

JEITA 一般社団法人 電子情報技術産業協会

平成 26(2014)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 —

2015 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 26 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 27（2015）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・開発者向けメイリングリスト（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 26（2014）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*	H26 (2014)	H25 (2013)	H24 (2012)	H23 (2011)	H22 (2010)
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	39 (+1)	38	36	22	24
	その他	表 A-2	1 (±0)	1	1	1	0
ミドルウェア PC 用	規則音声合成	表 A-3	29 (±0)	29	26	26	23
	その他	表 A-4	1 (±0)	1	2	0	0
ハードウェア	規則音声合成 (規則音声合成 LSI)	表 A-5	2 (±0)	2	2	2	1
	その他	表 A-6	2 (+1)	1	1	1	2
その他	サーバ製品	表 A-7	12 (+2)	10	13	18	8
	PCソフト製品・ツ ール製品	表 A-8	22 (-2)	24	25	23	18
	サービス	表 A-9	12 (+3)	9	10	7	3
合計			120 (+5)	115	116	100	79

*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 26 年度）の回答製品数は総計 120 件となり、平成 25 年度調査と比較して総数で 5 件の増加となった。

以下、平成 25 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分類でみると、製品件数が増加したのは、

- 「ミドルウェア組み込み向け」の「規則音声合成」製品 (1件増),
- 「ハードウェア」の「その他」製品 (1件増),
- 「サーバ製品」(2件増),
- 「サービス」(3件増)

の4項目であり、逆に製品件数が減少したのは、

- 「PCソフトウェア・ツール等」(2件減),

の1項目で製品数が減少した。それ以外の項目では製品数は同じであった。

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*	H26 (2014)	H25 (2013)	H24 (2012)	H23 (2011)
第1分類	第2分類		社数 (増減**)	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	11 (±0)	11	10	12
	その他	表 A-2	1 (±0)	1	1	1
ミドルウェア PC用	規則音声合成	表 A-3	13 (±0)	13	12	14
	その他	表 A-4	1 (±0)	1	2	0
ハードウェア	規則音声合成 (規則音声合成 LSI)	表 A-5	2 (±0)	2	2	2
	その他	表 A-6	2 (+1)	1	1	1
その他	サーバ製品	表 A-7	9 (+1)	8	9	8
	PCソフト製品・ツール製品	表 A-8	11 (±0)	11	12	13
	サービス	表 A-9	9 (+2)	7	8	6
合計(のべ)			59 (+4)	55	57	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で見ると、本年度の調査にて製品の提供から撤退するベンダは存在しなかった。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、

- 「ハードウェア」の「その他」製品 (1ベンダ増),
- 「サーバ製品」(1ベンダ増),
- 「サービス」(2ベンダ増),

の3項目であり、それ以外の項目ではベンダ数に変化はなかった。

「サーバ製品」で1ベンダ、「サービス」分野で2ベンダが増えているのはすでに他の項目でも製品を提供しているベンダが、これらの製品カテゴリでも新たに提供を開始した結果である。「ハードウェア(その他)」製品のカテゴリでは、平成25年度調査でいったん製品提供を中止したベンダが、今回の平成26年度調査時に製品の提供を再開しており、ベンダ数が増える結果となった。

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のべ数で59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 23 年度～26 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*	H26		H25	H24	H23
			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	表 B-1	13	+3	10	8	8
	PC 向け	表 B-2	11	+4	7	5	9
②ソリューション利用者 向け音声認識 ソリューション製品	IVR/CTI/CRM 分野	表 B-3	14	+2	12	12	13
	Web サービス	表 B-4	7	+3	4	3	1
	その他	表 B-5	6	-1	7	9	6
③エンドユーザ向け 応用製品	応用製品	表 B-6	2	-1	3	5	5
	応用ソフトウェア製品	表 B-7	14	+1	13	13	15
合計		-	67	+11	56	55	57

*表番号：認識製品一覧の対応表番号， **増減：平成 25 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*	H26		H25	H24	H23
			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	表 B-1	7	0	7	6	5
	PC 向け	表 B-2	7	+2	5	4	7
②ソリューション利用者 向け音声認識 ソリューション製品	IVR/CTI/CRM 分野	表 B-3	10	+1	9	9	8
	Web サービス	表 B-4	6	+2	4	3	1
	その他	表 B-5	4	0	4	6	4
③エンドユーザ向け 応用製品	応用製品	表 B-6	1	-1	2	4	5
	応用ソフトウェア製品	表 B-7	4	-1	5	6	10
合計		-	39	+3	36	38	40

*表番号：認識製品一覧の対応表番号， **増減：平成 25 年度からの増減

回答して頂いた 18 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も、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 からわかるここ数年の傾向を以下に示す。

- ・ システム開発者向けの音声認識エンジンについては、組込み向け製品、PC 向け製品が昨年度に比べてそれぞれ 3 製品、4 製品増加した。また、組込み向け製品として、新たに悲鳴認識エンジンが掲載された
- ・ ソリューション製品については、CTI 分野、Web サービス共に増加傾向にある。特に、クラウドコンピューティングを利用したものが増えている
- ・ 3 製品がバージョンアップされた。一方で、新規追加された製品は 19 製品であり、新製品開発が活発化していると考えられる
- ・ ベンダ数については、昨年度に比べて 3 社増加し、ここ数年の減少傾向に歯止めが掛かった

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、携帯端末のアプリケーション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 26 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)

製品動向調査に含まれていない応用製品での動向としては、2014 年 9 月にはソフトバンクモバイルが、クラウドサービス型の音声合成ソリューション「Rizbell (リズベル)」を発表し、同 10 月から販売パートナーを通じた提供を開始した（但し本調査への掲載の承諾を得られなかったためリストには不掲載）。このサービスは IVR での電話自動応答に用いられる合成音声だけでなく、電子カタログや e ラーニング教材でのナレーション、商業施設での案内放送などでの合成音声にも利用可能な音声合成ソリューションである。

Rizbell が注目に値するのは、音声合成エンジン自体を開発・提供しているベンダではなく、通信キャリアがサービス提供の担い手となっている点である。このことから、音声合成技術がエンジン開発ベンダによって提供されるだけでなく、より幅広い層に提供されるニーズが高まって来ていると見なすことができる。平成 24 年度（2012 年度）に国内での日本語サービスが始まった Apple 社の Siri、NTT ドコモ社のしゃべってコンシェルに代表されるスマートフォン向け音声対話アプリで採用された音声合成技術は、H26 年度（2014 年度）に入って着実に普及が進んでいると考えられる。

なお、この Rizbell の音声合成は、東芝の音声合成エンジンの技術が使われており、発表時点で 4 カ国語（17 話者）を提供した（2015 年 7 月時点では、日本語を含めて 8 か国 10 地域、計 22 話者を提供している）。さらに Rizbell は音声合成だけでなく、同時に音声認識、自然言語（対話型 FAQ）についても同様にクラウド型サービスを提供しており、今後の動向が注目される。

(音声認識)

音声認識については、新規追加された製品が 19 製品あり、新製品開発が活発化していると考えられる。また、製品動向調査の範囲外の応用として、エンドユーザ向けの音声対話アプリケーション、家電の対話的な操作などが増えている。また、Siri 等スマートフォンを利用したクラウド型音声対話サービスが普及し、音声対話が現実的になってきている。このようなサービスおよびそれを支える技術を活用する形で、Apple の CarPlay、トヨタの T-Connect、Android Auto などテレマティクスサービスにおける音声対話の実用化が進むと予想される。