

# 対話AI技術の金融サービスへの適用検証

## デジタル対話サービスによる業務改善アプローチ

コンシューマ市場でのスマートフォンやスマートスピーカーの台頭、企業でのチャットボットや音声認識サービスなどの試行導入など、AI技術を活用した音声／対話インタフェース「対話AI」が注目されている。対話AIは、対話情報をデジタル化することで新たなサービス創出や業務改善をめざすものである。

本稿では、この対話AIを金融機関向けの実ビジネスへ適用するにあたって、既存チャットボットで抱えている課題とそれらを解決するためのアプローチについて考察し、今後の対話AIに求められる技術・機能について述べる。

白井 剛 | Shirai Takeshi

山本 正明 | Yamamoto Masaaki

浅野 優 | Asano Yu

藤田 雄介 | Fujita Yusuke

津浪 克行 | Tsunami Katsuyuki

### 1. はじめに

近年、ディープラーニングに代表されるブレイクスルーや計算機の性能向上により、企業のAI (Artificial Intelligence) 技術活用に対する期待が高まり、各新聞・雑誌の紙面で注目を集めている。そのAI技術活用の一

つに対話AIがあり、対話AIは音声（アナログデータ）をテキスト（デジタルデータ）に変換する音声処理とテキストの文の意図を解釈する言語処理で構成されている（図1参照）。

この対話AIを、コマンドとして声を認識する機械や質疑応答を人間の代わりに行うFAQ (Frequently Asked Questions)チャットボットのような新しいインタフェースとして活用することで業務効率化が図れる。そのため、

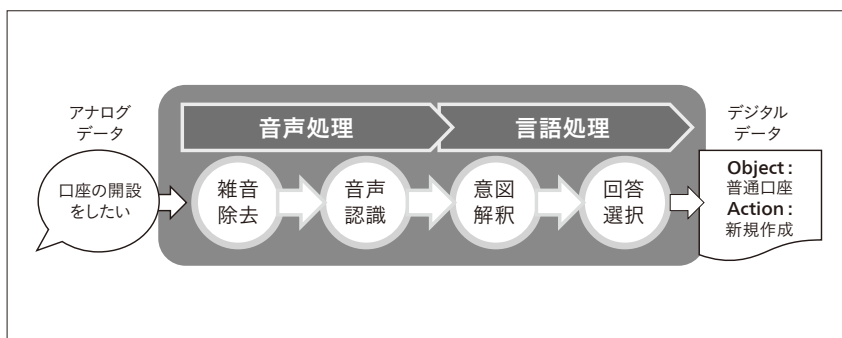


図1 | 対話AI技術の概要

音声処理と言語処理を使うことで、利用者の発話内容からシステムで処理できる単位のデジタルデータに変換する。

少子高齢化による労働人口の減少の問題に直面している日本の社会において、課題解決の手段として期待されている。

本稿では対話AIを金融機関の業務へ適用する際の課題を整理するとともに、これらの解決アプローチについて考察を述べる。

## 2. 対話AIの動向

近年、スマートスピーカー、チャットボット、コミュニケーションロボット、音声アシスタントなど、さまざまなシーンで対話をインタフェースとするデバイスが身近に広がり、対話インタフェースへの市場の期待も高まっている。

### 図2| インタフェースの変化

近年、目的達成の方法が申込書への記入から対話へと大きく変わってきている。この変化は、「書く」→「打つ」→「話す」というように、利用者の理解する手間や記入する手間が不要となる利用者本位のインタフェースへと進んでいる。



表1| チャットボットや音声認識を利用したサービスの市場での使われ方

チャットボットや音声認識を利用した業務サービスはさまざまな領域で使われ始めており、今後も拡大していくと想定される。

業務での位置づけ	使われ方	業務例	種別		価値
			CB	音声	
フロントエンド業務 (顧客折衝, 営業 ほか)	24時間化	電話サポート受付24時間化	○		量的サービスの向上
		コンシェルジュサービス		○	
	多言語対応	SNS/チャットによるリアルタイム翻訳	○		
		Skype <sup>※</sup> によるリアルタイム翻訳		○	
オムニチャネル化	SNS/チャットによる変更受付	○			
	音声による再注文		○		
コンプライアンスリスク軽減	マーケティング個別化	モニタリングやチェック支援		○	
		コーディネート提案	○		
バックエンド業務 (社内向け業務, 人手作業 ほか)	折衝記録	音声による注文指示		○	質的サービスの向上
		議事録の書き起こし		○	
	満足度向上	口頭注文のテキスト化		○	
		SNS/チャットで購入が完結	○		
コールセンター・ヘルプデスク支援	ハンズフリー化	コールセンターで解約抑止		○	業務負荷の軽減
		顧客問い合わせのFAQ前さばき	○		
		社内IT問い合わせ回答を自動化		○	
		営業日報入力		○	
		作業記録入力		○	

注：略語説明ほか  
 SNS (Social Networking Service), FAQ (Frequently Asked Questions), CB (Chatbot)  
 ※) Skypeは、Microsoftグループ企業の登録商標です。

## 2.2

### デジタル対話の広がり

対話AIのチャットボットによる問い合わせの自動応答や、音声処理のみを利用した議事録作成など、これらの技術はさまざまな用途に活用できる。すでに市場で使われているチャットボットや音声認識を利用したサービス例を表1に示す。金融機関の「フロントエンド」から「バックエンド」まで幅広い領域で使われており、提供できる価値は単純な業務負荷の軽減だけではなく、量的／質的サービスの向上にも寄与する技術だと考えられる。

## 3. 対話システムを業務へ適用する際の課題

現状普及している対話システムの一つにチャットボットがあるが、業務へ適用する際、応答精度が十分でないことや運用メンテナンスに多くの工数が必要になることなどの課題がある。また、音声認識は音声コマンドによる家電の操作や会議での議事録作成などでの活用の例があるが、十分な精度で認識するためには口元にマイクを近づける必要があることや、人数分のマイクを用意する必要があるなどの課題がある。

## 3.1

### 従来型（一問一答型）のチャットボット

従来型の対話システム（日立開発）の対話方式は、利用者の1つの質問文に対して1つの回答文を出力する一問一答型対話方式である。処理としては、利用者の質問文と事前登録しておいた想定質問文との類似度を計算する。類似度が高い場合、想定質問文とひも付けられた回答文を選択して出力する（図3参照）。

しかし、同じ質問内容に対して冗長な表現が多数あり、従来のシステムでは入力された質問に対応する想定質問文を抽出できないことが課題となっていた。そこで、さまざまな言い換えパターンが主に7種類に分類できることに着目して定義し、従来システムが質問文の言い換えパターンごとに正しく回答できる確率（正答率）を評価した（表2参照）。評価データとしては、FAQデータ（195件）から人手で質問文の言い換え（7種類×195件）を作成し、言い換えパターン別のシステムの正答率を評価した（図4参照）。

評価の結果、想定した情報（例：業務名、端末名、画面名など）が欠落した言い換え（パターン6、7）が正答率を著しく低下（39～44%）させていることが判明した。

図3 従来型（一問一答型）のチャットボット

1つの質問につき1つの回答を選び出し、回答を行う形式である。自然言語処理技術を入れることで、1つの質問で相手の意図の絞り込みが可能となる。

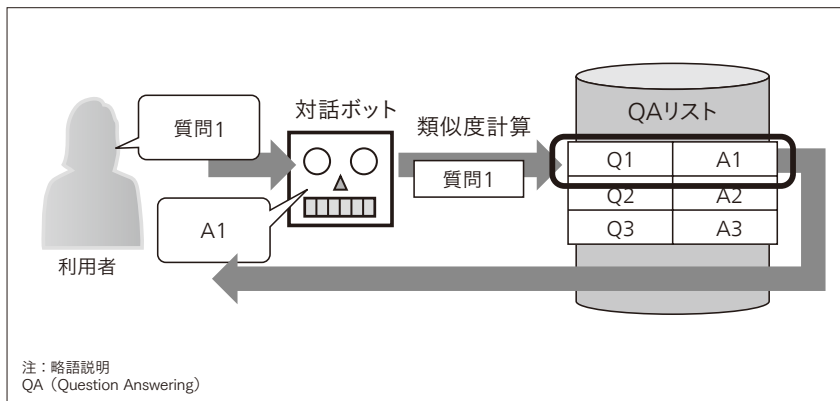


表2 言い換えパターン

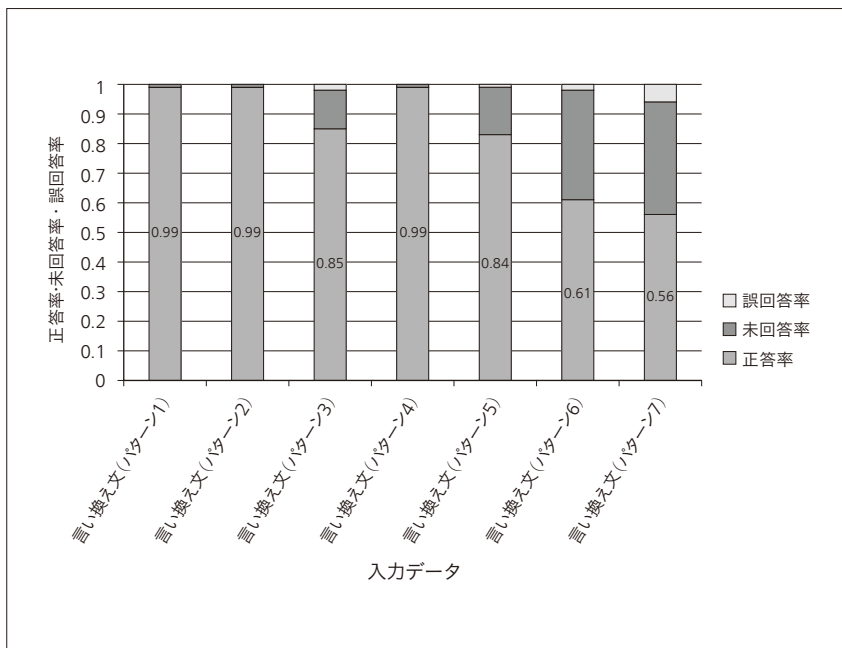
原文の内容を聞きたいときに、利用者が質問すると想定される言い換えパターンを列挙した。特にパターン6、パターン7ではさまざまなケースも含まれるため、回答を長文にして正答にみせるような応答文の作成が必要とされる。しかし、応答文が長文になると読む手間が増え、結局利用者には理解されなくなってしまう。

パターン	概要	例文
-	原文	運用管理端末を使って、サービスメニュー画面の印刷ボタンをクリックしても反応しない。
1	語尾の言い換え	運用管理端末を使って、サービスメニュー画面の印刷ボタンをクリックしても反応し <b>ません</b> 。
2	てにをはなどの言い換え	運用管理端末を使って、サービスメニュー画面で印刷ボタンの <b>クリックを</b> しても反応を <b>し</b> ない。
3	同義語の言い換え	<b>イントラウェブ</b> を利用して、メニュー画面の <b>プリント</b> ボタンを <b>押下</b> しても <b>無反応</b> 。
4	句の順番の入れ替え	印刷ボタンを <b>サービスメニュー画面</b> でクリックしても反応しなかった。 <b>運用管理端末を使った</b> 。
5	異なる語彙と構文	<b>イントラウェブのメニュー画面</b> で、印刷をクリックしても <b>何も起きなかった</b> 。
6	含意などの言い換え	<b>印刷を実行</b> しても反応しない。
7	目的と手段の言い換え	サービスメニュー画面で <b>印刷したい</b> 。

正答率が低くなるパターン6, 7

図4 言い換えパターン別の正答率

日立の意図理解エンジンを使用した際の正答率の精度を示す。パターン1やパターン4のように「てにをは」や句の入れ替えに関しては、ほぼ間違いなく吸収して正答を返すことができる。パターン3やパターン5は、語彙が登録されていなかったため、正答率が低下した（登録すれば正答を返した）ケースである。パターン6、パターン7などの含意ではあるが聞き方が違うケースや、必要な情報が欠落しているケースについて、正答率が著しく低下することがわかった。



3.2

システム主導による対話誘導

正答率の低い6と7のパターンに対応するため、システムから聞き返すことにより、言い換えの質問とそれに対応する内容を取得し蓄積することで正答率を向上させる方式（以下、「システム主導型対話」と記す。）を考案した。この方式は、対話の主導権をユーザーからシステムへと移行させ、回答の絞り込みに必要な情報をシステムが利用者に聞き返すことで取得／回答する。また、回答が複数存在する場合にはその絞り込みに必要な情報を得るため、主体的にシステムがユーザーに対して質問をする（図5参照）。

必要な情報がそろい、ほかの選択肢がなくなった時点で絞り込まれた質問文に回答する。なお、利用者からの情報抽出には機械学習CRF（Conditional Random

Fields）を使用した。これらの有効性を検証するため回答率の再評価をした結果、パターン6、7における対話失敗となる割合を約20%低減することが可能となった。

3.3

自発的な成長による運用コスト削減

チャットボットは運用メンテナンスにおいて膨大なコストが掛かるという課題がある。現状のチャットボットでは利用者からの問い合わせに回答するためには、さまざまな質問表現を登録する必要がある。しかし、最初にすべてを網羅的に用意しておくことは難しい。

そのため、システム稼働後に人手で対話ログを調べ、未登録の質問表現を後から追加する必要があった。

そこで、メンテナンスコストを抑えるため未登録の表現を自動で取得する成長型対話システム（以下、「成長

図5 提案方式(システム主導型対話)

図4のパターン6やパターン7の例のように、回答を特定するために必要な項目を利用者が質問文の中にすべて含めて質問するとは限らない。このとき、質問の中で不足している情報を対話ボットが判断して利用者に聞き返すことで、応答すべき回答を特定する。

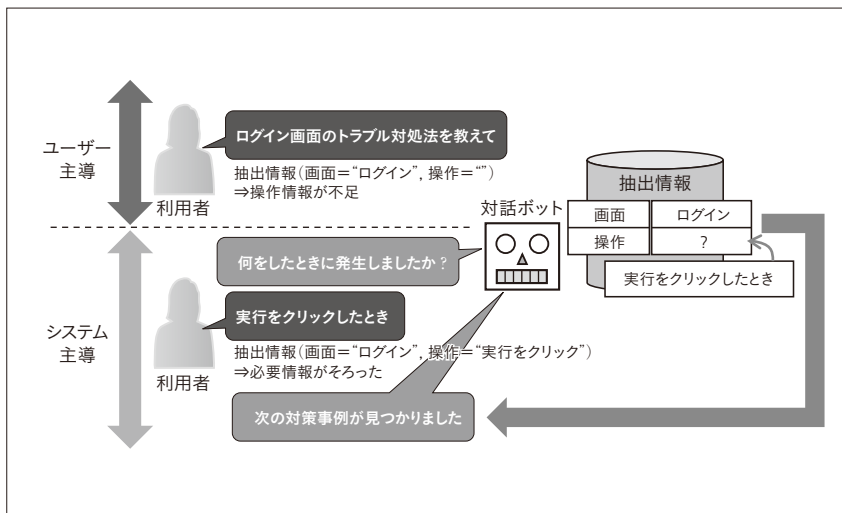
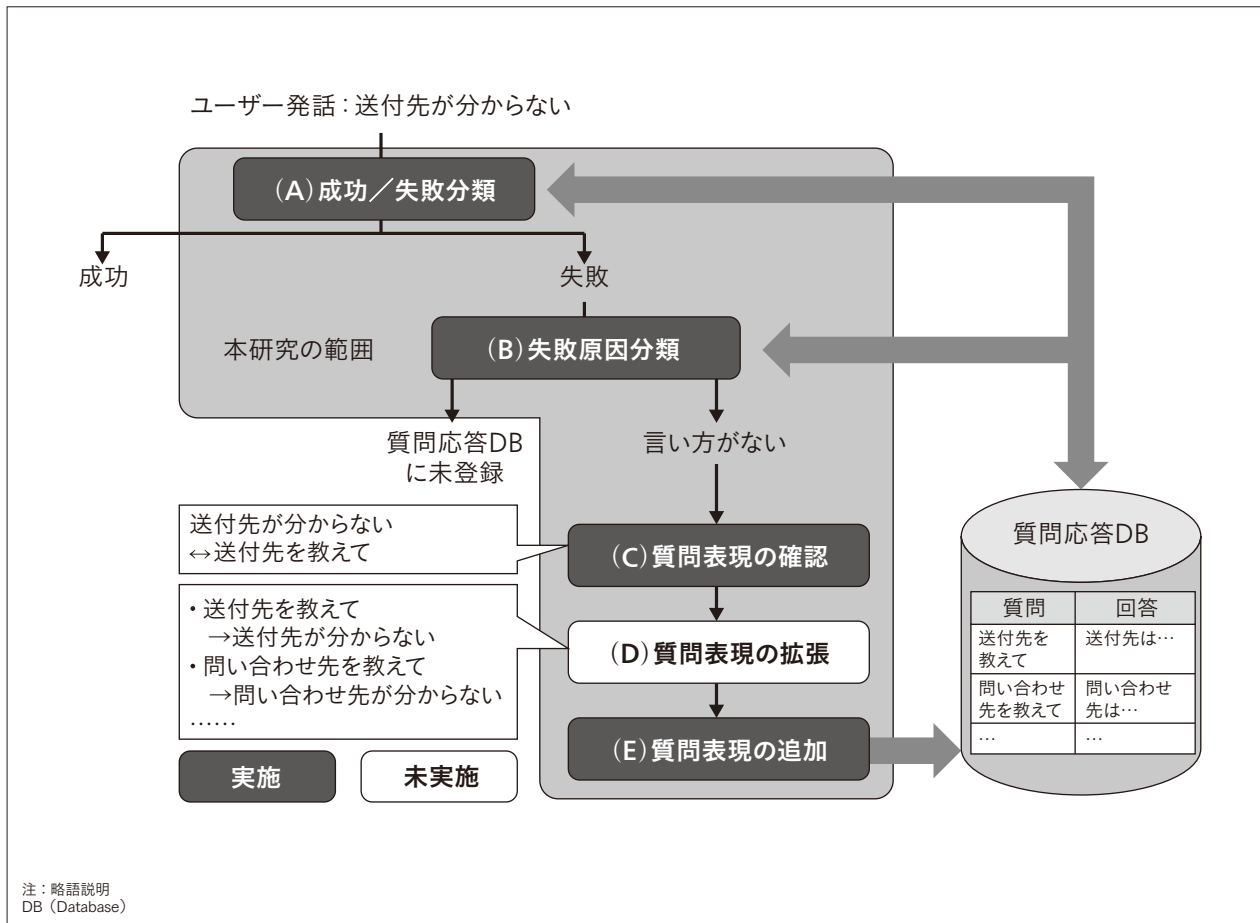


図6 | 成長型対話システムの仕組み

対話ログ内のユーザー発話を、(A) 正答できたものと正答できなかったものに分類する。その後、失敗と分類されたユーザー発話に対し、「(B) 失敗原因分類」を行い、失敗を「質問表現と回答が質問応答DBに登録されていない場合」と「回答は登録されているが適切な質問表現が登録されていない場合」に分類する。後者の場合は、「(C) 質問表現の確認」にて、ユーザー発話と質問応答DB内の表現が同じ意図かどうかを確認する。(C)が同じ意図の場合、「(D) 質問表現の拡張」で、同じ意図の2つの表現から言い換え規則を生成、質問応答DB内の全ての質問表現に規則を適用し、新たな質問表現を獲得する。最後に、「(E) 質問表現の追加」にて、ユーザー発話と(D)で拡張した質問表現を質問応答DBに追加する。



型対話」と記す。)の仕組みを取り入れた(図6参照)。この仕組みの運用コスト削減効果を検証するために日立のコールセンターにおけるオペレーターログを用いてシミュレートしたところ、質問表現を1件増やすコストが60%削減できた。

### 3.4

#### インタラクティブな対話の実現

従来のチャットボットに「システム主導型対話」や「成長型対話」を導入することで、チャットボットの課題である運用コストを抑えながら正答率を向上させることが実現できた。しかし、一問一答以外の対応が可能になったため、正しく回答するためには質問に応じて適切な対話方式を選ぶ仕組みが必要になった。

さらに、利用者には対話の途中でも分からない用語があれば質問をしたいというニーズがあるため、このニーズに対応する仕組みも必要になった。

このような理由から、対話タスク切り替え機能を実装

し、利用者に合わせてシステムが応答内容を切り替えることができる対話システムを考案した(図7参照)。

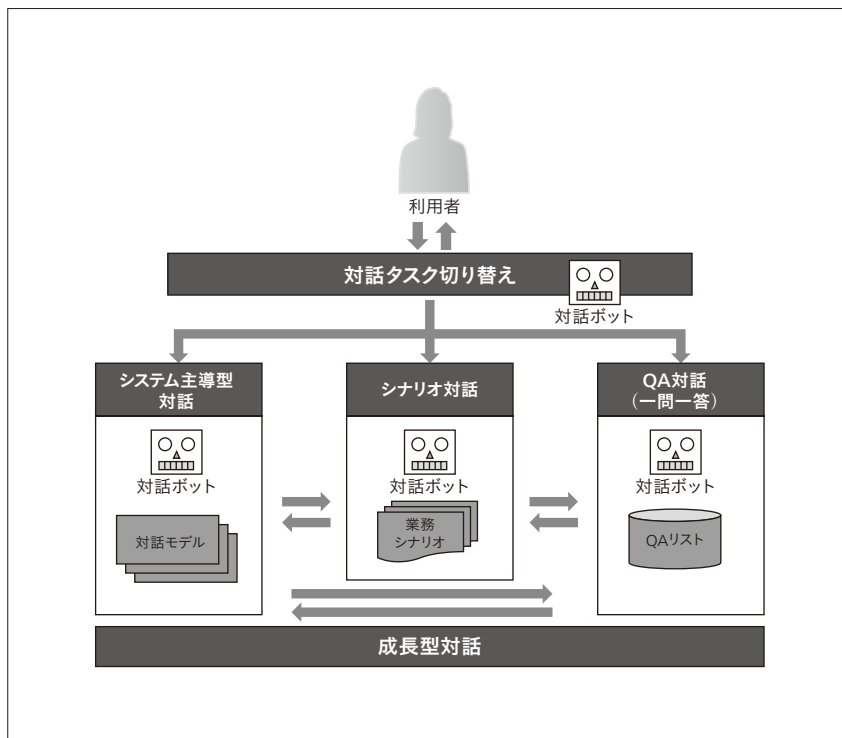
## 4. 音声処理の課題と音源分離技術

スマートスピーカーの市場投入や働き方改革の影響もあり、普段の会議や店頭での顧客との折衝において音声認識を気軽に活用することで、業務負荷の軽減や蓄積された対話情報を活用したいというニーズが高まってきている。しかし、現状では口元とマイクの距離が大きく離れると音声認識率が低下するため、電話機で音声を取得できるコールセンターや、グースネックマイクなど1人に1つ高精度のマイクを用意して音声取得が可能な大規模な会議での適用ケースが多かった。

さまざまな理由から、ビジネスシーンにおいては利用者がマイクを意識しなくても対話内容がテキスト化されることが望ましい。そのため1つのマイクを置くだけで、

図7 対話タスク切り替え機能

対話タスク切り替えでは、利用者は対話モデルを意識することなく、利用者の質問内容から対話ボットが最適な答えを持っている対話モデルへ問い合わせ、最適な回答を答えられるようにする。



その場に居る複数人の対話内容をテキスト化できる技術の活用を考案した (図8参照)。

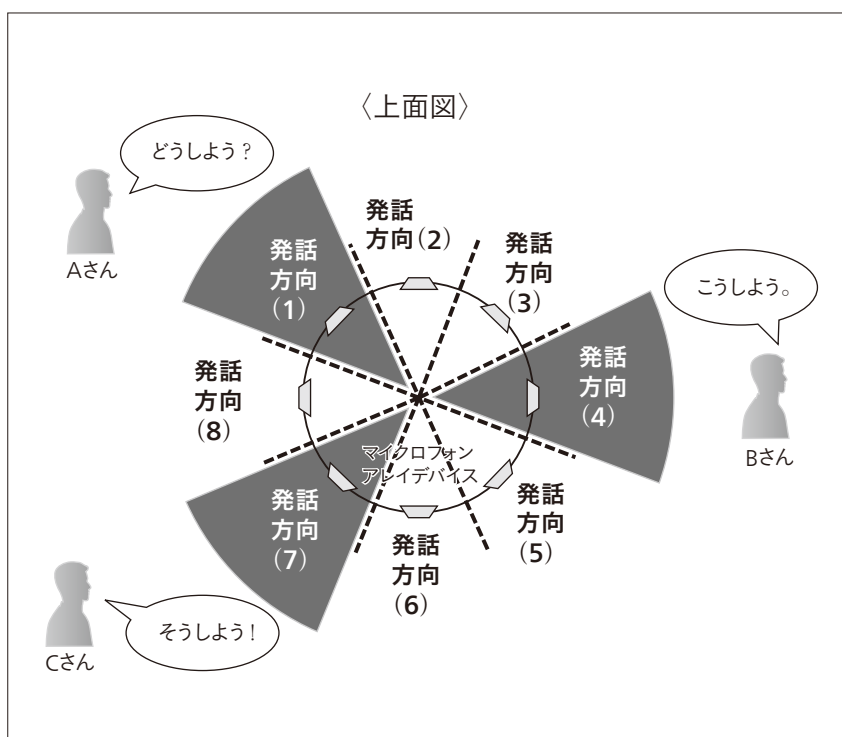
独自の音源分離技術により、複数のマイクロフォンを搭載したマイクロフォンアレイを使うことで話者の方向別に音声を分離することができるため、1つのマイクロフォンアレイで複数人の発話を分離してテキスト化することが可能になった。

これによりマイクを意識せずに対話内容の記録が可能

になり、数人の社内会議や営業店ローカウンターにおける顧客との商談などさまざまなシーンへの活用の可能性が広がった (図9参照)。さらに、取得データを分析することで商談内容のコンプライアンスチェックや顧客の嗜好 (し) 好把握が可能となり、サービスの質の向上へつながることが期待される。

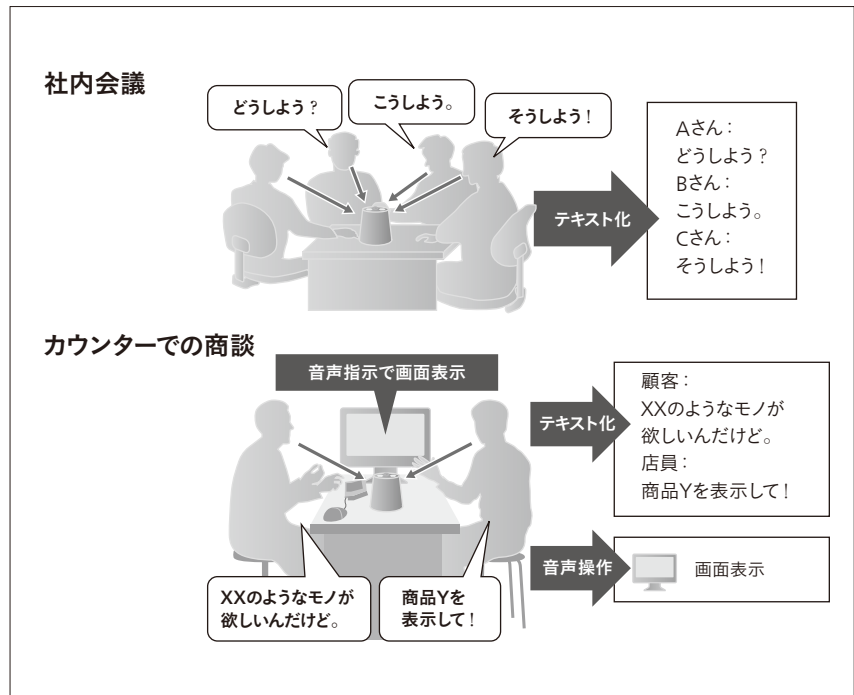
図8 発話方向で分けた音声認識

1つのマイクで発話の方向別に音声をテキストに変換することができ、Aさん、Bさん、Cさんがそれぞれの方向から発した音声をそれぞれ分けてテキストにすることが可能である。



## 図9| 音声活用範囲の拡大

社内会議において、会議中の宿題事項などをメモに取る代わりに発話することでテキスト状態で書き留め、書く手間を省くことにより会議を効率的に進める。また、商談中の折衝履歴の作成支援として音声のテキスト化の活用が可能である。



## 5. おわりに

本稿では、対話AIと音声認識を活用した業務の効率化について考察した。

今後、対話AI技術はユーザーとの単なるインターフェースの枠を越え、さまざまなビジネスシーンで多様な形で応用が期待できる。この取り組みを続け、デジタル対話を通じた顧客のビジネス革新へ貢献していきたいと考えている。

### 参考文献など

- 1) 日立ニュースリリース、自発的に成長する音声対話AI技術を開発 (2017.9), <http://www.hitachi.co.jp/New/cnews/month/2017/09/0928b.pdf>
- 2) 浅野優, 外: 不明点の確認により成長する対話システム, 言語・音声理解と対話処理研究会 81, p.66~71 (2017.10)

### 執筆者紹介



#### 白井 剛

日立製作所 金融ビジネスユニット 金融システム営業統括本部  
金融イノベーション推進センタ 所属  
現在、金融向けビッグデータ/AIに関わる企画業務に従事



#### 山本 正明

日立製作所 研究開発グループ  
デジタルテクノロジーイノベーションセンタ  
メディア知能処理研究部 所属  
現在、対話システムの研究開発に従事  
博士(工学)  
電子情報通信学会会員



#### 浅野 優

日立製作所 研究開発グループ  
デジタルテクノロジーイノベーションセンタ  
メディア知能処理研究部 所属  
現在、対話システムの研究開発に従事  
博士(情報科学)  
人工知能学会会員



#### 藤田 雄介

日立製作所 研究開発グループ  
デジタルテクノロジーイノベーションセンタ  
メディア知能処理研究部 所属  
現在、音源分離・音声認識システムの研究開発に従事  
日本音響学会会員、電子情報通信学会会員



#### 津浪 克行

日立製作所 サービス&プラットフォームビジネスユニット  
サービスプラットフォーム事業本部 IoT・クラウドサービス事業部  
アプリケーションサービス第1本部 所属  
現在、デジタル対話サービスの事業立ち上げに従事