

# スマートスピーカー

## 日本の視点：日本におけるスマートスピーカーの現状と今後の課題

「話す」「書く」「触る」「ジェスチャーする」など、人間の自然な行動に基づいて機械を操作するNUI (Natural User Interface) の利用が増えている。特に音声認識技術を活用したVoice UI (Voice User Interface) の利用は、AI (人工知能) の技術進歩と相まって、一般消費者の生活シーンにおいても身近なものになりつつある。

Voice UI を代表するのが、AIアシスタントサービスである。2011年にApple iPhone 4Sの新機能として導入されたSiri以降、対話型の音声操作に対応したAIアシスタントがスマートフォンなどのデバイスに搭載されるようになった。本文で取り上げている、AIアシスタントを搭載したスマートスピーカーについては、2014年末にAmazonが販売開始した「Amazon Echo」を皮切りとして英語圏での販売が拡大している。今後は非英語圏の言語対応が進み、人間の生活シーンや企業活動の様々な領域に合わせた利用の拡大が予想されている。

スマートスピーカーの市場は、Google Home、Amazon Echo等のハードウェアや、Google AssistantやAmazon Alexa、Apple Siri、LINE ClovaなどのAI音声プラットフォームを中心に語られることが多く、プラットフォーム上で使用するツールやコンテンツを提供する事業者を含めてエコシステムが構成されている(図表1)。これらの音声エコシステムに参加するスタートアップ企業が消費者の様々な生活シーンや業種に合わせたサービスを提供することも影響し、2019年には前年63%増の成長率で市場が拡大する見込みである<sup>1)</sup>。

音声プラットフォーム自体は今後も引き続きGoogle、Amazon、Appleといった海外事業者が中心となって主導権を握っていくと想定した場合、日本企業が競争優位性を発揮できる領域としては、海外企業にとって高い参入障壁となる日本語ユーザー向けのサービスの開発となる。日本の事業者は、既存の自社のハードウェアデバイスやソリューションと音声プラットフォームをつなげて日本向けサービスを開発・提供する形で、既存

ビジネスを活かした収益拡大の位置づけで音声エコシステムに関与するのが現実的であろう。

本稿では、需要が更に拡大するスマートスピーカーと音声アシスタント機能について、日本における現状の活用状況、課題等を整理し、今後の社会に与えるインパクト、サービスの付加価値、今後の発展の方向性について見ていきたい。

### 日本におけるスマートスピーカーの利用状況

日本においては、2017年にスマートスピーカーの発売が開始されたが、デロイトの調査では60%の認知度に対して普及率は3%にとどまっている<sup>2)</sup>。その理由としては、前述の音声認識技術の日本語対応に時間がかかっているということに加え、音声検索・音声操作そのものへの抵抗が未だ強い<sup>3)</sup>という市場の特性にも起因していると考えられる。

アジア勢では中国のAlibabaが2017年に「Tmall Genie」を発表しているほか、中国のLenovoグループ、韓国のLGやSamsungもスマートスピーカーに積極的に参入している<sup>4)</sup>。一方で日本では家電メーカーや住宅メーカーなどのプレイヤーが対応に時間を要しており、そもそもの対応サービス自体がまだ少ないのが現状である。

しかし今後、音声認識技術の日本語対応が進み、精度と機能が向上することと並行して、2019年から順次開始される5GサービスやIoTの進展に伴うインフラ面の高度化により、AIアシスタントの普及が進展すると考えられる。また、AIアシスタントのコネクテッドデバイスへの搭載が進み、5Gの高速、超低遅延、多数同時接続を利用してIoT機器間で膨大な情報データが常時やりとりされるようになると、利用が一気に拡大すると予想される。

利用環境が整う中で、ユーザーの音声サービスに対する心理的な障壁を和らげながら付加価値のあるサービスを企業や消費者側に提供することができれば、家庭内における家電等と連携した様々な生活シーンや企業

図表1 音声エコシステムの構成

分類	内容
音声プラットフォーム	グローバルプレイヤーが提供するプラットフォーム群 (Google、Amazon、Apple、LINE等)
ソフト開発ツール	プラットフォーム上や自社のWebサイト向けに音声・テキストチャットボットアプリをつくるためのAI SaaSプラットフォーム
分析ツール	開発した会話アプリの解析、アプリの利用状況、ユーザーとのセッション時間やエンゲージメント、自然言語分析、ユーザーの意図や購買ステージ、感情分析等を含む
エージェンシー	アプリ開発の代理店
メディア	業界情報メディア
エンターテインメント	コンテンツ・アプリの提供

活動において、多様なユースケースが出てくると想定される。その際に、AIアシスタントを手軽に利用するためのハブとして、各家庭やビジネスシーンにおいてスマートスピーカーの導入が進むことは想像に難くない。さらに一般消費者の家電製品から自動車まで様々なデバイスに音声認識型のAIサービスが搭載されるようになるのではないだろうか。

### スマートスピーカーの発展を支えるテクノロジーのブレークスルー

このようなサービス発展の見通しを立てる上では、技術的ブレークスルーの実現とそれを支えるいくつかの技術的要件がある。

#### 【技術的なブレークスルー(例)】

- ① 自然言語処理技術(要素技術含む)の更なる向上
  - AI・深層学習・自然言語処理技術の連携
  - マルチモーダル学習の進展 等
- ② 言語資源データの構築・オープン化
  - 重点取り組み分野のデータ整備
  - データ整備・提供を担う専門機関の強化
  - データ取得・ツール検証を加速する模擬環境、実証環境の整備
  - 産学連携によるデータ・ツールの集積の好循環(AIクラウドの提供、オープンツール開発支援)
  - 取得データのオープンデータ化
  - 民間等保有データの共有、横断的活用(データ取引ルール、API公開、データ連携・互換性の向上) 等
- ③ 翻訳機能の強化
  - 対応言語数の拡大(少数言語含む)
  - 多言語同時対応
  - 非母国語話者の音声認識精度の向上 等

- ④ 対話AIの仕組みと外部連携サービスの強化
  - APIを通じた他のクラウドサービスとの連携(フィンテック、シェアリングエコノミーなど) 等
- ⑤ 非タスク指向型システムの強化
  - コンテキストを考慮した非タスク指向型対話システムの構築 等
- ⑥ ヒューマンインターフェース(HMI)の進化
  - マルチモーダルインターフェースの開発 等
- ⑦ 複合領域で開発が進められるAI
  - 環境知能(Ambient Intelligence)の進化 等

上記のブレークスルーを支えるために求められる技術的な要件としては、まず、音声テクノロジーの技術領域における機能向上が挙げられる。AIアシスタントは、スマートスピーカーやスマートフォンなどのデバイスを通じて、ユーザーの発話を認識して家電操作や音楽・動画再生、情報検索などを実行する。そのメリットは、音声を使用するとキーボードやタッチ操作に比べて手早く簡単にデバイスの操作ができる点にある。一方で、音声のみに頼るといった特性ゆえに、現状では技術的な制約がある。音声テクノロジーは音声認識、話者認識、環境音認識、感情解析に大別されるが(図表2)、現時点においてスマートスピーカーで対応できているのは音声認識に留まる場合がほとんどで、環境音認識や感情解析を可能にするには機能が不十分である。環境音認識の精度が高まれば、雑多な音が混じる環境でも高い精度で必要な音声情報を抽出して適切な処理ができるようになる。さらに感情解析機能が深化すれば、使用者の気分を考慮したレコメンデーション等が可能になり、ECやコンテンツ視聴といった場面で活用できるようになるだろう。

図表2 音声テクノロジーの機能分類

分類	内容	例
音声認識	発話者の言葉を認識する	スマートフォン、スマートスピーカー
話者認識	複数人が話している中で話者を判別する	議事録
環境音認識	空間内で通常とは異なる音を検知する	家庭室内、工場機器の異常音検知
感情解析	声色や抑揚から発話者の感情を導き出す	メンタルヘルスケア

出所：デロイト トーマツ コンサルティング合同会社作成

二つ目の要件として挙げられるのは、人間との対話能力の向上である。現状のAIアシスタントの役割は、ヘルプデスクや日用品の購買など、ユーザーが何を求めているかをあらかじめ決めてある場合のコミュニケーションが前提になっている。それに対して、ユーザーの発言から推理を重ね、課題に対する新たなソリューションを提案する、といったより高度な会話を遂行する機能が研究が進んでいる。この機能が実現すれば、一般消費者や企業に対するAIアシスタントサービスの付加価値は更に向上することになる。

現状でもすでに以上のような要素を含めた音声エコシステムの構築が進んでおり、段階的にサービスの充実・更なる活用への道筋が見えてくると考えられる。

今までは一部の企業やユーザー向けのものであったAIだが、クラウド上での利用などを通してより広い範囲で適用できるようになりつつあることで<sup>5</sup>、AIアシスタントの活用の幅が広がり、機能向上のスピードアップにつながる事が予想される。

また近い将来、音声認識、話者認識、環境音認識、感情解析に加えて、画像認識等の様々なAI技術がクラウド上で相互連携することで、人間を取り巻くモノや空間全体において、連携された情報が知的に作用する環境知能 (Ambient Intelligence) が実現し、人々の安心・安全・快適な生活を実現するために必要な企業活動も含めた形で、社会全体の変化に大きな影響を与えることになるだろう。

### 想定されるユースケース

以上のような課題が段階的に解決すれば、BtoBとBtoCの双方でAIアシスタント/スマートスピーカーのユースケースは飛躍的に拡大するだろう。企業活動においては、音声認識システムが中小を含んだ多くの企業に導入されることで、業務効率の改善が見込めるのみならず、電話対応の自動化や無人化によって深刻な人手不足の解消につながることも想定される。

また、日本語を話さない人に対して、安心して日本での生活・観光を楽しむ環境を提供することができる可能性もある。具体的には、ホテルや観光案内などサービス業での翻訳機能を搭載した音声アシスタント利用などが想定される。特に2020年の東京オリンピックを控え、急増が予想される世界各国からの旅行者の利便性の向上や、東南アジア、アフリカなど今後さらなる発展が見込まれる国々との新たなビジネスの拡大にもつながると考えられる。

医療分野では、高齢化社会における地方在住高齢者に対する遠隔医療や、日常会話の音声の変化から精神疾患等の早期予測・診断補助をするといった用途での活用も検討が進んでいる<sup>6</sup>。これらに関しては患者本人のケアにとどまらず、医療費の公費負担などを考慮した際の公共的な側面での継続的な投資やサポートが重要となってくるであろう。

さらに将来的な可能性として、前述の音声・画像・動画データを統一的・複合的に処理するフレームワークの構築が可能になれば、音声だけでは判断が難しい人間の表情の変化や、画像だけでは判断が困難な声の変化を定量的に評価する方法をより詳細に認識して分析することが可能になり、医療、防犯、エンターテインメントなどの様々な分野でサービスイノベーションが起きることになるであろう。

これらのユースケースを鑑みると、オフィス、訪日外国人が利用するホテル等施設、医療施設、高齢者家庭を含めた一般住宅の各部屋・スペースといった多様な場所にスマートスピーカーを配置することが想定され、端末と対応機器を含めた普及拡大の可能性が見えてくるだろう。

### 今後のビジネス上の課題

#### マネタイズモデル

今後のビジネス上の課題としてまず、AIアシスタントサービスのマネタイズモデルについて考えておきたい。現在、Amazon、Google、Apple、LINEをはじめとする企業が提供しているスマートスピーカーは、海外で急速に普及しており、国内でも今後それなりに普及すると考えられる。とはいえ、スマートスピーカーのビジネスモデルは、単体事業では収益基盤としてはあまり期待することはできないと考えるのが妥当である。

AIアシスタントサービスによる家電コントロール機能はあくまでもその家電の付随機能であり、この機能だけで収益を上げることは容易ではない。現に、Amazonは本業のEC販売の延長として、プライム会員向けのワンクリック注文を音声で可能とすることで、EC販売を通じて収益を上げていると想定される。Googleにおいては本業の広告販売の延長として、ユーザーから音声入力情報を収集することで、広告に必要な情報を収集していると推察されるが、端末販売としては製造原価や流通コストなどを考えると利益に貢献しているとは考えにくい。Appleについては、本業のハードウェア販売の延長として、サービス自体は安価に提供する戦略を取っていると考えられる。このように、すでにスマートスピーカー販売で先行している企業は、端末販売そのものを目的にしているわけではなく、自社のビジネスに紐づくマネタイズポイントを設定している。

一方、日本企業でAIアシスタントサービスの収益モデルを見つけた企業

は未だ少ない。端末販売のみでの収益確保が難しいとした場合、スマートスピーカーを新たな顧客タッチポイントとしてEC販売拡大につなげる、音声プラットフォームを通じた新たな課金サービスを顧客に提供する、といった方法で収益確保を狙う方法が考えられる。また、昨今注目が高まっている音声広告ビジネスが拡大すれば、スマートスピーカーを通じて今まで他のフォーマットが取り逃がしていた顧客に新たにリーチできる可能性も出てくる。

この分野で日本企業の対応は遅れば遅れるほど、既存のビジネスモデルに与えるインパクトは大きいだろう。例えば、Amazonがスマートスピーカーを通じてホームIoT機器の販売、配送に加え、設定、運用を一元的に行うようになったと仮定した場合、Amazonが扱えるものしか売れなくなり、今後はモノではなくサービスに活路を見いだそうとしていた製造業企業の目論見が崩壊するというシナリオも考えられる。

### 中期的な方向性の検討

人間の会話を完全に理解するAIの登場は、遠い将来ではないだろう。企業は今すぐにこの新たなフロンティアに備えるべきである。AIアシスタントサービス対応製品の拡大に伴って、搭載端末を通じた音声操作を頼りにする消費者や企業は今後も増え続けることは間違いない。関連プレイヤー各社は、今のうちにこの変化を見据え、自社の開発・研究体制を整えて、EC取引やマーケティング戦略をこうした端末やサービスに連携させておかななくてはならないだろう。

かつてネットショッピングが実店舗での体験を一変させたように、音声AI技術は一般消費者の日々の生活スタイルや企業活動を次の段階に引き上げるであろう。人間とコンピューターのやり取りから生じる有意義な会話と、結果重視の解決策への期待は日を追うごとに高まっている。日本においては、日本語という言語バリアがある故に、海外プレイヤーが十分に対応できない部分で独自のビジネスモデルやサービスを構築できる可能性がある。その観点で、日本のプレイヤー各社が日本語処理をベースとしたテクノロジー開発で連携しつつ、こうしたテクノロジーの開拓と貢献に取り組む必要があるのではないかと。

今後、テクノロジーの進展により精度を向上させたAIアシスタントが、ユーザーの意図を適切に汲み取りながら行動を起こす知的なシステムへと進化し、ユーザーにとってもさらに有意義な存在となり、社会全体の変化へとつながっていくことになるであろう。そのような未来の姿を見据えたサービス設計とビジネスモデル構築が求められている。

### 著者



**菅原 幹太**  
**Sugahara, Yoshihiro**

デロイト トーマツ コンサルティング  
合同会社  
シニアマネジャー

外資系コンサルティングファーム複数社を経て現職。多数の大規模基幹システム導入経験を有し、昨今ではWeb/EC改革等、販売・マーケティング領域での知見を活かしたグローバルレベルでの業務・IT改革関連の企画推進・実行プロジェクトを推進。



**深本 大**  
**Fukamoto, Hiroshi**

デロイト トーマツ コンサルティング  
合同会社  
マネジャー

電機メーカーを経て現職。ハイテク、電子部品業界等を中心に、事業戦略、営業・マーケティング戦略、デジタル・トランスフォーメーション、オペレーション改革、経営管理基盤強化など、グローバルレベルで幅広いテーマのプロジェクトに従事。

1. TMT Predictions 2019「スマートスピーカー グローバル版：普及は過渡期に」参照(P17～)
2. 世界モバイル利用動向調査, デロイト トーマツ, 2018: <https://www2.deloitte.com/jp/ja/pages/technology-media-and-telecommunications/articles/com/mobile-consumer-survey-2018.html>
3. 日本人の音声操作に対する意識調査2017, KDDI, 2017/10/05: <http://news.kddi.com/kddi/corporate/newsrelease/2017/10/05/besshi2726.html>
4. 「主戦場」は指から「声」へ 拡大するAIスピーカー市場, 「週刊エコノミスト」第96巻 第20号, 2018/5/22
5. TMT Predictions 2019「AI グローバル版：利用の「民主化」が進む」参照(P9～)
6. Model can more naturally detect depression in conversations, MIT News, 2018/8/29: <http://news.mit.edu/2018/neural-network-model-detect-depression-conversations-0830>