

# 음성지원 챗봇 모바일 애플리케이션

최인경, 최윤정, 이예린  
동국대학교 멀티미디어공학과  
e-mail : ko9983@naver.com, winning\_13@naver.com, h2002y@naver.com

## A Voice-enabled Chatbot Mobile Application

In-Kyung Choi, Yun-Jeong Choi, Ye-Rin Lee  
Dept. of Multimedia Engineering, Dongguk University-Seoul

### 요 약

사회적 문제와 인공지능 기술의 발달로 챗봇 서비스에 대한 관심이 점점 증가하고 있으며, 그 결과 TTS(Text to Speech) 및 STT(Speech to Text) 기술을 기반으로 한 보조형 프로그램에 대한 개발이 다양한 모바일 환경에서 진행 중이다. 본 논문에서는 문자를 소리로 변환해주는 TTS(Text to Speech) 기술과 소리를 문자로 변환해주는 STT(Speech to Text) 기술을 사용하여 음성 지원 챗봇 시스템을 제작하고 이를 안드로이드 기반의 모바일 애플리케이션으로 구현한 ‘음성 지원 챗봇 모바일 애플리케이션’을 제안하고, 이와 관련하여 관련 기술 및 기대효과에 대해 소개한다.

### 1. 서론

급격한 최저임금 상승으로 인하여 고용주의 매장 직원 고용 부담감이 증가하고 있고, 비대면 서비스에 대한 소비자의 선호도 또한 점점 증가하고 있다. 이에 대한 결과로 사회 전반에서 무인화 및 비대면 서비스가 확산되고 있는데, 그중 하나가 챗봇 서비스이다. 인공지능 기술이 발달함에 따라 챗봇 기술의 발전 속도도 빠르게 증가하고 있고, 챗봇의 활용 분야가 다양해지면서 그 시장 규모도 점점 커지고 있다. 대한민국을 비롯하여 전세계의 많은 기업들이 챗봇 기술을 도입하여 이에 대한 서비스를 제공하고 있거나 서비스를 준비 중이다.

챗봇은 사용자가 메시지를 통해 채팅하듯 질문을 입력하면 인공지능이 빅데이터 분석을 바탕으로 사용자와 대화를 하듯 해답을 주는 메신저이다. [1] 챗봇은 크게 시나리오형 챗봇과 인공지능형 챗봇으로 나눌 수 있는데 먼저 시나리오형 챗봇은 미리 설정해 놓은 단어에 따라 정해진 답을 하는 방식이다. 시나리오형 챗봇의 경우 인공지능형 챗봇에 비해 구현이 상대적으로 용이하다는 장점이 있지만 미리 설정해 놓지 않은 단어에 대한 답변이 불가능하고, 질문의 복잡도에 따라 원하지 않은 답변을 내놓을 수 있다는 단점이 있다. 인공지능형 챗봇은 자기 학습도 가능할 뿐 아니라 인공지능 기술을 기반으로 복잡한 질문에도 그에 알맞은 답을 할 수 있어 시나리오형 챗봇보다 자연스러운 대화가 가능하다는 장점이 있다.

사회적 문제와 인공지능 기술의 발달로 챗봇 서비스에 대한 관심이 점점 증가하고 있다. 본 논문에서는 챗봇의 개념과 제안하는 시스템 및 관련 기술 그리고 기대효과에 대해 소개하고자 한다.

### 2. 제안하는 시스템

급격한 최저임금 상승으로 인한 고용주의 매장 직원 고용 부담감, 비대면 서비스에 대한 소비자의 선호도 증가와 같은 사회적 문제를 해결하기 위해 우리가 제안하고자 하는 시스템은 ‘음성지원 챗봇 모바일 애플리케이션’으로, 이는 문자를 소리로 변환해주는 TTS(Text to Speech) 기술과 소리를 문자로 변환해주는 STT(Speech to Text) 기술을 사용하여 음성지원 챗봇 시스템을 제작하고 이를 안드로이드 기반의 모바일 애플리케이션으로 구현하는 것이다. 모바일 애플리케이션으로 구현된 챗봇은 매장 내외에서 모두 이용이 가능해 시간과 장소에 제약 없이 언제 어디서든 서비스를 요청할 수 있고, 서비스를 제공받을 수 있다.

기존의 텍스트 중심 챗봇 서비스에서 벗어나 음성 인식 기능을 추가적으로 제공하는 ‘음성지원 챗봇 모바일 애플리케이션’을 개발함으로써 기대할 수 있는 효과는 고객의 대기시간 감소, 고용주의 인건비 감소, 소비자의 구매력 상승, 기업의 경쟁력 강화 등이 있다. 시간과 장소에 제약을 받지 않아 고객은 언제 어디서든 서비스 이용이 가능하고, 기업은 24 시간 고객을 응대할 수 있기 때문에 빠른 고객 서비스가 가능해진다. 이에 따라 서비스를 제공받기 위해 고객이 대기해야 하는 시간은 감소할 것이다. 또한 직원의 업무 일부를 챗봇이 대체함으로써 고용주는 인건비를 절감할 수 있고, 직원들은 다른 일에 집중할 수 있어 일에 대한 효율성도 증가할 것이다. 그 뿐만 아니라 구매에 대한 접근성도 좋아져 소비자의 구매력도 상승하게 될 것이다. 이를 통해 기업은 소비자에게 보다 나은 서비스를 제공할 수 있게 되고, 이는 기업의 경쟁력 강화로 이어질 것이다.

사용자는 해당 애플리케이션을 이용하여 상품에 대해 문의하거나 상담을 받을 수 있으며, 등록된 계정을 통해 고객 맞춤형 서비스를 제공받을 수 있다. 애플리케이션 이용을 통해 직원의 도움 없이도 일부 서비스 이용이 가능하게 되고, 그 결과 인건비도 감축할 수 있게 된다.

음성지원 챗봇 서비스와 함께 BLE(Bluetooth Low Energy)를 이용한 근거리 통신 기술인 비콘(Beacon)을 활용하여 이용자에게 위치 기반 서비스를 제공한다. 비콘은 전력 소모가 적고, 실내에서도 정확성이 높다는 장점이 있다. 매장 내 설치된 비콘이 실시간으로 스마트폰을 스캔하고 있다가 애플리케이션 이용자가 일정 거리 이내로 접근하면 그 신호를 인지하여 애플리케이션 실행에 대한 팝업 창을 띄워주거나 할인 쿠폰이나 행사에 대한 팝업 알림을 제공한다.

### 3. 관련 기술

‘음성지원 챗봇 모바일 애플리케이션’을 제작함에 있어 우리가 사용할 툴은 마이크로소프트(Microsoft)에서 제공하는 Cognitive Service 중 Speech Service, LUIS(Language Understanding), QnA Maker 이다.

#### 4.1 Cognitive Service

Cognitive Service 는 개발자가 지능형 애플리케이션을 빌드함에 있어 도움이 되도록 사용할 수 있는 API, SDK 및 서비스로, 주요 핵심 요소에 따라 Vision, Speech, Language, Search, Knowledge 로 분류된다. [2]

#### 4.2 Speech Service

Speech Service 를 이용하면 음성을 텍스트로, 텍스트를 음성으로 변환할 수 있다. [3] Speech Service 의 주요 기능인 Speech to Text API 를 이용해 사용자의 음성을 실시간으로 변환하여 텍스트로 제공할 수 있고, Text to Speech API 를 통해 텍스트를 실제 사람이 말하는 것처럼 자연스러운 음성으로 변환할 수 있다.

#### 4.3 LUIS(Language Understanding)

LUIS 는 사용자의 발언에서 의미를 예측하여 정보를 추출한 뒤, 그 의도를 파악하여 그에 따른 결과 개체를 제공해준다. [4] 예를 들어 사용자가 “가장 가까운 매장이 어디인가요?”라고 말하면 사용자 발언의 의도인 ‘매장의 위치’를 파악하고 그에 따라 결과 개체인 ‘위치’ 정보를 제공해준다. 또 사용자가 “가장 인기있는 제품 추천해주세요”라고 말하면 ‘베스트셀러’라는 사용자 발언의 의도를 파악하여 그에 따라 결과 개체인 ‘인기있는 상품의 이름’ 정보를 제공해준다.

<표 1> LUIS 모델

사용자 발언	의도	결과 개체
“가장 가까운 매장이 어디인가요?”	Store Location	Location
“가장 인기있는 제품 추천해주세요”	Best Seller	Popular Product Name
“매장 운영시간 알려주세요”	Store Hours	Time

#### 4.4 QnA Maker

QnA Maker 는 챗봇 기능을 구현하기 위한 도구로, 텍스트를 질문 및 답변 지식 베이스로 결합해준다. [5]

QnA Maker 의 지식 베이스는 사용자 발언의 의도를 가장 잘 나타내는 텍스트를 포함하는 질문과 사용자 질문이 관련 질문과 일치할 경우 반환되는 답변 그리고 선택적 메타데이터의 집합으로 구성된다.

메타데이터는 질문 및 답변과 연결된 태그로, 메타데이터 태그는 질문 및 답변을 필터링하고 질문 매칭이 수행되는 집합을 제한하는 데 사용된다. 여러 개의 대체 질문이 모두 단일 답변에 매핑 될 수도 있고, 질문-답변 쌍에 연결된 메타데이터 필드가 여러 개 있을 수도 있다.

QnA Maker 는 지식 베이스의 QnA 에서 사용자의 자연어 질문과 가장 적합한 답변을 매칭하는 방식으로 답을 한다.

### 4. 결론 및 향후과제

인공지능 및 챗봇 기술의 발달 속도는 점점 더 빨라지고 있으며, 이와 관련된 시스템은 앞으로 더 많아질 것이다. 인공지능 및 챗봇 기술의 발달은 우리 삶의 질을 향상시켜줄 것이다. 하지만 이에 대한 우려의 목소리도 존재한다. 최신 기술에 적응력이 높은 젊은층에 반해 노인층은 이러한 기술의 사용에 어려움을 겪는다. 이들을 지칭하는 말로 ‘디지털 소외계층’이라는 말이 있을 정도이다. 디지털 소외 계층은 스마트 기기 사용에 불편함을 느끼고, 정보 격차로 인해 소외감을 느끼기도 한다. 또 심할 경우 우울감을 느끼기도 하는데, 기술이 발달될수록 이러한 현상은 심화될 것이다.

디지털 소외 현상을 해결하기 위해서 디지털 소외 계층이 최신 기술에 적응할 수 있도록 지속적인 관심을 가져야 하고, 그들의 적응을 돕기 위해 사회적인 지원 및 교육 프로그램을 제공해야 한다. 무엇보다 이에 대한 디지털 소외계층의 적극적인 자세와 노력이 필요하다.

기술의 발달을 통한 삶의 질 향상도 중요하지만 소외계층 없이 모든 사람이 동일한 혜택을 누릴 수 있도록 하는 것이 더 중요하다.

#### Acknowledgments

"본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학지원사업의 연구결과로 수행되었음"(2016-0-00017)

#### 참고문헌

[1] 한경 경제용어사전, “챗봇”, <http://dic.hankyung.com/apps/economy.view?seq=13232>  
 [2] Microsoft Azure, “Cognitive Service”, <https://docs.microsoft.com/ko-kr/azure/cognitive-services/welcome>  
 [3] Microsoft Azure, “Speech Service”, <https://docs.microsoft.com/ko-kr/azure/cognitive-services/speech-service/>  
 [4] Microsoft Azure, “LUIS”, <https://docs.microsoft.com/ko-kr/azure/cognitive-services/luis/>  
 [5] Microsoft Azure, “QnA Maker”, <https://docs.microsoft.com/ko-kr/azure/cognitive-services/qnamaker/>