

Przetwarzanie mowy dla platformy Android

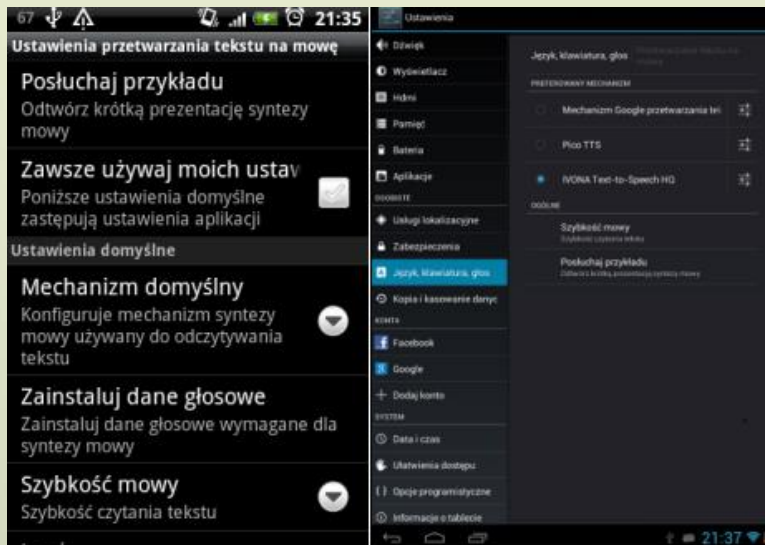
Przegląd dostępnych narzędzi i popularnych aplikacji

Agnieszka Słowik
Norbert Mieczkowski



Text-To-Speech dla platformy Android

- Android Donut – początek syntezy mowy
- Android 1.6 – klasa TextToSpeech będąca kontrolerem TTS i podklasa Engine zawierająca stałe parametry; 5 języków
- Mobilne odpowiedniki (np. IVONA)
- Android 4.0 – klasy umożliwiające implementację własnych systemów syntezy mowy wraz z implementacją głosów

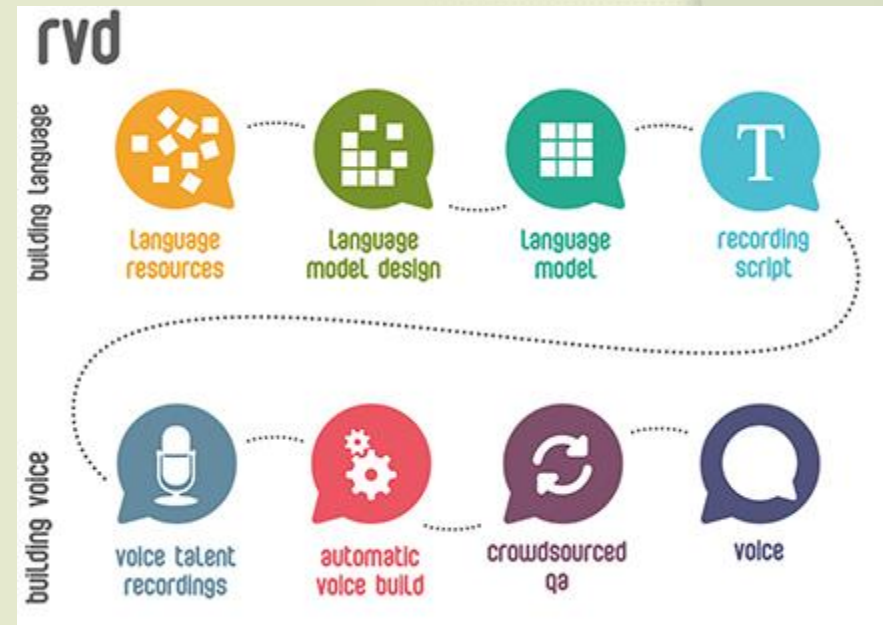


Działanie systemu Text-To-Speech

Na przykładzie IVONA TTS



- Użycie algorytmów NLP
- Radzi sobie z liczbami, datami i skrótami
- Rapid Voice Development odpowiada za fonetyzację, intonację i akcenty
- Na każdym etapie RVD jakoś danych jest sprawdzana przez rodzimych użytkowników języka



IVONA TTS

- 13 języków w tym polski
- Kompatybilny m.in. z Google Maps, aplikacjami do e-booków, przeglądarką internetową
- Duży wybór głosów męskich i kobiecych (Maja)
- IVONA for Developers – Speech Cloud Resources, SDK



Aplikacje TTS

- SpeakMe – powiadomienia
- Easy Text to Speech – działa w tle; dostęp do schowka; przychodzące smsy
- Talk–Text to Voice – strony internetowe; play/pause/stop; eksport plików audio; realistyczny głos
- CereProc – duży wybór głosów łącznie z odgłosami zwierząt w różnych językach, głosy polityków i celebrytów
- Acapela TTS Voice – 30 języków, 90 głosów



Android Lollipop – nowości dla TTS

- Precyzyjna kontrola syntezy mowy: profile głosowe pamiętające jakość, akcent i czas opóźnienia
- Wsparcie dla wykrywania błędów w syntezie



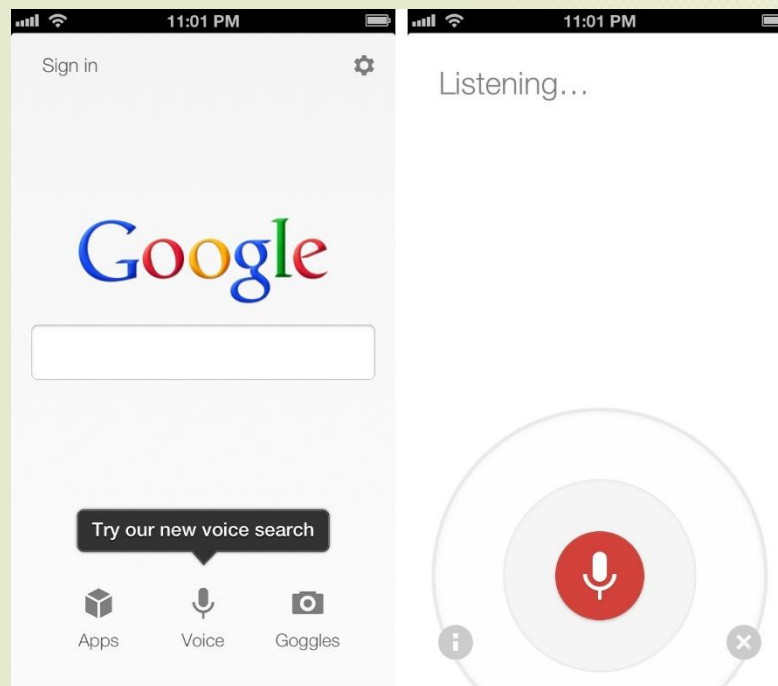
Speech-To-Text

- Początki STT na smartfonach to rok 2008
- android.speech - podstawowa biblioteka Android SDK
- Google Voice Search (połączone z Google Now w systemach Android 4.1 i wyższych)



Speech-To-Text dla języka polskiego

- Google Voice Search oferuje wsparcie dla języka polskiego od 2010 roku



Google Now

„Inteligentny” asystent

- Dostępny na Android, iOS, OS X i Microsoft Windows
- Stanowi rozszerzenie wyszukiwarki Google
- Potrafi odpowiadać na pytania oraz wykonywać polecenia w systemie
- Dostępny we wszystkich językach cywilizowanego świata



S VOICE

Osobisty asystent i nawigator

- Aplikacja w telefonach Samsung Galaxy S III i wyższych
- Wykonuje komendy głosowe pozwalające m.in:
 - wyszukiwać informacje w Internecie
 - wykonywać połączenia telefoniczne
 - otwierać aplikacje
 - ustawiać alarmy
 - aktualizować Facebooka i Twittera
- Dostępne języki:
 - angielski, francuski, hiszpański, niemiecki, włoski
 - arabski, koreański



android.speech

Biblioteka

- Dostępne klasy:
 - RecognitionListener
 - RecognitionService
 - RecognitionService.Callback
 - RecognizerIntent
 - RecognizerResultsIntent
 - SpeechRecognizer



Przetwarzanie mowy dla platformy Android

Przegląd dostępnych narzędzi i popularnych aplikacji

Agnieszka Słowik
Norbert Mieczkowski

