

JEITA 一般社団法人 電子情報技術産業協会

令和 4(2022)年度 音声認識・合成技術 応用製品の動向について

— 音声入出力方式標準化専門委員会 実施の製品動向調査結果より —

2023 年 8 月

音声入出力方式標準化専門委員会

## ○概要

音声認識・合成技術の適用分野の拡大、応用製品の普及を目的として、本委員会では継続して製品動向調査を実施している。本調査は応用製品の開発者、製品の導入・購入を考える企業・個人に有益な情報を提供するとともに、継続実施することにより、技術の進歩、利用領域の変遷を理解する貴重な資料となっている。

本年度（令和4年度）の調査は平成22年度以降の調査方法を基本的に踏襲し、以下の方法により実施した。

### (1) 情報提供の依頼

令和5（2023）年1月にJEITA本委員会ホームページに情報提供の依頼を掲示するとともに、音声研究・開発者向けメーリングリスト（[onsei-mail@aist.jpn.org](mailto:onsei-mail@aist.jpn.org)）に同一内容のメールを配信。なお、前年に情報を提供いただいた企業には直接メールを発信し、依頼した。

### (2) 掲載製品

情報提供企業が製造あるいは販売する音声認識・合成関連製品であり、令和4（2022）年12月末現在で日本国内で購入可能な製品とした。

### (3) 提供内容

委員会で作成した、製品分類ごとの製品特徴・仕様を記載するフォームを配布、記載いただき電子ファイルとして受領した。

### (4) 掲載内容

情報提供企業からいただいた内容は基本的に加工を行わず掲載している。ただし、内容の変更を伴わない書式の統一、情報提供者の明らかな誤解による分類誤りの修正は、委員会の判断にて行っている場合がある。

## ○音声合成のシステム開発者向け製品の動向調査結果

日本国内で販売されている音声合成製品を調査した結果について以下に概要をまとめる。

### 製品カテゴリの分類

今回の調査（2022年度調査）で採用している製品分類の7カテゴリを表1に示す。

表 1 音声合成製品の製品分類

ID	分類	
A-1	組込み用 ミドルウェア・SDK	組込み OS にて動作するライブラリ提供の音声合成ミドルウェア製品。
A-2	汎用 OS 用 ミドルウェア・SDK	汎用 OS 上で動作するライブラリとして提供される音声合成ミドルウェア製品。アプリケーションに組込むことで動作する製品。 対象 OS: windowsOS, MacOS, iOS, AndroidOS を始め, Linux OS も含む。
A-3	サーバ用 ミドルウェア	クライアントサーバ構成を実現できるなど、アプリ単体に組込んでランタイムのみで動作するというような制限がなく、外部からの呼び出しで動作する音声合成エンジンソフトウェア。 マルチスレッドで動作するなど、複数リクエストを受けることのできる製品。 対象 OS: windows Server OS, Linux OS など。
A-4	Web サービス	利用者が音声合成エンジンの動作するソフトウェアを入手することなく利用可能な提供形態の音声合成サービス・製品。webAPI にて提供されるクラウドサービスなど。
A-5	ツール製品	利用者が音声ファイルを生成したり、辞書記述を編集するなど、ツールとして利用可能な製品・サービス。
A-6	ハードウェア製品	ハードウェアとして提供される音声合成製品。LSI 提供, ボード提供。
A-7	その他	上記 6 カテゴリに当てはまらない音声合成関連製品・サービス。

### 音声合成製品の製品件数とベンダ数

調査結果を、「合成関連製品一覧表（別 pdf ファイル）」に示す（別冊 1 に表 A-1～表 A-3, 別冊 2 に表 A-4～表 A-7 を所収）。

以下では、今年度の調査結果をもとに、調査対象とした 7 カテゴリの分類ごとの製品件数と製品を提供しているベンダ数について述べる。表 2 に調査結果示す。

表 2 分類毎の音声合成製品の件数と製品を提供しているベンダ数

カテゴリ ID	分類カテゴリ	R4 (2022) 製品件数	R4 (2022) ベンダ数
表 A-1	組込み用 ミドルウェア・SDK	15	8
表 A-2	汎用 OS 用 ミドルウェア・SDK	21	12
表 A-3	サーバ用 ミドルウェア	7	6
表 A-4	Web サービス	12	9
表 A-5	ツール製品	20	12
表 A-6	ハードウェア製品	5	2
表 A-7	その他	16	10
合計		96	59

今年度の回答製品数は総計 96 件であった。製品分類カテゴリごとに製品件数をみると、

- 「汎用 OS 用ミドルウェア・SDK」製品（表 A-2, 21 件）
- 「ツール製品」（表 A-5, 20 件）

の製品数が多かった。

続いて

- 「その他」製品（表 A-7, 16 件）
- 「組込み用ミドルウェア・SDK」製品（表 A-1, 15 件）
- 「web サービス」製品（表 A-4, 12 件）

などの製品件数が 10 件以上であった。

また、今年度の提供ベンダ数は総計 59 社であった。ここで「提供ベンダ数」とは「のベンダ数」であり、複数の製品カテゴリに回答したベンダは重複してカウントされている。

製品分類カテゴリごとにベンダ数をみると、7 つの製品カテゴリの中では、

- 「汎用 OS 用ミドルウェア・SDK」製品（表 A-2, 12 社）
- 「ツール製品」（表 A-5, 12 社）

の提供ベンダ数が最多であった。

以下は

- 「その他」製品（表 A-7, 10 件）
- 「web サービス」製品（表 A-4, 9 社）
- 「組込み用ミドルウェア・SDK」製品（表 A-1, 8 社）
- 「サーバ用ミドルウェア」製品（表 A-3, 6 社）

のように続いた。

- 「ハードウェア」製品（表 A-6, 2 社）

は 2 社の回答が寄せられた。

なお、複数のカテゴリに製品を提供しているベンダを重複してカウントせずに集計すると、調査に回答したベンダ（異なり数）の総数は 25 社であった。

### 音声合成製品の製品件数と提供ベンダ数の変化

以下では、2022 年度調査までの 4 年間の製品数（表 3）とベンダ数（表 4）を示す。

今年度の回答製品数は総計 96 件であった。製品件数の総数を前年度の調査と比較すると、1 件ほど減っている（表 3）。「組込み用 MW SDK」製品（表 A-1）で 2 件が減、「汎用 OS 用 MW SDK」製品（表 A-2）で 4 件が減、「サーバ用 MW」製品（表 A-3）で 3 件が減、「web サービス」製品（表 A-4）で 1 件が減った一方、「ハードウェア製品」（表 A-6）で 1 件が増、「その他」製品（表 A-7）で 8 件が増えた。

但し「その他」製品（表 A-7）での 8 件の増加のうち 5 件は、前年度調査で「ツール製品」（表 A-5）に分類されていた歌声ソフトや編集機能を持たない読み上げソフトのカテゴリを「その他」製品（表 A-7）にカテゴリ変更したことによるものである。つまり、「その他」製品（表 A-7）では実質的に 3 件増え、「ツール製品」（表 A-5）では実質的に 5 件が増

表 3 分類毎の音声合成製品の 製品件数 と推移

ID	新しい分類	R4	R3	R2	R1
		(2022)	(2021)	(2020)	(2019)
		件数(±)	件数	件数	件数
A-1	組込み用 MW SDK	15 (-2)	17	17	23
A-2	汎用 OS 用 MW SDK	21 (-4)	25	25	30
A-3	サーバ用 MW	7 (-3)	10	10	10
A-4	Web サービス	12 (-1)	13	11	10
A-5	ツール製品	20 (±0)	20	17	19
A-6	ハードウェア製品	5 (+1)	4	3	1
A-7	その他	16 (+8)	8	9	8
合計		96 (-1)	97	92	101

えたことになる。ちなみに、「ツール製品」（表 A-5）での実質 5 件の増加は、今年度新規に回答のあったベンダ 3 社による 5 製品が追加されたことが増加の要因である。

その他、最も減少幅の大きかった「汎用 OS 用 MW SDK」製品（表 A-2、4 件減）については、今年度回答のなかった 1 社による 5 製品が減ったためであり、実質的な減少であるかの判断は難しい。一方、3 件減った「サーバ用 MW」製品（表 A-3）については、ベンダ 1 社の撤退によって 2 件減と、別のベンダによる 1 製品の削除による実質的な減少であった。

表 4 分類毎の音声合成製品の 提供ベンダ数 とその推移

ID	新しい分類	R4	R3	R2	R1
		(2022)	(2021)	(2020)	(2019)
		件数(±)	件数	社数	社数
A-1	組込み用 MW・SDK	8 (-1)	9	9	10
A-2	汎用 OS 用 MW・SDK	12 (±0)	12	12	13
A-3	サーバ用 MW	6 (-2)	8	8	8
A-4	Web サービス	9 (±0)	9	8	6
A-5	ツール製品	12 (+1)	11	10	9
A-6	ハードウェア製品	2 (±0)	2	2	1
A-7	その他	10 (+5)	5	6	5
合計		59 (+3)	56	55	52

今年度の提供ベンダ数は総計 59 社であった。提供ベンダ数を前年度の調査と比較すると、3 社ほど増えている（表 4）。「組込み用 MW SDK」製品（表 A-1）で 1 社、「サーバ用 MW」製品（表 A-3）で 2 社減った一方で、「ツール製品」（表 A-5）で 1 社、「その他」製品（表 A-7）で 5 社増えた。

一方、複数のカテゴリに製品を提供しているベンダを重複してカウントせずに集計したときの、調査に回答したベンダ（異なり数）の総数は 26 社で、前年度の調査より 2 社増えた。今年度新規に回答のあった 3 社と前年度回答がなかったが今年度復帰した 1 社の計 3 社が増えた一方、撤退した 1 社と回答のなかった 1 社で計 2 社が減ったことにより、トータルで 2 社の増加となった。

### ○音声認識の製品動向調査

本年度調査にあたって、その製品分類・記載項目について検討し、昨年同様の分類で調査を実施することとした。具体的には、製品分類は音声認識製品を使用する立場から、①システム開発者向けとして、組み込み向け（ボード等含む）と PC 向け、②ソリューション利用者向けとして、CTI ソリューション、Web API サービスソリューション、その他ソリューション、そして③エンドユーザ向けとして、応用製品とソフトウェア製品の、計 7 つに分類している。

各分類における令和元年度から令和 4 年度の製品件数とその推移を表 6 に示す。また、製品を提供しているベンダ数とその推移を表 7 に示す。令和 3 年度の調査時に分類の見直しを行ったため、令和 2 年度以前の Web サービスと令和 3 年度以降の Web API サービスの製品数を比較することは控えた。

表 6 分類毎の音声認識製品の件数と推移

分類		表番号*1	R4		R3	R2	R1
			件数	増減*2	件数	件数	件数
① システム開発者向け 音声認識エンジン	組み込み向け(ボード等含む)	表 B-1	9	0	9	9	11
	PC 向け	表 B-2	4	0	4	4	6
② ソリューション利用者 向け製品	IVR/CTI/CRM 分野	表 B-3	14	0	14	14	15
	Web サービス(~R2)	表 B-4				13	14
	Web API サービス(R3~)		7	0	7		
	その他	表 B-5	7	-1	8	8	7
③ エンドユーザ向け 応用製品	応用製品	表 B-6	10	+1	9	15	17
	応用ソフトウェア製品	表 B-7	34	+1	33	22	19
合計		-	85	+1	84	85	89

\*1 表番号：認識製品一覧の対応表番号。 \*2 増減：令和 3 年度からの増減。

表 7 分類毎の音声認識製品のベンダ数とその推移

分類		表番号*1	R4		R3	R2	R1
			件数	増減*2	件数	件数	件数
① システム開発者向け 音声認識エンジン	組み込み向け(ボード等含む)	表 B-1	4	0	4	4	5
	PC 向け	表 B-2	4	0	4	4	5
② ソリューション利用者 向け製品	IVR/CTI/CRM 分野	表 B-3	11	+1	10	10	11
	Web サービス(~R2)	表 B-4				9	8
	Web API サービス(R3~)		7	0	7		
	その他	表 B-5	5	-1	6	7	6
③ エンドユーザ向け 応用製品	応用製品	表 B-6	7	0	7	8	6
	応用ソフトウェア製品	表 B-7	8	+1	7	4	3
合計		-	46	+1	45	46	44

\*1 表番号：認識製品一覧の対応表番号。 \*2 増減：令和 3 年度からの増減。

回答していただいた 23 社に関する調査結果の詳細を表 B-1~B-7 に掲載する。

調査結果は、上記方法による本委員会からの依頼に対して回答のあった情報のみを掲載しているため、網羅性の視点からは完全な調査ではない。よって、周知であるにも関わらず

調査結果に記載されていない製品が存在する。

また、昨年度は回答があったが今年度は回答が得られなかったベンダについて、製品自体が継続して販売されていることが企業ホームページから確認できる場合でも、該当する製品情報を掲載していない。

したがって、本調査結果のみから、詳細な動向を判断することは困難であるが、表 6 と表 7 からわかるここ数年の傾向を以下に示す。

- ・ 一昨年度の調査以降、製品数、ベンダ数とも、大きな変動はない。
- ・ システム開発者向けの音声認識エンジンについては、組込み向け(ボード等含む)製品、PC 向け製品とも、一昨年度の調査以降、製品数とベンダ数の変動がない。
- ・ また、ソリューション利用者向け製品についても、IVR/CTI/CRM 向けソリューションと Web API サービスの製品数については、それぞれ 1 製品増加、1 製品減少で、トータルでは変動がなかった。その他製品の 1 件の減少についても、純粋な製品の減少ではなく、昨年度回答があったベンダのうち 1 ベンダについて、今年度は回答が得られなかったことによるものであった。
- ・ エンドユーザ向け製品については、応用ソフトウェア製品の製品数とベンダ数が増加傾向である。また、昨年度と比べて、全体で 2 製品増加した(上記と同様、今年度のみ回答が得られなかったベンダの 1 件減少を考慮すると、3 製品の増加)。新規製品としては、会議・議事録作成を支援する製品が昨年度に続きリリースされたほか、自動でテキスト化するボイスレコーダー製品、キャラクターと会話ができるエンターテイメント向け製品、工事現場の KY (危険予知) 活動を分析する製品があった。

## ○令和4年(2022年)度の音声認識・合成技術による応用製品の全般的な動向

## (音声合成)

製品動向調査の結果だけからは読み取りにくい、2022年の音声合成市場の全般的な動向について述べる。ひとつめは、前年(2021年)にも増して、商用利用も可能な、無料や安価な音声合成の製品・サービスのリリースが加速したことが挙げられる。2021年にリリースされた「VOICEVOX<sup>1</sup>」や「COEIROINK<sup>2</sup>」は2022年に入ってもボイスや機能の追加、改善が頻繁に行われた。さらに、2022年3月には、AHS社から「VOICEPEAK 商用可能6ナレーターセット」が発売<sup>3</sup>された。最新のAI音声合成技術が搭載され、喜怒哀楽の表現等にも対応した6種類の話者を低価格で商用利用可能なことで話題となった。この他、2022年には、オープンソースである「VOICEVOX」をベースにしたオープンソースの読み上げソフトとして、TALQu<sup>4</sup>、LMROID<sup>5</sup>、SHAREVOX<sup>6</sup>、ITVOICE<sup>7</sup>がリリースされた。2021年以來のこれらのリリースの特徴として、以下の2点が挙げられる。1) 提供元が個人であるものも多いこと、2) 個人利用および法人による商用利用の両方が許諾されているものが多いこと、である。これらの製品・サービスでは、利用許諾がボイス(キャラクター)ごとに定められており、個人利用での動画配信による収益化についての条件なども規定されている。音声合成の開発・提供の担い手が「個人」にまで広がってきているのに合わせて、利用者である「個人」に対しても、具体的なユースケースに応じた利用許諾条件の整備を通じて、権利の保護を求める動きが生まれていると言える。

ふたつめとして、提供側の盛り上がりだけでなく、YouTubeやニコニコ動画などでの「動画コンテンツ(解説動画、実況動画など)」で合成音声を利用する側の盛り上がりも顕著だったことが挙げられる。ネット上の動画における合成音声の利用をまとめた統計や調査報告こそ見つけられていないものの、たとえば、リリースされている多数の音声合成エンジンそれぞれの特徴を紹介するとともに、動画作成をする際の「合成音声の使い方(合成音声の選び方)」自体もコンテンツにしているYouTubeチャンネル<sup>8</sup>が存在するなど、合成音声の利用の裾野は確実に広がっていると言える。他にも、YouTube動画再生の合間に流れる「広告動画」のナレーションに合成音声が使われていると感じられた経験などは、誰にもあるの

1 VOICEVOX (ボイスボックス) . <https://voicevox.hiroshiba.jp/>

2 COEIROINK (コエイロインク) . <https://coeiroink.com/>

3 『VOICEPEAK 商用可能6ナレーターセット』3月11日発売。  
<https://prtimes.jp/main/html/rd/p/000000015.000060783.html>

4 TALQu (トーク) . <https://haruqa.booth.pm/items/2755336>

5 LMROID (レミロイド) . <https://lmroidsoftware.wixsite.com/nhoshio>

6 SHAREVOX (シェアボックス) . <https://www.sharevox.app/>

7 ITVOICE (イタボイス) . <http://itvoice.starfree.jp/>

8 「登録者1万人記念」合成音声界隈の半年間のふりかえり. 2022年6月  
<https://www.youtube.com/watch?v=uFC-AENfIQQ>  
(YouTubeチャンネル「ゆうぶろ」 <https://www.youtube.com/@yuupro>)

ではないだろうか。実際、広告・プロモーションやマーケティングに合成音声を使用することを指南する記事<sup>9</sup>、<sup>10</sup>なども散見される。

合成音声の利用が広がっているのは、ネット上の「動画コンテンツ」だけに限らない。テレビ番組中で、敢えて味わいのあるナレーションとして合成音声が使われている例も、もはや合成音声の採用自体を取り立てて話題にしないほどである。また、街中での公共アナウンスなどの音声合成に思われる経験ももはや珍しいことではない。また、2022年のニュースとして、CoeFont社による「おしゃべりひろゆきメーカー<sup>11</sup>」が公開された一方で、出所が明確でない団体による「AI 安倍晋三」プロジェクト<sup>12</sup>で故人の合成音声公開され（その後削除された）、賛否が分かれるなど、話題になった。

### （音声認識）

一昨年度・昨年度に続き、新型コロナウイルス感染症（COVID-19）が猛威をふるう中で、人々のライフスタイルや業務の様々な変化に対して、音声認識技術の利用が引き続き進んだ一年と言える。

非対面のコミュニケーション手段として定着したオンライン会議においては、リアルタイムなコミュニケーション支援や議事録作成支援などに音声認識が重宝されている。

また、新型コロナウイルス感染症に関する問い合わせへの対応や自宅療養者への健康確認の業務により業務負荷が高まるコールセンターでは、対応の自動化やオペレータ業務の効率化のために音声認識が貢献しており、今年度もコールセンター関連製品の統合などの動きがみられた。

本年度は、ChatGPTなどの生成系AIが盛り上がりを見せた年でもあった。近年の主流であるディープラーニングによる認識精度の向上などに加え、生成系AIの技術を活用した認識結果の改善、認識結果の新たな活用などによる音声認識製品の更なる発展が期待される。

---

<sup>9</sup> 合成音声マーケット プロモーションに合成音声を活用。2022年3月。

<https://www.trans.co.jp/column/trend/gouseionsei/>

<sup>10</sup> ウェビナーレポート「音マーケティング=音声広告」だけじゃない。2022年3月

<https://canvas.d2cr.co.jp/tips/8567/>

<sup>11</sup> CoeFont社「おしゃべりひろゆきメーカー」を期間限定で公開（2022年9月）

<https://prtimes.jp/main/html/rd/p/000000029.000078329.html>

<sup>12</sup> ITmedia「AI 安倍晋三」ネットで物議。2022年9月。

<https://www.itmedia.co.jp/news/articles/2209/26/news135.html>