통계청 보도자료 다시 대한민국! 새로운 국민의 나라

보도시점 2024. 1. 19.(금) 12:00 배포 2024. 1. 19.(금) 08:30

[2024년 통계청 민생정책 100% 활용하기]

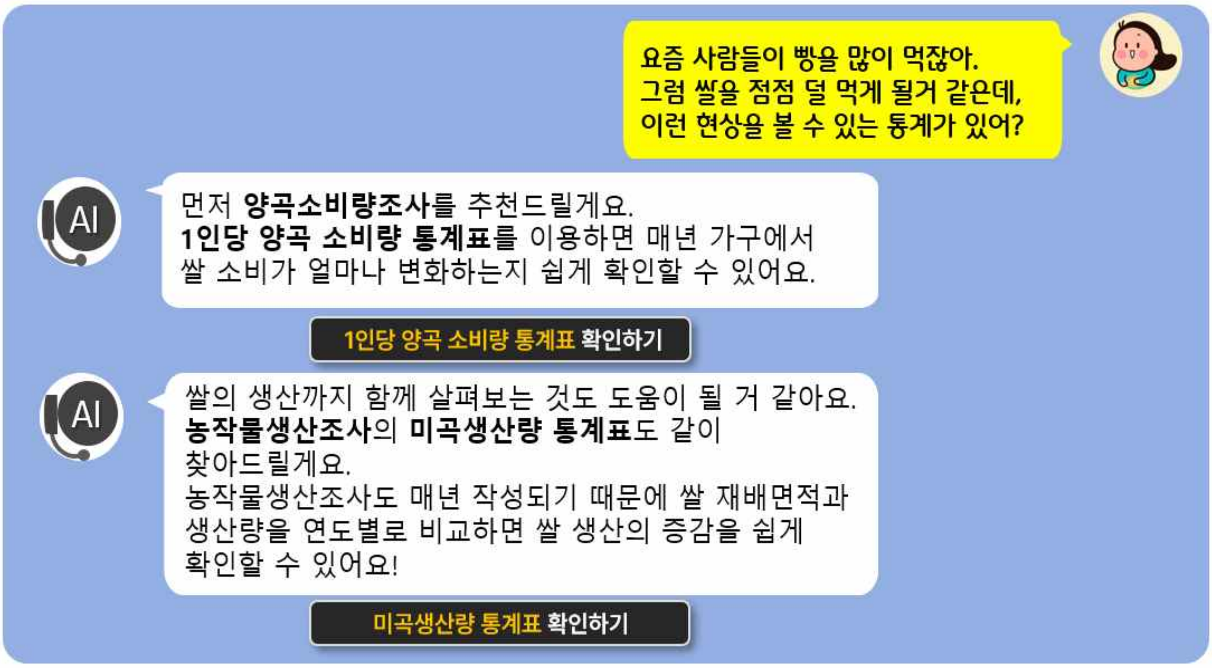
어려운 통계? 편리한 활용!
국가통계가 초거대 AI를 품다.

- 일상언어로 묻고 답하는 초거대 AI 통계 챗봇 서비스 제공 -

통계청(청장 이형일)은 국민들이 일상언어로 손쉽게 통계정보를 검색하고 어려운 통계용어와 개념을 알기 쉽게 풀어서 설명하는 **초거대 AI 기반 통계 챗봇 서비스**를 선보일 예정이다.

그동안 국가통계포털(KOSIS)에 챗봇 서비스가 제공되고 있으나, 정확한 통계용어를 사용하지 않는 경우 제대로 된 답변을 받기 어려웠다. 그러나 초거대 AI가 국가통계와 접목되면 자연어 이해 능력이 대폭 향상되어 **일상언어로 질문**하는 통계도 **맥락과 의도를 파악하여 이용자가 궁금한 사항을 척척 대답**할 수 있게 된다.

초거대 AI 통계 챗봇 예시



요즘 사람들이 빵을 많이 먹잖아.
그럼 쌀을 점점 덜 먹게 될거 같은데,
이런 현상을 볼 수 있는 통계가 있어?

먼저 **양곡소비량조사**를 추천드릴게요.
1인당 양곡 소비량 통계표를 이용하면 매년 가구에서 쌀 소비가 얼마나 변화하는지 쉽게 확인할 수 있어요.

쌀의 생산까지 함께 살펴보는 것도 도움이 될 거 같아요.
농작물생산조사의 **미곡생산량 통계표**도 같이 찾아드릴게요.
농작물생산조사도 매년 작성되기 때문에 쌀 재배면적과 생산량을 연도별로 비교하면 쌀 생산의 증감을 쉽게 확인할 수 있어요!

초거대 AI를 구성하는 핵심요소는 **대규모언어모델(LLM)**로, 수천억개의 매개변수를 학습했기 때문에 자연어의 이해능력이 뛰어날 뿐만 아니라 학습 데이터로부터 생성한 정보를 **쉬운 단어로 풀어서 설명하는 능력**도 탁월하다. **초거대 AI 기반 통계 챗봇**은 이 언어모델을 활용하므로 이용자가 통계에 대한 배경지식이 풍부하지 않더라도 경제·사회 분야 등 **우리 주변의 여러 현상에 대한 궁금점을 쉽게 파악**할 수 있는 채널을 제공한다.

한편 초거대 AI가 정확한 답을 찾지 못할 경우, 방대한 학습 내용 중에 비슷한 부분만 짜깁기하여 잘못된 정보를 주는 이른바 **환각 현상**에 대한 우려가 높다. 이러한 환각 현상을 예방하기 위해 **최신기법[*]이 적용된 팩트체크 로직을 구현**하여 정확한 정보 제공을 보장할 예정이다.

[*] **검색 증강 생성**(RAG: Retrieval Augmented Generation) : 언어를 이해하고 표현하는 능력은 민간의 검증된 초거대 AI 언어모델을 활용하되, 실제로 정보를 검색하고 답변을 생성하는 부분은 전문적인 통계문서를 고차원적인 벡터로 저장한 통계학습데이터베이스에서만 찾아오는 방식

통계청은 **2024년 하반기 대국민 서비스를 목표**로 주요 통계에 대해 **통계 추천과 전문적 질의응답**이 가능한 초거대 AI 기반 통계 챗봇 서비스를 구축한다. 이번 사업은 본격적으로 초거대 AI 모델을 통계행정의 각 영역에 도입하기 전 **시범사업의 성격**을 가지며, 검증된 최신기술의 반영과 신속한 추진을 위해 **민간 클라우드의 AI 디지털서비스를 활용**하여 구축할 예정이다.

이형일 통계청장은 "구슬은 꿰어야 보배가 되는 것처럼, 통계가 국민의 일상생활에서 빛을 발하기 위해 앞으로는 **초거대 AI가 통계라는 구슬을 꿰게 될 것**"이라며 "앞으로도 **국민의 눈높이에 맞춘 통계서비스를 제공**하기 위해 노력해 나가겠다"라고 밝혔다.

담당 부서	통계서비스정책관 통계정보플랫폼과	책임자	과 장	채관병 (042-481-3604)
		담당자	사무관	신우람 (042-481-3694)

OPEN 출처표시 공공누리 공공저작물 자유이용허락

참고1 초거대 AI 통계 챗봇 서비스 활용 사례

# 2023년(초거대 AI 통계 챗봇 도입 전) 대학원생 A씨의 연구일지

식품공학을 전공하는 대학원생 A씨는 빵의 소비 증가와 쌀 소비 변화의 관계가 궁금해졌다. 관련 통계를 찾기 위해 통계청의 **국가통계포털**에 접속해서 쌀을 검색했더니 쌀과 관련된 통계표가 무려 252개나 검색되었다. 쌀 소비량의 증감이 궁금했을 뿐인데, 쌀의 생산비나 물가지수, 관세 등 쌀과 관련된 통계가 무더기로 검색되다 보니 어느 통계를 참고해야 할지 난감했다. 국가통계포털의 Q&A 게시판에 질문글을 남겨보라는 동료의 조언에 따라 질문글을 등록하고 2일이 지난 후, 양곡소비량조사와 농작물생산조사를 추천한다는 통계청의 답글을 읽고 A씨는 비로소 분석에 들어갈 수 있었다.

# 2025년(초거대 AI 통계 챗봇 도입 후) 대학원생 B씨의 연구일지

대학원생 B씨는 기사를 통해 국가통계포털의 통계 챗봇에 초거대 AI가 도입되었다는 소식을 접했다. B씨는 국가통계포털에 접속해 질문창에 다음과 같은 문장을 입력했다.

“나는 요즘 사람들이 빵을 많이 먹게 되면서 자연스럽게 쌀밥은 덜 먹을거라고 생각해. 쌀 소비의 변화를 볼 수 있는 통계가 있을까?”

통계 챗봇은 다음과 같은 답을 내놨다.

“먼저 양곡소비량조사를 추천드릴게요. 1인당 양곡 소비량 통계표를 이용하면 매년 가구에서 쌀 소비가 얼마나 변화하는지 쉽게 확인할 수 있어요.”

B씨가 우선 궁금해했던 쌀 소비 통계를 추천한 챗봇은 답변을 이어나간다.

“쌀의 생산까지 함께 살펴보는 것도 도움이 될 거 같아서 농작물생산조사의 미곡생산량통계표도 함께 찾아드릴게요. 매년 작성되는 통계라서 쌀의 재배면적과 생산량을 연도별로 비교하면 쌀 생산의 증감까지 쉽게 확인할 수 있어요.”

B씨는 Q&A 게시판에 글을 남기거나 통계청 관련부서에 전화를 돌리지 않고도 통계 챗봇을 통해 연구에 필요한 통계자료를 손쉽게 얻을 수 있었다.