

발 간 등 록 번 호

11-1241140-100001-10



2025년 연구보고서

현장조사 질의응답 분야의 생성형 AI 활용 기초연구

2026. 3.



<https://mods.go.kr/dsri>



국가데이터처
국가데이터연구원

현장조사 질의응답 분야의 생성형 AI 활용 기초연구

안영민



발간사

“데이터의 가치는 분석과 활용을 통해 의사결정을 지원하고, 혁신과 효율성 향상 등 구체적인 성과를 창출하는 데서 비롯됩니다.”

급변하는 불확실성의 시대에 데이터는 더 이상 단순한 숫자의 기록이 아니라, 미래를 예측하고 사회 문제를 해결하는 핵심 나침반으로 자리매김하고 있습니다. 국가데이터연구원은 이러한 시대적 요구에 부응하여 국민의 삶을 실질적으로 개선하고 AI 기반의 공공 AX 대전환을 뒷받침하기 위한 데이터 기반 연구에 지속적으로 매진해 왔습니다.

2025년 연구보고서에는 우리 사회가 직면한 환경 변화에 능동적으로 대응하고자 첨단 기술을 국가통계에 접목하기 위해 치열하게 고민한 연구 성과를 담았습니다.

첫째, 인공지능(AI) 기반 국가통계 기술혁신을 선도하고자 노력하였습니다.

생성형 AI 기술을 현장조사에 적용하기 위한 기초연구를 통해 조사자료의 내용검토 및 자동분류, 질의응답에 활용 가능성을 모색하였으며, 이는 통계 생산의 신속성과 정확성을 획기적으로 제고하는 토대가 될 것입니다. 아울러 생성형 AI를 활용한 나우캐스트 지표 서비스 제공 방안 연구는 통계서비스의 새로운 가능성을 여는 의미 있는 첫걸음이라 할 수 있습니다.

둘째, 점차 열악해지고 있는 조사환경에 대응하기 위해 새로운 통계방법론 연구와 국가통계 품질제고를 위한 연구를 강화하였습니다.

확률표본과 자원자표본을 통합한 추정 방안 연구는 응답자 조사 부담을 완화하고 비확률표본의 병행 활용 가능성을 제시하였으며, 데이터 과학기술을 활용한 자료수집 개선 연구와 데이터 통합방법 연구는 다양한 데이터의 연계·통합 방법을 보다 체계화하였습니다.

셋째, 사회적 사각지대를 조명하고 지속가능한 미래를 지원하기 위한 데이터 기반 정책 연구에 집중하였습니다.

최근 심각한 사회 문제로 대두된 ‘고립·은둔 청년’의 실태 파악을 위한 조사 문항 개발 연구를 비롯하여, 돌봄 분야 국가통계 활용 방안과 국내 최초의 기후변화 통계·지표 분석 연구는 데이터가 사회안전망 강화에 기여할 수 있음을 보여줍니다. 또한 소득이동통계 심층 분석 연구와 생애과정 이행에 대한 중·고령기 비교 연구는 관련 정책의 실효성과 활용도를 한층 높일 것으로 기대됩니다.

아울러 가계동향조사의 소비지표 작성 연구와 퇴직연금 적립금 배분 방법 연구는 국민의 체감 경기를 보다 정확히 진단하고 합리적인 경제정책 수립을 지원하는 든든한 기반이 될 것입니다.

2025년 10월부터 새롭게 출발한 국가데이터처 국가데이터연구원은 앞으로도 최신 기술과 사람을 잇는 데이터 연구를 통해 국가통계의 지평을 지속적으로 확장해 나가겠습니다.

본 연구보고서가 통계 생산자와 이용자 모두에게 실질적인 도움이 되고, 각계각층의 의사결정자에게 깊이 있는 통찰을 제공하기를 기대합니다.

많은 관심과 성원을 부탁드립니다.

2026년 3월

국가데이터연구원장

가진

목 차

제1장 서론	1
제2장 현황 연구	3
제1절 AI 챗봇 사례	3
제2절 선행 연구 현황	6
제3장 현장조사 질의응답 업무 현황	9
제1절 현장조사 특성	9
제2절 질의응답 유형 분석	13
제4장 생성형 AI 챗봇 기술 연구	22
제1절 AI 챗봇과 RAG	22
제2절 단계별 기술 검토	24
제5장 요약 및 향후 과제	32
제1절 요약 및 시사점	32
제2절 향후 연구방향	33
참고문헌	35
Abstract	37

요 약

생성형 AI를 기반으로 하는 챗봇은 반복적인 질의응답 업무를 자동화하는 기술로 고객 민원응대의 신속성, 편의성에 기여할 수 있다. 공공부문에서는 세계적으로 AI 기술 중 챗봇이 가장 폭넓게 활용되고 있을 정도로 확산성이 크며 국내에서도 생성형 AI 챗봇을 통해 행정혁신을 도모하려는 움직임이 활발하다.

본 연구는 이러한 관점에서 생성형 AI 기반 챗봇을 실제 업무에 적용하기 위한 기초적 검토를 수행하였다. 이를 위해 첫 번째로 관련 사례 및 선행 연구를 통해 챗봇 기술의 지향점을 가늠해 보았다. 또한 실제 업무 현장에서 수집되는 질의응답 데이터를 구조적으로 검토하여 난이도 수준과 챗봇 설계 시 고려해야 할 요인을 도출하였다. 마지막으로 생성형 AI 챗봇의 효율적 활용을 위한 기술로 RAG 기반 검색·증강·응답 체계의 구축 절차를 정리하였다.

연구결과의 발전을 위해서는 단계적인 실증연구와, 챗봇 성능에 대한 사용자 만족도 등 챗봇에 대한 평가 연구가 이루어질 필요가 있다. 무엇보다도 최신 AI 기술에 대한 꾸준한 적용과 개선 연구를 통해 사용자 눈높이에 맞는 챗봇 개발을 지속해야 한다.

주요 용어 : 생성형 AI, 현장조사, 챗봇, RAG

제 1 장

서 론

인공지능(Artificial Intelligence, 이하 AI) 기술의 급속한 발전으로 사회 전반의 구조가 근본적으로 변화하고 있다. 특히 최근 등장한 생성형 AI(Generative AI)는 인간의 언어를 이해하고 자연스러운 대답을 수행할 수 있는 수준에 이르렀다. 이러한 기술을 빠르게 적용한 기업들은 고객응대, 제품상담, 실시간 서비스 등 다양한 영역에서 AI를 기반으로 업무를 전환하고 대체하여 비용 절감, 서비스 신속성 및 효율성을 동시에 달성하고 있다.

공공분야에서도 이 같은 흐름을 쫓아가고 있다. 민원응대는 국민과 최근접하여 이루어지는 핵심 행정서비스이다. 그러나 공공분야는 효율성뿐만 아니라 공정성·정확성·일관성까지 함께 고려해야 하는 복합적 특성을 지니므로 민간에 비해 시스템 도입이나 업무체계의 전환이 다소 늦은 편이었다. 그러나 최근 우리 정부는 AI 3대 강국을 목표로 AI 대전환을 선언하는 등 AI 기술의 적극적인 도입을 통한 행정서비스의 혁신을 강조하고 있다. 이에 힘입어, 각 정부기관에서도 행정 효율화를 위해 AI 기술의 도입과 활용을 발 빠르게 시도 중이다.

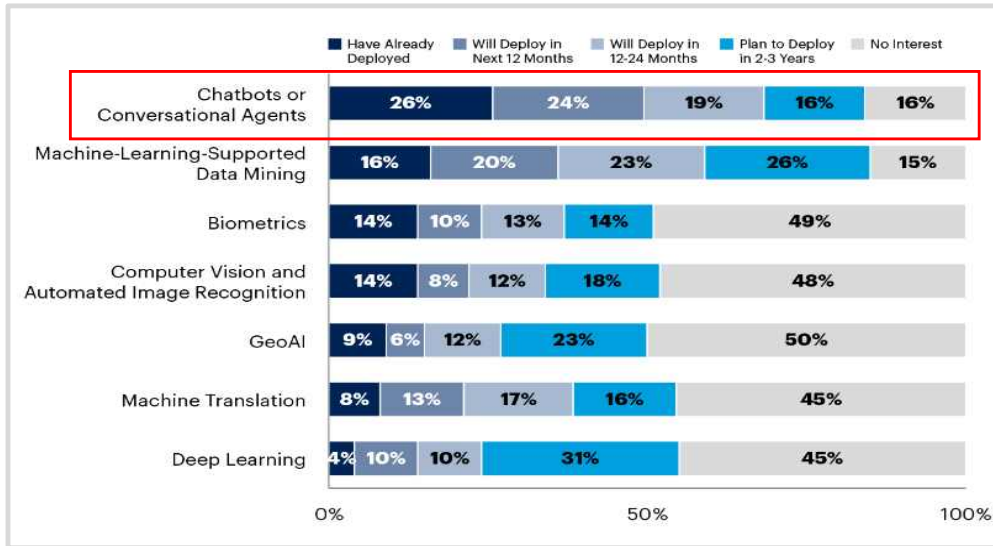
세계적으로 정부 기관 내에서 AI 기술 중 가장 많이 활용된 부문은 챗봇(Chatbot)을 비롯한 대화형 에이전트이다. 가트너(Gartner)에 따르면 전 세계 정부 기관의 26%¹⁾가 챗봇을 비롯한 대화형 에이전트를 도입하여 여러 AI 기술 가운데서도 가장 도입이 활발한 분야로 나타났다. 또한 AI 기술 중 대화형 에이전트를 통한 공공행정의 자동화와 효율화에 대한 기대치가 가장 높아, 이 기술이 향후 공공부문 AI 기술 채택의 핵심이 될 것으로 예상하였다.²⁾

AI 챗봇은 언어모델(Language Model)에 기반한 대화형 도구로서, 사용자의 질의 의도를 분석하고 맥락에 맞는 응답을 생성하는 기능을 가진다. 초기에는 단순히 사전 정의된 규칙(rule-based)에 따라 응답하는 수준에 머물렀으나, 최근에는 대규모 언어모델(Large Language Model, 이하 LLM)을 통해 복합적 상황을 이해하고 자연스러운 문맥을 유지할 뿐만 아니라 구체적인 근거자료를 참조하는 수준으로 발전하였다. 이에 따라 과거 콜센

1) “Gartner Says Government Organizations Are Increasing Investment in AI, But Their Workforce Remains Apprehensive”, 2021.10.5., <https://www.gartner.com>

2) “Gartner Reveals Top Technologies Shaping Government AI Adoption”, 2025.9.9., <https://www.gartner.com>

터 중심으로 운영되던 고객상담 업무가 AI 챗봇에 의해 자동화되고 있으며, 공공기관 역시 민원 관련 질의응답 업무의 효율화를 위해 챗봇 도입을 본격화하고 있다.



<그림 1-1> 각국 정부 내 AI 활용 사례(가트너, <https://www.gartner.com>)

통계작성에서도 AI 기술 활용이 효율적인 민원 분야가 있다. 현장조사 과정에서 통계담당자는 조사응답자나 현장조사 담당자와의 소통 과정에서 다양한 질문에 응답하고, 요청사항을 처리해야 한다. 조사응답자로부터는 조사 개요, 응답자 선정 과정 등에 대한 궁금증에 답변하고 응답 협조를 구해야 하며, 정확한 조사를 위해 현장조사 담당자의 실질적인 조사지침과 자료처리 기준 문의에 대해 답변해야 한다. 이러한 질의응답은 비교적 응답 기준이 명확하고 유사 문의가 자주 반복된다는 점에서 자동화에 적합하다. 따라서 AI 기반 챗봇을 활용해 자동화할 경우, 조사관리 효율성, 응답자 편의, 행정비용 절감 등의 효과를 기대할 수 있다.

따라서 본 연구에서는 국가통계 현장조사 질의응답 업무에 생성형 AI 챗봇 기술을 활용한 자동화 방안을 탐색하는 것을 목적으로 사례조사와 문헌조사, 현장조사 질의응답 데이터 검토 및 챗봇 기술 분석을 병행한 기초연구를 수행하였다. 연구의 구체적 구성은 다음과 같다.

2장에서는 챗봇에 대한 국내 행정기관의 주요 활용 사례를 분석하고, 관련 선행 연구 및 국가데이터처 내부에서 시도된 챗봇 구축 사례를 확인해 보았다. 3장에서는 주요 가구조사를 대상으로 현장조사 질의응답 업무의 특성과 난이도 등 챗봇이 응답할 업무의 특성에 대해 살펴보았다. 4장에서는 생성형 AI 챗봇의 단계별 구축 기술과 현장조사 질의응답 측면의 검토사항을 확인해 보았다. 마지막으로 5장에서는 본 연구의 결과를 토대로 시사점과 향후 연구방향을 제안하였다.

제 2 장

현황 연구

제1절 AI 챗봇 사례

1. 행정안전부 ‘구뻐’

행정안전부가 운영 중인 국민비서 ‘구뻐’는 대표적인 공공부문 AI 기반 상담 서비스이다. 국민들이 단일 플랫폼을 통해 다양한 민원을 처리할 수 있도록 행정안전부를 비롯한 8개 행정기관의 11종 행정서비스를 통합하여, 사용자가 다수 기관의 민원에 대해 단일채널로 상담받을 수 있도록 설계되었다. 공공행정의 서비스 품질을 향상을 위해 다기관이 통합된 챗봇의 선행 사례로 볼 수 있다.

사용자는 ‘구뻐’를 통해 상담 분야를 직접 선택하지 않아도, 챗봇이 입력된 키워드를 자동으로 분석하여 적합한 서비스 분야를 분류하고 이에 따라 해당되는 개별 기관의 챗봇 서비스와 연동되어 다양한 분야의 상담을 제공받을 수 있다.

행정안전부는 최근 ‘구뻐’의 상담서비스를 확장하여 AI 국민비서 서비스 구현에 착수하였다고 밝혔다. 기존의 단순한 질의응답을 넘어서 민간 AI 에이전트와 공공서비스를 연계하여 대화를 통해 증명서 발급, 자원 예약 등의 서비스가 가능하도록 발전된 기술 적용을 시도 중인 것으로 확인되었다.

2. 서울특별시 ‘서울톡’

서울특별시가 운영하는 ‘서울톡’은 AI 기술을 기반으로 서울시 행정 전반에 대한 질의응답과 공공서비스 이용을 지원하는 서비스이다. 이 서비스는 행정 안내, 민원 신청, 복지, 문화, 교통, 환경, 주택영역 등을 포괄하므로 사용자는 지방자치단체 대부분의 업무에 대한 상담과 정보를 제공받을 수 있다. 서울특별시에 따르면, ‘서울톡’은 총 552종의 행정정보와 54종의 민원신청(건의·불편·현장민원 등), 약 7,000여 종의 공공서비스 예약, 40종의 도서관 서비스, 900종의 평생교육 온라인 강좌 안내를 제공하고 있다.



<그림 2-1> '구뵤' 챗봇 연계



<그림 2-2> '서울톡' 위치기반 서비스

‘서울톡’은 별도의 애플리케이션 설치 없이 카카오톡 기반의 메신저 플랫폼을 통해 제공된다는 점에서 접근성이 우수하다. 특히 스마트폰을 활용한 실시간 위치 기반 기능을 연계하여 현장에서 쓰레기 무단투기, 금연구역 단속과 같은 현장민원을 즉시 신고·접수할 수 있어 이용 편의성이 높다. 서울특별시에 따르면 2024년 상반기 기준으로 기존 민원 처리 창구인 ‘서울시 120 다산콜센터’ 문의의 16.7%가 ‘서울톡’으로 대체된 것으로 확인되었다.

다만, ‘서울톡’의 응답 구조는 규칙 기반(rule-based) 또는 검색 기반의 고정 응답에 의존하고 있어서, 복합적인 질문의 맥락 이해나 유연한 대응을 하는 데 어려움이 있다. 이에 따라 서울특별시는 최근 생성형 LLM을 적용한 「챗봇 2.0」 사업을 추진 중이며 기존 규칙 기반 챗봇이 갖는 제약을 보완하고, 대민 행정서비스의 변화를 시도하고 있다.

3. 고용노동부 ‘AI 노동법 상담 서비스’

고용노동부는 생성형 AI 기술을 활용한 ‘AI 노동법 상담 서비스’를 통해 노동 관련 법률 상담을 지원하고 있다. 이 서비스는 임금체불, 근로계약, 퇴직금 등 다양한 노동 관련 문제에 대한 상담을 제공하며, 노동 법령·관례·행정해석 등 최신 데이터베이스를 기반으로 답변을 생성한다. 상담 결과는 필요시 노동분쟁의 참고자료로 활용할 수 있도록 설계되어 있다.

특히 이 서비스는 생성형 AI 기반으로 구축되어 24시간, 34개국 언어로 제공하고 있어서 야간 근로자나 외국인 근로자 등 기존 행정상담 접근이 어려웠던 사용자도 시

간·언어 제약 없이 상담을 받을 수 있다. 응답은 사용자의 질문 재확인, 답변, 법령 등 근거자료 링크로 구성되어 있으며 직전 질문과 답변 내용을 기억·연계하는 기술로 사용자가 추가 질문하는 경우 효율적인 답변이 가능하다. 이처럼 AI 노동법 상담 서비스는 생성형 AI 기반의 챗봇 구축을 통해 행정 상담서비스의 폭과 품질을 한층 강화한 사례이다.

한편, 고용노동부는 대국민 상담서비스 외에도 내부 행정업무 효율화를 위해 ‘근로감독관 AI 비서’를 함께 운영하고 있다. 근로감독관은 노동기준법을 포함한 주요 노동관계법령의 이행 여부를 감독하는 핵심 인력으로, 이 AI 비서는 감독관이 법령 해석, 사안별 유사판례 검색, 조치 절차 확인 등을 신속하게 수행할 수 있도록 지원하는 내부 행정보조 시스템이다. 이를 통해 근로감독관의 업무 부담 경감과 행정의 신속성과 일관성을 제고하고 있다.

고용노동부 상담서비스는 AI 기술이 행정기관 내부의 업무지원과 대국민 행정상담 서비스 양 측면에서 모두 효율성 있게 구축된 사례이다. 특히 공공분야 AI 도입으로 단순 민원 상담 차원을 넘어, 행정처리 근거자료로 활용 가능하게 정책 해석과 판단의 근거를 확장한 점이 의미 있다.

<표 2-1> 챗봇 사례별 특징 및 시사점

구분	국민비서 ‘구뵀’	‘서울톡’	AI 노동법 상담 / 근로감독관 AI 비서
서비스 성격	중앙정부 통합 알림·상담 서비스	지자체 통합 행정안내·민원·예약 챗봇	노동법 상담(AI 챗봇) 및 내부 행정지원(AI 비서)
운영 기관	행정안전부 + 8개 행정기관	서울특별시	고용노동부
주요 기능	행정 알림서비스, 11종 행정서비스 상담, 다수 기관 서비스 통합 상담	행정안내, 현장민원 접수, 공공서비스 예약, 도서관·교육정보 제공	노동법 상담, 판례·법령 기반 응답, 감독관 법령 조회·조치 절차 안내
기술 기반	규칙 기반/ 기관별 데이터 연계, 자동분류 기능	규칙 기반/검색 기반 챗봇	생성형 AI 기반 상담, 최신 법령·판례 데이터 연동, 내부업무 보조 LLM
언어 지원	13개 언어 지원	기본 한국어 중심	34개 언어 지원
접근 방식	정부24 카카오톡 등 메신저·포털	카카오톡 메신저 기반	공식 홈페이지
서비스 대상	대국민	서울시민 및 방문자	대국민(상담), 내부지원(근로감독관)
시사점	민간 AI 전환 연계 시도	위치 기반 서비스 생성형 AI 전환 시도	생성형 AI 챗봇 내외부 업무 지원

제2절 선행 연구 현황

1. AI 챗봇 선행 연구

AI 챗봇에 대한 연구는 약 10년간 AI 기술 발전과 더불어 심화·확장하며 발전되고 있다. 본 기초연구에서는 국내 연구에 한정하여 내용을 살펴보았다.

장준희(2016)는 챗봇 등장 초기 메신저 플랫폼을 기반으로 한 챗봇 서비스의 등장과 이로 인한 ICT 생태계 진화 및 부가가치 창출을 예측하였으며, 빅데이터 플랫폼 구축, AI 기술, 보안기술 등 기술 고도화의 필요성을 강조하였다. 김성근 외(2018)는 챗봇의 유형분류를 시도하였으며 자연어의 처리, 이해, 생성 기술 소개 및 국내외 도입 사례에 대한 연구를 수행하였다.

이후 공공행정에 챗봇을 도입하는 경우 검토해야 하는 여러 가지 사안에 대한 다양한 연구가 시도되었다. 박동아(2017)는 발전하고 있는 AI 챗봇 서비스에 대한 기술적 분석과 공공행정 분야 적용 사례, 상용서비스와는 차별되는 공공행정 서비스 제공을 위해 양질의 데이터 확보가 중요함을 제안하였다. 서교리 외(2017)는 AI를 기반으로 한 공공행정 가상비서는 수요자인 국민 중심의 서비스 발굴이 필요하며, 범부처 데이터 통합과 개인정보 유출대책을 강조하였다. 윤상오(2018)는 공공서비스에서 AI 챗봇을 도입하는 경우 기술오류, 편견 및 왜곡, 경쟁적인 챗봇 도입으로 서비스 중복 및 예산낭비, 기술·법·제도·윤리적 측면의 부작용에 초점을 두고 안전하고 단계적인 챗봇 서비스 도입이 필요함을 주장하였다.

최근에는 생성형 AI 기술 발전에 따라 RAG 기술과 접목한 챗봇 구축에 대한 연구가 활발해지고 있다. 조찬영 외(2023)는 생성형 AI의 환각문제 해결을 위해 RAG기반 챗봇을 제안하였으며, RAG기반 챗봇이 환각을 줄일 수 있음을 확인하였다. 정천수(2023)는 상용LLM과 기업 보유정보를 이용한 RAG기반 AI 서비스를 단계적으로 구현하였고 비용과 시간을 줄일 수 있는 효율성 있는 모델 선택이 필요하다는 결론을 도출하였다. 김휘군 외(2024)는 RAG와 참조정보를 관리하는 벡터데이터베이스의 최적화를 통해 챗봇의 답변 품질을 향상시킬 수 있음을 연구하였다. 최광미(2024)는 RAG기술과 오픈소스 AI 기반 맞춤형 챗봇 구축 절차와 활용방안을 제시하였다.

행정기관에서 챗봇 운영이 확산됨에 따라 챗봇 성능과 확산요인을 분석하는 실증연구에 대한 시도도 이루어지고 있다. 이혜림 외(2025)는 서울시 자치구별로 다양한 원인들이 복합적으로 챗봇 활용과 확산에 미치는 영향을 실증적으로 규명하였다.

최근에는 AI 챗봇의 고도화 등 기능 개선과 평가에 관한 연구도 시도되고 있다.

김혜윤 외(2025)는 RAG의 검색 정확도 향상을 위해 다양한 임베딩 모델과 AdvancedRAG

기술을 조합하고 한국어 성능을 평가하였다. 유희경 외(2025)는 단일 출처로는 답변이 어려운 질문에 대해 Tree-of-Retrieval RAG(ToR-RAG) 구조를 제안하고 이를 LLM-as-a-Judge 모듈을 활용하여 평가하였다.

<표 2-2> AI 챗봇 관련 선행 연구

선행 연구	연구 의의
장준희(2016), 인공지능 기반 챗봇 서비스의 발전 동향, 한국정보화진흥원	·챗봇 기술 도입 초기, 그 의의와 발전 방향 연구
김성근, 신민철, 강주영(2018), 챗봇 기술 소개 및 사례 분석, 한국통신학회	·자연어 처리, 이해, 생성 기술현황과 업종별 사례 연구
박동아(2017), 인공지능 기반 대화형 공공 행정 챗봇 서비스에 관한 연구, 멀티미디어학회	·공공행정 챗봇 발전 과정과 사례 연구
서교리, 강동석, 박선주, 어재경, 김재민, 김효정(2017), 인공지능 기반의 공공 지능형 가상 비서 서비스 발전 모델 연구, 한국통신학회	·공공행정 가상비서 모델 및 발전 방향에 대한 연구
윤상오(2018), 인공지능 기반 공공서비스의 주요 쟁점에 관한 연구: 챗봇(ChatBot)서비스를 중심으로, 한국공공관리학보	·인공지능 챗봇의 부작용과 쟁점 연구
조찬영, 강성준, 정현준(2023), RAG 기반 랭체인을 이용한 생성형 AI 챗봇 구현, 한국정보기술학회	·상용 AI 기반 RAG 챗봇 구축, RAG 성능 단순비교
정천수(2023), LLM애플리케이션 아키텍처를 활용한 생성형 AI 서비스 구현 : RAG 모델과 LangChain 프레임워크 기반, 한국지능정보시스템학회	·오픈소스 GTP4ALL 기반, 내부지식활용 RAG 구축 ·챗봇 성능평가 부재
김휘균, 이지은, 박상현(2024), 인공지능 챗봇을 위한 검색증강생성 및 벡터 데이터베이스 최적화 연구 동향, 정보과학회지	·RAG 및 벡터DB 최적화에 관한 기술동향 문헌연구
최광미(2024), 검색증강생성(RAG) 기반 기업 맞춤형 챗봇(Chatbot) 시스템 구축 및 활용, 한국전자통신학회	·오픈소스 모델(Gemma2 9B) 기반 RAG 구축
이혜림, 김성현, 엄석진(2025), AI 챗봇 도입 현황과 확산 요인 분석: 서울특별시 25개 자치구를 중심으로, 한국사회와 행정연구	·챗봇 활용과 확산에 영향을 미치는 외부요인 실증연구
김혜윤, 노연수, 박종혁, 박민정, 양병욱, 정윤서(2025), 임베딩 모델 및 Advanced RAG 기법을 활용한 한국어 텍스트 분류, 응용통계연구	·검색정확도 향상을 위한 Advanced RAG 성능 실증분석
유희경과 문남미(2025), ToR-RAG ToR기반 RAG를 활용한 복합 질문 처리방안 연구, 한국컴퓨터정보학회논문지	·복합 질의 해결을 위한 ToR-RAG 구조 제안 및 JudgeLLM평가

2. 기관 내 생성형 AI 챗봇 시도

연구 이외에 최근 국가데이터처 내부에서도 생성형 AI를 활용한 챗봇에 대한 관심과 시도를 확인할 수 있었다.

주로 통계인재개발원(현 국가데이터인재개발원) ‘AI 적용 데이터사이언스과정³⁾의 과제 프로젝트로 2024년 과정에서는 생성형 AI를 활용한 가계동향조사 분류코드 챗봇에 대한 구축이 수행되었다. 2025년에도 역시 가계동향조사 담당자를 위한 AI 수행비서에 조사지침에 대한 대화형 Q&A 챗봇을 구축하면서 생성형 AI를 활용하였고, 또다른 과제로 OCR·RAG 기반 가계동향조사 항목 자동분류 챗봇 구축 프로젝트가 진행되기도 하였다. 모두 가계동향조사의 현장조사 단계에 생성형 AI 및 RAG 기술 활용 챗봇을 구축하여 현장조사 효율화를 시도한 사례로 가계동향조사 업무의 자동화·효율화를 위한 기관 내 관심이 높다는 것을 확인할 수 있다.

교육과정 외에 실제 업무에 적용된 사례도 있다. 국가데이터처 산업통계과에서는 2025년 경제통계통합조사 콜센터 운영과정에 생성형 AI를 기반으로 한 RAG 챗봇을 구축하여 시범 운영하였다. 조사지침서를 참고자료로 최신 상용모델을 활용하였으며 업무 혁신성을 인정받아 혁신 우수사례 금상을 수상하였다.

위와 같은 시도들을 통해 최신 생성형 AI 모델 및 RAG 기술을 활용하여 정확하고 편의성 높은 챗봇 개발에 대한 관심이 지속되고 있음을 확인하였다. 그러나 실제 업무에 도입하기 위한 체계적인 검토 및 다양한 실험, 성능평가 또는 그에 준하는 기준 연구는 다소 부족한 점이 사실이다.

<표 2-3> 현장조사 생성형 AI 챗봇 시도

구분	연도	과제명	의의
교육과정	2024	생성형 AI를 활용한 가계동향조사 분류코드 챗봇	·오픈소스 AI모델 ·분류코드에 특화된 모델
	2025	가계동향조사 담당자를 위한 AI 수행비서 제작	·오픈소스 모델(llama-3.1-8B) ·RAG기반 대화형 챗봇
	2025	OCR·RAG 기반 가계동향조사 항목 자동분류 챗봇	·상용 모델(GPT-4o) OCR/ RAG 항목 자동분류
실무적용	2025	경제통계국 산업통계과 경제통계통합조사 RAG 챗봇	·상용 모델(GPT-5.0) ·RAG기반 대화형 챗봇

3) 국가데이터처 통계인재개발원에서 내부 직원 및 외부기관 공무원을 대상으로 데이터과학 전문가(데이터사이언티스트) 양성을 위해 장기간(25년 기준 16주) 운영되는 과정

제 3 장

현장조사 질의응답 업무 현황

본 장에서는 생성형 AI 챗봇이 응답해야 하는 통계 현장조사의 업무 특성을 살펴보고자 한다. 챗봇이 현장조사 질의응답 업무를 기존 방식과 유사하게, 정확하게 처리하기 위해서는 실제 현장에서 발생한 질의응답 사례를 분석하는 연구가 선행되어야 한다. 유형별 질의 빈도를 분석하고, 응답에 활용되는 지침 및 기준정보의 유무, 응답 난이도, 응답 방식 등을 연구하여 챗봇에 기능을 부여하고 정확성과 일관성 높은 답변을 구축하기 위해서 필요한 과정이다.

제1절 현장조사 특성

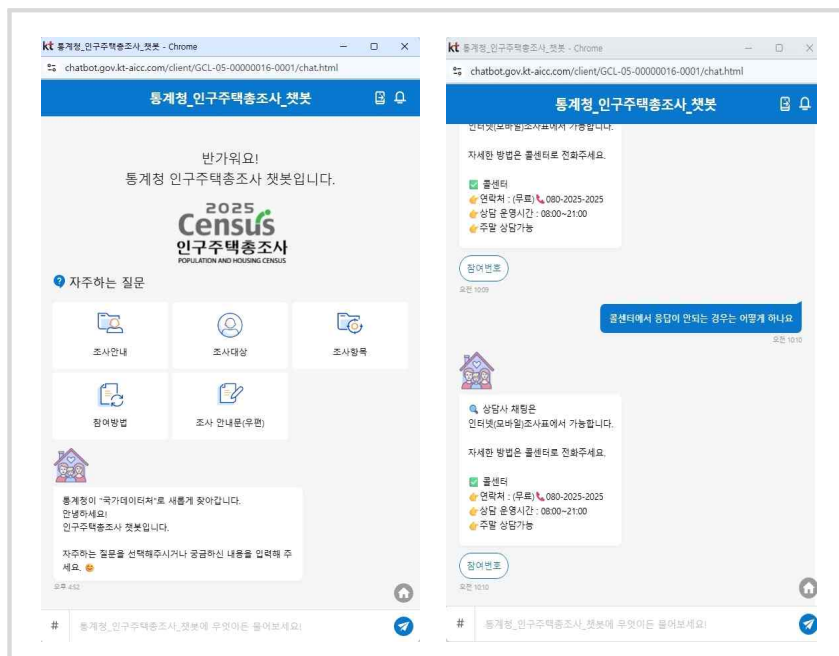
국가데이터처에서는 가구 또는 사업체를 대상으로 조사통계를 작성하고 있다. 조사통계는 조사담당자가 현장방문조사 또는 비대면조사를 통해 응답가구(사업체)로부터 응답받은 내용을 조사시스템에 입력하고 그 내용에 대한 오류나 누락 여부 등에 대한 검토단계를 거쳐 작성된다.

<표 3-1> 통계 현장조사 단계별 절차

주체	→	대상가구 /사업체	→	조사 담당자	→	지방청(사무소)	→	본부 담당부서
역할		·응답		·대상확인, 설득 ·현장조사 수행 ·입력·내검		·행정사항 조치 ·입력자료 총괄내검 ·현장 관리		·조사계획 수립 ·내용검토·질의 ·집계·분석
질 의	대상	·조사담당자 ·지방청 ·본부 담당부서		·지방청 ·본부 담당부서		·본부 담당부서		
	내용	·표본선정 이유 ·조사개요 ·응답방법·기준 ·비밀보호 등		·표본관리 ·시스템 사용 ·항목문의 ·내용검토		·표본관리 ·시스템 사용 ·항목문의 ·내용검토		

현장조사 단계에서는 관련된 주체에 따라 다양한 민원과 질의응답이 발생한다. 먼저 응답가구 또는 사업체로부터 발생하는 민원과 질문은 주로 대상선정 이유, 목적 및 주기 등의 조사개요, 조사항목에 대한 상세내용과 응답기준, 비밀보장 여부 등에 대한 내용으로 방문한 조사담당자 또는 지방청, 본부 담당부서가 응답한다. 조사담당자는 조사 현장에서 발생한 응답대상의 변동 관리, 조사항목 작성요령, 시스템 사용법, 내용검토 등에 대한 다양한 질문을 하고 지방청 또는 본부 담당부서에서 이에 대해 답변한다. 지방청은 현장 총괄관리 역할로 응답자 또는 조사담당자로부터 접수한 질문이 해결되지 않는 경우나 지방청 차원의 자료검토 단계에서 발생하는 질문을 본부 담당부서에 문의하게 된다. 본부 담당부서가 처리해야 하는 현장조사 질의응답은 질의 주체에 따라 조사대상 가구(사업체), 조사담당자, 지방청으로 구분된다.

대규모 예산이 투입되는 총조사에서는 이러한 질의에 대응하기 위해 별도의 콜센터를 구축하여 운영한다. 2025년 기준 인구주택총조사는 전국 약 500만 가구를 대상으로 약 124억 원의 예산이 투입되는 대규모 조사로 응답자와 조사원의 민원과 질의는 상담 콜센터를 통해 24시간 처리가 가능하도록 전문위탁기관이 맡아 운영하였다. 콜센터는 AI 기반의 챗봇(보이스봇)이 1차로 민원응대를 하고 부족하거나 추가 답변이 필요한 경우 2차로 상담원을 통한 전화 또는 채팅상담 방식으로 이원화되었다. 2025년 인구주택총조사 콜센터 구축비용은 약 5.8억 원⁴⁾에 이른다.



<그림 3-1> 2025년 인구주택총조사 규칙 기반 AI 챗봇

4) 콜센터 인프라 구축비 및 상담프로그램 개발비(같은 해 농림어업총조사와 연계 운영)

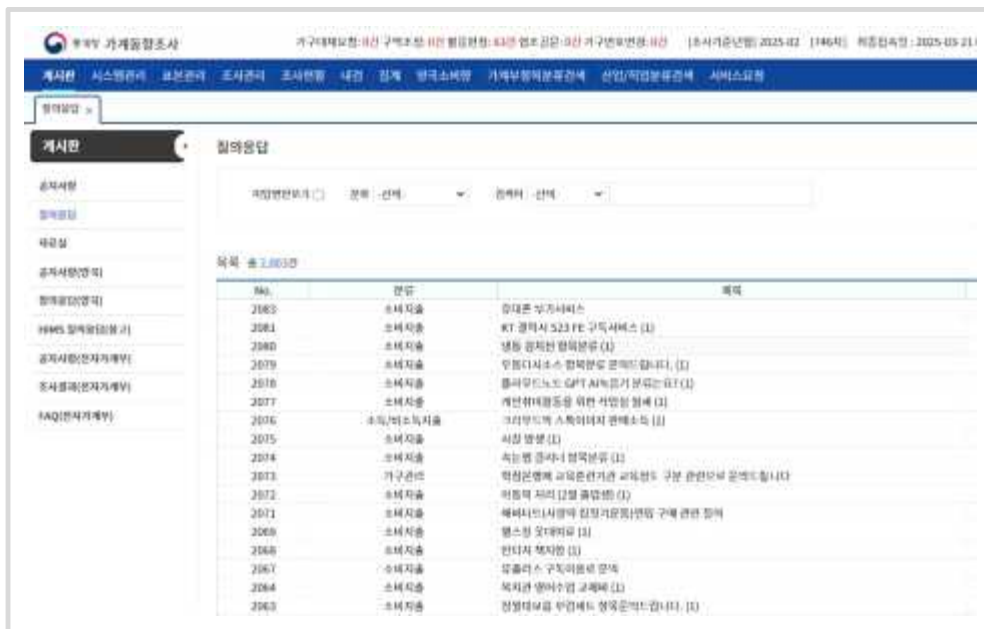
총조사를 제외한 나머지 통계조사는 별도의 예산을 투입하여 이러한 질의응답을 해결하기가 어려운 상황이므로 대부분의 부서에서는 응답자로부터 접수된 질의는 우선으로 해결하고, 조사담당자와 지방청으로부터 발생하는 질의는 이미 구축되어 있는 내부 자료처리 시스템의 묻고답하기 게시판 등을 통해서 주로 응답이 이루어지고 있다.

현장조사 질의응답의 특성을 분석하기 위해서는 앞서 언급된 응답자, 조사담당자, 지방청의 모든 질의와 응답을 검토해야 그 특징을 완전히 이해할 수 있다. 그러나 통계작성 담당부서의 주된 업무는 응답 데이터의 관리와 집계에 집중되어 응답자로부터 접수된 질의응답을 별도로 보관하거나 기록하는 사례는 거의 없다.

본 연구에서는 이러한 한계로 인하여 내부 자료처리 시스템 게시판을 통한 조사담당자와 지방청으로부터 요청된 질의응답 사례에 대해서만 살펴보기로 한다. 이를 위해 국가데이터처의 대표적인 가구 대상 조사 중 현장조사 수행주기가 월간인 가계동향조사와 연간인 가계금융복지조사의 자료를 통해 현장조사 질의응답 현황과 그 특징을 살펴보았다.

1. 가계동향조사

가계동향조사는 가구에 대한 가계수지 실태를 파악하여 국민의 소득과 소비 수준의 변화를 측정하고 분석 등에 필요한 자료를 제공하기 위해 작성하는 통계이다. 우리나라 소비자물가지수 작성에 필요한 가중치 산출, 소득수준 측정 및 소득동향 파악을 통한 정책수립, 국민소득 추계 등 경제·사회통계 작성 등에 필요한 기초자료를 제공한다.



<그림 3-2> 가계동향조사 조사시스템 질의응답 화면(이용자용 통계정보보고서)

가계동향조사는 별도의 임시 조사원 채용 없이 정규직 공무원 또는 공무원들이 상시적으로 조사를 하고 있다. 대부분 수년간 통계조사 유경험자인 조사담당자들이 매년 온라인, 집합교육, 워크숍 등을 통해 업무에 대한 교육을 받고 매달 현장조사를 수행하고 있다.

현장조사에서 발생하는 질의와 응답은 가계동향조사 조사시스템의 질의응답게시판을 활용하여 운영되고 있다. 게시판을 이용하여 조사담당자, 지방청 총괄담당자가 질문내용을 작성하면 본부 담당부서의 직원이 직접 게시판을 통해 응답을 작성하는 체계이다.

2. 가계금융복지조사

가계금융복지조사는 가구의 자산, 부채, 소득 등의 규모, 구성, 분포와 미시적 재무건전성과 소득분배지표를 파악하여 사회경제 관련 정책과 연구에 활용하기 위하여 작성

<표 3-2> 가계금융복지조사 조사업무 흐름도

단계	일정	본부 담당부서	지방청(조사요원)
조사준비	1~3월	·시스템 점검 ·조사표류 준비 ·조사용품 준비	·조사원 채용 ·조사용품 수령확인
조사안내/ 업무설명회	3월	·공문, 인사장 발송 ·공무원 업무설명	·업무설명 참석
교육	3월 중순 ~ 말	·조사요원 교육	·교육장소 일정확인 ·조사요원 교육
준비조사	3월 말	·조사대상 질의응답 ·홈페이지 Q&A 운영 ·콜센터 운영	·홍보물품 부착, 배부 ·가구연락 방문확인 ·가구명부 보완
본조사	3월 말 ~ 4월 중순	·항목 질의응답 ·홈페이지 Q&A 운영 ·콜센터 운영	<조사원> ·조사표 작성, 답례품 지급 ·항목연관성 점검 <조사관리자> ·대상가구 명부보완 ·조사표 보완검토 ·항목 질의응답 <지도공무원> ·업무량 조정 ·지침전달 및 내검지도
입력 내검	4월 중순 ~5월 중순	·입력지도 ·내검요령 설명	·조사표 입력 ·입력자료 점검 및 보완
조사표 제출	5월 중순	·조사표 수령 및 정리	·조사표 취합, 제출

되는 통계이다. 가계생활 수준의 정도 및 변화 요인 등을 종합적으로 파악하여 재정 및 복지 관련 정책과 연구에 활용되고 있다.

가계금융복지조사는 지방청에서 도급조사원을 채용하여 현장조사를 수행한다. 2024년 기준 채용규모는 도급조사원 700명 이상, 조사관리자 150명 이상, 입력내검원 120명 이상이다. 동일 조사 유경험자에 대해 채용우대 조건으로 조사원을 선발하고 있다.

현장조사 기간 중 질의응답은 콜센터와 조사입력 시스템 게시판의 두 가지 방법을 이용하여 이루어지고 있다. 조사기간 중 한시적으로 운영되는 콜센터는 응답자와 조사원들이 전화를 해서 질문을 하면 채용된 콜센터 요원이 유선상으로 답변해 주는 방식이다. 내부 시스템을 활용할 수 있는 조사원, 조사관리자 등은 조사기간 또는 내용검토 기간에 시스템 ‘묻고답하기’ 게시판을 이용해서 질문을 할 수 있다. 게시된 질문은 본부 업무 담당자가 검토 후 직접 응답한다. 가계금융복지조사는 매년 조사시스템 질의응답 코너를 운영하고 사례를 축적하여 관리한다. 주요 질의응답 사례는 차년도 조사 지침서 및 조사사례집에 반영하여 조사원 교육에도 활용된다.

제2절 질의응답 유형 분석

1. 가계동향조사

2023년부터 2025년까지 가계동향조사 조사시스템 묻고답하기 게시판에 등록된 게시물 건수는 총 997건⁵⁾이며 1개 게시물에 2개 이상의 질의가 있는 경우를 분리한 개별 질의응답 건수는 총 1,082건이다. 조사담당자의 질의정보와 본부 담당부서의 응답정보의 수준 및 난이도 등이 서로 상이하므로 질의와 응답을 각각 구분하여 특성을 파악하였다.

가. 질의 유형 분석

묻고답하기 게시판은 조사담당자가 문의내용을 등록할 때, 사전에 설정된 카테고리⁶⁾ 중 질문내용과 가장 부합하는 항목을 선택하도록 구성되어 있다. 이러한 분류체계를 참고하여 질의자료의 유형을 재구성하였다.

5) 2025년 자료는 자료입수 기준일(2025년 10월 20일) 이전까지의 자료

6) 가구관리, 검색어등록요청, 소득/비소비지출, 소비지출, 지침서류, 개선의견

<표 3-3> 가계동향조사 질의 유형별 분석(건, %)

구 분	2023년	2024년	2025년	합계	비율(%)
1) 가구관리	12	16	34	62	5.7
2) 소득/비소비지출	60	67	67	194	17.9
3) 소비지출	377	194	230	801	74.0
4) 시스템 오류/개선	1	17	7	25	2.3
총합계	450	294	338	1,082	100.0

1) 가구관리

가구관리 업무에 대한 비율은 5.7%로 나타났다. 가계동향조사 가구관리 업무는 조사대상 및 제외 확인, 조사구 연동 또는 부적합으로 인한 조사구·표본가구 교체 등의 과정을 의미한다.

현장조사에서는 가계동향조사 가구표본 관리지침서를 기준으로 가구관리 업무를 처리하게 되는데, 선정된 응답가구 및 가구원의 조사대상 또는 제외 여부가 모호하거나 지침에 명시된 기준만으로 명확히 판단되지 않을 때 질의응답 게시판을 통해 질문을 하게 된다. 또 조사대상 가구와 가구원이 확정된 후 가구원별로 학력 정도, 취업사항 등 특징을 관리하는 표본가구 관리명부 작성과정과 주요품목 구입 여부, 반려동물 양육 여부 등 가계부 표지항목 조사기준에 대한 질문도 가구관리 유형으로 분류하였다.

이 유형의 정확한 응답을 위해서는 기존에 참고 가능한 정보 이외에 추가 정보가 필수인 사례도 있다. 예를 들어 가구원의 소속 교육기관이 학위 취득 및 학력인정 여부에 따라 교육정도가 달라지는 경우에는 챗봇이 추가 정보를 직접 검색하거나, 추가 정보를 질문하고 기준과 정보를 종합하여 판단해야 한다.

2) 소득/비소비지출

소득/비소비지출 관련 질의는 전체의 17.9%로 나타났다.

가구에서 지원받는 사회보장 차원의 정부 지원금 및 환급금이 새로 발생하는 경우 이에 대한 항목분류 관련 질문이 대부분이며, 이러한 질문 응답에는 지급되는 소득 지급 주체(지자체·중앙정부, 가구 내·외)와의 관계를 통해 소득 여부 및 분류를 파악해야 하는 경우가 다수 있다. 기존 정보로 분류하기 어려운 경우에는 지원금 유형과 지급 주체 등에 대한 추가 정보가 필요하다.

또 다른 유형은 조사지침서, 항목분류집 및 사례집의 기존 정보를 확인하였으나 응답가구의 개별 사례와 불일치하여 재질의하는 유형으로 응답 난이도가 높은 유형이다. 질문에 세부 정보가 담기지 않은 경우에는 항목 분류를 위한 추가 정보를 요청해야 한다.

3) 소비지출

가구 소비지출과 관련된 질의는 74.0%로 대부분을 차지하였다. 조사담당자는 가구에서 응답한 가계부상 지출품목을 사전에 정의된 품목별 소비지출분류 항목에 따라 구분하여야 한다. 각 지출항목과 항목에 해당하는 품목유형, 제품명 등은 사전에 배부되는 항목분류집 또는 조사입력시스템의 검색어 사전에 등록되어 있어 현장조사 과정에서는 조사담당자들이 조사표 입력단계에서 참고자료로 활용하게 된다.

그러나 분류 기준이 다소 복잡하거나, 신제품 출시 등으로 항목분류집에 실시간 반영될 수 없는 품목이나 여러 성분이 혼합된 품목 등은 현장에서 정확한 세부항목 분류에 어려움을 겪을 수 있기 때문에 이에 대한 질의가 빈번하다.

소비지출에 대한 질문은 가장 많은 부분을 차지하므로 세부유형을 한 번 더 구분하여 살펴보았다. 가구에서 구매한 상품의 항목분류에 대한 문의는 전체 질의유형 중 53.5%로 세부항목 중에서도 가장 많았다. 질문과 응답이 비교적 단순하여 기존 지침과 항목분류 기준으로도 응답이 가능한 유형으로 판단된다.

다만, 기존 검색어에 존재하지 않는 신제품 또는 신제품에 대한 질문이 6.9%로, 챗봇의 경우 사용자가 최초 질문에서 용도, 상품의 목적 등을 상세히 설명하지 않은 채 질의한다면 챗봇은 정확하기 응답하기 어렵다. 따라서 챗봇이 사용자에게 추가 질문을 하거나, 또는 외부 자료 검색을 통해 부가정보를 확보하는 방안이 필요하다.

<표 3-4> 소비지출 질의 세부 유형(2023~2025년, 건, %)

소비지출	건수	비율*(%)	챗봇 구축 검토사항
신제품	75	6.9	
분류기준	181	16.7	·질의 발생 모니터링 및 관리 ·상세 사례 및 다양한 품목 예시
재질의	41	3.8	
합계	801	74.0	

*전체 질의 건수 1,082건 대비 비율

또 다른 유형으로 기존 항목분류 기준에 대하여 조사담당자들이 상세설명을 요청하는 경우가 있다. 항목분류 기준은 가계동향조사 항목분류집에 상세설명이 있으나 현장에서 개별 사례가 지침상 기준과 정확히 일치하지 않아 분류가 어려운 경우에 질의가 발생한다. 챗봇이 단일 참고자료만으로 응답이 가능한 경우와 여러 종류의 참고자료를 복합적으로 추론하여야 하는 상황을 구분할 필요가 있다.

마지막 소비지출 유형은 기존 지침자료와 질의응답을 참고하였으나 현장에서 여러 번 재질의된 경우이다. 질문의 비율은 3.8%로 높지 않지만 재질의될 정도로 난이도가 다소 높은 질문이라 볼 수 있다. 챗봇 응답 시 조사지침서와 항목분류집의 기준과 상세사례 및 품목예시를 복합적으로 제안하는 형태로 응답방식을 설계한다면 현장에서 유용하게 활용될 것으로 보인다. 또한 챗봇을 통한 질문 및 응답내용과 빈도를 담당부서에서 모니터링해서 기준을 구체화하여 피드백해 준다면 유용할 것으로 기대된다.

4) 시스템 오류 및 개선

가계동향조사 입력시스템은 매월 정상적으로 운영되는 시스템으로서 연간조사 대비 오류 상황이 비교적 드물 것으로 예상된다. 실제 자료를 통해서도 해당 비율은 2.3%로 낮게 나타났다. 오류 상황에서는 문제 해결을 위해 시스템 관리자가 조치해야 할 필요가 있으며 챗봇에서 시스템 오류에 대한 질의 발생을 신속하게 체크할 수 있도록 하는 기능이 필요할 것으로 보인다. 실제 오류가 아닌 시스템 사용법 숙지가 필요한 경우도 있으며 이 경우 챗봇이 시스템 사용 매뉴얼을 안내하여 신속하게 조치될 수 있어서 유용할 것으로 기대된다.

나. 답변 유형 분석

1) 단일 항목 코드 안내

답변 유형은 단일 항목분류코드에 대한 안내가 전체 응답의 57.3%로 가장 많았다. 응답 내용이 비교적 단순한 형태로 응답 난이도가 높지 않은 것으로 판단된다. 항목분류집을 참고한 응답을 설계하면 챗봇을 통한 업무 효율화 효과가 클 것으로 예상되는 유형이다.

2) 상황별 코드 안내

응답 내용이 사례에 따라 다른 경우도 13.8%이다. 가구의 특성, 지출의 목적 등에 따라 처리방법이 달라지기 때문이다. 이 경우에는 A상황인 경우 항목분류코드(A), B 상황인 경우 항목분류코드(B) 같은 방식으로 응답이 작성되었다. 항목분류코드가 달라지는 기준은 식품, 공산품 등 품목에 따라 매우 복잡하다, 예를 들어 식품인 경우

세부 성분, 구입처 등에 따라, 공산품인 경우 용도, 전기사용 여부 등에 따라 분류코드가 달라진다. 따라서 챗봇이 질문에 포함된 정보만으로 단일 항목분류코드만 제시하는 것보다 상황에 따라 코드를 선택할 수 있도록 안내할 필요가 있다. 가장 일반적이고 가능성 높은 코드를 우선 제시하고 상황별로 후순위 코드를 제안하는 방법도 효과적으로 보인다.

<표 3-5> 가계동향조사 응답 유형 분석(건, %)

구 분	2023년	2024년	2025년	합계	비율(%)
1) 단일 항목 코드 안내	255	165	200	620	57.3
2) 상황별 코드 안내	58	40	51	149	13.8
3) 시스템(오류) 수정	36	24	37	97	9.0
4) 추가정보 필요	25	4	4	33	3.0
5) 기존 사례 안내	52	27	24	103	9.5
6) 응답 없음	24	34	22	80	7.4
총합계	450	294	338	1,082	100.0

3) 시스템 오류 수정

응답 유형 중 시스템 오류 수정 후 수정완료 답변이 9%로 나타났다. 질의 유형의 시스템 오류 문의 비율이 2.3%인데 비해 답변 유형에서의 비율이 높게 나타났다. 구체적인 사례를 통해 살펴본 결과, 조사담당자가 지출품목명으로 코드를 검색한 결과 해당 품목의 분류코드가 2가지 이상 나타난 경우나 기존 지침 또는 기준과 검색어 분류코드가 맞지 않는 경우에 분류코드를 보완한 사례가 다수인 것으로 확인되었다. 항목 분류 기준은 분류코드 개편이나 신품목 등장 등 상황에 따라 변동될 수 있으므로 과거의 검색어-항목분류코드 조합의 수정 사항이 현재 시스템에 미처 반영되지 않은 상황으로 예상할 수 있다. 이와 같은 경우 챗봇은 수정되지 않은 기존 검색어-항목분류코드를 정답으로 인지하고 잘못된 응답을 할 가능성이 있으므로 챗봇 구축 사전 단계에서 검색어와 항목분류코드가 올바르게 매치되어 있는지 확인하는 절차가 필요한 것으로 보인다.

4) 추가정보 필요

게시물 상의 질의만으로는 충분한 분류코드 안내나 작성요령 설명이 어려워 추가정

보를 요청한 응답도 3.0%이다. 사용자의 질문 내용이 충분하지 않거나 한정된 정보만을 가지고 질문한 경우 챗봇은 질문내용만으로는 잘못된 항목분류 또는 작성요령을 설명하는 오류를 발생할 가능성이 있다. 이러한 경우 챗봇도 실제 질의응답 과정과 마찬가지로 사용자에게 추가적인 정보를 요청할 수 있도록 세부적인 설계가 필요하다.

5) 기존 사례 안내

조사지침서, 사례집, 항목분류집의 특정 페이지 또는 기존 게시판 질의응답 내용을 참고하도록 안내한 응답이 9.5%를 차지하였다. 이미 안내된 사항이나 반복되는 질문에 대한 난이도가 낮은 응답으로, 챗봇을 통한 상세 설명과 응답으로 효율적인 현장조사 관리에 효과적일 것으로 판단된다.

6) 응답 없음

응답 유형 중 ‘유선답변’ 등 참고할 응답내용이 없는 경우도 7.4%로 나타났다. 이 같은 경우는 게시물로 응답할 경우 해석이 복잡하거나 오해의 소지가 있어서 전화 등 질문자에게 직접 답변한 경우이다. 이러한 유형은 실제로 챗봇이 참고할 수 있는 응답내용이 없으므로 챗봇 구축 전에 응답내용을 데이터화하는 작업이 필요하다.

<표 3-6> 가계동향조사 질의 및 응답 유형별 챗봇 검토사항

구 분		비율(%)	챗봇 구축 검토사항
질의	가구관리	5.7	정보가 부족한 경우 추가정보 요청
	소득/비소비지출	17.9	정보가 부족한 경우 추가정보 요청 / 외부 자료검색 필요
	소비지출	74.0	
	항목분류/신품목	53.5	외부 자료검색 기능
	분류기준/재질의	20.5	질문 모니터링/ 챗봇 피드백 필요
	시스템 오류	2.3	실시간 모니터링 필요
응답	단일 항목 코드 안내	57.3	응답 난이도 낮음
	상황별코드 안내	13.9	정보가 부족한 경우 추가정보 요청 / 상황별 우선 순위 코드 안내
	시스템 오류 수정	9.0	검색어 사전 재정비
	추가정보 필요	3.0	정보가 부족한 경우 추가정보 요청
	기존 사례 안내	9.5	응답 난이도 낮음
	응답 없음	7.4	응답부재 자료 데이터화
합 계		100.0	

2. 가계금융복지조사

2023년부터 2025년까지 가계금융복지조사 묻고답하기 게시판에 등록된 질의 건수는 총 691건이다. 가계금융복지조사 자료는 부득이하게 질의게시물의 제목만 확보되었다. 상세한 질의 및 답변 내용을 확인할 수 없으므로 질문 제목 정보만을 통해 유형을 유추하여 구분하였다. 2025년 가계금융복지조사 시스템부터는 질문을 게시할 때 사용자가 직접 유형을 선택하도록 변경되어서 미리 분류된 유형을 일부 참고하여 구분하였음을 밝혀둔다.

<표 3-7> 가계금융복지조사 질의응답 유형별 분석(건, %)

구분	2023년	2024년	2025년	합계	비율(%)
1) 인터넷조사	5	6	37	48	6.9
2) 작성 요령	50	53	63	166	24.0
3) 자료 입력	43	27	129	199	28.8
4) 시스템 관련	7	53	96	156	22.6
5) 내용 검토	18	16	88	122	17.7
총합계	123	155	413	691	100.0

1) 인터넷조사 관련

인터넷조사 관련 질의는 전체 6.9%로 비교적 낮은 비율을 차지하고 있다. 가계금융복지조사는 현장조사와 더불어 응답 편의를 위해 온라인을 통한 인터넷조사를 병행하여 실시하고 있다. 가구에서는 응답을 위한 가구별 아이디와 초기비밀번호를 부여받아서 인터넷조사에 접속한다. 부여된 초기비밀번호는 응답내용 보호를 위해 가구에서 최초 접속 후 자체적으로 변경하여 이용하는데 변경된 비밀번호를 분실한 사례가 잦아 조사담당자를 통해 비밀번호 초기화를 요청하는 경우가 주요 내용이다. 이와 같은 질의 발생 시 시스템관리자를 통한 비밀번호 초기화 조치가 필요하므로 챗봇과 조사입력시스템이 상호 연동되어 조치가 취해지는 등 실시간 처리가 가능할 경우 매우 유용할 것으로 판단된다.

2) 작성요령

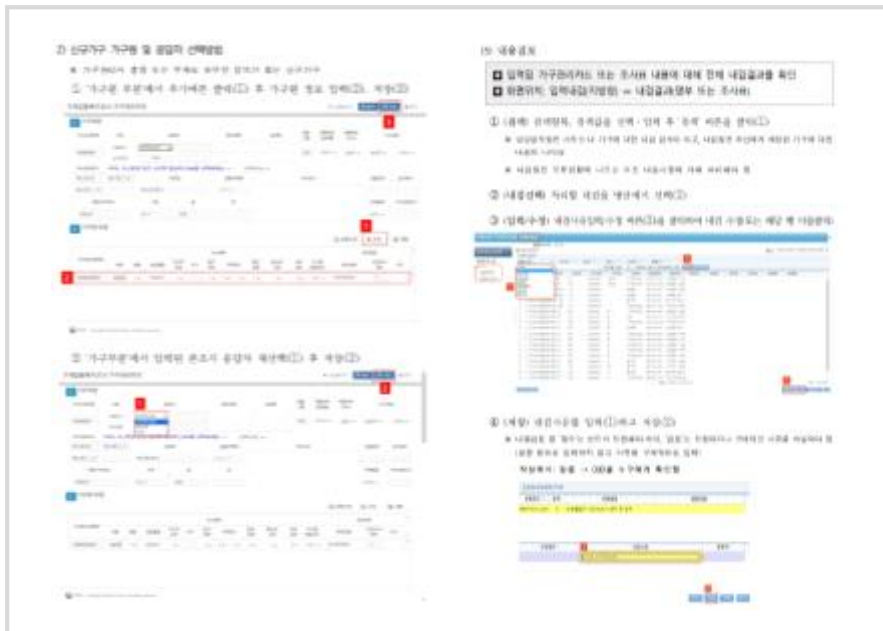
가구의 자산, 부채, 소득, 지출 등 응답 항목에 대한 작성요령의 질문이 24%로 많은 비중을 차지하였다. 가계금융복지조사는 가구의 금융 상태를 세부적으로 파악하는 항

목이 있고 전체 조사 항목이 약 190개, 평균 응답 소요시간이 42.2분7)에 달하는 난이도가 높은 조사이다. 조사항목이 많고 시간이 많이 소요되는 복잡한 조사이므로, 챗봇 시스템이 조사지침서와 사례집, 기존 질의응답 내용 등 조사요령에 해당하는 자료를 참고하여 활용한다면 효율적인 현장조사 수행에 도움이 될 것으로 예상되는 부분이다.

3) 자료입력

자료입력 관련 질의는 전체의 28.8%로 가장 많은 부분을 차지하였다. 가계금융복지 조사는 응답가구의 주소, 가구원, 가구원별 특성 등 조사문항 이외에 일반사항을 가구 관리카드에 입력하고 관리한다. 자료입력은 주로 가구관리카드의 입력사항을 수정하고 싶은 경우, 초기화 또는 수정보완을 요청하는 질의가 대부분이다. 시스템에서 관리자의 조치가 필요한 부분으로 이는 인터넷조사 유형과 마찬가지로 챗봇과 조사입력시스템 간 연동조치가 가능하다면 즉시 해결 가능할 것으로 보인다.

자료입력과 관련된 또 다른 질의 유형은 조사표 입력방법 및 시스템 이용방법에 대한 문의이다. 가계금융복지조사는 매년 조사담당자를 채용하므로 시스템 사용에 익숙하지 않은 사용자의 경우 자료 입력방법에 대한 질의가 발생한다. 따라서 조사시스템 사용요령을 챗봇이 참조하여 응답하게 될 경우 반복된 질문에 효율적으로 대응이 가능한 부분이다.



<그림 3-3> 가계금융복지조사 시스템 입력 요령(가계금융조사 조사지침서)

7) 2024년 가계금융복지조사 이용자용 통계정보보고서

4) 시스템 관련

시스템 관련 질의는 22.6% 비중으로 나타났다. 가계금융복지조사는 조사항목 수정, 시스템 개선·보완 등으로 인해 조사입력시스템 변동이 매년 발생하게 된다. 따라서 변경된 조사입력시스템의 기능이 원활하지 않은 경우 현장조사 초기 조사담당자로부터 시스템 기능사항에 대한 문의가 발생한다. 문제 해결을 위해 관리자가 시스템을 직접 수정해야 할 필요가 있는 경우는 본부 담당자가 질의내용을 실시간 체크할 수 있도록 하는 기능이 유용하다. 조사담당자 PC의 시스템 조치만으로 해결 가능한 경우에는 챗봇이 문제 해결 절차를 자동으로 안내할 수 있다. 시스템 관련 유형 중 일부는 자료입력 유형과 마찬가지로 시스템 사용에 익숙하지 않은 경우이므로 챗봇이 조사시스템 사용 매뉴얼을 참고하여 응답하도록 설계하는 것이 유용하다.

5) 내용검토

내용검토에 대한 질의유형은 17.7%로 나타났다. 조사표 입력이 완료되면 조사항목 간의 정합성과 전년도 금액과의 비교검토를 통해서 내용검토가 필요한 사항에 대한 점검메시지가 입력시스템상에 표출된다. 가계금융복지조사의 내용검토 사항은 코드로 구분하여 관리하고 있다. 입력내용의 확인과 사유가 필요한 검토(CHK)코드는 입력항목이 이상치에 가까운 반면, 응답가구의 특별한 사유가 확인되는 경우 사유를 입력하면 전체 조사표 입력이 완료되는 코드이다. 입력내용의 명백한 오류로 판단되는 필수(ERR) 코드가 표출되면, 해당 항목의 변경을 위한 수정 없이는 조사표 입력이 완료되지 않는다. 이 유형의 질의 대부분은 필수코드가 표출되어 삭제를 요청하는 경우로, 입력자는 정상 입력하였는데 시스템 오류가 발생하였거나, 입력자가 필수코드 내용을 제대로 이해하지 못하여 오류로 판단하는 경우이다. 마찬가지로 시스템에 대한 사용설명이나 내검코드에 대한 재설명, 또는 관리자 차원의 시스템 수정이 필요한 부분이다.

<표 3-8> 가계금융복지조사 질의응답 유형별 챗봇 검토사항

구 분	챗봇 구축 검토사항
1) 인터넷조사	·조사입력시스템 연동 ·질의 발생 실시간 체크
2) 작성요령	·조사지침서, 사례집 검색 및 안내
3) 자료입력	·조사입력시스템 연동 ·조사시스템 입력요령 검색 및 안내
4) 시스템 관련	·질의 발생 실시간 체크 ·조사시스템 입력요령 검색 및 안내 ·문제 해결 절차 안내
5) 내용검토	·질의 발생 실시간 체크 ·내검코드 검색 및 안내

제 4 장

생성형 AI 챗봇 기술 연구

본 장에서는 앞서 살펴본 현장조사 질의응답 업무를 대신할 생성형 AI 챗봇의 구축 기술을 검토해 보았다. 기초연구이므로 세부적인 기술과 적용보다는 챗봇 구축과정을 단계별로 살펴보고 각 단계별 필수적인 요소와 검토사항에 대해서 중점적으로 살펴보았다.

제1절 AI 챗봇과 RAG

1. 챗봇의 정의 및 유형

챗봇이란 사람과의 문자나 대화를 통해 질문에 알맞은 답이나 각종 연관정보를 제공하는 AI 기반의 커뮤니케이션 소프트웨어를 지칭한다. 챗봇이라는 용어는 대화를 뜻하는 채팅(Chatting)의 'Chat'과 로봇(Robot)의 'Bot'이 결합된 합성어로 여기서 'Bot'은 사용자의 행동 또는 다른 프로그램의 동작을 모방하거나 자동화하여 처리하는 소프트웨어를 의미한다. 챗봇은 기술 단계에 따라 세 가지 유형으로 구분할 수 있다.

규칙 기반 챗봇(Rule-based Chatbot)은 사전에 정의된 질문-응답의 규칙과 키워드에 의존하여 작동한다. 사용자의 입력과 데이터베이스에 저장된 규칙 간 일치 여부에 따라 응답이 결정되므로 구현이 용이하고 특정 분야에서는 높은 정확도를 달성할 수 있다. 반대로 사용자의 표현이 규칙 패턴과 일치하지 않거나 복잡한 경우, 또 미리 정의된 답변 유형과 다른 경우에는 응답의 정확성이 떨어지는 한계를 가진다. 단순 FAQ 자동응답, 고객센터 기본 안내 등이 대표적인 활용 사례이다.

두 번째 기계학습 기반 챗봇(Machine Learning Chatbot)은 대량의 학습 데이터를 이용하여 사용자 질문 패턴과 의도를 학습하고, 통계적 모델 또는 지도학습(Supervised Learning)·비지도학습(Unsupervised Learning) 기반 접근법을 통해 응답을 생성한다. 이는 규칙 기반 시스템보다 유연성이 높고, 질문표현의 다양성이나 문맥 이해 측면에서 앞선 규칙 기반 챗봇에 비해 상대적으로 개선된 성능을 보여준다. 다만 학습 데이터 품

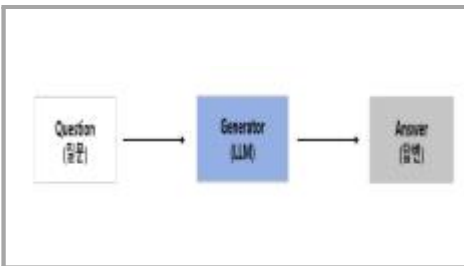
질과 양에 크게 의존하며, 데이터에 의존한 편향이나 오답 가능성이 존재한다. 활용 사례로는 자동화 고객상담, 상품 추천, 사용자 맞춤형 안내 서비스 등이 있다.

최근에는 LLM 등 생성형 AI 기술이 발전함에 따라 세 번째 유형인 생성형 AI 기반 챗봇(Generative AI Chatbot)이 등장하였다. 생성형 AI 챗봇은 사전에 정의된 정답이 아니라 기존 데이터에 없는 새로운 문장을 스스로 생성할 수 있는 강력한 언어능력 및 추론능력을 가진다. 이와 같은 챗봇의 발전 단계는 단순히 기능이 확장되었다는 측면보다 인간처럼 동작하기 위한 발전 과정이라고 할 수 있다.

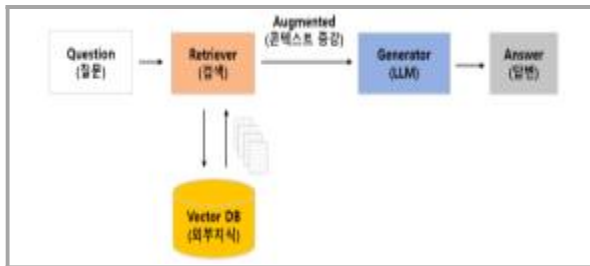
앞서 대부분 공공기관에서 생성형 AI 챗봇으로 기존 챗봇의 개선을 추진하고 있으며 현장조사 질의의 다양성과 응답에 필요한 복합적인 추론 과정 등을 고려할 때 본 연구를 통해 달성하고자 하는 목표로 생성형 AI 챗봇을 정하는 것이 바람직하다.

2. 검색증강생성(RAG)

이와 같이 생성형 AI가 대량의 텍스트를 사전 학습하여 강력한 언어 능력을 갖추고 있음에도, 학습된 데이터에 존재하지 않는 최신 정보나 특정한 분야의 전문적인 지식을 제공하기 어렵다는 한계가 있다. 이 경우 LLM이 부정확한 응답 또는 허위 정보를 답변할 가능성이 있으며 이에 기반한 챗봇도 동일한 위험을 내재하고 있다. 이와 같은 LLM의 구조적 한계를 보완하기 위해 검색증강생성(Retrieval-Augmented-Generation, RAG)기술이 등장하였다.



<그림 4-1> 일반적 LLM 답변 방식



<그림 4-2> RAG-LLM 답변 방식

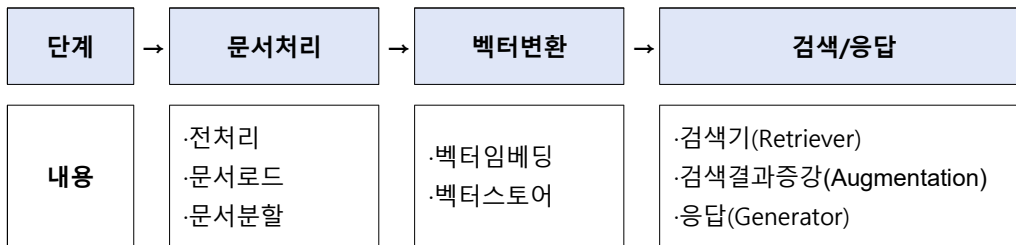
RAG 기술은 LLM의 환각현상 등 부정확한 응답을 해결하기 위해 제안되었다. RAG는 LLM이 응답을 생성하기 전에, 외부 지식저장소(FAQ, 문서, 규정 등)에서 관련 정보를 참조하고, 이를 토대로 답변을 생성한다. 참고해야 하는 정보가 수시로 변경되거나 모델을 통한 내부정보의 유출이나 보안에 대한 우려가 있는 경우 유용한 기술로, 모델을 재학습시키지 않아도 최신 정보와 전문 지식을 실시간으로 반영할 수 있어 시간과 비용을 크게 절약할 수 있다. LLM의 한계를 극복하면서도 그 장점을 살릴 수 있는 가

장 효율적인 접근 방식으로 최근에는 LLM 활용에서 환각현상 제거를 위한 필수 기술로 인식되며 널리 활용되고 있다.

제2절 단계별 기술 검토

본 절에서는 본격적으로 생성형 AI를 활용한 RAG 기반 챗봇을 구축하는 절차를 살펴보겠다. 이를 위해 3장에서 확인한 현장조사 질의응답 자료의 특성을 감안하여 주요 단계별 검토사항과 선택이 필요한 기술요소들을 정리하였다.

<표 4-1> RAG 챗봇 구축 단계



1. 문서 처리

가. 데이터 전처리

첫 단계는 데이터 전처리 과정이다. RAG가 참조할 문서 내에 오류, 중복 자료를 제거하여 분석과 검색에 적합한 형태로 정제하고 구조화해 주는 작업이다. 이후 문서로드 단계에서 불필요한 문자 제거나 공백·줄바꿈 처리 등의 단순하고 반복적 수준의 전처리는 쉽게 수행할 수 있다. 그러나 파일 형식 통합, 오류 내용의 보완, 불필요한 자료에 대한 제거 등 자료 내용에 대한 판단이 필요한 전처리는 문서 로드 단계 이전에 우선적으로 수행되어야 한다. RAG의 정확도는 검색 대상 문서의 품질에 크게 의존되므로 잘 설계된 전처리를 통해 데이터의 품질을 확보하고, 이후 단계에서 챗봇이 정확하고 신뢰성 있는 답변을 제공할 수 있는 기반을 마련해야 한다. 현장조사 질의응답 업무용 챗봇 구축을 위한 참고자료는 조사지침서, 조사사례집과 질의응답(Q&A) 사례이며, 가계동향조사의 경우 추가로 가구관리지침서, 항목분류집 및 품목분류 검색어DB가 있다. 이 중 3장에서 살펴본 바와 같이 응답부재 자료에 대한 데이터화, 검색어 DB에 대한 재정비 등이 전처리 단계에서 수행되어야 하는 작업이다.

나. 문서 로드(document load)

전처리를 통해서 데이터 품질을 확보한 이후에는 RAG가 읽어 들이기 위해 문서 로드 과정을 거쳐야 한다. 이 단계에서는 참조정보의 손실이나 오류 없이 충분한 정보를 활용할 수 있도록 로드하는 방안의 연구가 필요하다. 특히 대다수의 참조자료 형태인 비정형 데이터의 경우 문서 내 표기 방식, 삽입된 표 또는 이미지 등에 따라 고려 사항이 많다. 따라서 전처리 절차와 더불어 문서 로드 단계는 RAG 검색 정확도에 큰 영향을 미칠 수 있다.

문서의 형태별 로드 시 검토사항을 살펴보면 우선, PDF 파일은 내용별로 단락구성이 되어 있어서 문서 로드 시 메타데이터 생성이 가능한 장점이 있으나 표 내부 문자 추출을 손실 없이 처리하는 방안이 필요하다. 엑셀(CSV) 형태의 데이터는 제목 셀에 대한 정보 처리, 행 단위의 레코드 생성과정이 수행되어야 한다. 또한 셀 내에 긴 문장이 들어 있는 경우의 정보 손실 여부를 확인하는 절차가 요구된다. 자료가 충분하지 않은 경우에는 웹상의 정보를 추가로 확인해야 정확한 답변이 가능한 경우도 있으므로 이러한 정보의 로드 방안도 검토 대상이다.

<표 4-2> 참고 문서별 로드 시 검토사항

문서 종류	유형	검토사항
조사지침서/조사사례집 /항목분류집	PDF	표 내부 문자 추출
검색어사전	EXCEL, CSV	제목 셀에 대한 정보 처리
질의응답자료	EXCEL, CSV	질문과 답변을 하나로 구성하여 행 단위의 레코드를 생성 *경우에 따라 웹(web)상 정보 로드 필요

다. 문서 분할(Document Split)

일반적으로 업무에 참고하는 문서는 작게는 수십 페이지 이상으로 이와 같이 분량이 큰 경우 그대로는 RAG 시스템에서 효율적인 검색과 처리가 어렵다. 따라서 로드 과정을 거친 문서는 크기를 줄이고 복합적인 정보를 분리하여 빠른 검색과 처리가 가능하도록 작게 분할하여야 한다. 이와 같은 과정을 문서 변환 또는 분할이라고 칭하며 분할된 작은 단위를 청크(Chunk)라고 한다. 이 분할을 통해 대규모 데이터를 효율적으로 관리하고, 사용자의 질문에 적절한 범위의 정보를 빠르게 검색할 수 있어 챗봇의 응답 속도와 정확도를 향상시킬 수 있다.

문서 분할 단계에서 주의할 점은 잘라내는 방법과 잘라낸 조각의 크기를 어떻게 규

정해서 처리시간과 검색품질 사이의 균형을 얼마나 효율적으로 설계하는가이다. 분할 방법에는 물리적인 글자 수로 분할하는 문자(Character) 단위 분할과, 문맥에 따라 분할하는 의미(Semantic) 단위 분할로 크게 나눌 수 있다. 너무 크게 쪼개면 불필요한 정보가 섞이거나 처리 시간과 비용이 증가할 수 있고, 반대로 너무 작게 쪼개면 맥락 정보가 유실되어 RAG 검색 품질이 저하되게 된다. 여러 가지 시뮬레이션을 통해 적합한 방안을 선택하는 것이 중요하다.

현장조사 질의과정에서 단순 항목코드 자료는 문자 단위 분할이 효율적일 수 있으나 지침서의 복합적인 기준 또는 가우관계 등 여러 정보를 활용해야 하는 측면에서는 의미 단위 분할이 유용하므로 이를 적절히 조정해야 한다.

<표 4-3> 문서 분할 유형

분할도구	구분	내 용
Character TextSplitter	문자 단위	단순하고 빠른 속도, 의미정보 손상 우려 옵션: seperator, chunk_size, chunk-overlap
Recursive Character TextSplitter	문자 단위	문단, 문장 단위를 유지하면서 단계적 분할, 의미 보존 옵션: seperator, chunk_size, chunk-overlap
Semantic Chunker	의미 단위	LLM 또는 Embedding모델로 유사도 분석, 의미 단위 분할, 연산량 많아 속도가 느림

2. 벡터 변환

가. 벡터 임베딩(Embedding)

청크 단위로 분할된 문서는 텍스트를 수치 벡터 형태로 변환하는 임베딩 과정이 필요하다. 벡터 임베딩이란 텍스트의 의미를 벡터로 수치화하여 컴퓨터가 이해하고 유사도 검색을 수행할 수 있도록 변환하는 과정으로 RAG 챗봇이 사용자의 질문과 참조 데이터 간의 의미적 유사성을 기반으로 적절한 문서를 선택하게 하는 핵심 절차이다.

텍스트를 수치로 바꾸기 위해서는 대량의 텍스트로 사전 학습된 LLM을 임베딩 모델로 사용한다. 일반적인 LLM은 사용자 질문 이해와 답변 생성이 목적이지만 임베딩에 사용되는 LLM은 텍스트를 수치로 변환하여 의미적 유사도를 계산하는 것에 중점을 두어 RAG 시스템의 검색에 주로 사용된다는 차이점이 있다.

임베딩 모델에서는 차원(dimension)이라는 개념이 중요한 요소로 임베딩 모델의 차

원은 텍스트를 수치 벡터로 변환할 때 변환된 벡터의 차원을 의미한다. 차원이 크면 텍스트의 의미를 세밀하게 표현할 수 있지만 연산량과 메모리 요구량이 커지고, 반대로 차원이 낮으면 속도와 저장공간에 효율성이 있지만 문장의 미세한 의미를 구분하기 어려운 단점이 있다.

<표 4-4> 다양한 임베딩 모델

구분	임베딩 모델명	특징	차원
공개 모델	multilingual-e5-small-ko	다국어 지원, 한국어에 특화. 작은 크기로 경량화 및 빠른 처리 가능	384
	paraphrase-multilingual-MiniLM-L12-v2	다국어 지원에 최적화. MiniLM 아키텍처로 경량화 및 효율성 높음	384
	ko-sroberta-multitask	한국어에 특화. 다양한 작업(분류, 감정 분석 등)에 적용 가능	768
상용 모델	text-embedding-3-large	대규모 텍스트 임베딩 모델. 높은 정확성과 풍부한 표현력	3,072
	text-embedding-004	최신 텍스트 임베딩 모델. 다양한 언어 및 컨텍스트에 대한 우수한 성능	768
	embed-multilingual-v3.0	다국어 지원. 다양한 언어 간의 유사도 계산에 적합	1,024

<표 4-4>와 같이 다양한 임베딩 모델이 있으며 모델 간 차원 차이가 존재한다. 공개모델은 속도, 비용, 한국어 지원 관점에서 유리한 경우가 많지만, 매우 긴 문서나 미세한 의미 구분이 필요한 고정밀 RAG 환경에서는 차원이 큰 상용 모델이 품질 면에서 더 안정적인 부분이 있다. 복잡하고 복합된 의미를 가지는 질문에 잘 응답하기 위해서는 임베딩 모델을 선택할 때 이러한 차원, 임베딩 품질, 지원 언어 등을 복합적으로 검토해야 한다.

나. 벡터 스토어(Vector Store)

임베딩 과정을 거친 문서는 벡터 형태로 벡터 스토어에 저장하여 관리하게 된다. 벡터 스토어는 고차원 벡터의 저장과 검색이 효율적으로 수행될 수 있도록 설계된 시스템으로, 사용자의 질문이 입력되면 질문을 벡터로 변환하여 관련성이 높은 문서를 빠르게 찾아 RAG 챗봇이 답변을 생성할 수 있게 지원하므로 단순한 데이터의 저장소가 아니라 검색 성능을 좌우하는 핵심 인프라이다.

비정형 데이터는 길이나 구조가 일정하지 않고 단어들의 배열만으로 의미나 특성의 유사성 확인이 어렵다. 벡터 스토어는 텍스트 데이터를 의미 기반으로 다룰 수 있도록 저장하는 특수한 방식으로 비정형 데이터를 저장하고 고차원 벡터 데이터를 효율적으로 관리하는 기능을 한다. 또 유사도 계산⁸⁾을 통해 복잡한 텍스트 간의 의미 기반 검색(semantic search)이 가능하다. 이를 통해 챗봇은 대규모 참조 데이터를 기반으로 정확한 정보를 제공하며 답변할 수 있다.

<표 4-5> 유형별 벡터DB

구분	특징	대표DB	활용	오픈소스
클라우드 관리형	클라우드 기반, 실시간 검색	Pinecone	기업용 고품질 서비스 환경	×
설치형 오픈소스	비용 절감, 유연성	Weaviate	보안, 규제 환경 (서버 직접 운영)	○
대규모 분산형	대규모 벡터, 고성능 분산시스템	Milvus	대규모 데이터셋	○
내장형	가벼움, 로컬 사용	ChromaDB	소규모 프로젝트, RAG 개발	○
인메모리형	메모리 기반	Redis VSS	실시간 응답	○

3. 검색 및 응답

가. 검색기(Retriever)

지금까지 RAG에 활용할 문서를 처리하고 시스템에서 검색 가능한 형태로 저장하는 단계를 거쳤다. 이제부터는 챗봇에서 사용자가 질문을 입력하면, 질문을 기반으로 저장된 문서에서 유사한 자료를 찾아서 질문과 근거 문서를 잘 연결해야 한다. 이 구조를 잘 구축하는 과정이 검색기(Retriever) 구성 단계이다. 이 단계에서는 질문과 문서를 연결하고 문서를 잘 검색하는 구조를 구축하는 것이 핵심이다. 사용자가 질문하는 단어를 확장하여 참조 문서의 특정 단어와 연결해야 사람이 답변하는 것처럼 작동하게 만들 수 있으므로 사용자의 질문을 다양한 방식으로 해석하여 답변하도록

8) 대표적으로 코사인 유사도(cosine similarity) 사용. 두 벡터 간 방향 각도에 대한 코사인함수를 계산한 값으로 -1에서 1의 값을 가짐. 각도가 0(방향이 유사)이면 코사인 값은 1에 가까워지고, 두 데이터의 의미가 유사한 것으로 판단. 텍스트 이미지 등 고차원 데이터에 적합

구축해야 한다.

검색 방식은 기본유사도(similarity) 방식, 최대한계관련성(Max Marginal Relevance, 이하 MMR) 방식, 하이브리드(Hybrid) 방식으로 구분할 수 있다. 기본유사도 방식은 사용자의 질문과 청크 간 의미적 유사성을 수치로 평가하는 방식으로 기본적인 검색방식이다. 구현이 쉽고 속도가 빨라 매뉴얼 기반 챗봇 시스템에 적합하다. MMR은 다양성을 고려하고 중복을 최소화하는 방식으로 기본유사도(일치성)와 문서집합들 간 유사도(다양성) 간 가중치(λ)를 조정함으로써 산출된다.⁹⁾ 문서에 중복이 많아서 중복검색을 줄이거나 요약형 검색이 필요할 때 유용한 방식이다. 하이브리드 방식은 의미 기반 벡터 검색과 키워드 검색을 결합한 방식이다. 언어적 의미와 단어 매칭을 함께 반영하므로 신뢰성 높은 검색이 가능하여 검색 품질을 충분히 확보할 수 있다.

<표 4-6> Retriever 검색 방식 구분

구분	개념	장/단점	활용
기본유사도 (similarity)	질문과 청크 간 의미적 유사도	(장) 구현 용이, 신속 (단) 반복적 검색	매뉴얼 기반 챗봇, 정답이 명확한 경우
최대한계 관련성 (MMR)	문서의 유사도와 다양성 동시 고려	(장) 검색 다양성 확보 (단) 계산량 증가, λ 에 의존	높은 문서 중복도, 다양한 답변 필요
하이브리드 (Hybrid)	의미 검색과 키워드 검색 결합	(장) 검색품질 우수 (단) 구현 난이도 증가	질문에 키워드와 의미가 혼합되는 경우

질의응답 유형별로 적용해 보면 단순 항목 분류문의는 답변이 명확하므로 기본유사도 검색이 충분히 가능해 보이나, 응답 난이도가 높은 질문도 존재하므로 키워드를 복합적으로 고려하는 하이브리드 검색 방안이 효율적으로 판단된다. 또한, 구체적인 상황에 따라 품목분류 등 답변이 달라지는 경우에는 다양한 응답결과를 제시할수 있는 MMR이 적합하다.

나. 증강(Augmentation)

다음은 RAG 단계 중 증강(Augmentation) 과정으로 검색으로 찾아온 문서를 LLM이 효과적으로 활용할 수 있는 정보로 확장·강화하여 구성하는 과정이다. 사용자의 질문과 유사하다고 판단되는 청크들 중에 중복과 불필요한 내용을 제거하고 하나의 텍스트

9) λ 는 0-1값을 가지며 $MMR = \lambda \times (\text{질문과 유사도}) - (1-\lambda) \times (\text{문서집합 간 유사도})$ 로 계산

로 묶어서 LLM이 이해하기 좋은 구조로 가공하는 처리를 수행한다. 검색된 문서 간의 통합 과정과 이후 단계에서 응답할 LLM이 질문에 잘 답변할 수 있도록 응답 형식을 지정하는 프롬프트 관리를 통해서 응답 품질을 향상시키는 과정도 수행한다.

프롬프트 관리란, 검색된 문서를 LLM이 해석하기 쉽도록 문서 정보를 체계적으로 배치하고, 질문과 통합하며, 모델에게 수행할 역할과 제약 조건을 명확히 지시하는 과정을 의미한다. 시스템(System), 사용자(Human), 인공지능(AI)으로 구분하여 각각의 역할을 지정하는 역할로 AI 메시지는 AI가 이전에 응답한 내용을 메시지로 기억하여 대화의 맥락을 유지하며 응답하는 기능도 부여할 수 있다.

실제로는 시스템 메시지를 통해 질문 내용을 정리 언급하여 질의 내용을 명확히 하고, 문서 출처를 표시해서 사용자가 직접 근거를 확인할 수 있는 방안도 필요하다. 또 AI 메시지로 복잡한 질문 또는 정보 재확인이 필요한 경우를 대비하여 이전 대화를 기억하도록 구축하는 것이 유용하다.

다. 응답 생성(Generator)

증강 단계에서 정리된 문서 내용과 프롬프트가 준비되면, LLM은 이를 기반으로 최종 답변을 생성(Generate)한다. 이 단계에서는 사용자의 질문과 통합된 문서 정보가 하나의 프롬프트로 전달되며, LLM은 제공된 자료를 분석하여 핵심 내용을 추출하고 적절한 순서로 재구성한다.

응답 생성 단계는 이 과정의 핵심 역할을 하는 LLM에 대해서 살펴보는 것으로 대신한다. LLM은 RAG 시스템에서 증강된 정보를 사용자 질문에 맞춰 의미 있게 변환하는 역할을 담당한다. 앞선 단계의 문서 검색과 구조화가 제대로 이루어지더라도 LLM의 이해·추론 능력이 부족하면 검색된 문서 내용을 정확하게 이해하는 능력이 부족하여 최종 답변의 신뢰성과 정확성이 저하될 수 있다. 따라서 LLM은 RAG 시스템의 핵심으로서, RAG 챗봇 구축 시에는 어떤 LLM을 사용할지 선택하는 작업이 우선되어야 한다. LLM 모델은 유형별 특징점이 매우 다르고 여러 가지 고려 요소가 있지만 크게 상용모델과 공개모델 두 유형을 살펴보기로 한다.

상용모델은 개발사가 사전학습과 지속적 업데이트를 수행하며, 사용자는 이를 플랫폼 형태로 제공받는다. 최신 AI 기술이 적용되어 있으므로 성능에서 비교적 우위에 있고 공개모델 대비 활용 관점에서 접근성이 높다. 그러나, 서비스 환경이 대부분 클라우드 기반이므로 내부자료 유출 등 보안상 문제에 대한 우려도 존재한다.

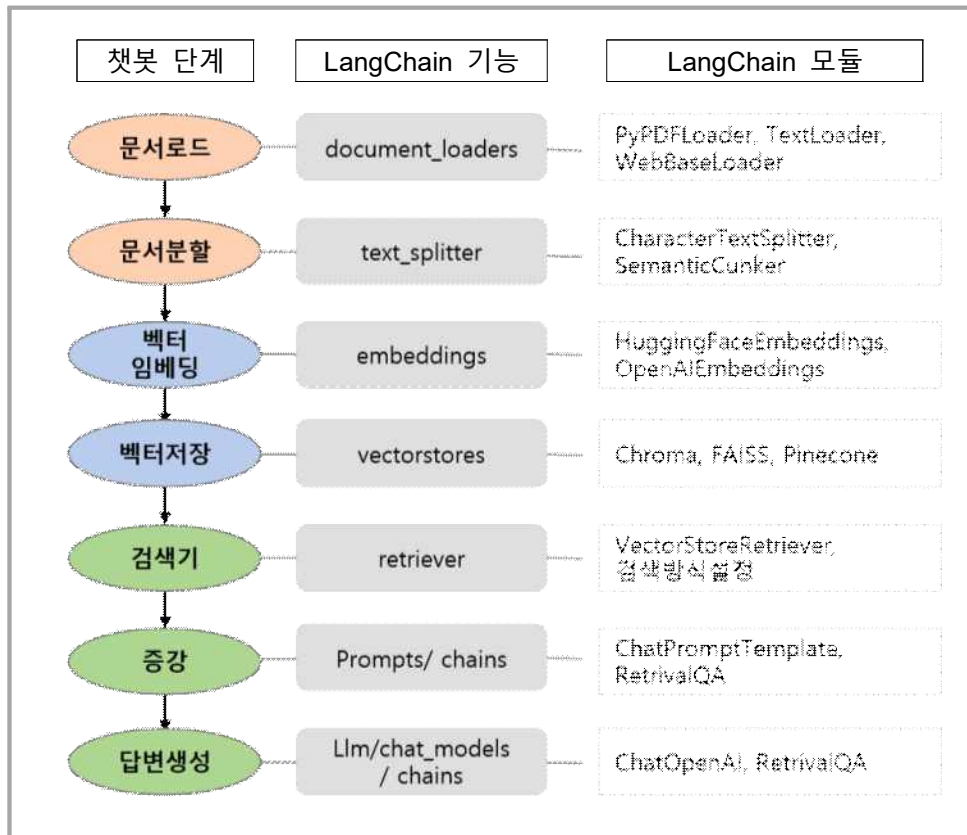
공개모델은 오픈소스 형태로 제공되어 기관 자체에서 모델을 로컬 또는 폐쇄망 환경에서 운영할 수 있다는 점에서 보안성과 통제력이 뛰어나다. 최근에는 공개모델과 성능 차이가 크지 않을 정도로 기술 수준이 발전했으며 무료로 사용이 가능하고 내부자료 활용체계 구축 등 튜닝의 확장성이 높아서 기관 특화형 챗봇 구축이 가능하고 활

용 범위가 넓다. 따라서 모델 선택 시에는 성능뿐만 아니라, 서비스 환경·보안 요건·운영 비용·유지관리 등을 종합적으로 판단하여야 한다.

4. LangChain 프레임워크

AI를 활용한 각종 어플리케이션 개발 과정에서 널리 활용되는 개발도구로 LangChain이 있다. LangChain은 다양한 상용·공개 AI 모델과 개발도구를 연결하고 사용자 정의 모듈을 추가할 수 있도록 지원한다. 또한 복잡한 프로세스를 단순화하여 설계하도록 도와주고 각 단계의 구체적인 옵션 제공과 전체 단계의 유기적인 결합을 통해 비교적 쉽게 어플리케이션 구축이 가능하다.

또한 공식문서, GitHub, 커뮤니티 등 다양한 채널로 활용 사례와 오류수정 방법을 학습할 수 있으므로 실제 AI 챗봇 구축을 위해서 필수적인 요소라 할 수 있다. 앞서 살펴본 생성형 AI RAG 챗봇 구축 절차를 LangChain에서 지원하는 모듈을 활용하여 각 단계별로 도식화한 모습이다.



<그림 4-3> RAG 단계별 LangChain 모듈

제 5 장

시사점 및 향후 과제

제1절 요약 및 시사점

본 연구에서는 국내 행정 챗봇 사례로 ‘구뎀’과 ‘서울톡’, ‘AI 노동법 상담 서비스’를 살펴보았다. 그중 ‘구뎀’과 ‘서울톡’은 일상적으로 사용하는 메신저를 활용하여 서비스 접근성을 향상시켰으며 ‘서울톡’은 위치기반 서비스와 연계하여 현장민원 접수가 가능한 점이 특징이다. ‘AI 노동법 상담 서비스’는 생성형 AI 기술 활용한 챗봇으로 ‘구뎀’, ‘서울톡’ 모두 생성형 AI 기반 고도화 작업을 추진 중이다. 생성형 AI 기반 챗봇이 최신기술 활용과 자원 활용 측면에서 효율적이라는 점을 확인하였고 조사관리, 조사원 안전 등 다양한 현장조사 분야 확장성을 위해 위치 기반 서비스 도입도 유용하게 검토해 볼 수 있다.

선행 연구 분석을 통해서 챗봇 기술의 발전 단계를 가늠해 보고 향후 연구의 방향까지 미리 예측해 보았다. 챗봇 도입 초기 정의와 개념 규정을 거쳐 공공기관 챗봇의 구축 검토사항과 RAG 기술 연구를 넘어서 현재는 응답 정확성 향상을 위한 다양한 연구가 수행되고 있다. 국가데이터처 내부에서도 챗봇 기능의 효용성과 필요성에 대한 방증으로 현장조사 질의응답 분야 챗봇에 대한 시도가 꾸준함을 살펴보았다.

주요 가구 조사인 가계동향조사와 가계금융복지조사의 실제 현장조사 질의응답 데이터를 통해 챗봇이 수행해야 할 업무에 대한 내용 분석도 수행하였다. 응답 난이도가 규칙 기반의 쉬운 수준부터 복잡한 가구관계 및 다양한 상황을 고려해야 하는 복합적인 경우까지 매우 폭넓다는 것을 분석을 통해 확인하였다. 실제 응답자료 분석을 통해서 참조자료의 보완과 정비가 필요하며 사용자에게 재질의를 하거나 상황별 복수 응답으로 사용자에게 선택지를 제공하는 방안도 챗봇 구축 시 검토해야 한다는 사실을 도출하였다. 향후 실제 챗봇 구축 시에는 본 분석에 따라 질의와 응답 난이도별로 기술을 구분하여 세분화할 필요가 있다고 판단하였다.

마지막으로 생성형 AI 챗봇의 환각문제 해결을 위한 RAG 기술과 이를 적용한 챗봇 구축기술을 단계별로 서술하였다. 우선 참조문서의 오류와 중복을 제거하는 전처리를 거쳐 RAG가 참고할 수 있도록 문서 로드를 수행하고, 문서를 작은 청크 단위로 분

할한다. 분할된 문서는 벡터 변환 단계인 임베딩 절차를 거쳐 효율적 의미 검색을 위한 벡터 스토어에 저장된다. 이후 사용자의 질문이 입력되면 질문과 참조 정보 간 검색을 통해 검색기가 문서를 검색하고 증강단계를 통해 검색된 문서정보를 재구성한 뒤 LLM에 전달하면 LLM이 이를 적절한 응답형태로 변환하여 출력된다. 개별 단계는 LangChain에서 지원하는 모듈을 활용하여 설계하고 전체적인 체인을 구축하는 것이 효과적이다. 각 단계별 선택옵션이 매우 다양하고 각 단계가 응답성능에 직접적인 영향을 미치므로 개별 기술 및 데이터와 가용 인프라 간 상호 최적의 절충점을 찾아 구축하는 것이 관건이다.

제2절 향후 연구방향

본 연구를 기반으로 향후 추진될 필요가 있다고 판단되는 추가 연구 방향을 3가지 정도로 제안할 수 있다.

첫 번째는 단계적 실증연구이다. 본 연구에서 살펴본 RAG 기반 생성형 AI 챗봇 구축 기초 기술을 활용하여 실제 데이터를 적용한 현장조사 질의응답 챗봇을 구축하기 위해 다양한 기술 조합을 실험하고 최적의 방안을 찾기 위한 시도가 필요하다. 실증연구에서 또 하나 검토해야 할 점은 단계적 구축이 필요하다는 것이다. 본 기초연구에 언급된 기술만으로 실제 현장에서 발생하는 모든 질문을 효과적으로 응답하기에는 무리가 있다. 살펴본 바와 같이 사용자의 질문과 요구되는 응답 수준이 단순한 것부터 복잡한 것까지 다양하기 때문이다. 최근 생성형 AI RAG 챗봇의 고도화를 위한 기술에 대해 다양한 연구와 실험이 계속되고 있으므로 향후 실증 연구의 방향은 단순질의-단순응답을 해결하기 위한 기본 단계 수준부터 복합질의-복합응답의 높은 단계 수준까지 기술 범위와 깊이를 확장해 가는 단계적인 실증연구가 필요하다. 실제 실무업무의 적용과 검토를 위해서도 이러한 방안이 효율적일 것으로 판단된다.

두 번째는 구축된 챗봇의 성능 평가에 대한 연구가 필요하다. 공공기관에서 AI 기술 도입 여부를 판단할 수 있는 요인은 기술의 안정성과 정확성, 효과성이다. 이와 같은 측면의 검토와 평가는 챗봇에 한정되지 않고 AI 기술을 통한 업무영역 확장에 근거 마련을 위해 꼭 수행될 필요가 있는 분야이다. 향후 실증연구 단계에서 우선적으로 비교적 간단한 설문 작성과 결과 분석을 통한 사용자 만족도 평가를 검토해 볼 수 있다. 챗봇 응답 성능에 대한 주관적인 감정을 수량화하는 방식으로 기술의 안정성과 정확성을 수치화하기 어렵지만 효과성에 대한 판단 근거는 충분하다. 이 방식은 일정 기간 챗봇을 시범 운영하고 사용자의 평가를 근거로 실무부서에서 도입 여부를 결정하는 참고자료로 활용될 수 있다.

마지막 세 번째는 지속적인 연구이다. AI에 대한 접근 장벽으로 매우 빠른 발전 속도를 쫓는 경우가 많다. 자고 일어나면 새로운 기술과 연구가 탄생하여 그 속도를 따라가기가 숨 가쁜 상황이며 단편적인 연구로 모든 최신 기술을 다 적용하기 어려운 한계가 있다. 지속적인 연구를 통해 최신 기술을 부지런히 학습하여 적용하고 개선하는 꾸준한 과정을 통해 기술의 안정성과 신뢰성을 확보해야 한다. 본 연구에서는 자료의 한계로 내부 조사담당자와 본부 담당직원 간의 업무로 한정하였지만, 최종적인 챗봇 서비스는 조사응답자 그리고 통계데이터 이용자인 국민 대상 서비스로 구축하여 현장 조사 전반에 대한 질의를 한 번에 처리하는 구조로 그려볼 수 있다. 이를 위해서는 AI 기술 수준에 대해 국민들의 높아진 눈높이를 만족시킬 수 있도록 오류를 최소화하고 편의를 향상시킬 수 있는 최신 기술에 대한 연구와 적용이 지속적으로 필요하다.

참고문헌

- 고용노동부. (2025.3.4.). “노동약자 보호를 위한 AI 혁신에 고용노동부와 공인노무사회가 힘을 모은다.” 보도자료.
- 고용노동부. (2025.8.31.), “고용노동부, 인공지능(AI)으로 국민 일상 바꾼다.” 보도자료.
- 국가데이터처. (2025). “경제통계 AI도입: RAG챗봇과 AI분류시스템”. **2025 국가데이터처 혁신 우수사례 경진대회 자료집**.
- 김성근, 신민철, 강주영. (2018). “챗봇 기술 소개 및 사례 분석”, **한국통신학회**, 21-28.
- 김정민. (2024). **국가통계 분야 인공지능 활용 체계 연구**. 국가통계연구원.
- 김혜윤, 노연수, 박종혁, 박민정, 양병욱, 정운서. (2025). “임베딩 모델 및 Advanced RAG 기법을 활용한 한국어 텍스트 분류”. **한국통계학회. 응용통계연구**, 571-588.
- 김휘균, 이지은, 박상현. (2024). “인공지능 챗봇을 위한 검색증강생성 및 벡터 데이터베이스 최적화 연구 동향”. **정보과학회지**, 8-15.
- 니시미 마사히로, 요시다 신고, 오시마 유키. (2025). **랭제인과 랭그래프로 구현하는 RAG.AI 에이전트 실전 입문**. 위키북스.
- 박동아. (2017). “인공지능 기반 대화형 공공 행정 챗봇 서비스에 관한 연구”, **멀티미디어학회**, 1347-1356.
- 서교리, 강동석, 박선주, 어재경, 김재민, 김효정. (2017). “인공지능 기반의 공공 지능형 가상 비서 서비스 발전 모델 연구”. **한국통신학회**, 890-891.
- 서울특별시. (2025.8.8.). “서울시, 행정 분야 자체 LLM 도입... '글로벌 AI 행정도시' 시동 건다.”. 보도자료.
- 서울특별시. (2024.7.10.). “AI시대 성큼, '120다산콜'도 똑똑하게 진화 중! 보이는 ARS·챗봇 편리”. 서울정보소통광장.
- 서울특별시. (2024.2.13.). “내 손안에 AI비서 '서울톡'으로 복지혜택부터 민원신청까지 똑딱!”. 서울시 홈페이지.
- 유희경, 문남미. (2025). “ToR-RAG ToR기반 RAG를 활용한 복합 질문 처리방안 연구”. **한국컴퓨터정보학회논문지**, 23-31.
- 윤상오. (2018). “인공지능 기반 공공서비스의 주요 쟁점에 관한 연구: 챗봇(ChatBot)서비스를 중심으로”. **한국공공관리학보**, 83-104.
- 윤성재. (2024). **RAG시스템 구축을 위한 랭제인 실전가이드**. 루비페이퍼.
- 이혜림, 김성현, 엄석진. (2025). “AI 챗봇 도입 현황과 확산 요인 분석: 서울특별시 25개 자치구를 중심으로”. **한국사회와 행정연구**, 73-106.
- 장준희. (2016). **인공지능 기반 챗봇 서비스의 등장과 발전 동향**. 한국정보화진흥원 “ICT융합 I&T”
- 정천수. (2024). “LLM애플리케이션 아키텍처를 활용한 생성형 AI 서비스 구현 : RAG 모델과 LangChain 프레임워크 기반”. **한국지능정보시스템학회**, 29(4), 129-164.

- 조찬영, 강성준, 정현준. (2023). “RAG 기반 랭체인을 이용한 생성형 AI 챗봇 구현”. **한국정보기술학회**, 460-463.
- 최광미. (2024). “검색증강생성(RAG) 기반 기업 맞춤형 챗봇(Chatbot) 시스템 구축 및 활용”. **한국전자통신학회**, 19(4), 1281-1292.
- 통계교육원(현 국가데이터인재개발원). (2024). “생성형 AI를 활용한 가계동향조사 분류코드 챗봇”. 데이터사이언티스트 전문가 양성 과정.
- 통계인재개발원(현 국가데이터인재개발원). (2025). “가계동향조사 담당자를 위한 AI 수행비서 제작”. 데이터사이언티스트 전문가 양성 과정.
- 통계인재개발원. (2025). “OCR·RAG 기반 가계동향조사 항목 자동분류 챗봇”. 데이터사이언티스트 전문가 양성 과정.
- 통계청(현 국가데이터처). (2024). **가계금융복지조사 이용자용 통계정보보고서**.
- 통계청. (2025). **가계동향조사 이용자용 통계정보보고서**.
- 통계청. (2025). “2025 인구주택총조사 종합시행계획”.
- 행정안전부. (2025.10.27.). “말 한마디로 이용하는 공공서비스, 행안부-네이버·카카오가 만든다”. 보도자료.
- Gartner Press Releases. (2021.10.5.). “Gartner Says Government Organizations Are Increasing Investment in AI, But Their Workforce Remains Apprehensive”.
- Gartner Press Releases. (2025.9.9.). “Gartner Reveals Top Technologies Shaping Government AI Adoption”.

Abstract**A Research on the Use of Generative AI
in Q&A Tasks of On-site Surveys****Youngmin Ahn**

Generative AI-based chatbots, which automate repetitive question and answer tasks, play a significant role in providing faster and more convenient customer services. Among AI-driven technologies, chatbots are widely used in the public sector, and rapidly expanding in the world. Korean administration also endeavors to promote administrative innovation through the deployment of generative AI-based chatbots.

From this perspective, this study examined practical applicability of generative AI-based chatbots. To this end, first of all, the development directions of chatbot technologies were assessed through case studies and prior research. In addition, through a systematic analysis of Q&A data collected during on-site survey processes, this study identified difficulty levels and derived considerations to be considered when designing chatbots. Finally, this study summarized the process of developing a RAG-based retrieval, augmentation and response system as a technology for the efficient use of generative AI-based chatbots.

To further advance these findings, step-by-step empirical studies and evaluations on chatbots, including assessment of user satisfaction with their performance, should be conducted. Above all, generative AI-based chatbots should be continuously refined to meet the needs of users through continued application and improvement in latest AI technologies.

Key words. Generative Artificial Intelligence (AI), on-site survey, chatbot, RAG

연구진

○ 안영민(국가데이터처 국가통계연구원 통계방법연구실 주무관)

* 연구진의 소속 및 직급은 연구과제 완료 시 기준임을 알려드립니다.

연구보고서 2025-20

현장조사 질의응답 분야의 생성형 AI 활용 기초연구

인 쇄 2026년 3월

발 행 2026년 3월

발 행 인 김 진

발 행 처 국가데이터처 국가데이터연구원
35220 대전광역시 서구 한밭대로 713
TEL.(042)366-7100 Fax.(042)366-7123

홈페이지 <https://mods.go.kr/dsri/>

ISSN(Online) 2733-4120





국가데이터처
국가데이터연구원

