

JEITA 一般社団法人 電子情報技術産業協会

平成 28(2016)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 —

2017 年7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（平成 28 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

平成 29（2017）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，平成 28（2016）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

提供いただいた情報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については，委員会にて変更をしている．

記載されている内容については，委員会にて確認をしておらず，内容は，すべて情報提供企業にその責任がある．

## ○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。  
当製品調査は、調査対象の製品分類を大きく

- 「ミドルウェア組込み用」,
- 「ミドルウェア PC 用」,
- 「ハードウェア製品」,
- 「その他」

の4カテゴリ（第1分類）に分けた上で、さらに細分類（第2分類）を設けて製品情報をまとめている。

調査結果を表A-1～9に示す。

以下では、その調査結果をもとに、分類ごとの製品件数とその推移について述べる。

表1に、調査対象とした9カテゴリについて製品件数と推移を示す。

表1 分類毎の音声合成製品の件数と推移

分類		表番号 *	H28 (2016)	H27 (2015)	H26 (2014)	H25 (2013)	H24 (2012)
第1分類	第2分類		件数 (増減**)	件数	件数	件数	件数
ミドルウェア 組込み用	規則音声合成	表 A-1	32 (-7)	39	39	38	36
	その他	表 A-2	1 (±0)	1	1	1	1
ミドルウェア PC 用	規則音声合成	表 A-3	27 (-4)	31	29	29	26
	その他	表 A-4	2 (±0)	2	1	1	2
ハードウェア	規則音声合成 (規則音声合成 LSI)	表 A-5	1 (±0)	1	2	2	2
	その他	表 A-6	3 (±0)	3	2	1	1
その他	サーバ製品	表 A-7	14 (±0)	14	12	10	13
	PCソフト製品 ・ツール製品	表 A-8	22 (±0)	22	22	24	25
	サービス	表 A-9	15 (+2)	13	12	9	10
合計			117 (-9)	126	120	115	116

\*表番号：合成製品一覧の対応表番号， \*\*増減：前年度からの増減

今年度（平成28年度）の回答製品数は総計117件となり、平成27年度調査と比較して総数で9件の減少となった。

以下、平成27年度と比較した今年度の製品件数の推移をまとめる。

第2分類で見ると、製品件数が増加した項目は

- 「その他」の「サービス」カテゴリ (2 件増)

逆に、製品件数が減少したのは、- 「ミドルウェア組み込み用」の「規則音声合成」製品 (7 件減)

- 「ミドルウェア PC 用」の「規則音声合成」製品 (4 件減)

の 2 項目で製品数が減少した。それ以外の項目では製品数の合計は同じであった。

但し、製品数はベンダ各社ごとに何を 1 製品と数えるかの基準が異なり、上記の製品件数の減少も言語ごとに別製品としていた製品をひとつに統合したケースも影響している。また、製品カテゴリごとに集計した増減では製品数の増加と減少が相殺されるケースも存在するため、製品件数の増減の絶対値はあくまで参考程度に捉えるのが妥当である。

表 2 では、該当製品が存在した 9 カテゴリについて、製品件数ではなく、製品を提供しているベンダ数とその推移を示す。

表 2 分類毎の音声合成製品の提供ベンダ数とその推移

分類		表番号*	H28 (2016)	H27 (2015)	H26 (2014)	H25 (2013)	H24 (2012)
第 1 分類	第 2 分類		社数 (増減**)	社数	社数	社数	社数
ミドルウェア 組み込み用	規則音声合成	表 A-1	12 (+1)	11	11	11	10
	その他	表 A-2	1 (±0)	1	1	1	1
ミドルウェア PC 用	規則音声合成	表 A-3	14 (+1)	13	13	13	12
	その他	表 A-4	1 (±0)	1	1	1	2
ハードウェア	規則音声合成 (規則音声合成 LSI)	表 A-5	1 (±0)	1	2	2	2
	その他	表 A-6	3 (±0)	3	2	1	1
その他	サーバ製品	表 A-7	10 (±0)	10	9	8	9
	PC ソフト製品 ・ツール製品	表 A-8	10 (±0)	10	11	11	12
	サービス	表 A-9	9 (+1)	8	9	7	8
合計(延べベンダ数)			61 (+3)	58	59	55	57

\*表番号：合成製品一覧の対応表番号， \*\*増減：前年度からの増減

音声合成製品の提供ベンダ数を合計（延べ数）で見ると、本年度の調査では 3 ベンダ増加した結果となった。

製品を提供しているベンダ数が増加した製品カテゴリは、

- 「ミドルウェア組み込み用」の「規則音声合成」製品 (1 ベンダ純増)
- 「ミドルウェア PC 用」の「規則音声合成」製品 (1 ベンダ純増)
- 「その他」の「サービス」(1 ベンダ純増)

の 3 項目であった。

逆に本年度の調査では、提供ベンダ数が減少した製品カテゴリは存在しなかった。

それ以外の項目では製品カテゴリのベンダ総数での増減はなかったが、「サーバ製品」「PCソフト・ツール製品」のカテゴリでは、製品提供を取りやめるベンダと新たに動向調査に参加するベンダが相殺されて、見かけのベンダ数に増減がなかった。

今回の音声合成の製品動向調査に製品情報を提供したベンダ数は、全カテゴリを合わせた延べ数で 61 社、複数のカテゴリに製品を提供しているベンダの重複を考慮すると、調査に回答したベンダ（異なり数）は現在 25 社に上る。

### ○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、昨年同様の分類ならびに記載項目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 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。各分類における平成 25 年度～28 年度の製品件数とその推移を表 3 に示す。また、製品を提供しているベンダ数とその推移について表 4 に示す。

表 3 分類毎の音声認識製品の件数と推移

分類		表番号*	H28		H27	H26	H25
			件数	増減**	件数	件数	件数
① システム開発者向け 音声認識エンジン	組込み向け(ボード等含む)	表 B-1	13	0	13	13	10
	PC 向け	表 B-2	12	+1	11	11	7
② ソリューション利用者 向け製品	IVR/CTI/CRM 分野	表 B-3	13	-5	18	14	12
	Web サービス	表 B-4	9	+2	7	7	4
	その他	表 B-5	5	+1	4	6	7
③ エンドユーザ向け 応用製品	応用製品	表 B-6	8	+2	6	2	3
	応用ソフトウェア製品	表 B-7	17	+2	15	14	13
合計		-	77	+3	74	67	56

\*表番号：認識製品一覧の対応表番号， \*\*増減：平成 27 年度からの増減

表 4 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*	H28		H27	H26	H25
			件数	増減**	件数	件数	件数
① システム開発者向け 音声認識エンジン	組込み向け(ボード等含む)	表 B-1	7	0	7	7	7
	PC 向け	表 B-2	6	0	6	7	5
② ソリューション利用者 向け製品	IVR/CTI/CRM 分野	表 B-3	9	-1	10	10	9
	Web サービス	表 B-4	6	+1	5	6	4
	その他	表 B-5	4	0	4	4	4
③ エンドユーザ向け 応用製品	応用製品	表 B-6	3	0	3	1	2
	応用ソフトウェア製品	表 B-7	3	-1	4	4	5
合計		-	38	-1	39	39	36

\*表番号：認識製品一覧の対応表番号， \*\*増減：平成 27 年度からの増減

回答して頂いた 15 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、上記手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機関について、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、該当製品情

報を掲載していない。

従って、本調査結果のみから、動向を判断することは困難であるが、表 3 と表 4 からわかるここ数年の傾向を以下に示す。

- ・ 昨年度からベンダ数は減ったが、製品数は増加しており、1社あたりの製品数は増えた。
- ・ 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向け製品ともに昨年度と大きな変化はなかった。
- ・ ソリューション利用者向け製品については昨年度に比べ、IVR/CTI/CRM分野が5製品減少し、一方で、Webサービスが2製品、その他が1製品増加した結果、全体で2製品減少した。この中には、昨年度のIVR/CTI/CRM分野の2製品がその他の分野に移動した分を含むが、この移動分を除いても、IVR/CTI/CRM分野が3製品減少した。
- ・ エンドユーザ向け製品については、応用製品と応用ソフトウェア製品の両方ともそれぞれ2製品増加した結果、全体で4製品増加した。
- ・ ベンダごとの内訳をまとめると、24製品がバージョンアップされた。また、新規に11製品が登録された一方で、昨年度に登録されていた4製品が今回は登録されなかった。既存製品の見直しが進むとともに、1社あたりの製品数が増えたことから新製品開発が依然として活発であると考えられる。
- ・ ベンダ数については、昨年度から2社減少した。一方で、分類のカバー範囲を増やすベンダが1社あったことから、ベンダの延べ数は昨年度とほぼ同じであった。

その他、本調査に含まれない委員会の独自調査からも、スマートフォンやロボットや専用端末からアクセスするサービス、また、2020年の東京オリンピックに向けた音声翻訳としての実証実験や製品化応用サービスが増加していることが本年度の大きな傾向であり、今後の動向が注目される。

なお、今後も、携帯端末上のアプリケーションや、サーバ上での音声認識利用サービスが増加することが予想され、製造企業、販売企業、サービス提供企業などが複雑化するとともに、販売製品という形態ではなく、有償・無償とにかかわらずエンドユーザに対する音声翻訳や介護福祉支援などのサービスとして提供される応用システムの増加も考えられる。次年度以降も、なるべく多くの情報を収集し、当該分野の動向を把握するとともに、関連製品開発者、導入検討者、エンドユーザに有益な情報を提供できるよう、調査方法について検討を続ける予定である。

## ○平成 28 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)

平成 28 (2016) 年度の製品動向調査の結果の集計には含まれない、全般的な動向のひとつとして、音声合成技術の活用が、ニュースの読み上げなど日常的に使われる実際のサービスで本格的に拡大し始めた点が挙げられる。2016 年 4 月に提供が開始された、朝日新聞社のスマートフォンアプリ「アルキキ<sup>1</sup>」では新聞記事を音声合成技術により聞くことができる。また、日本経済新聞社も 2016 年 10 月、スマートフォンアプリ「もっと日経<sup>2</sup>」で「新聞読み上げサービス」の提供を開始した。

また、もうひとつの動向として、前年 (2015 年) にエーアイによるタレントのマツコ・デラックスのカスタムボイスがテレビ番組で使われる事例<sup>3</sup>があったのに引き続き、2016 年にはカスタムボイスのさらなる活用が進んだ点も挙げられる。この背景として、TTS ベンダが提供している既存の声色を利用するのでは、同一 TTS ベンダを利用している他社サービスと声色が同一になってしまうのを避けたい、さらには、有名声優を採用した音声合成によりサービスの一層の差別化を図りたい、などの動機が働いたと考えられる。たとえば、HOYA サービスは「音声合成の声優事務所<sup>4</sup>」と題したサイトを立ち上げて顧客に声色の選択肢を提示し、その中でカスタムボイス (オリジナルボイス) の作成を提案している。また、東芝<sup>5</sup>はカプコンやソニー・ミュージックコミュニケーションズとの取り組みに於いて、カスタムボイスが採用されている他、FM NACK5 での取り組みでは番組パーソナリティーの小林克也の似声で話す「コバカツ君<sup>6</sup>」を登場させた。

その他、2016 年時点の研究フェーズでは、DNN (Deep Neural Network) を利用した音声合成に関する研究発表が急増しており、2017 年以降、実サービスでも DNN による音声合成が登場してくることが予想される。

---

<sup>1</sup> HOYA サービス. 朝日新聞社アプリ「アルキキ」  
[https://voicetext.jp/case/application/html/casestudy20160407\\_asahi.html](https://voicetext.jp/case/application/html/casestudy20160407_asahi.html)

<sup>2</sup> 東芝デジタルソリューションズ. 日本経済新聞社アプリ「もっと日経」  
<http://www.toshiba.co.jp/cl/news/news20161019-2.htm>

<sup>3</sup> エーアイ. 「マツコロイド」の音声に AITalk®採用 <http://www.ai-j.jp/archives/7889>

<sup>4</sup> HOYA サービス 音声合成の声優事務所 <http://voicetext.jp/voiceactor/>

<sup>5</sup> 東芝デジタルソリューションズ.  
事例紹介 <https://www.toshiba.co.jp/cl/pro/recaius/lineup/tospeak02.html>  
めがみスピークエンジン [http://www.toshiba.co.jp/recaius/2016\\_12\\_08.html](http://www.toshiba.co.jp/recaius/2016_12_08.html)

<sup>6</sup> コバカツ君 <http://www.toshiba.co.jp/cl/news/news20161207.htm>



**(音声認識)**

音声認識については、引き続き既存製品の見直しが進むとともに、新製品開発が依然として活発であると考えられる。また、製品動向調査の範囲外の応用として、スマートフォンや専用端末からアクセスするクラウド型音声対話サービスの実用化が進んでいる。例えば、このようなサービスを活用して家庭からの物品購入や情報検索などを容易にする音声アシスタント端末として、国内は未適用だが、Amazon 社の Amazon Echo の米国での普及が進んでおり。また、Apple 社の HomePod の年内販売開始が発表された。また、国内外においては、Google 社の Google Home、国内においては LINE 社の WAVE、NTT ドコモ社の petoco の年内販売開始が発表されている。

他には、オリンピックに向け、インバウンド対応を目的とした音声翻訳の実証実験や製品化が活発になってきている。また、自動車関連では、安全性と便利さの実現のために、トヨタ社などが感情認識の研究を開始している。さらに、音声認識を支える背景技術として深層学習 (Deep Learning) が多く用いられている。