

# Interfejs głosowy dla aplikacji mobilnej Radio1.7

Michał Kleszczyński, 292430, Lider projektu  
Beata Kotra, 292433  
WIMiR, Inżynieria Akustyczna

## Cel projektu oraz jego szacowany efekt

Celem projektu jest zaprojektowanie funkcjonalnego i intuicyjnego systemu obsługi istniejącej już aplikacji mobilnej na telefony z systemem Android umożliwiającej słuchanie internetowej studenckiej rozgłośni radiowej Radio1.7. Aplikacja umożliwia również głosowanie w rankingu radiowym zwanym „Gorąca 17”, a także przeglądanie wpisów z aktualności oraz bloga muzycznego. Każda z tych funkcjonalności będzie możliwa do aktywacji oraz obsługi z poziomu interfejsu głosowego.

Efektom projektu ma być sprawny, niepowodujący dyskomfortu u użytkownika interfejs głosowy, będący rozszerzeniem istniejącej aplikacji, wprowadzający kolejny poziom nowoczesności dla jej użytkowników.

## Wykorzystywane systemy

1. **System Android**, jako platforma, na której udostępniana jest aplikacja, oraz **Kotlin** jako wykorzystywany język programowy – ponieważ projekt jest rozwinięciem istniejącej już aplikacji i posiadamy pełny dostęp do kodu źródłowego, wykorzystane będzie to samo środowisko w celu uniknięcia niekompatybilności oraz późniejszych problemów przy wdrożeniu systemu. Poza tym jest to ułatwienie dla implementacji systemu, gdyż nie trzeba wykorzystywać zewnętrznych narzędzi do ingerencji w program.
2. **Android Speech API**, jako nakładka programowa na interfejs **Cloud Speech-to-Text** od firmy Google –ponieważ Aplikacja działa na telefonach z systemem Android oczywistą decyzją jest wybranie systemu rozpoznawania mowy od firmy nim się zajmującej, czyli Google. Speech API jest systemem rozwijanym od dawna, co powoduje, że wybranie go zwiększa prawdopodobieństwo jego skuteczności. System jest już używany przez dużą społeczność, tak dużą, iż w razie wystąpienia problemu, znalezienie jego rozwiązania będzie łatwiejsze.

## Osobowy plan organizacji pracy

1. Opis architektury
2. Opis dialogów, komend i poleceń
3. Projekt możliwych wariantów inicjalizacji systemu
4. Wykonanie potrzebnych zmian w interfejsie użytkownika
5. Podpięcie komend głosowych do głównego ekranu odtwarzania
  - a. Rozpoczynanie, pauzowanie odtwarzacza
  - b. Zmiana głośności
6. Podpięcie komend głosowych do innych ekranów aplikacji

Na podstawie powyższych punktów będzie realizowana praca przez najbliższe tygodnie.

Plan na tydzień 1:

- Beata Kotra
  - Opis dialogów, komend i poleceń
  - Projekt możliwych wariantów inicjalizacji systemu
- Michał Kleszczyński
  - Opis architektury

Plan na następne tygodnie będzie zaktualizowany o kolejne tygodnie na podstawie wykonanych prac, i będzie dynamicznie dostosowywał się do potrzebnych zadań i poprawek.

## Tydzień 1

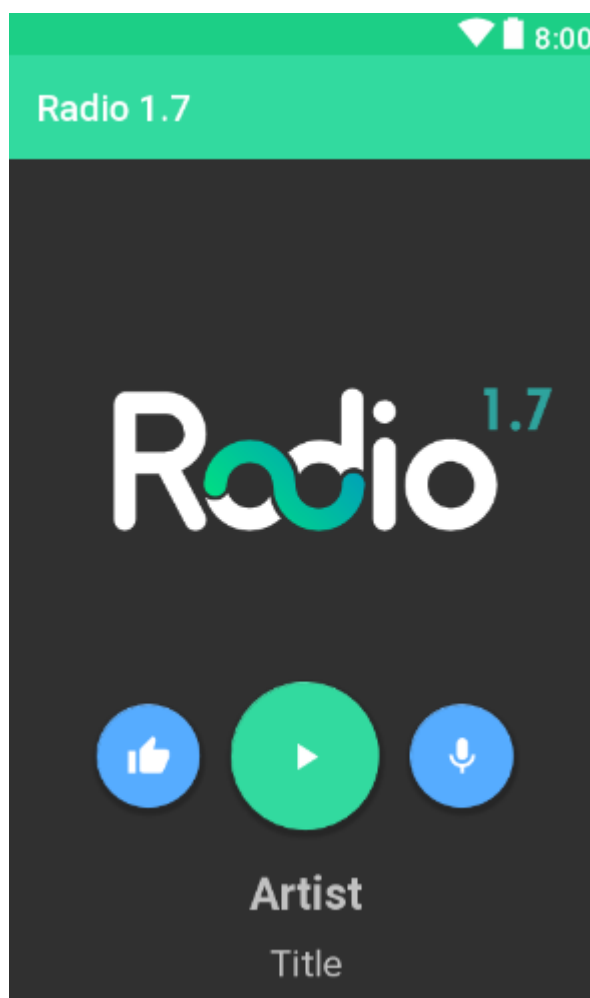
### Możliwe warianty inicjalizacji systemu

W trakcie planowania interfejsu głosowego brane były pod uwagę dwie opcje inicjalizacji systemu:

- poprzez specjalny przycisk aktywujący sterowanie głosowe
- poprzez wypowiedzenie przez użytkownika odpowiedniej sekwencji słów

Obie opcje mają swoje wady i zalety oraz wiążą się z różnymi trudnościami. Pierwsza z nich wymaga konkretnej akcji użytkownika, czyli wciśnięcia przycisku na ekranie, przez co usługa nie jest dostępna cały czas np. wtedy gdy użytkownik znajduje się w pewnej odległości od telefonu lub nie ma możliwości wykonania koniecznej akcji np. podczas jazdy samochodem. Znacznie ogranicza to łatwość jej użycia. Natomiast druga z opcji wiąże się z trudniejszą implementacją gdyż konieczne jest ciągłe nasłuchiwanie w celu identyfikacji hasła aktywującego usługę.

Ostatecznie wybrana została pierwsza z opcji. Jest ona łatwiejsza do zaimplementowania a jednocześnie wystarczająca na tym etapie prac. W przypadku pomyślnego zakończenia projektu istnieje możliwość rozwinięcia systemu i drugiej z opcji aktywacji. Użytkownik w menu opcji aplikacji



Rysunek 1. Poglądowy zrzut ekranu z ekranu odtwarzacza aplikacji

będzie mógł wybrać, który ze sposobów chce wykorzystać.

Po wciśnięciu przycisku aktywowana zostaje funkcja sterowania odtwarzaczem za pomocą komend głosowych. Aktywna jest aż do momentu ponownego wciśnięcia danego przycisku. To czy funkcja jest aktywna czy nie można łatwo dostrzec dzięki zmianie ikony, która się na nim znajduje.



Rysunek 2. Funkcja nieaktywna. Wcisnąć aby aktywować.



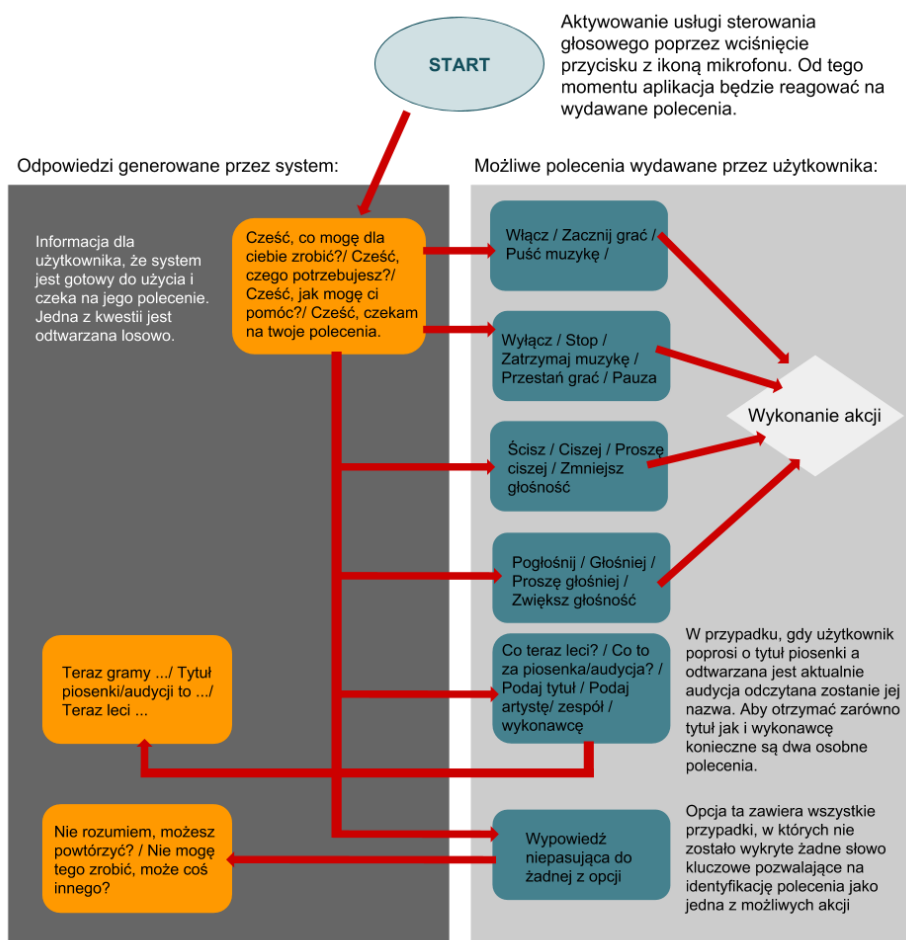
Rysunek 3. Funkcja aktywna. Wcisnąć aby wyłączyć.

## Projekt konwersacji

Kim są użytkownicy oraz jakie są ich potrzeby

Aplikacja głównie skierowana jest do młodych użytkowników, studentów w przedziale wiekowym między 19 a 30 lat. Interfejs głosowy ma służyć ułatwieniu sterowaniem playera. Zawierać musi opcję włączenia, wyłączenia, przyciszenia, pogłośnienia oraz podania informacji o aktualnie odtwarzanym utworze lub audycji.

## Schemat dialogu



Rysunek 4. Projekt konwersacji

## Architektura aplikacji

Aplikacja jest oparta o wzorec MVVM (Model-View-ViewModel), stąd dodanie nowej funkcjonalności jaką jest interfejs głosowy nie stanowi problemu. Do tej pory utworzono odpowiednie metody, które zostały podpięte pod odpowiednie zdarzenia w aplikacji (wciśnięcie guzika). W przeciągu następnych tygodni zostanie zaimplementowane zachowanie aplikacji według podanego wyżej schematu.

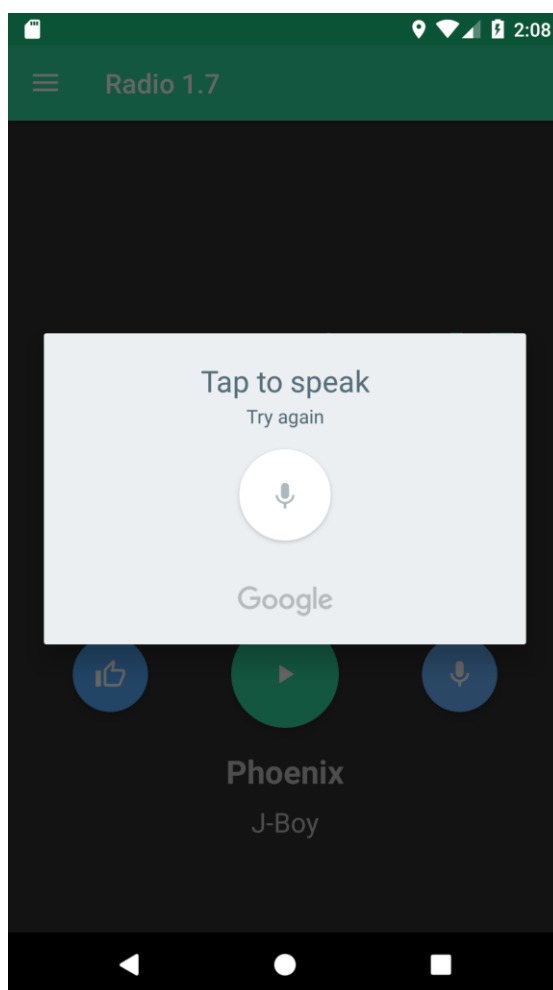
## Tydzień 2

### Obsługa prostych komend

Wykorzystując gotowe komponenty Android Speech API stworzono pierwszą wersję systemu obsługi odtwarzacza, która reaguje na proste komendy:

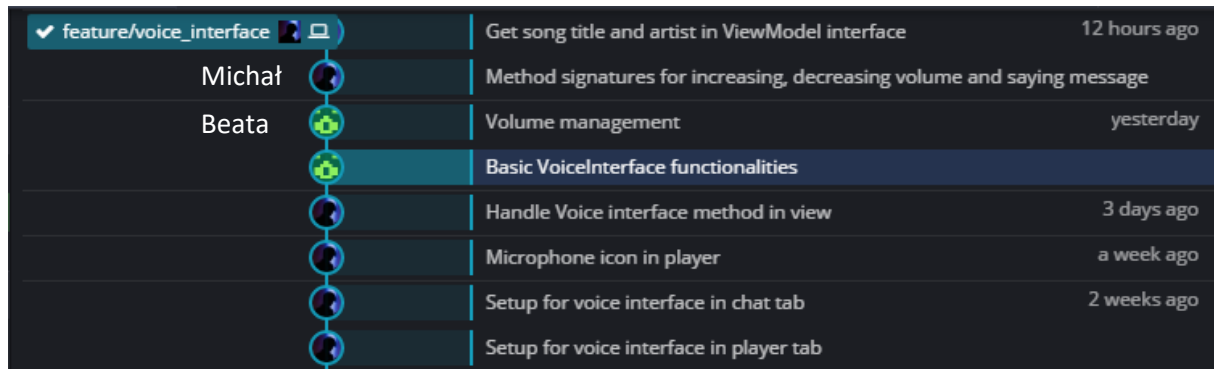
- Start – rozpoczyna odtwarzanie
- Stop – zatrzymuje odtwarzanie
- Głośniej – zwiększa głośność systemową o jeden stopień
- Ciszej – zmniejsza głośność systemową o jeden stopień

Po wciśnięciu przycisku mikrofonu aktualnie odpala się okno widoczne na obrazku poniżej (wiadomość jest automatycznie lokalizowana do języku systemu, stąd angielski tekst wiadomości)



Rysunek 5. Wygląd okienka informacyjnego

## Podział wykonanych prac



✓ feature/voice_interface	Get song title and artist in ViewModel interface	12 hours ago
Michał	Method signatures for increasing, decreasing volume and saying message	
Beata	Volume management	yesterday
	Basic VoiceInterface functionalities	
	Handle Voice interface method in view	3 days ago
	Microphone icon in player	a week ago
	Setup for voice interface in chat tab	2 weeks ago
	Setup for voice interface in player tab	

Rysunek 6. Schemat udziału prac w repozytorium projektu

Michał:

Z racji znajomości struktury aplikacji był odpowiedzialny za stworzenie sygnatur metod w odpowiednich plikach i podpięcie ich pod interfejs aplikacji.

Beata:

Odpowiedzialna za implementacje oraz testy wyżej wymienionych metod.

## Plan na następny tydzień

Na następny tydzień wyznaczono cele, które ulepszą działanie aplikacji

- Ukrycie okna, które informuje o rozpoczętym rozpoznawaniu mowy, zastąpienie ikonami mikrofonu na przycisku
- Rozpoznawanie poleceń ciągle, do momentu przerwania przez użytkownika
- Rozpoznawanie polecenia na podstawie słowa kluczowego, a nie na podstawie całej frazy
- Dodanie syntezy mowy, w celu dostarczenia odpowiedzi na komendy użytkownika oraz włączeniu i gotowości funkcji