

# 음성 기반 대화형 키오스크 애플리케이션 개발

안세홍<sup>1</sup>, 김언지<sup>1</sup>, 임호준<sup>1</sup>, 허대영<sup>2</sup>

<sup>1</sup>국민대학교 소프트웨어학부

<sup>2</sup>국민대학교 소프트웨어학부

andy7634@kookmin.ac.kr, ki6z5u1m@kookmin.ac.kr, joken@kookmin.ac.kr, dyheo@kookmin.ac.kr

## Development of a Voice-Based Conversational Kiosk

Sehong Ah<sup>1</sup>, Eonji Kim<sup>1</sup>, Hojun Lim<sup>1</sup>, Daeyoung Heo<sup>2</sup>

<sup>1</sup>Dept. of Software, Kookmin University

<sup>2</sup>Dept. of Software, Kookmin University

### 요 약

본 연구는 기존 키오스크의 사용자 접근성의 한계를 극복하고 사용자에게 직관적이고 개선된 경험을 제공하는 음성 대화형 키오스크 애플리케이션 개발에 중점을 둔다. 음성 인식을 통해 사용자와 상호작용하고, 대화 내용을 텍스트로 표시해 재확인할 수 있도록 하는 인터페이스를 제공하여 키오스크 이용자의 접근성과 사용성을 향상시켰다. 특히 키오스크 사용에 어려움을 겪는 고령층의 편의를 증진하여, 소비자층을 원활히 확대하는 데 기여할 것으로 기대한다. 사용자는 대화를 통해 주문과 조작을 할 수 있도록 하여, 기존에 복잡한 인터페이스를 단순화하고, 메뉴 추천, 재료 조회 등의 기능을 통해 사용자의 접근성을 높인다. 또한, 가게 운영 측면에서도 효율성을 높일 수 있다. 혼잡 시간대에 대한 추가 인력 소요비용을 절감할 수 있을 것이다.

### 1. 서론

본 연구는 기존 키오스크 사용에 대한 어려움과 접근성 문제를 해결하기 위해 음성 대화 방식의 인터페이스를 연구하여 실제 애플리케이션을 개발하는 것에 중점을 둔다. 시장에서는 효율성과 편의성을 위해 키오스크를 도입한다. 그러나 복잡한 인터페이스 때문에 고령층을 포함한 다수의 이용자가 어려움을 겪는다. 이러한 문제를 해결하기 위해 사용자에게 친근한 대화형 인터페이스를 도입하였다. 제안하는 키오스크는 음성 인식 기술을 활용하여 사용자가 실제 대화를 통해 주문할 수 있는 환경을 제공하여, 고령층, 장애인, 어린이 등 키오스크 사용에 어려움을 겪는 사용자들의 접근성과 편의성을 극대화하는 것에 중점을 두었다.

### 2. 연구 배경 및 필요성

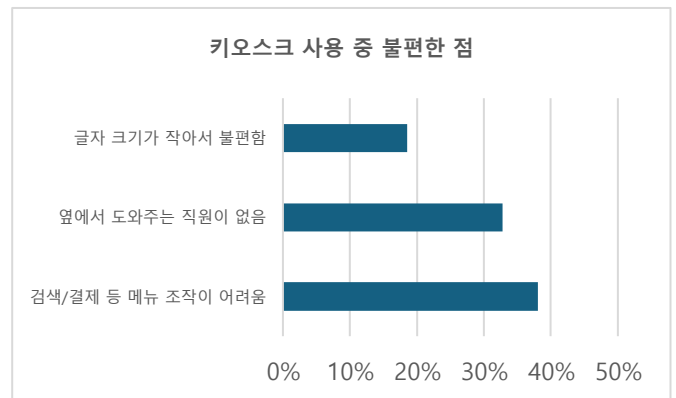


그림 1. 키오스크 이용 중 불편한 점 (중복응답) <sup>[1]</sup>

많은 가게에서 키오스크는 효율성과 편의성을 위해 도입하였지만, 서울디지털재단의 조사에 따르면 55 세 이상 고령층 중 45.8%만이 키오스크를 사용해본 경험이 있다. 이는 55 세 미만의 94.1%와 비교해 절반에도 미치지 않는 수치로, 고령층의 사용 불편을 시사한다.

이들은 ‘복잡한 단계 및 화면 조작’, ‘주문 상품에 대한 문의 불가’ 문제를 겪고 있다. [그림 1] [1][2]

우리 사회가 점점 고령화 사회로 접어드는 만큼, 고령층의 편의를 증진할 수 있는 인터페이스를 연구하고, 이를 도입하는 것의 중요성이 커지는 실정이다.

### 3. 시스템 구조 및 개발

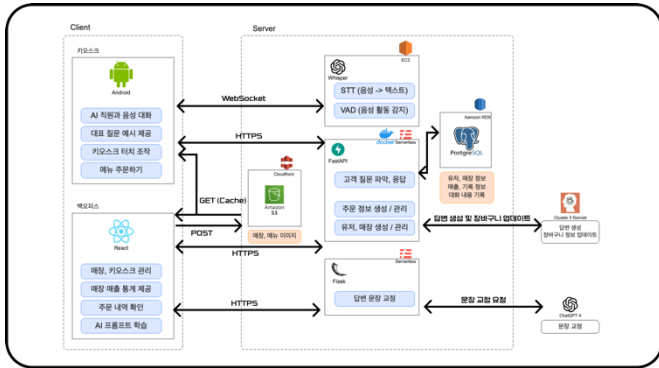


그림 2. 시스템 구조도

제안하는 키오스크 시스템은 그림 2 와 같이 안드로이드 기반의 (1) 키오스크 애플리케이션, (2) 웹 기반 백오피스, OpenAI 의 음성 인식 모델인 Whisper [3] 를 사용하는 클라우드 환경의 (3) 백엔드 서버로 구성된다. 키오스크 애플리케이션은 대화형 인터페이스의 표면 계층을 제공하며, 웹 기반 백오피스는 매장 관리자가 메뉴와 정보를 설정할 수 있는 관리 페이지를 제공한다. 백엔드 서버는 데이터 저장 및 처리를 클라우드에서 담당하고, 음성 인식 및 대화 생성을 위한 핵심 기능을 수행한다. 백엔드에서 음성 대화를 위해서, 단말기에서 수신된 음성 데이터를 텍스트로 변환이 필요한 데, Whisper 모델은 실시간 자막 생성과 같은 인공지능 태스크를 수행할 수 있는데, 다양한 언어 및 다양한 모델 크기를 제공하여 이를 선택하였다.

다. 대화 내용은 실시간으로 텍스트로 변환되어 화면에 채팅 형식으로 표시되며, 사용자는 이를 통해 주문 과정에서 했던 대화 기록을 확인할 수 있다. 대화 내용은 익숙한 챗봇 채팅 형식으로 제공하여 기존의 대화나 화면에서 확인할 수 있게 하였다.

대표 예시 질문을 버튼으로 제공받을 수 있다. 또한, 터치형 키오스크가 친숙한 사용자들을 위해 대화형 모드와 터치형 모드를 분리 및 전환되게끔 하여 메뉴를 터치 주문 가능하게 하였다.

백오피스는 그림 3 의 (b)와 같이 웹 기반으로 인터페이스를 제공하고, 매장 정보 입력, LLM 응답 설정, 주문 내역 및 매출 통계 확인 기능을 포함한다. 매장의 관리자는 키오스크가 매장 및 메뉴에 대해 잘 이해할 수 있도록 설명을 덧붙일 수 있어 특정 식재료에 알레르기가 있거나 이해가 필요한 사용자에게 도움을 줄 수 있다.

### 4. 결론 및 향후 과제

본 연구는 키오스크 사용에 어려움을 겪는 이들을 위해 사용자의 직접적인 터치를 최소화하고, 대화를 통해 주문과 조작을 할 수 있게 하였다. 채팅 대화와 같은 인터페이스와 LLM 기반의 대화형 기능을 통해 기존 키오스크가 가지고 있는 한계를 극복하고 다양한 사용자 계층의 요구를 충족시킬 수 있을 것으로 기대한다. 음성 인식 성능 개선, 사용자 피드백을 반영한 UI/UX 개선, 데이터 기반의 맞춤형 서비스 제공 기능 추가 등을 계획하고 있다.

### Acknowledgement

본 연구는 2024 년 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학사업의 연구결과로 수행되었음(No. 2022-0-00964)

### 참고문헌

- [1] 롯데멤버스 라임, 키오스크 이용경험 설문조사 실시  
<https://www.discoverynews.kr/news/articleView.html?idxno=977413>
- [2] 키오스크(무인정보단말기) 이용실태 조사  
<https://www.kca.go.kr/smartconsumer/synapviewer.do?menukey=7301&fno=10036497&bid=00000146&did=1003409523>
- [3] OpenAI. Whisper: Inroducing Whisper  
<https://openai.com/index/whisper/>

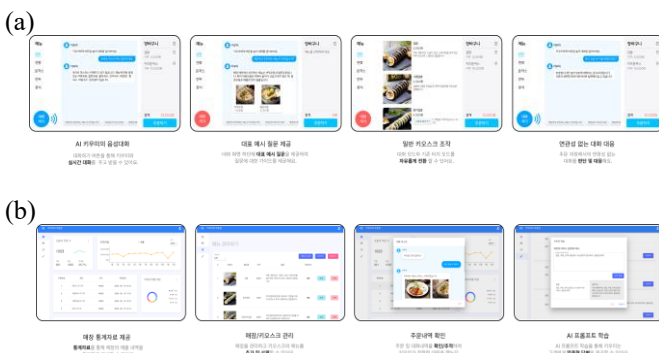


그림 3. 애플리케이션에서 제공하는 기능들

키오스크 애플리케이션은 그림 3 의 (a) 에서 보여 주는 것과 같이 음성 인식 및 대화 기능을 통해 사용자가 음성으로 메뉴를 선택하고 주문을 진행할 수 있