

電話での受付や見守りを自動化する「AI電話サービス」

サービスイノベーション部

かわせ ともこ
川瀬 智子

5G・IoTビジネス部

おぐり しん
小栗 伸

サービスデザイン部

さいとう ゆうき
斉藤 優樹

クラウド型コンタクトセンタシステムが普及し、コンタクトセンタを導入する事業者が増加しているが、オペレータの人材不足が問題となっている。そこで、定型の電話応対および応対後の事務処理の自動化を目的に、音声認識による本人確認を特徴とする電話応対自動化サービス「AI電話サービス」を開発した。これにより、予約や申込みの受付、高齢者の見守りといった業務の自動化が可能となる。

1. まえがき

顧客満足度向上を目指す上で、コンタクトセンタなどの電話窓口は、ITリテラシーを要さない顧客接点として意義が大きい。近年、クラウド型コンタクトセンタ^{*1}が普及し、コンタクトセンタを導入する事業者が増加している。しかし、多様化する顧客に対して分かりやすく迅速に対応し、顧客満足度を向上させるためには、業務知識に加え、コミュニケーションスキルや通話後業務を遂行するためのITスキルがオペレータに要求されるが、それを満足するだけの人材が不足している。また人員配置の観点からは、繁忙期・閑散期に柔軟に人員を調整す

る必要がある。それに加え、新型コロナウイルス感染への対策として、コンタクトセンタのオフィスの人数低減も求められている。このような中、人工知能（AI）を活用した電話対応業務の自動化への需要が高まっている。海外では2018年ごろからAIで電話業務を支援するサービスの提供が進み、国内でも2020年には電話での申込み受付にAIを導入する事例が相次いで発表された。従来オペレータが担っている対応のうち、定型の対応をAIが代行することで、オペレータ稼働を非定型の対応に集中させることができる。また、通話後業務をAIが代行することで、オペレータに求められるITスキルを限定できるため、ITスキルを持ったオペレータの人材

©2021 NTT DOCOMO, INC.

本誌掲載記事の無断転載を禁じます。

本誌に掲載されている社名、製品およびソフトウェア、サービスなどの名称は、各社の商標または登録商標。

^{*1} クラウド型コンタクトセンタ：顧客対応のためのシステムで、自社でサーバを保有するのではなく、ネットワーク上のサーバを利用して運用するもの。

不足の軽減に繋がる。

ドコモでは、総合お問合せの電話窓口（総合センタ）の自動音声応答装置（IVR：Interactive Voice Response）^{*2}に、「音声認識IVR」[1]の機能をすでに実装している。これは、用件を話すと、適切な専門センタへ自動接続するものであり、AIエージェントサービス「しゃべってコンシェル^{*3}」および「my daiz^{*4}」[2]で蓄積された音声対話サービスのノウハウが活用されている。音声認識IVRの導入により、オペレータに接続されるまでの待ち時間削減や、総合センタでの電話応対時間削減、専門センタへの転送業務削減の効果が出ている。

音声対話技術を、さらに電話対応業務を効率化するソリューションとし、かつ顧客事業者に提供するために、ドコモは「AI電話サービス」というクラウドサービスを新たに開発した。本稿では、AI電話サービスのサービス概要とその仕組み、多様なユースケースに対応するための音声認識技術に関す

る仕組みについて解説する。

2. サービス概要

AI電話サービスは、自治体や小売店、飲食店、コールセンタをもつ企業といった事業者での利用を想定したサービスである。図1に示すように、受電でのユースケースとしては、サービス申込み／変更や、よくある問合せへの対応、飲食店や配車などの予約受付などがある。受電だけでなく架電にも対応しており、高齢者の在宅確認や体調確認による見守り、サービスの案内、予約や入金案内の確認・リマインドなど、コンタクトセンタの業務にとどまらない広い用途に適用できる。また、これまでオペレータが対応していた内容のうち、定型の対応業務を自動化するのみならず、音声対話技術とRPA（Robotic Process Automation）^{*5}の連携により、通話後の業務も自動化する。例えば、対話内容のログに基づい

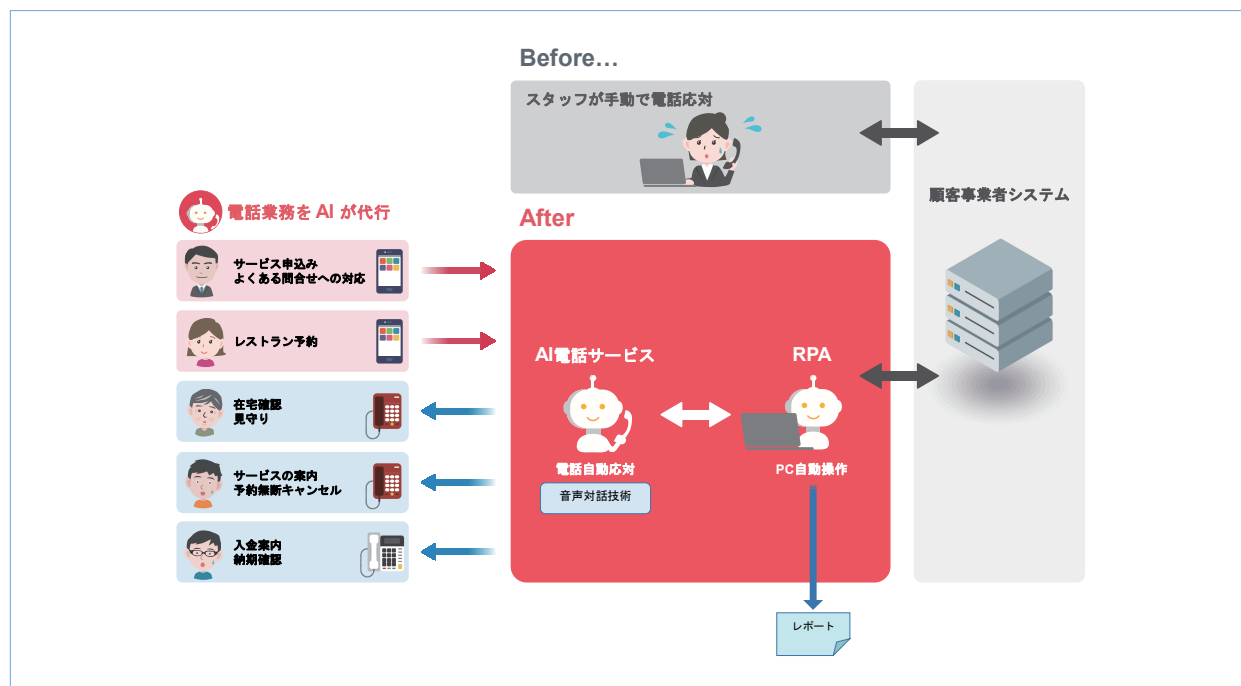


図1 AI電話サービスの概要とユースケース

^{*2} 自動音声応答装置（IVR）：電話で「〇〇の場合は〇番を押してください」のように音声で案内するシステム。

^{*3} しゃべってコンシェル：キャラクターとの会話や、会話を通じた電話発信、アラーム設定、乗換検索、占いを提供する、スマートフォンやタブレット上で動作する音声対話エージェント。

^{*4} my daiz：ユーザーに合わせた幅広い情報を提供する、スマート

フォンやタブレット上で動作する音声対話エージェント。

^{*5} RPA：定型業務を自動化する仕組み。

たレポートの自動作成や、顧客事業者のシステムとの連携が可能である。AI電話サービスの導入には、人材確保の問題の軽減に加え、24時間365日対応可能といった効果もある。

3. システム構成

AI電話サービスのシステム構成を図2に示す。「AI電話コアアプリケーション」が、ドコモの対話技術を提供する「ドコモAIエージェントAPI（Application Programming Interface）^{*6}」[2]とクラウド型のコールセンタサービス「Amazon Connect」を連携させることにより、電話応対自動化機能を実現している。また、電話を利用するため、音声インタフェースの機能を音声認識エンジンにより提供している。

ドコモAIエージェントAPIは、あらかじめ決められた対話シナリオに沿ってAIが受け答えする機能を提供しており、顧客事業者自身で柔軟に対話シナリオを作成できる点が特長の1つである。

Amazon Connectは呼制御機能を提供しており、これによりAI電話サービスを利用する顧客事業者は、呼制御サーバの管理が不要になり、拡張も容易というメリットが得られる。オペレータへの転送機能もあるため、AIによる自動応対が難しいケースが発生した場合にはオペレータによる対応に切り替えられる。一方、Amazon Connectを採用する制約として、使用できる音声合成^{*7}エンジンは「Amazon Polly」のみとなる。Amazon Pollyでは、AIの音声として特定の人物の合成音声を再現することはできず、日本語の場合、提供されている話者は男女1名ずつであるため、選択肢はそのどちらかのみとなる。ただし、選択した話者の声の範囲で話速や間、音量のチューニングにより、例えば重要な言葉をAIにゆっくり大きく話させるといったことは可能である。

AI電話コアアプリケーションは、Amazon Connectから取得したユーザ音声を音声認識エンジン^{*8}に逐次送り、認識結果として発話テキストを逐次受け取る。音声認識エンジンでユーザの発話の終端を

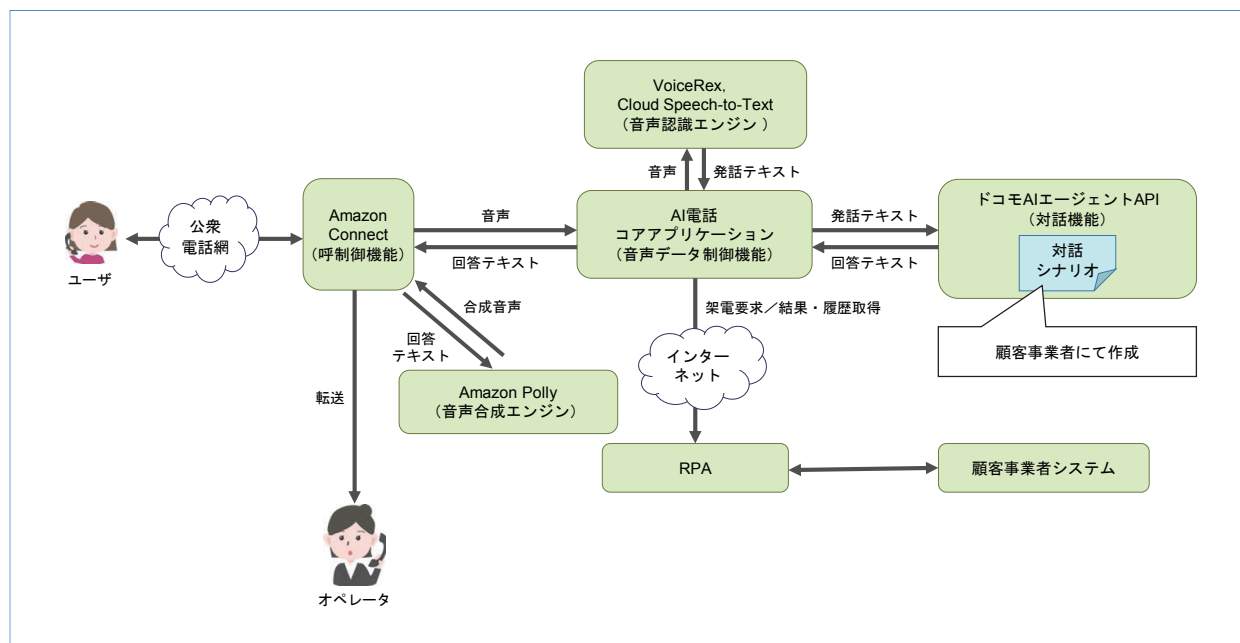


図2 AI電話サービスのシステム構成

^{*6} API：ソフトウェアの機能を他のプログラムから利用できるように切り出したインタフェース。

^{*7} 音声合成：テキストから人工的に音声データを作り出し、テキストを読上げできるようにする技術。

^{*8} 音声認識エンジン：音声データを入力し、発話内容をテキスト化する装置。

検出すると、AI電話コアアプリケーションがドコモAIエージェントAPIやRPAと連携してその後の処理を実行する。

音声認識エンジンには主にNTTの「VoiceRex^{*9}」とGoogleのCloud Speech-to-Textを使用しており、ユーザ発話ターン単位で選択できる。エンジンごとに強みの領域が異なるため、聞き取る内容に応じて、あらかじめシナリオでエンジンを指定できるようにしている。例えば、電話口で本人確認が生じる場合、氏名の聞き取りが必要となるが、VoiceRexでは音声認識結果を漢字と読み仮名の両方で出力でき、単語の分類が「姓」や「名」であるという情報も出力できるため、日本人の氏名を正確に認識するために有用である。

4. 多様なユースケースに対応するための音声認識技術に関する取り組み

AI電話サービスでは、VoiceRexに、しゃべってコンシェルやmy daiz、音声認識IVRで実績のあるドコモ独自の言語モデル^{*10}を適用している。その上で、VoiceRexを多様なユースケースで電話応対自動化に適用するために、次の3つのことに取り組んだ。

4.1 氏名の音声認識

サービス申込み／変更や予約の受付といったユースケースでは、本人確認のシナリオを完遂させることが重要であり、そのためには、氏名を正確に音声認識させる必要がある。そこで、氏名を学習データとして前述の言語モデルに追加するチューニングを実施した。個人情報の制約により、実在の氏名は利用できないため、架空の日本人の氏名を生成して学習データとした。チューニング前後の氏名認識性能を評価したところ、誤り率は、チューニング前の誤り率の7割以下まで減少した。

4.2 シナリオ特有の単語の音声認識

対話シナリオは顧客事業者ごとに異なり、シナリオ中で想定されるユーザ発話も対話シナリオごとに異なる。あるシナリオで特有の頻出単語が、一般的なAIエージェントとの対話ではまれな場合、誤認識することも珍しくない。例えば、飲食店予約のユースケースでは、ユーザが「個室」という単語を発話することが頻繁にある。しかし、前述の言語モデル内部では「個室」の出現頻度がそれほど高くないため、「保湿」「皇室」といった単語に誤認識されることがある。また、言語モデルに含まれていないサービス名などは原理上、認識できない。そのため、一般的には前述したように、事前に言語モデルをチューニングする必要がある。しかし、対話シナリオを追加するたびに言語モデルをチューニングして音声認識エンジンに実装することは、計算量の観点からも運用の観点からも現実的ではない。そこで、VoiceRexでは、音声認識リクエストごとに想定単語のリストを指定することで、言語モデルを変更することなく、指定された単語が出力されやすくなる機能を備えている。この機能を活用することで、音声認識性能を向上させることができた。

4.3 発話内容に応じた適切なタイミングの応答

AIエージェントによる音声対話では、応答の早さ、すなわち、ユーザ発話が終端してから短時間でAIによる発話を再生することが、高いユーザ体験に繋がる一要素である。一方で、住所や連続番号、自由回答など、ユーザがポーズ（呼吸や間）をおいて発話する場合、そのポーズを発話の終端と見なしてしまうと、AIがポーズより後の発話を聞き取れなかったり、ユーザ発話を遮って応答し始めてしまったりすることになる。つまり、ポーズを含む発話の聞き取りの成功率と応答の早さがトレードオフの関係となる。

^{*9} VoiceRex：NTTメディアインテリジェンス研究所が開発した音声認識エンジン。

^{*10} 言語モデル：単語の並び方の頻度を表現したモデル。

そこでAI電話サービスでは、シナリオ内のユーザの発話ターンごとに、想定される発話内容に応じて許容ポーズ長を設定し、AI電話コアアプリケーションが音声認識エンジンに対して動的に指定している。例えば、「ご氏名はドコモ太郎様でよろしいでしょうか」というAIの発話後は、「はい」または「いいえ」といった短いユーザ発話が想定されるため、許容ポーズ長を数百ミリ秒と短く設定する。一方、「健康のために気を付けていることはありますか」というAIの発話後は、ユーザは考えながら長く発話することが想定されるため、許容ポーズ長を1秒以上と長く設定する。これにより、前述のトレードオフを解消し、短いユーザ発話にはテンポ良く応答しつつ、ポーズを含むユーザ発話も最後まで聞き取れるという効果が得られる。

5. 実証実験

ドコモは、AI電話サービスを商用提供する前に、

試験用環境を構築し、2つのユースケースで実証実験を実施した。

5.1 本人確認を伴う申込み受付

月額サービスを提供する事業者における、電話応対業務削減の効果を確認するために、受電での申込み受付のユースケースで実証実験を実施した。ここでの申込み受付は顧客事業者システム内のユーザデータベースと連携した本人確認を伴うため、図3に示すようなシナリオを設計、適用した。名義をユーザデータベースで検索するだけでユーザが一意に特定できれば、受付完了となる。名義確認、お客様番号確認、住所確認まではユーザが一意に特定できなくても、料金支払いに関する確認と合わせることで一意に特定できる場合は受付完了となる。本実証実験では、受付対話完結率77%という結果が得られた。なお、実証実験では音声入力のみで対話を進めたが、商用システムはダイヤルキー入力にも対応しており、ダイヤルキー入力と併用した場合の対

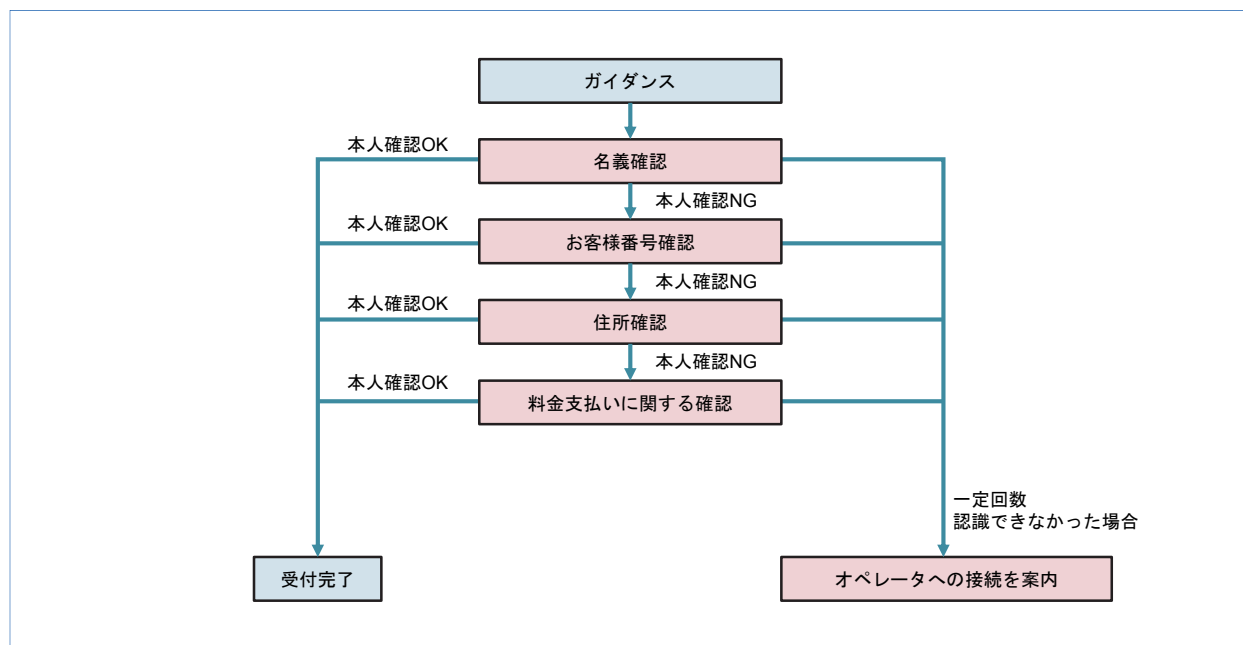


図3 本人確認を伴う申込み受付のシナリオイメージ

話完結率は88%となる見込みである。

5.2 高齢者見守り

独居高齢者は他者とのコミュニケーションが少なくなりやすく、毎日声をかけてもらうなどの支援を必要とするが、地域の支援機関も高齢者1人ひとりをケアするだけの人手が足りない、という問題が起

きている。そこで、AI電話サポートが独居高齢者や支援機関の課題解決に繋がるかを検証するために、独居高齢者に定時に自動で架電し、体調や安否を確認する実証実験を実施した。電話では、AIから表1のような質問をし、高齢者と会話する。高齢者見守りの効果は定量的な測定が難しいため、対象者へのヒアリングを実施した。ヒアリング結果を基に、

表1 高齢者見守りでの質問

項 目	質 問
睡眠について	昨夜はよく眠れましたか。
	昨日は何時ごろに寝ましたか。
	寝ている途中に目が覚めることはありましたか。
	何かほかに睡眠で気になることはありますか。
食事について	昨日は三食食べられましたか。
	今週お肉やお魚、卵などのタンパク質は食べましたか。
	食欲はありますか。
	何かほかに食事で気になることがありますか。
活動について	昨日は外出されましたか。
	今日外出される予定はありますか。
	家族や友人とお話しされましたか。
	来週やってみたいと思っていることがありますか。
	どんなことをやってみたいですか。
体調について	今日の体調はどうですか。
	お通じはありますか。
	最近、病院は受診されましたか。
	健康のために気を付けていることはありますか。
	どんなことに気を付けていますか。
身の回りのことについて	昨日は入浴されましたか。
	湯船には浸かっていますか。
	歯のお手入れは毎日できていますか。
	何かほかに身の回りのことで気になることはありますか。

2021年2月から2次実証実験を実施している。

6. あとがき

本稿では、電話応対を自動化するAI電話サービスについて解説した。本サービスでは、電話応対業務に従事する人材不足の問題を軽減する効果が期待できる。2020年12月からAI電話サービスの商用提供が開始されており、申込みや予約の受付、高齢者の見守りといったユースケースで、実証実験の継続

や正式なサービス開始が計画されている。今後は、音声認識性能のさらなる向上という技術的課題に継続的に取り組む。

文 献

- [1] 橋本, ほか: “AIによるコールセンタお客様満足度向上とオペレータ業務効率—音声認識IVRの開発—,” 本誌, Vol.25, No.4, pp.6-11, Jan. 2018.
- [2] 大庭, ほか: “ドコモAIエージェント・オープンパートナーイニシアティブ,” 本誌, Vol.26, No.3, pp.6-11, Nov. 2018.