

# LLM 기반 프라이빗 센싱 데이터 모니터링을 위한 엣지 컴퓨팅용 음성 챗봇 구현

김영준, 오명훈  
호남대학교 컴퓨터공학과

20210047@my.honam.ac.kr, mhoh@honam.ac.kr

## Implementation of an LLM-based voice chatbot for edge computing to monitor private sensing data

Young-Jun Kim, Myeong-Hoon Oh  
Dept. of Computer Engineering, Honam University

### 요 약

본 논문에서는 실시간 환경 센싱 데이터와 LLM 기반 챗봇을 결합한 엣지 컴퓨팅용 음성 기반 대화 시스템을 구현하였다. 사용자가 날씨와 같은 환경 데이터를 요구하는 질문을 음성으로 했을 시, 시스템은 STT API 를 통해 질의를 텍스트로 변환하고, 환경 측정 센서들이 실시간으로 전송한 실내 온도, 습도, 가스 농도 등의 데이터를 수신한다. OpenAI 에서 제공하는 LLM 모델인 GPT 는 센서 데이터를 포함한 자연어 응답을 생성하고, 최종적으로 TTS API 텍스트를 음성으로 변환하여 사용자에게 응답한다. 라즈베리파이의 Docker 환경 위에 Mosquitto 브로커를 구축하고, ESP8266, 아두이노 등으로 구성된 본 시스템 환경 센싱 데이터와 대화형 AI 모델을 통합함으로써 사용자의 질의에 실시간 데이터를 반영한 응답이 가능함을 보이고, 정확한 응답 제공 여부를 확인하였다.

### I. 서 론

최근 음성으로 대화를 나눌 수 있는 음성 기반 지능형 비서가 보편화되면서, 일상 언어로 정보 질의나 기기 제어를 자연스럽게 수행하게 되었다. 동시에 IoT 기술의 발전으로 다양한 환경 센서를 통해 실내외 데이터를 수집할 수 있게 되었지만, 일반 사용자가 직접 활용하기엔 여전히 불편함이 따른다. 이에 음성으로 실시간 센싱 데이터를 질의하고, AI 챗봇이 자연어로 응답하는 센싱 챗봇 시스템을 제안한다. Kush[1]는 기존 대화형 AI 가 정형화된 데이터에 의존하는 반면, 센서 데이터를 실시간 통합하여 반응하는 시스템은 최근 연구되고 있다고 설명했다. GPT 모델과 음성 입출력을 결합하여 자연스러운 상호작용을 지원하며, 실시간 센서 데이터를 기반으로 정확한 응답을 제공할 수 있다. 본 논문에서는 라즈베리파이와 아두이노 기반 환경 센서로부터 수집된 데이터를 GPT 기반 챗봇과 연동하여 사용자가 음성 질의를 통해 실시간 환경 정보를 편리하게 얻을 수 있는 시스템을 구현할 것이다.

### II. 시스템 구성

본 시스템은 각종 환경 측정 센서, 네트워크 모듈, 챗봇 서버, 음성-텍스트 변환 API 로 구성된다. 그림 1 은 각 구성 요소와 데이터 흐름을 나타낸다. 아두이노 보드에 각종 센서들을 연결하여 수집된 데이터를 ESP8266 Wi-Fi 모듈을 통해 전송한다. 데이터 전송에는 MQTT(Message Queuing Telemetry Transport) 프로토콜이 사용되며, 저전력 IoT 환경에 최적화된 경량 메시지 기반 통신 방식이다. 전송된 센싱 데이터는 라즈베리파이 내의 MQTT 브로커인 Mosquitto 를 통해

수신된다. MQTT 메시지를 중개하는 역할을 하며, 발행된 메시지를 저장하거나, 토픽을 구독한 장치 또는 서버에 발행(Publish)하여 전달한다. 챗봇 서버는 Python 기반 애플리케이션으로, MQTT 브로커의 토픽을 구독(Subscribe)하여 최신 센싱 데이터를 주기적으로 수신하며, OpenAI GPT API 는 사용자의 질의에 응답할 때, 사전에 정의된 함수 호출 조건에 따라 실시간 데이터를 반영한 응답을 생성한다. 음성 질의는 [2] 네이버 클로바의 STT(Speech To Text) API 를 이용하여 텍스트로 변환되어 챗봇 서버로 전달되며, 생성된 응답은 구글 클라우드 플랫폼[3]의 TTS(Text To Speech) API 를 통해 음성으로 변환된다.

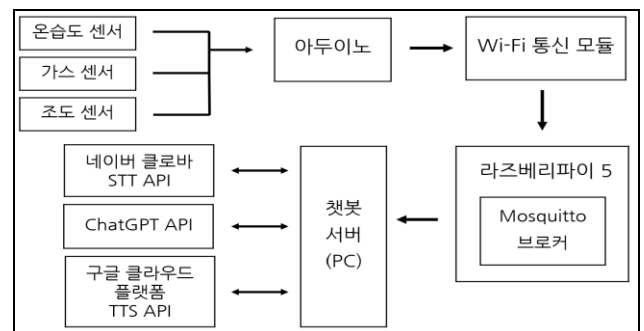


그림 1. 센싱 챗봇 시스템의 데이터 구성

### III. 동작 절차

본 시스템은 사용자의 음성 질의를 통해 실시간 환경 센싱 데이터를 활용한 자연어 응답을 생성하여 전달하는

방식으로 동작한다. 사용자의 음성 질의는 STT API 를 통해 텍스트로 변환된다. 변환된 텍스트는 OpenAI 의 GPT-4o-mini 모델에 전달되고, 챗봇 서버에서는 MQTT 토픽 구독을 통해 실시간으로 수신되는 센싱 데이터(그림 2)를 활용하여 사용자의 질의에 적합한 자연어 응답을 생성한다[4]. 이후 생성된 텍스트 응답은 TTS API 를 통해 음성으로 변환되어 사용자에게 전달된다.

구현된 시스템의 성능 및 안정성 검증을 위해 각종 센서들을 이용한 환경에서의 실험을 진행하였다. 그림 3 과 같이 “현재 사무실 온도가 어떻게 되나요?”와 같은 실시간 환경 정보 질의 시, GPT 모델은 MQTT 로 수신한 센싱 데이터를 기반으로 “현재 사무실 온도는 24.5 도입니다.”와 같은 응답을 정확하고 자연스럽게 제공하였다. 또한 센싱 데이터가 필요하지 않은 질의에는 그림 4 와 같이 GPT 모델이 센싱 데이터를 사용하지 않고, 기본적인 API 동작 또한 안정적으로 수행됨을 확인하였다.

```
수신된 센서 데이터: {'l': 123, 't': 24.5, 'h': 60.0, 'g': 350}
수신된 센서 데이터: {'l': 123, 't': 24.5, 'h': 60.0, 'g': 350}
```

그림 2. 챗봇 시스템에서 수신 받은 데이터 예시

```
인식된 텍스트: 현재 사무실 온도가 어떻게 되나요?
GPT 응답: 현재 사무실 온도는 24.5도입니다.
```

그림 3. 센싱 데이터 기반 응답 예시

```
인식된 텍스트: 영어 hello를 한국어로 번역해주세요
GPT 응답: Hello를 한국어로 번역하면 "안녕"이에요.
```

그림 4. 기본적인 GPT API 기반 응답 예시

#### IV. 한계 및 개선 방안

OpenAI GPT 모델이 외부 데이터를 필요로 할 때 사전에 정의된 함수의 description(설명)을 기반으로, 모델이 사용자 질의와 관련된 함수를 의미적으로 유사한지 판단하여 자동으로 호출한다. 본 시스템은 실시간 센싱 데이터를 활용하여 GPT 모델이 자연어 응답을 생성하는 구조이나, 사용자의 질문이 “좀 덥다.”, “공기가 탁하다”와 같이 모호하거나 주관적인 경우 GPT 모델은 함수를 호출의 필요성을 정확하게 판단하지 못해 그림 5 와 같이 기본 응답만 제공하는 한계가 있었다. 이를 개선하기 위해, GPT 모델에서 함수를 호출하는 기준인 description 문구를 수정하여, 주관적인 질문에도 함수를 호출할 수 있도록 description 이 더 포괄적인 의미를 갖도록 수정하였다. 표 1 는 description 문구에 따른 호출 변화를 정리하였다. 최종적으로 챗봇 시스템을 구성할 때 사용한 Python 기반 코드의 description 은 그림 6 과 같이 구성하였다.

```
인식된 텍스트: 지금 방이 덥지 않아?
GPT 응답: 저는 AI로서 온도를 느낄 수 없습니다. 죄송합니다.
```

그림 5. 기본 응답만 제공하는 한계 예시

```
"type": "function",
"function": {
  "name": "get_temperature",
  "description": (
    "실내 온도, 습도, 공기 상태 등 환경과 관련된 "
    "표현이 감지되면 이 함수를 호출하세요."
  ),
}
```

그림 6. 센싱 데이터 함수 description

표 1. Description 에 따른 결과

단계	Description	결과
1	실내 온도를 물을 때 호출하세요.	미 호출
2	실내 온도와 관련된 질문이 감지되면 호출하세요.	호출
3	실내 온도, 습도, 공기 상태 등 환경과 관련된 표현이 감지되면 이 함수를 호출하세요.	모호한 질의에도 호출

#### V. 결론

본 논문에서는 프라이빗 환경의 실시간 센싱 데이터를 기반으로 음성 질의를 받고 대답하는 챗봇 시스템을 구현하였다. 각종 환경 측정 센서에서 받은 데이터를 라즈베리파이, 아두이노, Mosquitto 등을 통해 수집하여 GPT 기반 언어 모델에 연동함으로써, 사용자는 음성 질의만으로 실내외 환경정보를 자연스러운 응답으로 받을 수 있었다.

본 시스템은 GPT 언어 모델이 외부 도구를 호출할 수 있는 구조를 활용하여, 로컬 센서에서 수집한 프라이빗 실시간 데이터를 직접 응답에 반영할 수 있도록 구현했다. 기존 LLM 기반 챗봇이 사전 학습된 지식이나 외부 API 에 의존하던 방식과 달리, 사용자의 질의에 따라 실시간 센싱 데이터를 동적으로 호출하여 보다 상황에 따라 현실적인 응답을 제공할 수 있었다.

#### ACKNOWLEDGMENT

이 논문은 2025 년도 과학기술정보통신부의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임. (No.2022-0-010000, 이동형 맞춤형 의료서비스 지원을 위한 유연의료 5G 엣지 컴퓨팅 SW 개발)

#### 참 고 문 헌

- [1] J. C. Kush, "Integrating Sensor Technologies with Conversational AI: Enhancing Context-Sensitive Interaction Through Real-Time Data Fusion," *Sensors*, vol. 25, no. 1, p. 249, Jan. 2025
- [2] Naver Cloud Platform, "CLOVA Speech Recognition (CSR)," Naver Cloud Docs. Available: <https://api.ncloud-docs.com/docs/en/ai-naver-clovaspeechrecognition>
- [3] Google Cloud, "Text-to-Speech," Google Cloud Docs. Available: <https://cloud.google.com/text-to-speech/docs>
- [4] OpenAI, "GPT-4o-mini," OpenAI API Docs. Available: <https://platform.openai.com/docs/models/gpt-4o-mini>