

令和5(2023)年度

# 音声合成 関連製品の動向調査

音声合成・システム開発者向け製品

この分冊に掲載されている調査結果

他の分冊に掲載されている調査結果

(第1分冊) ミドルウェア・SDK

(第2分冊) サービス・ツール・その他製品

製品カテゴリ		
A-1	A-2	A-3
組込用 ミドルウェア・SDK	汎用OS用 ミドルウェア・SDK	サーバ用 ミドルウェア

製品カテゴリ			
A-4	A-5	A-6	A-7
webサービス	ツール製品	ハードウェア製品	その他 製品サービス

カテゴリ 開始ページ数

p. 2	p. 9	p. 21
------	------	-------

p. 27	p. 33	p. 42	p. 45
-------	-------	-------	-------

ベンダ名

AHS				p.34		
Cerence Japan		p.10				
Fairy Devices				p.28		
Gatebox						p.46
HOYA MD部門 ReadSpeaker	p.3	p.11	p.22	p.29	p.35	p.47
NTTテクノクロス		p.12	p.23	p.30	p.36	p.48
アクエスト	p.4	p.13				p.43
エアアイ	p.5	p.14	p.24	p.31	p.37	p.49
オープンソース(名古屋工業大学)						p.50
クリエートシステム開発	p.6	p.15				
スカイフィッシュ					p.38	p.51
セイコーエプソン						p.44
東芝デジタルソリューションズ	p.7	p.16		p.32	p.39	
日立ケーイーシステムズ		p.17				
日立ソリューションズ・クリエイト		p.18			p.40	
日立ソリューションズ・テクノロジー	p.8	p.19	p.25			
日立産業制御ソリューションズ		p.20			p.41	
ヤマハ						p.52

実施: JEITA 音声入出力方式標準化専門委員会

<https://www.jeita-speech.org/survey/survey-top.html>

調査実施: 2023年12月末時点

発行: 2024年10月

## 音声合成・システム開発者向け製品

### 製品カテゴリ A-1

	ページ
組込み用ミドルウェア・SDK	2
HOYA MD部門 ReadSpeaker	3
アクエスト	4
エーアイ	5
クリエートシステム開発	6
東芝デジタルソリューションズ	7
日立ソリューションズ・テクノロジー	8

表A-1. 組み込み用ミドルウェア・SDK

HOYA株式会社 MD部門 ReadSpeaker

製品分類・製品名	speechEngine SDK Embedded
製品情報のURL	<a href="https://readspeaker.jp/products/development/speechengine-emb/">https://readspeaker.jp/products/development/speechengine-emb/</a>
製品概要 (100字程度)	スマートフォン・組み込み端末のアプリケーション開発に適したスタンドアロン型の音声合成SDKです。
音声合成の機能・内部仕様に関する主な特徴・機能 (200字程度)  ・言語処理機能の特徴 ・ユーザ辞書への対応の有無 ・合成方式の特徴 など。	DNN音声合成をベースにした、Android、iOSや各種組み込み機器向けのアプリケーションソフトの開発に適したスタンドアロン型の音声合成SDK製品です。最先端の「ディープラーニング」の導入により、全体的な品質が向上し、感情表現がより繊細で自然になりました。また従来の技術では難しかった相槌・語尾・言い直しなどの細かいニュアンスの表現力が大幅に向上しました。ディープラーニングは高いマシンスペックを要求されますが、ReadSpeakerは独自の改良を重ねリソースが少ない環境でもご利用いただけます。
音声合成の詳細仕様  ・入力テキストのフォーマット (文字コード・JEITA規格・SSML対応、など) ・合成可能な言語 ・声色のバリエーション ・制御可能な項目 (話速・ピッチ、など) ・出力音声の形式 など。	・入力テキストのフォーマット:UTF-8、SSML対応 ・合成可能な言語 44言語 ・制御可能な項目:話速・ピッチ・音量・アクセント ・出力音声の形式:16bit LPCM、16bit LPCM Wave
動作要件(装置仕様)  ソフトウェア構成(対応OSなど) ハードウェア構成(メモリ要件など) など。	Android 10.x以降 iOS 14.x以降 Embedded-Linux  RAM 10MB～ ROM 11MB～
発表・発売時期	2010年
前年(2022年度調査)からの製品の更新内容	動作要件
2023年中のプレスリリースURL	<a href="https://readspeaker.jp/news/">https://readspeaker.jp/news/</a>
価格(税別)	個別相談
連絡先	<a href="https://readspeaker.jp/">https://readspeaker.jp/</a> TEL:03-6479-7182
備考	

表A-1. 組み込み用ミドルウェア・SDK

株式会社アクエスト

製品分類・製品名	AquesTalk10 (音声合成部)	AquesTalk pico (音声合成部)
製品情報のURL	<a href="https://www.a-quest.com/products/aquestalk10.html">https://www.a-quest.com/products/aquestalk10.html</a>	<a href="https://www.a-quest.com/products/aquestalk_pico.html">https://www.a-quest.com/products/aquestalk_pico.html</a>
製品概要 (100字程度)	AquesTalk1、AquesTalk2の後継エンジンです。これまでの小型軽量の特徴を維持したまま、クリアな音質、高い明瞭性、各種パラメータで様々な声質に調整が可能などの特徴を持っています。 ARM Cortexなどの32bitのシステムへの組み込みに最適です。	世界最小サイズの音声合成エンジンです。8bitマイコンでもリアルタイムの音声合成ができます。リソース制約の厳しいワンチップマイコン構成などの小型の組み込み機器に適しています。
音声合成の機能・内部仕様に関する主な特徴・機能 (200字程度)  ・言語処理機能の特徴 ・ユーザー辞書への対応の有無 ・合成方式の特徴 など。	言語処理機能は別途AqKanji2Koe-A/Mライブラリを使用 ユーザー辞書機能有	言語処理機能は別途AqKanji2Koe-A/Mライブラリを使用 ユーザー辞書機能有
音声合成の詳細仕様  ・入力テキストのフォーマット (文字コード・JEITA規格・SSML対応、など) ・合成可能な言語 ・声色のバリエーション ・制御可能な項目 (話速・ピッチ、など) ・出力音声の形式 など。	入力テキスト かな表記音声記号列 音声出力仕様 16KHzサンプリング, 16bitPCM 声種 基本素片: 女声2種, 男性1種から選択(複数可)+ パラメータ調整によるオリジナル声種 制御項目 話速、ピッチ	入力テキスト ローマ字表記音声記号列 音声出力仕様 8KHz/10KHzサンプリング 16bitPCM 声種 女声4種/男声1種/ロボ声2種から選択
動作要件(装置仕様)  ソフトウェア構成(対応OSなど) ハードウェア構成(メモリ要件など) など。	対応OS 各種OS(OSに非依存) 関数インターフェース C関数呼び出し 処理量(CPU) 20DMIPS(各種32/64bit CPU) メモリ容量(RAM) 11KB(ワーク:9KB, 出力バッファ:640B, Stack:700B) メモリ容量(ROM) 134KB~321KB(プログラム+音素片データ1つ~3つ)	対応OS 各種OS(OSに非依存) 関数インターフェース C関数呼び出し 処理量(CPU) 8bit 4MHz以上 メモリ容量(RAM) 500byte(stack含む) メモリ容量(ROM) 23Kbyte(最小構成)
発表・発売時期	2017/10/10	2011/10/17
前年(2022年度調査)からの製品の更新内容		
2023年中のプレスリリースURL		
価格(税別)		
連絡先	(株)アクエスト infoaq@a-quest.com	(株)アクエスト infoaq@a-quest.com
備考		

表A-1. 組み込み用ミドルウェア・SDK

株式会社エーアイ

製品分類・製品名	AITalk micro	
製品情報のURL	<a href="https://www.ai-j.jp/products/micro/">https://www.ai-j.jp/products/micro/</a>	
製品概要 (100字程度)	様々なOS、CPFで組み込み可能な高品質音声合成エンジンです。	
音声合成の機能・内部仕様に関する主な特徴・機能 (200字程度)	<ul style="list-style-type: none"> <li>●iOS/Android/組込Linux に対応</li> <li>●選べる話者</li> <li>※AITalk CustomVoiceで作成したオリジナルの話者も利用可能</li> <li>●感情調整機能</li> <li>●低容量・低リソース</li> <li>●汎用PC向けSDKとの統一API</li> <li>●単語辞書・記号ポーズ機能</li> <li>●ルビ機能</li> <li>●ブックマーク機能</li> <li>●音声ファイル挿入機能</li> </ul>	
音声合成の詳細仕様	<ul style="list-style-type: none"> <li>●入力形態 漢字仮名混じり文 韻律記号付き文字列(AI 独自規格orJEITA TT-6004 規格)</li> <li>●文字コード 日本語 Shift_JIS(CP932) 日本語 UTF-8</li> <li>●合成可能な言語/声色のバリエーション 標準語: 女性8名、男性4名、男の子2名、女の子2名 関西弁: 女性1名、男性1名</li> <li>※AITalk CustomVoiceで作成したオリジナルの話者も利用可能</li> <li>●制御可能な項目 音高: 可変(標準の0.5~2.0倍) 話速: 可変(標準の0.5~2.0倍) 抑揚: 可変(標準の0.5~2.0倍) 音量: 可変(無音0.0~2.0)</li> </ul>	
動作要件(装置仕様)	<p>[ソフトウェア構成]</p> <ul style="list-style-type: none"> <li>・Android 4.1 (API レベル 16) 以上</li> <li>・iOS 10.0 以上</li> <li>・Raspbian Buster (kernel version 5.4) 以降</li> <li>・Ubuntu 18.04 (64bit) (kernel version 5.4) 以降</li> </ul> <p>※上記以外のプラットフォームの対応についてはお問い合わせください。</p>	
	<p>[ハードウェア構成]</p> <p>【Android・Raspberry-Pi】</p> <ul style="list-style-type: none"> <li>●RAM(日本語解析あり) 約 30 MB ~ 45 MB (1 話者使用時) ※話者1名につき約 10 ~ 20 MBの追加</li> <li>●RAM(日本語解析なし) 約 10 ~ 20 MB (1 話者使用時) ※話者1名につき約 10 ~ 20 MBの追加</li> </ul> <p>【Android・Raspberry-Pi】</p> <ul style="list-style-type: none"> <li>●ROM/HDD(日本語解析あり) 75 ~ 110 MB (1話者の場合) ※話者1名につき約 40 MBの追加</li> <li>●ROM/HDD(日本語解析なし) 60 MB (1話者の場合) ※話者1名につき約 40 MBの追加</li> </ul>	<p>【iOS】</p> <ul style="list-style-type: none"> <li>●RAM(日本語解析あり) 約 30 MB ~ 45 MB (1 話者使用時) ※話者1名につき約 10 ~ 20 MBの追加</li> <li>●RAM(日本語解析なし) 約 10 ~ 20 MB (1 話者使用時) ※話者1名につき約 10 ~ 20 MBの追加</li> </ul> <p>【iOS】</p> <ul style="list-style-type: none"> <li>●ROM/HDD(日本語解析あり) 65 ~ 100 MB (1話者の場合) ※話者1名につき約 40 MBの追加</li> <li>●ROM/HDD(日本語解析なし) 50 MB (1話者の場合) ※話者1名につき約 40 MBの追加</li> </ul>
発表・発売時期	2021/4/27	
前年(2022年度調査)からの製品の更新内容		
2023年中のプレスリリースURL	<a href="https://www.ai-j.jp/topics/">https://www.ai-j.jp/topics/</a>	
価格(税別)	開発環境:50万円~(税別) ※別途サポート費用が発生 ※再配布のロイヤリティは個別相談	
連絡先	TEL: 03-6801-8461 Mail: info@ai-j.jp HP: https://www.ai-j.jp	
備考	上記仕様以外の環境で利用する場合は、ボーディング作業(別途費用)が発生します。	

表A-1. 組み込み用ミドルウェア・SDK

クリエイトシステム開発株式会社

製品分類・製品名	ドキュメントトーカー開発システム 日本語音声合成エンジン for WindowsCE	ドキュメントトーカー たかし ドキュメントトーカー けいこ
製品情報のURL	<a href="https://www.createsystem.co.jp/main/downloadg.html">https://www.createsystem.co.jp/main/downloadg.html</a>	<a href="https://www.createsystem.co.jp/main/dtalkerAndroidSDK.html">https://www.createsystem.co.jp/main/dtalkerAndroidSDK.html</a>
製品概要 (100字程度)	Microsoft Corp. のWindows CE 上で動作する日本語音声合成ライブラリです。	アンドロイド標準TTSに準拠した高品質日本語音声合成エンジンです。
音声合成の機能・内部仕様に関する主な特徴・機能 (200字程度)  ・言語処理機能の特徴 ・ユーザ辞書への対応の有無 ・合成方式の特徴 など。	<ul style="list-style-type: none"> <li>日本語テキストから表音文字を生成する言語処理ライブラリ</li> <li>表音文字から音声波形データ(PCM16ビット/8kHz/16kHz/モノラル)を生成する波形処理ライブラリ</li> <li>生成された音声波形データを音として出力する音声合成出力ライブラリ</li> <li>ユーザー辞書登録機能</li> </ul>	アンドロイド標準TTSに準拠しているため、標準TTSを使用したアプリケーションで高品質な音声を利用できます。
音声合成の詳細仕様  ・入力テキストのフォーマット (文字コード・JEITA規格・SSML対応、など) ・合成可能な言語 ・声色のバリエーション ・制御可能な項目 (話速・ピッチ、など) ・出力音声の形式 など。	<p>入力形態:漢字仮名混じり文/単語/英単語 合成可能言語:日本語 文字コード:UNICODE JEITA TT-6004対応:準拠 JEITA IT-4006対応:準拠 辞書:基本:16万語 英単語:27000語 記号:380種 住所:73000語 漢字詳細読み:6600種 ユーザ辞書機能:有 合成単位:1ピッチ波形編集 男女各約15000種 声:男声/女声 声の高さ:可変(5段階) 発声速度:可変(10段階) 音量:10段階 トーン: その他の機能:歌声合成, 詳細読み機能 出力音声の形式:16kHz,16ビット,モノラル</p>	<p>入力形態:漢字仮名混じり文/単語/英単語 合成可能言語:日本語 文字コード:UNICODE JEITA TT-6004対応:準拠 JEITA IT-4006対応:準拠 SSML:準拠 辞書:基本:16万語 英単語:27000語 記号:380種 住所:73000語 漢字詳細読み:6600種 ユーザ辞書機能:有 合成単位:可変長音素列 男女各約15000種 コーパスベース 波形編集 声:男声:たかし 女声:けいこ 声の高さ:可変(5段階) 発声速度:可変(10段階) 話速変換:1.0倍~0.5倍 音量:10段階 トーン: その他の機能:歌声合成, 詳細読み機能 出力音声の形式:16kHz,16ビット,モノラル</p>
動作要件(装置仕様)  ソフトウェア構成(対応OSなど) ハードウェア構成(メモリ要件など) など。	<p>ソフトウェア環境: Windows Embeddd 開発環境: Windows Embedded Compact 7" ハードウェア環境:各種PDA 辞書:8MB</p>	<p>開発環境: Android Studio ソフトウェア環境:Android ハードウェア環境:Android端末 辞書:34MB</p>
発表・発売時期	2002年11月	2012年11月
前年(2022年度調査)からの製品の更新内容		2023/06 V2.3.5 Android13対応
2023年中のプレスリリースURL		
価格(税別)	SDK:31,500円(税込み)	ロイヤリティ:100万円/プロジェクト
連絡先	web-info@createsystem.co.jp, <a href="https://www.createsystem.co.jp/">https://www.createsystem.co.jp/</a>	web-info@createsystem.co.jp, <a href="https://www.createsystem.co.jp/">https://www.createsystem.co.jp/</a>
備考	音声合成波形生成エンジンは富士通(株)製をWindowsCE用に移植 音声合成言語処理は独自の処理方式を採用 PDAに合わせて対応	音声合成波形生成エンジンは富士通(株)製をAndroid用に移植 音声合成言語処理は独自の処理方式を採用 Google Play よりダウンロード可能

表A-1. 組み込み用ミドルウェア・SDK

東芝デジタルソリューションズ株式会社

製品分類・製品名	音声合成MW ToSpeakG3 TSPG3-TTS	日本語音声合成MW ToSpeakG4 TSPG4-TTS JP	音声合成MW ToSpeakGx Neo TSPGX-TTS
製品情報のURL	<a href="https://www.global.toshiba/jp/products-solutions/ai-iot/recaius/lineup/tospeak.html">https://www.global.toshiba/jp/products-solutions/ai-iot/recaius/lineup/tospeak.html</a>	<a href="https://www.global.toshiba/jp/products-solutions/ai-iot/recaius/lineup/tospeak.html">https://www.global.toshiba/jp/products-solutions/ai-iot/recaius/lineup/tospeak.html</a>	<a href="https://www.global.toshiba/jp/products-solutions/ai-iot/recaius/lineup/tospeak.html">https://www.global.toshiba/jp/products-solutions/ai-iot/recaius/lineup/tospeak.html</a>
製品概要 (100字程度)	自然で安定した声質の合成音声、機器ローカルの処理でテキストから生成できる組み込みミドルウェア。組み込み環境に好適で多言語に対応。	自然で安定した声質の合成音声、機器ローカルの処理でテキストから生成できる組み込みミドルウェア。感情、声質パラメータの制御が可。	自然で安定した声質の合成音声、機器ローカルの処理でテキストから生成できる組み込みミドルウェア。肉声感の高い高音質版。
音声合成の機能・内部仕様に関する主な特徴・機能 (200字程度)	<ul style="list-style-type: none"> <li>・ユーザー辞書機能：可能</li> <li>・SSML対応：可能(言語により対応機能の範囲は異なります)</li> </ul>	<ul style="list-style-type: none"> <li>・ユーザー辞書機能：可能</li> <li>・SSML対応：可能</li> </ul>	<ul style="list-style-type: none"> <li>・ユーザー辞書機能：可能</li> <li>・SSML対応：可能(言語により対応機能の範囲は異なります)</li> </ul>
音声合成の詳細仕様	<p>[入力形態] プレーンテキスト/表音文字列</p> <p>[合成可能言語] 日本語/アメリカ英語(米語)/中国語(北京語)/韓国語/ほか。</p> <p>標準製品は最大11言語に対応可能。 (個別対応の場合最大30言語に対応可能)</p> <p>[声色のバリエーション] 男・女。日本語・アメリカ英語(米語)はオプションボイスも用意あり</p> <p>[制御可能な項目] 声の高さ、発話速度、など</p> <p>[出力音声の形式] モノラル、22.050kHz、16bit</p>	<p>[入力形態] プレーンテキスト/表音文字列</p> <p>[合成可能言語] 日本語</p> <p>[声色のバリエーション] 男・女</p> <p>[制御可能な項目] 声の高さ、発話速度、感情、声質パラメータなど。</p> <p>[出力音声の形式] モノラル、22.050kHz、16bit</p>	<p>[入力形態] プレーンテキスト/表音文字列</p> <p>[合成可能言語] 日本語/アメリカ英語(米語)/中国語(北京語)</p> <p>[声色のバリエーション] 男・女。日本語・アメリカ英語(米語)はオプションボイスも用意あり</p> <p>[制御可能な項目] 声の高さ、発話速度、など</p> <p>[出力音声の形式] モノラル、44.100kHz、16bitなど</p>
動作要件(装置仕様)	<p>[ソフトウェア環境] Linux, iOS, Androidなど 個別相談</p> <p>[ハードウェア環境] ARM系ほか 個別相談</p>	<p>[ソフトウェア環境] Linux, iOS, Android</p> <p>[ハードウェア環境] ARM系</p>	<p>[ソフトウェア環境] Linux, iOS, Androidなど</p> <p>[ハードウェア環境] ARM系</p>
発表・発売時期	発売中	発売中	発売中
前年(2022年度調査)からの製品の更新内容	特になし	特になし	特になし
2023年中のプレスリリースURL			
価格(税別)	(個別相談)	(個別相談)	(個別相談)
連絡先	東芝デジタルソリューションズ株式会社 デジタルエンジニアリングセンター リカياسビジネス推進部  <a href="https://www.toshiba-sol.co.jp/pro/recaius/contact.html">https://www.toshiba-sol.co.jp/pro/recaius/contact.html</a>	東芝デジタルソリューションズ株式会社 デジタルエンジニアリングセンター リカياسビジネス推進部  <a href="https://www.toshiba-sol.co.jp/pro/recaius/contact.html">https://www.toshiba-sol.co.jp/pro/recaius/contact.html</a>	東芝デジタルソリューションズ株式会社 デジタルエンジニアリングセンター リカياسビジネス推進部  <a href="https://www.toshiba-sol.co.jp/pro/recaius/contact.html">https://www.toshiba-sol.co.jp/pro/recaius/contact.html</a>
備考	ご指定CPU/OSへの移植はご相談ください。		

表A-1. 組み込み用ミドルウェア・SDK

株式会社日立ソリューションズ・テクノロジー

製品分類・製品名	Ruby Talk 6.0	Micro Ruby Talk
製品情報のURL	<a href="https://www.hitachi-solutions-tech.co.jp/iot/solution/voice/Ruby_Talk/index.html">https://www.hitachi-solutions-tech.co.jp/iot/solution/voice/Ruby_Talk/index.html</a>	<a href="https://www.hitachi-solutions-tech.co.jp/iot/solution/voice/Ruby_Talk/index.html">https://www.hitachi-solutions-tech.co.jp/iot/solution/voice/Ruby_Talk/index.html</a>
製品概要 (100字程度)	「Ruby Talk」は、漢字かな交じりテキスト文から、肉声並みの高品位音声メッセージを生成します。クラウドサーバから低リソースの組み込み機器まで、幅広い用途でご使用いただける音声合成ソリューションです。	「Ruby Talk」は、漢字かな交じりテキスト文から、肉声並みの高品位音声メッセージを生成します。超低リソースの組み込み機器向け音声合成ミドルウェアです。
音声合成の機能・内部仕様に関する主な特徴・機能 (200字程度)  ・言語処理機能の特徴 ・ユーザー辞書への対応の有無 ・合成方式の特徴 など。	ユーザー辞書への対応 有 肉声感の高い自然な音声品質 DNN(Dep Neural Network)の導入により、安定した自然な韻律を実現 会話調読み上げ 語りかけ言葉の発話を合成音声で実現 カーナビゲーション専用音源 カーナビゲーションで使用される誘導音声に特化した高音質な音源を提供 ドメイン特化音声合成技術 特定分野で頻出するフレーズについて高音質な音声合成を実現	ユーザー辞書への対応 有 肉声感の高い自然な音声品質 DNN(Dep Neural Network)の導入により、安定した自然な韻律を実現 会話調読み上げ 語りかけ言葉の発話を合成音声で実現 カーナビゲーション専用音源 カーナビゲーションで使用される誘導音声に特化した高音質な音源を提供 ドメイン特化音声合成技術 特定分野で頻出するフレーズについて高音質な音声合成を実現
音声合成の詳細仕様  ・入力テキストのフォーマット (文字コード・JEITA規格・SSML対応、など) ・合成可能な言語 ・声色のバリエーション ・制御可能な項目 (話速・ピッチ、など) ・出力音声の形式 など。	入力テキストのフォーマット 漢字仮名交じり文/発話表記コード(JEITA TT-6004準拠テキスト、日立韻律テキスト) 合成可能な言語 日本語 声色のバリエーション 11話者(女性7(内子供1)、男性4(内子供1)) アナウンス、エンターテインメント読み上げなど、幅広い用途をカバー 制御可能な項目 声の高さ(10段階)、話速(10段階)、ピッチ(10段階)、アクセント(10段階)、抑揚(10段階) 出力音声の形式 22.05kHz, 16bit, モノラル	入力テキストのフォーマット 漢字仮名交じり文/発話表記コード(JEITA TT-6004準拠テキスト、日立韻律テキスト) 合成可能な言語 日本語 声色のバリエーション 11話者(女性7(内子供1)、男性4(内子供1)) アナウンス、エンターテインメント読み上げなど、幅広い用途をカバー 制御可能な項目 声の高さ(10段階)、話速(10段階)、ピッチ(10段階)、アクセント(10段階)、抑揚(10段階) 出力音声の形式 22.05kHz, 16bit, モノラル
動作要件(装置仕様)  ソフトウェア構成(対応OSなど) ハードウェア構成(メモリ要件など) など。	ハードウェア環境 CPU:SH、ARM等の32ビット、64ビットCPU以上推奨 メモリ(RAM):8MB以上 メモリ(ROM):3MB以上	お問い合わせください
発表・発売時期	2018年10月1日	2008年10月1日
前年(2022年度調査)からの製品の更新内容	無し	無し
2023年中のプレスリリースURL	無し	無し
価格(税別)	弊社営業にお問い合わせください。	
連絡先	<a href="https://www8.hitachi.co.jp/inquiry/hitachi-solutions-tech/products/jp/form.jsp">https://www8.hitachi.co.jp/inquiry/hitachi-solutions-tech/products/jp/form.jsp</a>	
備考	日立ソリューションズ・テクノロジーは、音声合成ミドルウェアのトータル・ソリューションとして、組み込み機器で音声合成を使用するために必要なソフトウェア/ハードウェアを提供します。また、機器への組み込み、独自素片データの作成、コンサルテーションなどの技術サービスも提供致します。	

製品カテゴリ A-2

	ページ
汎用OS用ミドルウェア・SDK	9
Cerence Japan	10
HOYA MD部門 ReadSpeaker	11
NTTテクノクロス	12
アクエスト	13
エーアイ	14
クリエートシステム開発	15
東芝デジタルソリューションズ	16
日立ケーイーシステムズ	17
日立ソリューションズ・クリエイト	18
日立ソリューションズ・テクノロジー	19
日立産業制御ソリューションズ	20

表A-2. 汎用OS用ミドルウェア・SDK

Cerence Japan株式会社

製品分類・製品名	Cerence SDK v6
製品情報のURL	
製品概要 (100字程度)	Cerence SDK (CSDK) は、自動車向け音声アシスタントをはじめとする、さまざまな音声アプリケーション開発に最適化されたシンプルな開発スイートです。TTSエンジン (Cerence TTS) と音声認識エンジンを梱包しています。
音声合成の機能・内部仕様に関する主な特徴・機能 (200字程度)  ・言語処理機能の特徴 ・ユーザ辞書への対応の有無 ・合成方式の特徴 など。	ユーザ辞書、正規表現によるテキスト前処理ルールをサポートしています。 音声合成の基本方式は、目的・リソースに応じて、波形接続方式、からニューラルネットワーク方式まで種々選択いただけます。 また、辞書編集ツール、録音音声とTTSの統合プロンプタ生成機能など豊富な機能を提供可能です。
音声合成の詳細仕様  ・入力テキストのフォーマット (文字コード・JEITA規格・SSML対応、など) ・合成可能な言語 ・声色のバリエーション ・制御可能な項目 (話速・ピッチ、など) ・出力音声の形式 など。	入力フォーマット：テキスト、SSML対応、日本語は、JEITA音声合成記号をサポート。独自の制御記号による制御も可能。欧州言語ではマルチ言語もサポート 言語：60言語以上 (お問合せください) 声色：各言語、男女など複数のボイスを用意しています。 形式：22.05 kHz sampling, 16 bit PCM
動作要件(装置仕様)  ソフトウェア構成(対応OSなど) ハードウェア構成(メモリ要件など) など。	Windows, iOS, Android, Linux その他OS にも対応可能 種々のメモリサイズのボイスデータを用意していますので、高音質システムから、省資源システムまで用途に応じた実現が可能です。
発表・発売時期	以前よりの製品を順次アップデートしております。
前年(2022年度調査)からの製品の更新内容	
2023年中のプレスリリースURL	
価格(税別)	お問合せください。
連絡先	<a href="mailto:japan@cerence.com">japan@cerence.com</a>
備考	サーバー型TTSも用意しております。

表A-2. 汎用OS用ミドルウェア・SDK

HOYA株式会社 MD部門 ReadSpeaker

製品分類・製品名	speechEngine SDK
製品情報のURL	<a href="https://readspeaker.jp/products/development/speechengine/">https://readspeaker.jp/products/development/speechengine/</a>
製品概要 (100字程度)	Windows・Linuxのアプリケーション開発に適したスタンドアロン型の音声合成SDKです。
音声合成の機能・内部仕様に関する主な特徴・機能 (200字程度)  ・言語処理機能の特徴 ・ユーザ辞書への対応の有無 ・合成方式の特徴 など。	高速処理な処理を実現しており高パフォーマンスなシステム構築が可能です。マルチスレッド対応により、複数の音声を同時に処理するシステムの構築に利用いただけます。 シンプルで分かりやすいAPI群は、言語・話者を問わず共通化され、効率の良い組み込み開発を実現します。高音質を追求した波形接続型や感情音声合成が扱えるDNN合成方式のエンジンをご利用いただけます。開発言語はC/C++/.NETに対応し、システムやユーザーに合わせ最適な環境を選択できます。
音声合成の詳細仕様  ・入力テキストのフォーマット (文字コード・JEITA規格・SSML対応、など) ・合成可能な言語 ・声色のバリエーション ・制御可能な項目 (話速・ピッチ、など) ・出力音声の形式 など。	・入力テキストのフォーマット:UTF-8、SSML対応 ・合成可能な言語 44言語 ・制御可能な項目:話速・ピッチ・音量・アクセント ・出力音声の形式:16bit LPCM,16bit LPCM Wave,8bit A-law PCM,8bit A-law PCM Wave,8bit $\mu$ -law PCM,8bit $\mu$ -law PCM Wave,8bit $\mu$ -law PCM SUN AU,8bit unsigned LPCM Wave,4bit Dialogic ADPCM
動作要件(装置仕様)  ソフトウェア構成(対応OSなど) ハードウェア構成(メモリ要件など) など。	Windows 10 / 11 RHEL 7.x / 8.x / 9.x CentOS 7.x / 8.x  RAM 50MB～ ROM 180MB～800MB
発表・発売時期	2004年
前年(2022年度調査)からの製品の更新内容	
2023年中のプレスリリースURL	<a href="https://readspeaker.jp/news/">https://readspeaker.jp/news/</a>
価格(税別)	個別相談
連絡先	<a href="https://readspeaker.jp/">https://readspeaker.jp/</a> TEL:03-6479-7182
備考	

表A-2. 汎用OS用ミドルウェア・SDK

NTTテクノクロス株式会社

製品分類・製品名	FutureVoice Crayon SDK	
製品情報のURL	<a href="https://www.futurevoice.jp/lineup/sdk/">https://www.futurevoice.jp/lineup/sdk/</a>	
製品概要 (100字程度)	パッケージソフトやロボットやデジタルサイネージ等に最適なスタンドアロン型の音声合成SDK(ソフトウェア開発キット)	
音声合成の機能・内部仕様に関する主な特徴・機能 (200字程度)	<ul style="list-style-type: none"> <li>・人の声と遜色ない自然な音声</li> <li>・バリエーション豊富な話者を選択可能、複数話者の音声を混合し独自話者の作成も可能</li> <li>・少量の音声データから誰の声でも音声として再現</li> <li>・音声のチューニング・導入が簡単</li> <li>・様々な機器へ搭載可能</li> <li>・クロスリンガル音声合成技術に対応(話者の声質を損なわずに一言語の音声から多言語の合成音声を生成できる技術)</li> </ul>	
音声合成の詳細仕様	<ul style="list-style-type: none"> <li>・入力テキストのフォーマット (文字コード・JEITA規格・SSML対応、など)</li> <li>・合成可能な言語</li> <li>・声色のバリエーション</li> <li>・制御可能な項目 (話速・ピッチ、など)</li> <li>・出力音声の形式 など。</li> </ul>	<ul style="list-style-type: none"> <li>・入力フォーマット: プレーンテキスト・読み指定テキスト・SSML</li> <li>・文字コード: UTF-8</li> <li>・ユーザ辞書機能: 単語辞書、文辞書、テキスト置換辞書、音声置換辞書、韻律テンプレート辞書</li> <li>・対応言語: 日本語・英語・中国語(普通話/簡体字・広東語/繁体字)・韓国語</li> <li>・出力サンプリングレート: 8,000Hz・22,050Hz・44,100Hz</li> <li>・出力音声ファイル形式: 16bit PCM(wav/raw)・8bit <math>\mu</math>-law(wav/raw)</li> </ul>
動作要件(装置仕様)	<p>ソフトウェア構成(対応OSなど) ハードウェア構成(メモリ要件など) など。</p>	<p>【サーバ・デスクトップ等】</p> <ul style="list-style-type: none"> <li>・OS: Rocky 9.x・RHEL 9.x</li> <li>・メモリ: 4GB以上の空き推奨</li> <li>・システム空き容量: 10GB以上の空き推奨</li> </ul> <p>【省リソース・組み込み機器等】</p> <ul style="list-style-type: none"> <li>・OS: Rocky 9.x・RHEL 9.x</li> <li>・メモリ: 30MB～256MB 程度</li> <li>・システム空き容量: 30MB～200MB 程度</li> </ul>
発表・発売時期	2017年12月	
前年(2022年度調査)からの製品の更新内容	—	
2023年中のプレスリリースURL	—	
価格(税別)	個別相談	
連絡先	E-mail: <a href="mailto:futurevoice.info-ml@ntt-tx.co.jp">futurevoice.info-ml@ntt-tx.co.jp</a>	
備考		

表A-2. 汎用OS用ミドルウェア・SDK

株式会社アクエスト

製品分類・製品名	AqKanji2Koe-A (言語処理部)	AquesTalk10 (音声合成部)
製品情報のURL	<a href="https://www.a-quest.com/products/aqkanji2koe.html">https://www.a-quest.com/products/aqkanji2koe.html</a>	<a href="https://www.a-quest.com/products/aquestalk10.html">https://www.a-quest.com/products/aquestalk10.html</a>
製品概要 (100字程度)	漢字かな混じり文のテキストを、AquesTalk用の音声記号列に変換します。漢字に“よみ”を付与するだけでなく、適切なアクセントや文節の句切りを付与することができます。	AquesTalk1、AquesTalk2の後継エンジンです。これまでの小型軽量の特徴を維持したまま、クリアな音質、高い明瞭性、各種パラメータで様々な声質に調整が可能などの特徴を持っています。
音声合成の機能・内部仕様に関する主な特徴・機能 (200字程度)  ・言語処理機能の特徴 ・ユーザ辞書への対応の有無 ・合成方式の特徴 など。	ユーザー辞書機能有	言語処理機能は別途AqKanji2Koe-A/Mライブラリを使用
音声合成の詳細仕様  ・入力テキストのフォーマット (文字コード・JEITA規格・SSML対応、など) ・合成可能な言語 ・声色のバリエーション ・制御可能な項目 (話速・ピッチ、など) ・出力音声の形式 など。	入力データ形式 漢字かな混じり文テキスト (UTF8/UTF16/UTF32/SJIS) 出力データ形式 AquesTalk用音声記号列(かな表記) 関数I/F C関数呼び出し	入力テキスト ローマ字表記音声記号列 音声出力仕様 8KHz/10KHzサンプリング 16bitPCM 声種 女声4種/男声1種/ロボ声2種から選択
動作要件(装置仕様)  ソフトウェア構成(対応OSなど) ハードウェア構成(メモリ要件など) など。	OS Windows, Linux, Mac, iOS, Android コードサイズ 約200KB *Windows版での実測値 辞書サイズ 10MB(約38万語) 処理速度: 約20万字/秒 (Windows PCで計測)	対応OS 各種OS(OSに非依存) 関数インターフェース C関数呼び出し 処理量(CPU) 8bit 4MHz以上 メモリ容量(RAM) 500byte(stack含む) メモリ容量(ROM) 23Kbyte(最小構成)
発表・発売時期	2022/12/6	2017/10/10
前年(2022年度調査)からの製品の更新内容		
2023年中のプレスリリースURL		
価格(税別)		
連絡先	(株)アクエスト infoaq@a-quest.com	(株)アクエスト infoaq@a-quest.com
備考		

表A-2. 汎用OS用ミドルウェア・SDK

株式会社エーアイ

製品分類・製品名	AITalk SDK	
製品情報のURL	<a href="https://www.ai-i.jp/products/sdk/">https://www.ai-i.jp/products/sdk/</a>	
製品概要 (100字程度)	AITalk SDKは、人間らしく自然な音声で自由に音声合成をすることができ、ライブラリ(dllまたはso)でご提供する組み込み用途向け音声合成エンジンです。	
音声合成の機能・内部仕様に関する主な特徴・機能 (200字程度)	<ul style="list-style-type: none"> <li>●60種類以上の言語に対応</li> <li>●選べる日本語話者</li> <li>※AITalk CustomVoiceで作成したオリジナルの話者も利用可能</li> <li>●選べる合成方式(波形接続音声、新DNN方式音声)</li> <li>●感情調整機能</li> <li>●イントネーション調整登録機能</li> <li>●音量・話速・声の高さ・抑揚設定</li> <li>●ポーズ長設定</li> <li>●単語辞書登録機能</li> <li>●マルチスレッド機能</li> </ul>	
音声合成の詳細仕様	<ul style="list-style-type: none"> <li>●入力形態 漢字仮名混じり文 韻律記号付き文字列(AI独自規格orJEITA TT-6004規格)</li> <li>●文字コード 日本語 Shift_JIS(CP932) 日本語 UTF-8</li> <li>●合成可能な言語/声色のバリエーション 標準語:女性8名、男性4名、男の子2名、女の子2名 関西弁:女性1名、男性1名 外国語:60種類以上</li> <li>※AITalk CustomVoiceで作成したオリジナルの話者も利用可能</li> <li>●制御可能な項目 音高:可変(標準の0.5~2.0倍) 話速:可変(標準の0.5~2.0倍) 抑揚:可変(標準の0.5~2.0倍) 音量:可変(無音0.0~2.0)</li> </ul>	
動作要件(装置仕様)	<p>[ソフトウェア構成]</p> <p>【Windows】 Windows 11 日本語版 (64 bit) Windows 10 日本語版 (64 bit) Windows Server 2022 日本語版 (64 bit) Windows Server 2019 日本語版 (64 bit) Windows Server 2016 日本語版 (64 bit)</p>	<p>【Linux】 Red Hat Enterprise Linux 9 Red Hat Enterprise Linux 8 Debian 12 Debian 11 Ubuntu 22.04 LTS Ubuntu 20.04 LTS</p>
	<p>[ハードウェア構成]</p> <ul style="list-style-type: none"> <li>●CPU 日本語(波形接続合成方式): Intel Core i3 または AMD 同等以上のプロセッサ 日本語(新DNN音声合成方式): Intel Core i7-8700 または AMD 同等以上のプロセッサ 外国語: Intel Core i3 または AMD 同等以上のプロセッサ</li> <li>●メモリ 日本語(波形接続合成方式): 約 250 MB 日本語(新DNN音声合成方式): 約 1,100 MB 外国語: 約 100 MB</li> <li>●HDD 製品単体: 約 450 MB 日本語 1 話者: 波形接続合成方式: 約 900 MB 新DNN音声合成方式: 約 200 MB 外国語 1 話者: 約 250 MB</li> </ul>	
発表・発売時期	2020/5/7	
前年(2022年度調査)からの製品の更新内容	新エンジンAITalk6に対応し、新しいDNN方式でより人間らしい日本語音声を実現。また、60以上の国と地域の言語に対応しました。	
2023年中のプレスリリースURL	<a href="https://www.ai-i.jp/topics/">https://www.ai-i.jp/topics/</a>	
価格(税別)	開発環境:500,000円~(税別)/1システムor1サービス ※別途サポート費用が発生 ※再配布のロイヤリティは個別相談	
連絡先	TEL:03-6801-8461 Mail:info@ai-i.jp HP:https://www.ai-i.jp	
備考		

表A-2. 汎用OS用ミドルウェア・SDK

クリエートシステム開発株式会社

製品分類・製品名	ドキュメントトーカー macOS, iPhone, iPad	ドキュメントトーカー たかし ドキュメントトーカー けいこ	ドキュメントトーカー 日本語音声合成エンジン for Windows
製品情報のURL	<a href="https://www.createsystem.co.jp/main/seihin/seihin.html">https://www.createsystem.co.jp/main/seihin/seihin.html</a>	<a href="https://www.createsystem.co.jp/main/dtalkerAndroidSDK.html">https://www.createsystem.co.jp/main/dtalkerAndroidSDK.html</a>	<a href="https://www.createsystem.co.jp/main/DTalkerSapi1.html">https://www.createsystem.co.jp/main/DTalkerSapi1.html</a>
製品概要 (100字程度)	macOS, iOS のSpeech Synthesizerに準拠した高品質日本語音声合成エンジンです。	アンドロイド標準TTSに準拠した高品質日本語音声合成エンジンです。	マイクロソフト社 Windows SAPI5に対応した日本語音声合成エンジンです。
音声合成の機能・内部仕様に関する主な特徴・機能 (200字程度)	macOS, iOS のSpeech Synthesizerに準拠していますので、システムの声として使用できる他、選択項目の読み上げ機能や、VoiceOverの声として、また、AVSpeechSynthesizerを使用した開発用に利用できます。	アンドロイド標準TTSに準拠しているため、標準TTSを使用したアプリケーションで高品質な音声を利用できます。	マイクロソフト社のSAPI5に準拠しているため、SAPI5を使用したアプリケーションで高品質な音声を利用できます。
音声合成の詳細仕様	<p>入力形態:漢字仮名混じり文/単語/英単語                      合成可能言語:日本語                      文字コード:UNICODE                      JEITA TT-6004対応:準拠                      SSML:macOS,iOS独自                      辞書:基本:16万語 英単語:27000語 記号:380種 住所:73000語 漢字詳細読み:6600種                      ユーザ辞書機能:有                      合成単位:可変長音素列 男女各約15000種 コーパスベース波形編集 1ピッチ波形編集                      声:男声:たかし、太郎 女声:けいこ、花子                      声の高さ:可変(5段階)                      発声速度:可変(10段階)話速変換:1.0倍~0.5倍                      音量:10段階                      トーン:                      その他の機能:歌声合成, 詳細読み機能                      出力音声の形式:16kHz,32ビット,モノラル</p>	<p>入力形態:漢字仮名混じり文/単語/英単語                      合成可能言語:日本語                      文字コード:UNICODE                      JEITA TT-6004対応:準拠                      JEITA IT-4006対応:準拠                      SSML:準拠                      辞書:基本:16万語 英単語:27000語 記号:380種 住所:73000語 漢字詳細読み:6600種                      ユーザ辞書機能:有                      合成単位:可変長音素列 男女各約15000種 コーパスベース波形編集                      声:男声:たかし、太郎 女声:けいこ、花子                      声の高さ:可変(5段階)                      発声速度:可変(10段階)話速変換:1.0倍~0.5倍                      音量:10段階                      トーン:                      その他の機能:歌声合成, 詳細読み機能                      出力音声の形式:16kHz,16ビット,モノラル</p>	<p>入力形態:漢字仮名混じり文/単語/英単語                      合成可能言語:日本語                      文字コード:シフトJIS                      JEITA TT-6004対応:準拠                      JEITA IT-4006対応:準拠                      SSML:準拠                      辞書:基本:16万語 英単語:27000語 記号:380種 住所:73000語 漢字詳細読み:6600種                      ユーザ辞書機能:有                      合成単位:可変長音素列 男女各約15000種 コーパスベース波形編集                      声:男声:たかし、太郎 女声:けいこ、花子                      声の高さ:可変(5段階)                      発声速度:可変(10段階)話速変換:1.0倍~0.5倍                      音量:10段階                      トーン:                      その他の機能:歌声合成, 詳細読み機能                      出力音声の形式:16kHz,16ビット,モノラル</p>
動作要件(装置仕様)	開発環境: XCODE ソフトウェア環境: macOS 13.1以上、iOS16.0以上 ハードウェア環境: Mac, iPhone, iPad 辞書:34MB	開発環境: Android Studio ソフトウェア環境: Android ハードウェア環境: Android端末 辞書:34MB	開発環境: Microsoft SAPI5, Visual Studio ソフトウェア環境: WindowsXP/Vista/Win7/Win8/Win10/Win11 ハードウェア環境: Windows x86,x64, Arm64 辞書:34MB
発表・発売時期	2023年5月	2012年11月	2005年9月
前年(2022年度調査)からの製品の更新内容	2023/5月新規	2023/06 V2.3.5 Android13対応	
2023年中のプレスリリースURL			
価格(税別)	ロイヤリティ:100万円/プロジェクト	ロイヤリティ:100万円/プロジェクト	パッケージ価格 5,800円(税別) ベクタープロレジサービス 4,800円(税別) 再配布のロイヤリティは別途ご相談
連絡先	web-info@createsystem.co.jp, <a href="https://www.createsystem.co.jp/">https://www.createsystem.co.jp/</a>	web-info@createsystem.co.jp, <a href="https://www.createsystem.co.jp/">https://www.createsystem.co.jp/</a>	web-info@createsystem.co.jp, <a href="https://www.createsystem.co.jp/">https://www.createsystem.co.jp/</a>
備考	音声合成波形生成エンジンは富士通(株)製をmacOS,iOS用に移植 音声合成言語処理は独自の処理方式を採用 App Storeよりダウンロード可能	音声合成波形生成エンジンは富士通(株)製をAndroid用に移植 音声合成言語処理は独自の処理方式を採用 Google Play よりダウンロード可能	音声合成波形生成エンジンは富士通(株)製を移植 音声合成言語処理は独自の処理方式を採用

表A-2. 汎用OS用ミドルウェア・SDK

東芝デジタルソリューションズ株式会社

製品分類・製品名	音声合成MW ToSpeakG3 TSPG3-TTS	日本語音声合成MW ToSpeakG4 TSPG4-TTS JP	音声合成MW ToSpeakGx Neo TSPGX-TTS
製品情報のURL	<a href="https://www.global.toshiba.jp/products-solutions/ai-iot/recaius/lineup/tospeak.html">https://www.global.toshiba.jp/products-solutions/ai-iot/recaius/lineup/tospeak.html</a>	<a href="https://www.global.toshiba.jp/products-solutions/ai-iot/recaius/lineup/tospeak.html">https://www.global.toshiba.jp/products-solutions/ai-iot/recaius/lineup/tospeak.html</a>	<a href="https://www.global.toshiba.jp/products-solutions/ai-iot/recaius/lineup/tospeak.html">https://www.global.toshiba.jp/products-solutions/ai-iot/recaius/lineup/tospeak.html</a>
製品概要 (100字程度)	自然で安定した声質の合成音声、機器ローカルの処理でテキストから生成できる組み込みミドルウェア。組み込み環境に好適で多言語に対応。	自然で安定した声質の合成音声、機器ローカルの処理でテキストから生成できる組み込みミドルウェア。感情、声質パラメータの制御が可。	自然で安定した声質の合成音声、機器ローカルの処理でテキストから生成できる組み込みミドルウェア。肉声感の高い高音質版。
音声合成の機能・内部仕様に関する主な特徴・機能 (200字程度)	<ul style="list-style-type: none"> <li>・ユーザー辞書機能: 可能</li> <li>・SSML対応: 可能(言語により対応機能の範囲は異なります)</li> </ul>	<ul style="list-style-type: none"> <li>・ユーザー辞書機能: 可能</li> <li>・SSML対応: 可能</li> </ul>	<ul style="list-style-type: none"> <li>・ユーザー辞書機能: 可能</li> <li>・SSML対応: 可能(言語により対応機能の範囲は異なります)</li> </ul>
音声合成の詳細仕様	<p>[入力形態] ブレンテキスト/表音文字列</p> <p>[合成可能言語] 日本語/アメリカ英語(米語)/中国語(北京語)/韓国語/ほか。</p> <p>標準製品は最大11言語に対応可能。 (個別対応の場合最大30言語に対応可能)</p> <p>[声色のバリエーション] 男・女。日本語・アメリカ英語(米語)はオプションボイスも用意あり</p> <p>[制御可能な項目] 声の高さ、発話速度、など</p> <p>[出力音声の形式] モノラル、22.050kHz、16bit</p>	<p>[入力形態] ブレンテキスト/表音文字列</p> <p>[合成可能言語] 日本語</p> <p>[声色のバリエーション] 男・女</p> <p>[制御可能な項目] 声の高さ、発話速度、感情、声質パラメータなど。</p> <p>[出力音声の形式] モノラル、22.050kHz、16bit</p>	<p>[入力形態] ブレンテキスト/表音文字列</p> <p>[合成可能言語] 日本語/アメリカ英語(米語)/中国語(北京語)</p> <p>[声色のバリエーション] 男・女。日本語・アメリカ英語(米語)はオプションボイスも用意あり</p> <p>[制御可能な項目] 声の高さ、発話速度、など</p> <p>[出力音声の形式] モノラル、44.100kHz、16bitなど</p>
動作要件(装置仕様)	<p>[ソフトウェア環境] Windows(x64)、Linux、iOS、Androidなど 個別相談</p> <p>[ハードウェア環境] ARM系ほか 個別相談</p>	<p>[ソフトウェア環境] Linux、iOS、Android</p> <p>[ハードウェア環境] ARM系</p>	<p>[ソフトウェア環境] Windows(x64)、iOS、Androidなど</p> <p>[ハードウェア環境] ARM系ほか</p>
発表・発売時期	発売中	発売中	発売中
前年(2022年度調査)からの製品の更新内容	特になし	特になし	特になし
2023年中のプレスリリースURL			
価格(税別)	(個別相談)	(個別相談)	(個別相談)
連絡先	東芝デジタルソリューションズ株式会社 デジタルエンジニアリングセンター リカياسビジネス推進部 <a href="https://www.toshiba-sol.co.jp/pro/recaius/contact.html">https://www.toshiba-sol.co.jp/pro/recaius/contact.html</a>	東芝デジタルソリューションズ株式会社 デジタルエンジニアリングセンター リカياسビジネス推進部 <a href="https://www.toshiba-sol.co.jp/pro/recaius/contact.html">https://www.toshiba-sol.co.jp/pro/recaius/contact.html</a>	東芝デジタルソリューションズ株式会社 デジタルエンジニアリングセンター リカياسビジネス推進部 <a href="https://www.toshiba-sol.co.jp/pro/recaius/contact.html">https://www.toshiba-sol.co.jp/pro/recaius/contact.html</a>
備考			

表A-2. 汎用OS用ミドルウェア・SDK

株式会社 日立ケーイーシステムズ

製品分類・製品名	Windows版音声合成機能ソフトウェア開発キット	iOS版音声合成機能ソフトウェア開発キット
製品情報のURL	<a href="https://www.hke.jp/products/voice/voice_index.htm">https://www.hke.jp/products/voice/voice_index.htm</a>	<a href="https://www.hke.jp/products/voice/voice_index.htm">https://www.hke.jp/products/voice/voice_index.htm</a>
製品概要 (100字程度)	ユーザープログラムに対して音声合成機能のAPIを提供するソフトウェア開発キットです。 テキストデータを音声出力、音声ファイルに出力することができます。 声の高さ、速さ、記号などの読み方を指定することができます。	ユーザープログラムに対して音声合成機能のAPIを提供するソフトウェア開発キットです。 テキストデータを音声出力、音声ファイルに出力することができます。 声の高さ、速さ、記号などの読み方を指定することができます。
音声合成の機能・内部仕様に関する主な特徴・機能 (200字程度)	<ul style="list-style-type: none"> <li>・肉声感が高くなめらかで、自然な韻律を付与し、自然な抑揚を実現しています。</li> <li>・また、新しい音声データでハキハキした音声を合成します。</li> <li>・最適な音声単位(素片)を効率よく選択し、選ばれた素片をなめらかに接続することで、非常に高い肉声感となめらかな高品位音声の合成を実現します。</li> <li>・単語間の関連性データを参照することで、複数の読み方のある語句であっても、前後の文脈をもとに読み方を正確に判定し、読み分けます。</li> </ul>	<ul style="list-style-type: none"> <li>・肉声感が高くなめらかで、自然な韻律を付与し、自然な抑揚を実現しています。</li> <li>・また、新しい音声データでハキハキした音声を合成します。</li> <li>・最適な音声単位(素片)を効率よく選択し、選ばれた素片をなめらかに接続することで、非常に高い肉声感となめらかな高品位音声の合成を実現します。</li> <li>・単語間の関連性データを参照することで、複数の読み方のある語句であっても、前後の文脈をもとに読み方を正確に判定し、読み分けます。</li> </ul>
音声合成の詳細仕様	<ul style="list-style-type: none"> <li>・合成可能言語:日本語</li> <li>・文字コード:シフトJIS</li> <li>・合成音声の種類</li> <li>【高品位版】 女声1種類/男声1種類</li> <li>【標準版/コンパクト版】 女声5種類/男声3種類 女児1種類/男児1種類</li> <li>・声の高さ:可変(10段階)</li> <li>・発声速度:可変(10段階)</li> <li>・その他:エコー、音量(10段階)</li> <li>・出力音声の形式:8/11/16/22kHz, 8/16bit, モノラルPCM/μ-Law</li> </ul>	<ul style="list-style-type: none"> <li>・合成可能言語:日本語</li> <li>・文字コード:シフトJIS</li> <li>・合成音声の種類</li> <li>【コンパクト版】 女声5種類/男声3種類 女児1種類/男児1種類</li> <li>・声の高さ:可変(10段階)</li> <li>・発声速度:可変(10段階)</li> <li>・その他:エコー、音量(10段階)</li> <li>・出力音声の形式:8/11/16/22kHz, 8/16bit, モノラルPCM/μ-Law</li> </ul>
動作要件(装置仕様)	<ul style="list-style-type: none"> <li>【高品位版】</li> <li>・対応OS <ul style="list-style-type: none"> <li>・Microsoft® Windows® 10 (日本語版)</li> <li>・Microsoft® Windows® 11 (日本語版)</li> <li>・Microsoft® Windows Server® 2019 (日本語版)</li> <li>・Microsoft® Windows Server® 2022 (日本語版)</li> </ul> </li> <li>※ユニバーサルWindowsプラットフォーム(UWP)向けアプリには非対応です。</li> <li>・CPU 使用OSに準拠</li> <li>・メモリ容量 使用OSに準拠(空き容量として200MB以上)</li> <li>・ストレージ容量 使用OSに準拠(空き容量として1.4GB以上)</li> <li>【標準版/コンパクト版】</li> <li>・対応OS <ul style="list-style-type: none"> <li>・Microsoft® Windows® 10 (日本語版)</li> </ul> </li> <li>※64bit版OSではWOW64上で動作します。</li> <li>※ユニバーサルWindowsプラットフォーム(UWP)向けアプリには非対応です。</li> <li>・CPU 使用OSに準拠</li> <li>・メモリ容量 使用OSに準拠(標準版:空き容量として140MB以上、コンパクト版:空き容量として20MB以上)</li> <li>・ストレージ容量 使用OSに準拠(標準版:空き容量として450MB以上、コンパクト版:空き容量として15MB以上)</li> </ul>	<ul style="list-style-type: none"> <li>・対応OS iOS9.2以降(日本語版64bit版)</li> <li>・ハード iPhoneシリーズ、iPadシリーズ</li> <li>・メモリ容量 使用OSに準拠(空き容量として20MB以上)</li> <li>・ストレージ容量 使用OSに準拠(空き容量として32MB以上)</li> </ul>
発表・発売時期	2011年10月	2011年10月
前年(2022年度調査)からの製品の更新内容	更新無し	更新無し
2023年中のプレスリリースURL	無し	無し
価格(税別)	開発環境360,000円～ ライセンスは別途相談	開発環境360,000円～ ライセンスは別途相談
連絡先	営業統括本部 03-5627-7191 webmaster@hke.jp <a href="https://www.hke.jp/products/voice/voice_index.htm">https://www.hke.jp/products/voice/voice_index.htm</a>	営業統括本部 03-5627-7191 webmaster@hke.jp <a href="https://www.hke.jp/products/voice/voice_index.htm">https://www.hke.jp/products/voice/voice_index.htm</a>
備考		

表A-2. 汎用OS用ミドルウェア・SDK

株式会社日立ソリューションズ・クリエイト

製品分類・製品名	自動音声合成API ボイスソムリエ ネオ マリアージュ
製品情報のURL	<a href="https://www.hitachi-solutions-create.co.jp/solution/voice_neomari/index.html">https://www.hitachi-solutions-create.co.jp/solution/voice_neomari/index.html</a>
製品概要 (100字程度)	ユーザーアプリケーションで利用している漢字かな混じりテキストデータから、高品位な音声を作成することができます。 音声合成機能はマイクロソフト社のVisual C++やC#などから利用可能なライブラリ(DLL、LIBファイル)として提供します。また、Microsoft Speech API 5.3準拠のライブラリ(DLLファイル)(SAPI対応は日本語のみ)も用意しています。
音声合成の機能・内部仕様に関する主な特徴・機能 (200字程度)	ユーザー辞書機能あり
音声合成の詳細仕様	<p>入力形態 : 漢字仮名混じり文／英語／中国語(簡体字)                  合成可能言語 : 日本語／英語(米国語)／中国語(北京語)                  文字コード : (日本語)シフトJIS／(英中)UTF-16LE                  JEITA TT-6004対応: 非対応                  JEITA IT-4006対応: 非対応                  SSML対応 : 非対応                  出力音声の形式 :                  PCMデータ(サンプリング周波数 8~48kHz/                  量子化ビット 8/16bit(μ-Law形式は8bitのみ))、                  MP4(AAC-LC)データ</p>
動作要件(装置仕様)	<p>■ソフトウェア環境                  Windows 11 Pro 64bit版                  Windows 10 Pro／Enterprise 32／64bit版                  Windows Server 2019 Standard Edition                  Windows Server 2016 Standard Edition                  Windows Server 2016 Datacenter Edition                  ■ハードウェア環境                  CPU: Core2Duo 2GHz以上推奨                  メモリ: 2GB以上推奨</p>
発表・発売時期	2010/10/1
前年(2022年度調査)からの製品の更新内容	なし
2023年中のプレスリリースURL	なし
価格(税別)	438,000円～(詳細は、HPをご参照下さい)
連絡先	営業企画本部 Tel: 0120-954-536 Fax: 03-5780-7628 E-mail: hsc-contact@mlc.hitachi-solutions.com
備考	

表A-2. 汎用OS用ミドルウェア・SDK

## 株式会社日立ソリューションズ・テクノロジー

製品分類・製品名	Ruby Box2
製品情報のURL	<a href="https://www.hitachi-solutions-tech.co.jp/iot/solution/voice/Ruby_Box/index.html">https://www.hitachi-solutions-tech.co.jp/iot/solution/voice/Ruby_Box/index.html</a>
製品概要 (100字程度)	音声による入力・操作・確認が可能で、現場の作業効率向上やサービス向上を実現します。 音声による遠隔指示や自動音声ガイドにより、手を使わずに操作が可能となります。
音声合成の機能・内部仕様に関する主な特徴・機能 (200字程度)  ・言語処理機能の特徴 ・ユーザ辞書への対応の有無 ・合成方式の特徴 など。	多言語の音声認識・合成に対応 日本語、英語に対応。 音声認識(37言語)および音声合成(65言語)まで拡張が可能 肉声に近い流暢な音声合成を実現 文脈を自動的に推測し、人間らしい自然なイントネーションを実現 通信ができない環境でも利用可能 スタンドアロン型で搭載が可能。オフラインで動作するため、 ネットワークに接続することが難しい環境下で利用可能
音声合成の詳細仕様  ・入力テキストのフォーマット (文字コード・JEITA規格・SSML対応、など) ・合成可能な言語 ・声色のバリエーション ・制御可能な項目 (話速・ピッチ、など) ・出力音声の形式 など。	合成可能な言語:65言語 入力テキストのフォーマット 日本語:漢字仮名交じり文/発話表記コード(JEITA TT-6004準拠テキスト) 日本語以外:各言語の文/単語/発話表記コード(NT-SAMPA、L&H+) 声色のバリエーション 65以上(各言語1話者以上) 制御可能な項目 音量:101段階 話速:351段階 ピッチ:151段階 音色:151段階
動作要件(装置仕様)  ソフトウェア構成(対応OSなど) ハードウェア構成(メモリ要件など) など。	CPU arm64/x86_64 1.5GHz以上 メモリ容量 約70MB~1GB ストレージ 約40MB 動作OS Windows/Android *その他のOSはお問い合わせ下さい
発表・発売時期	2023年3月
前年(2022年度調査)からの製品の更新内容	無し
2023年中のプレスリリースURL	無し
価格(税別)	弊社営業にお問い合わせください。
連絡先	<a href="https://www8.hitachi.co.jp/inquiry/hitachi-solutions-tech/products/jp/form.jsp">https://www8.hitachi.co.jp/inquiry/hitachi-solutions-tech/products/jp/form.jsp</a>
備考	

表A-2. 汎用OS用ミドルウェア・SDK

株式会社日立産業制御ソリューションズ

製品分類・製品名	<p>合成音声組込み用API 「美音工房®DLL」</p> <p>●美音工房は、(株)日立産業制御ソリューションズの登録商標です。</p>
製品情報のURL	<p><a href="https://info.hitachi-ics.co.jp/product/seihin-k/bion/bion_top.html">https://info.hitachi-ics.co.jp/product/seihin-k/bion/bion_top.html</a></p>
製品概要 (100字程度)	<p>人間の肉声に極めて近い品質の音声をパソコン上で編集し、音声ファイルとして再生・出力することができるソフトウェア。</p>
音声合成の機能・内部仕様に関する主な特徴・機能 (200字程度)	<p>ユーザ辞書への対応:読み込み、登録修正、削除、単語表記、単語読み機能あり</p> <p>合成方式の特徴:(株)日立製作所中央研究所の音声合成技術を利用。</p> <p>・言語処理機能の特徴 ・ユーザ辞書への対応の有無 ・合成方式の特徴 など。</p>
音声合成の詳細仕様	<p>入力形態:テキストファイルまたはメモリ情報を読み込み</p> <p>合成可能な言語:日本語</p> <p>音声データサイズ:特に制限なし</p> <p>声のバリエーション:女性/男性</p> <p>声の高さ:10段階の設定</p> <p>発生速度:10段階の設定</p> <p>音量:10段階の設定</p> <p>抑揚:10段階の設定</p> <p>高域強調:高域強調する/しない</p> <p>無声化:無声化する/しない</p> <p>文末ポーズ長:10段階の調節</p> <p>出力音声の形式: 16bit PCM(モノラル):8kHz、11.025kHz、16kHz、22.05kHz μ-Law(モノラル):8kHz、11.025kHz、16kHz、22.05kHz</p>
動作要件(装置仕様)	<p>ソフトウェア構成:Windows10(32bitOS/64bitOS)</p> <p>ハードウェア構成: CPU:2GHz以上推奨 メモリ:1GB以上 ディスク領域:1GB以上</p>
発表・発売時期	<p>2012年2月15日</p>
前年(2022年度調査)からの製品の更新内容	<p>特になし</p>
2023年中のプレスリリースURL	<p><a href="http://info.hitachi-ics.co.jp/product/seihin-k/bion/bion_top.html">http://info.hitachi-ics.co.jp/product/seihin-k/bion/bion_top.html</a></p>
価格(税別)	<p>500,000円(税抜)</p> <p>※「美音工房DLL」の動作に必要な「美音工房」を併せた価格</p>
連絡先	<p>営業統括本部 産業営業本部 産業営業第一部 流通・サービスグループ 山本 仁</p> <p>TEL :090-9003-7365</p> <p>アドレス:jin.yamamoto.qr@hitachi.com</p>
備考	

## 製品カテゴリ A-3

	ページ
サーバ用ミドルウェア	21
HOYA MD部門 ReadSpeaker	22
NTTテクノクロス	23
エーアイ	24
日立ソリューションズ・テクノロジー	25

## HOYA株式会社 MD部門 ReadSpeaker

製品分類・製品名	speechServer
製品情報のURL	<a href="https://readspeaker.jp/products/development/speechserver/">https://readspeaker.jp/products/development/speechserver/</a>
製品概要 (100字程度)	ネットワーク型のシステム構築に適した音声合成SDKです。
音声合成の機能・内部仕様に関する主な特徴・機能 (200字程度)  <ul style="list-style-type: none"> <li>・言語処理機能の特徴</li> <li>・ユーザ辞書への対応の有無</li> <li>・合成方式の特徴</li> <li>など。</li> </ul>	高速処理を実現しており、高パフォーマンスなシステム構築が可能です。マルチスレッド対応により、複数の音声を同時に処理するシステムの構築に利用いただけます。 WindowsServer、RHEL、CentOSに対応しています。開発言語もC/C++/Java/.NETと幅広く対応し、システムやユーザーに合わせ最適な環境を選択できます。 シンプルで分かりやすいAPI群は、言語・話者を問わず共通化され、効率の良い組み込み開発を実現します。SSML (Speech Synthesis Markup Language) にも対応しています。
音声合成の詳細仕様  <ul style="list-style-type: none"> <li>・入力テキストのフォーマット (文字コード・JEITA規格・SSML対応、など)</li> <li>・合成可能な言語</li> <li>・声色のバリエーション</li> <li>・制御可能な項目 (話速・ピッチ、など)</li> <li>・出力音声の形式</li> <li>など。</li> </ul>	<ul style="list-style-type: none"> <li>・入力テキストのフォーマット:UTF-8、SSML対応</li> <li>・合成可能な言語 44言語</li> <li>・制御可能な項目:話速・ピッチ・音量・アクセント</li> <li>・出力音声の形式:16bit LPCM,16bit LPCM Wave,8bit A-law PCM,8bit A-law PCM Wave,8bit <math>\mu</math>-law PCM,8bit <math>\mu</math>-law PCM Wave,8bit <math>\mu</math>-law PCM SUN AU,8bit unsigned LPCM Wave,4bit Dialogic ADPCM</li> </ul>
発表・発売時期	2004年
前年(2022年度調査)からの製品の更新内容	
2023年中のプレスリリースURL	<a href="https://readspeaker.jp/news/">https://readspeaker.jp/news/</a>
価格(税別)	個別相談
連絡先	<a href="https://readspeaker.jp/">https://readspeaker.jp/</a> TEL:03-6479-7182
備考	

表A-3. サーバ用ミドルウェア

## NTTテクノクロス株式会社

製品分類・製品名	FutureVoice Crayon Server
製品情報のURL	<a href="https://www.futurevoice.jp/lineup/server/">https://www.futurevoice.jp/lineup/server/</a>
製品概要 (100字程度)	電話自動音声応答や対話システム向けにネットワークを利用して、マルチタスクによる高速音声合成処理を行うサーバシステム
音声合成の機能・内部仕様に関する主な特徴・機能 (200字程度)  ・言語処理機能の特徴 ・ユーザ辞書への対応の有無 ・合成方式の特徴 など。	<ul style="list-style-type: none"> <li>・人の声と遜色ない自然な音声</li> <li>・バリエーション豊富な話者を選択可能、複数話者の音声を混合し独自話者の作成も可能</li> <li>・少量の音声データから誰の声でも音声として再現</li> <li>・音声のチューニング・導入が簡単</li> <li>・様々な機器へ搭載可能</li> <li>・クロスリンガル音声合成技術に対応(話者の声質を損なわずに一言語の音声から多言語の合成音声を生成できる技術)</li> </ul>
音声合成の詳細仕様  ・入力テキストのフォーマット (文字コード・JEITA規格・SSML対応、など) ・合成可能な言語 ・声色のバリエーション ・制御可能な項目 (話速・ピッチ、など) ・出力音声の形式 など。	<ul style="list-style-type: none"> <li>・入力フォーマット: プレーンテキスト・読み指定テキスト・SSML</li> <li>・文字コード: UTF-8</li> <li>・ユーザ辞書機能: 単語辞書、文辞書、テキスト置換辞書、音声置換辞書、韻律テンプレート辞書</li> <li>・対応言語: 日本語・英語・中国語(普通話/簡体字・広東語/繁体字)・韓国語</li> <li>・出力サンプリングレート: 8,000Hz・22,050Hz・44,100Hz</li> <li>・出力音声ファイル形式: 16bit PCM(wav/raw)・8bit <math>\mu</math>-law(wav/raw) ※拡張可</li> </ul>
発表・発売時期	2017年12月
前年(2022年度調査)からの製品の更新内容	クロスリンガル音声合成技術に対応、国内初の商用提供開始
2023年中のプレスリリースURL	<a href="https://www.ntt-tx.co.jp/whatsnew/2023/230120.html">https://www.ntt-tx.co.jp/whatsnew/2023/230120.html</a>
価格(税別)	個別相談
連絡先	E-mail: futurevoice.info-ml@ntt-tx.co.jp
備考	

表A-3. サーバ用ミドルウェア

## 株式会社エーアイ

製品分類・製品名	AITalk Server
製品情報のURL	<a href="https://www.ai-j.jp/products/server/">https://www.ai-j.jp/products/server/</a>
製品概要 (100字程度)	AITalkは、人間らしく自然な音声で自由に音声合成をすることができます。 AITalk Serverは、電話自動応答やインターネットサービスなど、ネットワークを利用し、マルチタスクで合成処理を行う場合に最適なエンジンです。
音声合成の機能・内部仕様に関する主な特徴・機能 (200字程度)  ・言語処理機能の特徴 ・ユーザー辞書への対応の有無 ・合成方式の特徴 など。	<ul style="list-style-type: none"> <li>●人の声をベースとした高品質音声合成</li> <li>●感情調整機能</li> <li>●ユーザー辞書機能搭載</li> <li>●RESTインターフェース (HTTP/1.1)</li> <li>●マルチスレッド対応</li> <li>●SSMLサポート</li> </ul>
音声合成の詳細仕様  ・入力テキストのフォーマット (文字コード・JEITA規格・SSML対応、など) ・合成可能な言語 ・声色のバリエーション ・制御可能な項目 (話速・ピッチ、など) ・出力音声の形式 など。	<ul style="list-style-type: none"> <li>●入力形態 漢字仮名混じり文 韻律記号付き文字列 (AI 独自規格orJEITA TT-6004 規格)</li> <li>●文字コード 日本語 Shift_JIS(CP932) 日本語 UTF-8</li> <li>●合成可能な言語/声色のバリエーション 標準語: 女性8名、男性4名、男の子2名、女の子2名 関西弁: 女性1名、男性1名</li> </ul> <p>※AITalk CustomVoiceで作成したオリジナルの話者も利用可能</p> <ul style="list-style-type: none"> <li>●制御可能な項目 音量、話速、ピッチ、抑揚</li> </ul>
発表・発売時期	2020/10/30
前年(2022年度調査)からの製品の更新内容	
2023年中のプレスリリースURL	<a href="https://www.ai-j.jp/topics/">https://www.ai-j.jp/topics/</a>
価格(税別)	開発環境: 500,000円～(税別)/1システムor1サービス ※別途サポート費用が発生 ※再配布のロイヤリティは個別相談
連絡先	TEL: 03-6801-8461 Mail: info@ai-j.jp HP: <a href="https://www.ai-j.jp">https://www.ai-j.jp</a>
備考	

表A-3. サーバ用ミドルウェア

株式会社日立ソリューションズ・テクノロジー

製品分類・製品名	Ruby Talk Server 3.0	Ruby Talk Server 2.1
製品情報のURL	<a href="https://www.hitachi-solutions-tech.co.jp/iot/solution/voice/Ruby_Talk/index.html">https://www.hitachi-solutions-tech.co.jp/iot/solution/voice/Ruby_Talk/index.html</a>	<a href="https://www.hitachi-solutions-tech.co.jp/iot/solution/voice/Ruby_Talk/index.html">https://www.hitachi-solutions-tech.co.jp/iot/solution/voice/Ruby_Talk/index.html</a>
製品概要 (100字程度)	オンプレミスサーバ向け製品です。Windows、Linux等主要OSに対応しています。詳細は、弊社営業にお問い合わせください。	オンプレミスサーバ向け製品です。Windows、Linux等主要OSに対応しています。詳細は、弊社営業にお問い合わせください。
音声合成の機能・内部仕様に関する主な特徴・機能 (200字程度)  ・言語処理機能の特徴 ・ユーザ辞書への対応の有無 ・合成方式の特徴 など。		
音声合成の詳細仕様  ・入力テキストのフォーマット (文字コード・JEITA規格・SSML対応、など) ・合成可能な言語 ・声色のバリエーション ・制御可能な項目 (話速・ピッチ、など) ・出力音声の形式 など。	合成可能な言語:46言語 入力テキストのフォーマット 日本語:漢字仮名交じり文/単語 日本語以外:各言語の文/単語 声色のバリエーション 46話者(各言語1話者) 制御可能な項目 ピッチ:10段階 アクセント:10段階 抑揚:10段階	合成可能な言語:3言語 入力テキストのフォーマット 日本語:漢字仮名交じり文/単語 英語:米語文/単語 中国語:北京語簡体字文/単語 声色のバリエーション 日本語:複数話者より選択可 英語:1話者 中国語:1話者 制御可能な項目 ピッチ:10段階 アクセント:10段階 抑揚:10段階
発表・発売時期	2015年8月	2013年8月
前年(2022年度調査)からの製品の更新内容	無し	無し
2023年中のプレスリリースURL	無し	無し
価格(税別)	弊社営業にお問い合わせください。	
連絡先	<a href="https://www8.hitachi.co.jp/inquiry/hitachi-solutions-tech/products/jp/form.jsp">https://www8.hitachi.co.jp/inquiry/hitachi-solutions-tech/products/jp/form.jsp</a>	
備考		