



## QiTalk micro

AITalk® microは、Android・iOSに対応した組み込み用小型音声合成SDKです。  
小型ながらも充実した機能を提供します。

### 提供形態

ライブラリ (dll、so)

### 主な用途

スマートフォンアプリケーション／オートモーティブ／ロボット／ゲーム

## 機能紹介



### 自由文音声合成・保存

任意の文章を人間の声に近い自然な音声で合成し、音声ファイルとして保存します。



### モバイルOS対応

iOS / Android / 組込Linux に対応



### 低容量・低リソース

さまざまな機器への組み込みが可能な低容量・低リソースを実現



### 感情調整

シチュエーションや用途に合わせた感情表現を実現しました。

※一部の話者のみ利用可能です。



### 入力対応

入力は、漢字かな交じり日本語テキスト及び、中間表記 (AI独自規格、JEITA TT-6004規格) に対応。



### 汎用PC向けSDKとの統一API

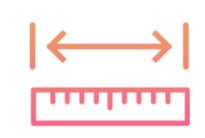
AITalk5 SDKとAPIを統一することで同等の機能を実現できるようになりました。

※一部制限あり



### 日本語解析を分離した省サイズ辞書

親機 (PC上の声の職人など) で高精度、大容量の日本語解析を行い、子機 (micro SDK) で日本語なし辞書を使うことで省メモリ・ディスクで高品質な合成



### フレーズ辞書・記号ポーズ辞書機能

声の職人で作成したフレーズ辞書・記号ポーズ辞書を読み込めるようになりました。



### ユーザー辞書機能

地名や名前などの特殊な読み方をする単語を辞書として登録・保存が可能です。



### ユーザー辞書機能の複数ロード機能

ユーザ単語辞書を同時に適用できるようになりました。



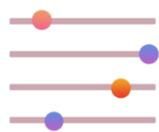
### ルビ機能

同じ漢字でも読み方が文章内で違う物等、ルビ記法で読み方を指定できます。



### ブックマーク機能

文中任意の位置にブックマークを設定できます。  
コールバック関数にてブックマークに関する情報を取得できます。



## 音声効果タグ機能

日本語テキスト中で、各種パラメーター値の指定や、ブックマークを付けることが可能です。



## 音声ファイル挿入機能

制御タグにより、入力テキスト中に任意のwavファイルを挿入できるようになりました。



## ポータビリティ

多彩なOS向けにポータビリティ可能です。

## 動作環境



### iOS

OS バージョン (iOS)	iOS 12.0 以上 (Xcode 15) その他バージョンについてはお問い合わせください。
RAM	日本語解析あり 約30MB~45MB (1話者使用時) ※話者1名につき約10~20MBの追加  日本語解析なし 約10~20MB (1話者使用時) ※話者1名につき約10~20MBの追加
ROM (Flash ROM)	日本語解析あり 65~100MB (1話者の場合) ※話者1名につき約40MBの追加  日本語解析なし 50MB (1話者使用時) ※話者1名につき約40MBの追加
開発言語	Swift/Objective-C からの C 関数呼び出しによる利用
ファイルフォーマット	16bit リニア PCM 形式 8bit リニア PCM 形式 8bit $\mu$ -Law PCM 形式
サンプリング周波数	24 kHz *1
対応中間言語フォーマット	AI 独自規格 JEITA TT-6004 規格

## 動作環境



### Android

OS バージョン (Android)	Android 4.1 (API レベル 16) 以上
メモリ	日本語解析あり 約30MB~45MB (1話者使用時)  ※話者1名につき約10~20MBの追加  日本語解析なし 約10~20MB (1話者使用時)  ※話者1名につき約10~20MBの追加
HDD	日本語解析あり 75~110MB (1話者の場合)  ※話者1名につき約40MBの追加  日本語解析なし 60MB (1話者の場合)  ※話者1名につき約40MBの追加
開発言語	Java
ファイル フォーマット	16bit リニア PCM 形式 8bit リニア PCM 形式 8bit $\mu$ -Law PCM 形式
サンプリング 周波数	24 kHz *1
対応中間言語 フォーマット	AI 独自規格 JEITA TT-6004 規格

# 動作環境



## Raspberry-Pi

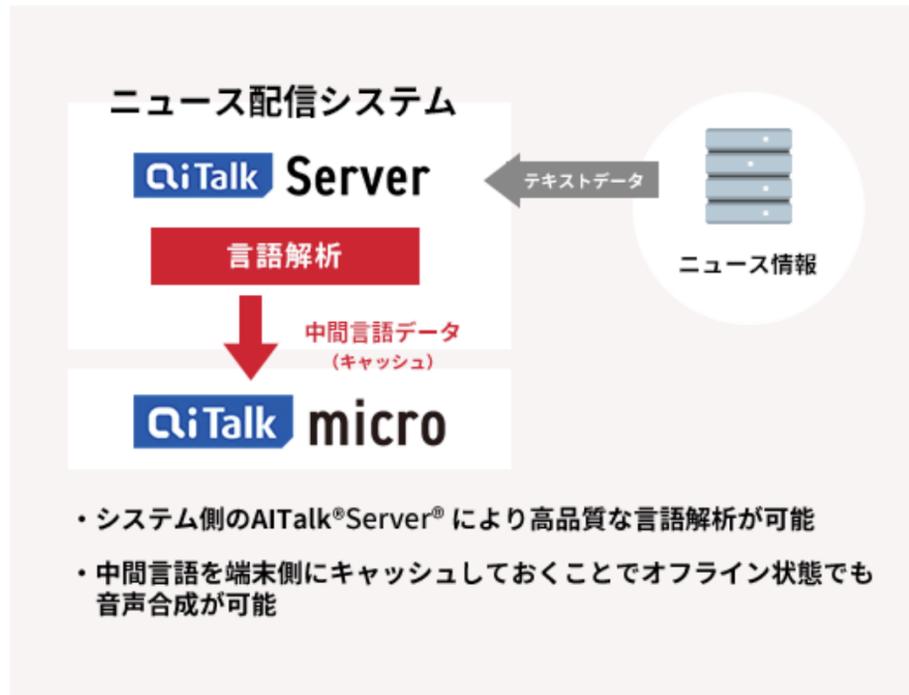
OS バージョン (Raspberry-Pi)	Raspbian Buster (kernel version 5.4) 以降 Ubuntu 18.04 (64bit) (kernel version 5.4) 以降 これより古い OS、kernel バージョンについてはお問い合わせください。
メモリ	日本語解析あり 約30MB～45MB (1話者使用時) ※話者1名につき約10～20MBの追加  日本語解析なし 約10～20MB (1話者使用時) ※話者1名につき約10～20MBの追加
HDD	日本語解析あり 75～110MB (1話者の場合) ※話者1名につき約40MBの追加  日本語解析なし 60MB (1話者の場合) ※話者1名につき約40MBの追加
開発言語	C/C++ C# (Mono)  ※Java は動作実績がございません。別途お問い合わせください。
ファイル フォーマット	16bit リニア PCM 形式 8bit リニア PCM 形式 8bit $\mu$ -Law PCM 形式
サンプリング 周波数	24kHz *1
対応中間言語 フォーマット	AI 独自規格 JEITA TT-6004 規格

\*1：記載以外のサンプリング周波数の音声辞書については別途お問い合わせください。付属するリサンプリング処理機能によって 48 kHz, 44 kHz, 24kHz, 22 kHz, 16 kHz, 8 kHz に変換可能です。

※ 記載以外の開発環境は別途お問い合わせください。

## 開発時のシステム構成イメージ

### Serverとmicroを活用したハイブリッドシステム（ニュース配信）



### 外部サービスとの連携（音声対話システム）

