

# STT 기반 카카오톡 이벤트를 TTS 음성으로 안내하는 IoT 어플 설계

2017104033 최세현

## 요약

위드코로나 시대로 접어들며 비대면 소통과 디지털 커뮤니케이션이 익숙해지고 있는 가운데 음성기술 AI 시장이 고속 성장하고 있다. 본 연구에서는 사람들이 많이 사용하고 있는 SNS인 카카오톡과 음성 합성 기술인 TTS를 사용하여 카카오톡에서 수신되는 메시지의 내용을 음성으로 알려주어 이동 중이나 운전 중일 때 스마트폰을 확인하지 않고도 메시지를 내용과 발신인을 확인할 수 있는 IoT 어플리케이션을 제안하고 구현한다.

## 1. 서론

### 1.1. 연구배경

위드코로나 시대로 접어들며 비대면 소통과 디지털 커뮤니케이션이 익숙해지고 있는 가운데 많은 사람들이 SNS를 통해 지인들과 소통한다. SNS에서 볼 수 있는 텍스트 위주의 정보들은 일반적인 상황에서는 유용하지만 이동중에는 확인하기엔 불편하기 때문에 많은 사람들이 음성으로 듣는 것을 선호한다. 다른 무언가를 하면서 동시에 화면에 집중하는 것은 쉽지 않으며 때로는 사고를 초래하기도 한다. 이러한 이유들로 인해 음성기술 AI 시장이 고속 성장하고 있다. 음성기술 AI는 스마트폰이나 자율주행 차량, 각종 전자기기를 비롯한 다양한 산업군에서 빠른 속도로 진화하며 생활 속에 녹아들고 있다.

본 연구에서는 사람들이 많이 사용하는 SNS인 카카오톡과 음성기술 시장에서 큰 비중을 차지하고 있는 음성합성 기술을 활용하여 카카오톡 메시지를 음성으로 확인할 수 있도록 하여 이동 중에 스마트폰 화면을 확인하여 생길 수 있는 사고를 방지하는 IoT 어플리케이션을 구현하고자 한다.

## 1.2. 연구목표

본 연구에서는 STT기반의 카카오톡의 다양한 이벤트와 음성합성 기술을 활용한 시스템을 구현할 것이다. 사용자가 스마트폰을 보고 있지 않아도 카카오톡으로 수신되는 메시지를 음성합성을 통해 사용자에게 음성으로 알려줌으로써 스마트폰을 볼 수 없는 상황에서도 편하게 사용할 수 있도록 구현하고자 한다. 또한, 블루투스 연결을 가능하게 함으로써 블루투스 스피커나 이어폰, 헤드셋을 통해서도 음성을 들을 수 있도록 할 것이다.

카카오톡으로 전송된 메시지를 어플리케이션으로 전달받기 위해 브로드캐스트 수신자를 등록하고 sms 수신 권한을 추가하여 sms를 수신 가능하도록 한다. 음성합성 기술인 TTS(Text to Speech)를 통해 전달받은 메시지를 음성으로 전달하며 음성의 속도와 크기를 사용자가 조절할 수 있다. 읽어오는 텍스트는 내용 뿐만 아니라 발신자와 발신시각을 사용자의 선택에 따라 읽어올 수 있다. 어플리케이션을 실행시키면 기능이 무조건 동작하는 것이 아니라 사용자의 필요에 따라 사용여부를 선택할 수 있도록 구현할 것이다.

## 2. 관련 연구

### 2.1 블루투스

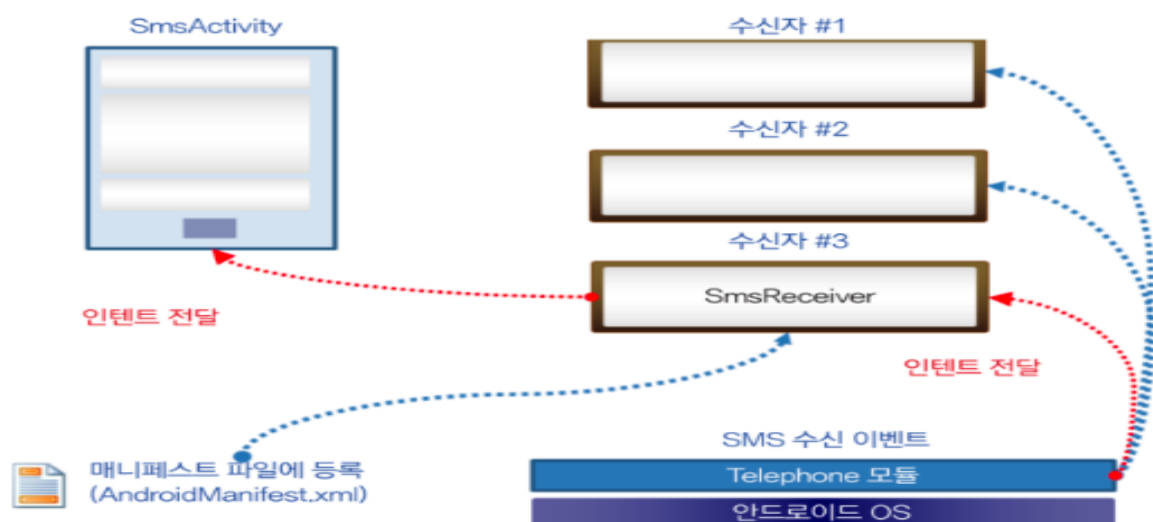
블루투스(Bluetooth)는 스마트폰, 노트북, 이어폰·헤드폰 등의 휴대기기를 서로 연결해 정보를 교환하는 근거리 무선 기술 표준을 의미하는 것으로, 주로 10미터 안팎의 초단거리에서 저전력

무선 연결이 필요할 때 쓰인다. 블루투스를 이용한 무선 네트워크는 언제 어디서나 모든 정보기기 간의 자유로운 데이터 교환이 이뤄질 수 있다.

Android 플랫폼은 블루투스 네트워크 스택에 대한 지원을 포함하므로 기기가 다른 블루투스 기기와 데이터를 무선으로 교환할 수 있다. 애플리케이션 프레임워크는 Android Bluetooth API를 통해 블루투스 기능에 대한 액세스 권한을 제공한다. 애플리케이션은 이러한 API를 통해 다른 블루투스 기기에 무선으로 연결하여 지점 간 무선 기능 및 다중 지점 무선 기능을 활성화할 수 있다.

## 2.2 Broadcast Receiver

안드로이드에서 브로드캐스팅(Broadcasting)이란 메시지를 여러 객체에 전달하는 것을 말한다. 여러 앱 구성 요소에 메시지를 전달할 때 브로드캐스팅을 사용한다. 개발하는 어플에서 브로드캐스팅 메시지를 받으려면 Broadcast Receiver를 만들어 앱에 등록해야한다. 기기 안에서 동작하는 카카오톡으로부터 메시지를 받기 위해 개발하는 어플에 Broadcast Receiver를 등록하면 카카오톡의 메시지가 개발하고 있는 어플로 전달된다.



[그림 1] 브로드캐스트 수신자의 동작방식

단말에서는 다른 사람으로부터 SMS 문자를 받았을 때 텔레포니(Telephony) 모듈이 처리하도록 한다. 이렇게 처리된 정보는 인텐트에 담겨 브로드캐스팅 방식으로 다른 앱에 전달된다. 브로드캐스트 수신자를 manifest 파일에 등록을 해놓은 상태라면 개발하고 있는 앱으로 인텐트를 전달할 수 있다. 브로드캐스트 수신자를 사용하면 어플이 실행되어 있지 않아도 원하는 브로드캐스트 메시지가 도착하면 다른 어플을 실행하고 있는 도중에도 인텐트 안에 들어있는 메시지를 받아볼 수 있다.

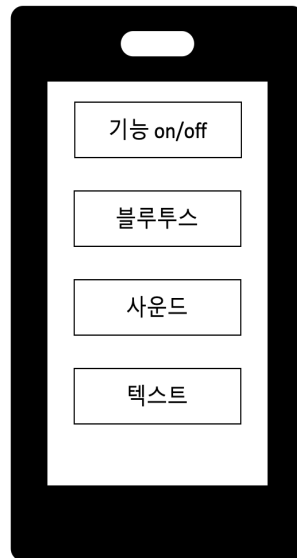
## 2.3 Text-to-Speech

음성합성(TTS)이란 입력문장(Text)을 사람이 청취할 수 있는 음성(Speech)으로 변환시키는 기술이라는 뜻으로 'Text To Speech'의 약자인 'TTS'라는 용어로 오늘날 더 많이 사용되고 있다. 이 때 입력문장은 데이터베이스의 자료일 수도 있고, 교통안내와 같은 문장일 수도 있다. 입력된 문장을 음성으로 변환시키는 방법은 음성의 기본 단위를 먼저 데이터베이스로 구축하여 놓은 뒤 이들을 조합하여 다시 구성하는 제한 어휘 음성합성과 어휘에 제한 없이 어떤 문장이라도 합성하는 무제한 어휘 음성합성 이 두 분류로 나뉜다. 이와 같은 방식을 TTS라고 한다. 이는 억양의 고저나 장단 등 운율에 관한 정보를 결정하는 기능을 갖는다. TTS의 등장은 시각장애인, 외국인, 아이, 노인과 같이 문자를 원활하게 읽지 못하는 상황에 처한 사람들에게 큰 도움을 줄 것이라는 기대를 받았으며, 오늘날 다양한 분야에서 사용되며 주목받는 기술로 부상하고 있다.

안드로이드에서는 구글 SpeechToText, TextToSpeech 기술인 stt, tts를 내장 API로 지원하고 있다. 별도의 설치 없이 gradle에 특별한 세팅이 필요하지 않고 manifest에 몇 가지 권한을 주면 바로 불러와서 사용가능 하다.

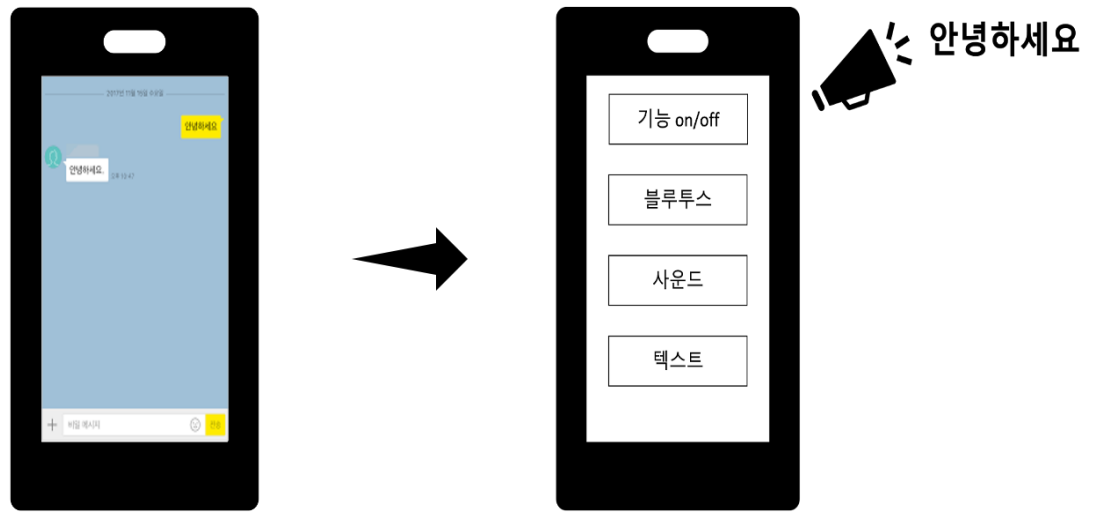
### 3. 프로젝트 내용

#### 3.1 시나리오



[그림 2] 메인 화면

[그림 2]는 애플리케이션의 메인 화면이다. 메인에서는 이 기능의 사용을 정하는 on/off 스위치를 통해 사용자에게 사용여부를 물어본다. 사용을 원한다면 블루투스와 사운드, 텍스트 설정을 가능하게 한다. 메인 화면에서 블루투스 부분은 블루투스의 사용여부와 연결된 기기의 정보를 알려주고 블루투스 설정 화면으로 들어갈 수 있게 해준다. 사운드 부분에서는 음성합성되는 소리의 볼륨을 조절할 수 있고 속도를 0.5배 부터 2배까지 선택할 수 있다. 그리고 카카오톡에서 오는 메시지 알림을 진동으로 받을지 여부를 선택할 수 있다. 텍스트 부분에서는 메시지의 내용 뿐만 아니라 발신자와 발신 시간을 사용자의 선택에 의해 얻을 수 있다.



[그림 3] 카카오톡 메시지를 받았을 때

카카오톡 메시지를 받으면 메인화면에서 설정했던 사운드와 텍스트 방식을 사용하여 메시지를 음성으로 알려준다. 메시지를 받으면 기본은 속도 1.0배에 메시지 내용만 읽어주지만 사용자가 지정한 설정에 따라 읽어주는 속도를 다르게 할 수 있으며 발신자와 발신 시간까지 읽어줄 수 있다. 또한 블루투스 스피커나 이어폰 등을 연결한 경우에는 음성이 연결한 장치에서 나온다.

### 3.2 요구사항

블루투스 연결은 Android Bluetooth API를 사용하여 블루투스 기능에 대한 액세스를 가능하게 만든다. 어플 내에서 블루투스의 on/off를 가능하게 하고 연결된 기기의 정보를 출력할 수 있게 한다.

카카오톡 메시지 알림을 받고 그 메시지의 내용과 발신자와 발신시간을 얻기 위해 브로드캐스트 수신자를 등록해야 한다. 그리고 그 정보를 얻어와 화면에 보여주는 것이 아니라 그 정보를 음성으로 전달해야 한다.

음성으로 전달하기 위해서는 TTS, 즉 음성합성 기술을 사용하여야 하고 안드로이드에서 구글 TTS API를 내장하고 있기 때문에 그것을 통해 구현한다.

Expandable RecyclerView와 ScrollerView를 사용하여 메인 화면을 구성할 때, 리사이클러뷰를 접고 펼칠 수 있도록 하여 사용자가 직관적으로 사용할 수 있도록 만든다.

#### 4. 향후 일정

진행 주차	내용
~4주차	인트로와 메인 화면 기획
~6주차	블루투스 연결 및 설정 구현
~8주차	브로드캐스트 리시버 등록 및 수신 권한 구현
~9주차	사운드 설정 구현
~12주차	TTS 구현 및 어플리케이션 테스트
~13주차	논문 작성
최종 발표	발표

## 5. 결론 및 기대효과

바빠서 스마트폰을 확인하지 못하거나 운전을 하고 있는 경우에 중요한 메시지가 오거나 업무상 필요한 메시지가 오면 스마트폰 화면을 보지 않아도 발신자와 발신 내용을 음성으로 알 수 있기 때문에 편리하게 사용가능하다. 그리고 혹시나 운전 중에 스마트폰을 확인하다가 생기는 사고를 방지할 수 있다.

블루투스 연결이 가능하기 때문에 이어폰이나 헤드셋을 착용하고 있어도 음성을 들을 수 있고 음성의 속도와 볼륨이 조정 가능해서 사용자의 필요에 맞출 수 있다.

최근에는 음성인식 기술을 탑재한 AI 인공지능 스피커를 많이 사용하는데 그걸 구매하지 않고도 스마트폰의 어플만으로도 메시지를 읽어주기 때문에 장비를 구매해야 하는 비용적인 측면을 감소시킬 수 있다.

## 6. 참고문헌

[1] 정재곤 저, 안드로이드 앱 프로그래밍

[2] Android Reference: <https://developer.android.com/>

[3] 이세환, &김봉현, &김승연. TTS를 이용한 음성지원 일정관리 시스템의 구현