수 계획 수립을 돕는 **[재난 예방 에이전트]** 등은 행정 서비스의 질을 한 단계 높이는 데 기여할 것입니다.

5. 단계적 도입을 통한 안정적인 공공 AI 진화 로드맵

온톨로지는 장기적으로 공공 AI 가 지향해야 할 가치 있는 기술 방향입니다. 다만, 스캐터엑스는 기술 도입의 '순서와 방법'에 대해 보다 안정적인 접근을 제안합니다.

도입 초기부터 대규모 온톨로지를 구축하기보다는, 메타데이터 RAG 를 선제적으로 도입하여 실무 에이전트를 운영하는 것이 예산과 시간 측면에서 합리적입니다. 이 과정을 통해 AI 가 이해할 수 있는 양질의 정제된 데이터가 자연스럽게 내부 시스템에 축적될 것입니다. 향후 이렇게 모인 데이터를 바탕으로 상향식(Bottom-up) 온톨로지를 점진적으로 구성해 나가는 것이 실패 부담을 줄이고 시스템의 실효성을 높이는 현실적인 방안입니다.

스캐터엑스는 공공기관이 가장 안정적이고 신뢰할 수 있는 AI 업무 환경을 구축할 수 있도록, 겸손한 자세로 최선을 다해 기술적 지원을 아끼지 않겠습니다.