

# 人を理解するAI技術を活用したMDIS音声ソリューションの高度化

西 健太郎\*  
Kentarou Nishi  
白浜広彬\*  
Hiroaki Shirahama  
羅 玉てい\*  
Gyokutei Ra

田口進也\*  
Shinya Taguchi  
藤田喜広\*  
Yoshihiro Fujita

Advanced Enhancement of MDIS Voice Solutions Utilizing AI Technologies that Understands People

\*三菱電機インフォメーションシステムズ㈱

## 要 旨

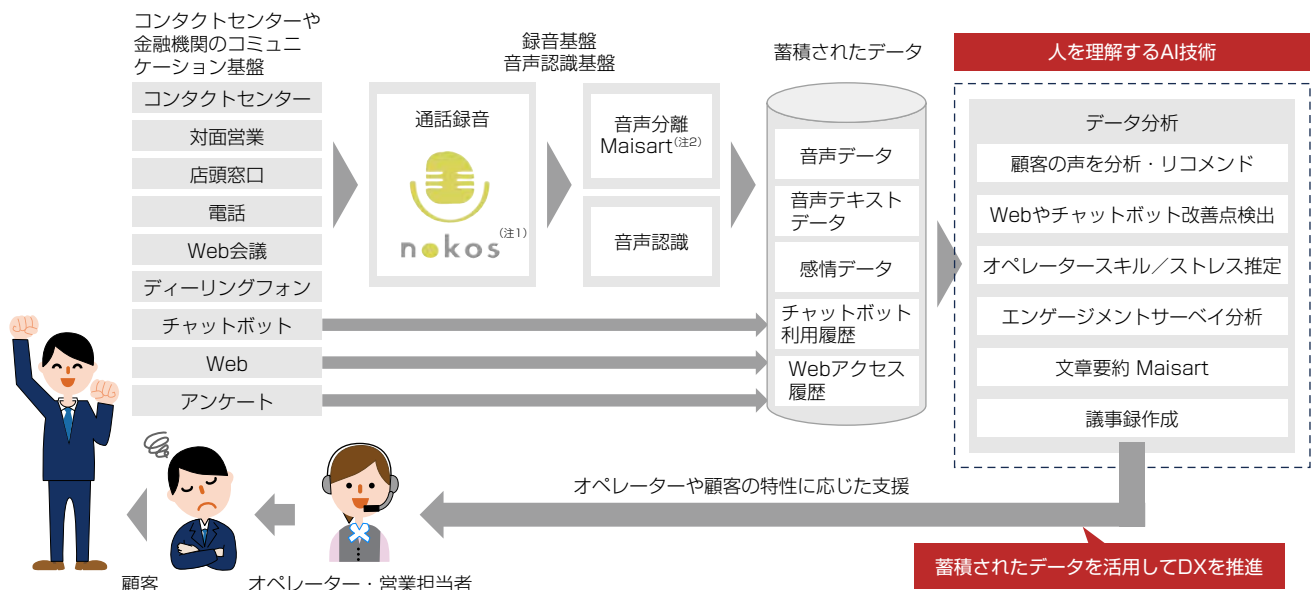
コンタクトセンターや対面営業では音声データの効果的な利活用が注目されている。三菱電機インフォメーションシステムズ㈱(MDIS)の音声データの感情分析技術は、音声だけでなく、音声から認識した感情値や顧客満足度なども含めた分析を行うため、オペレーターの対応スキル向上による顧客満足度向上と、オペレーターの離職傾向の推定による離職率低下が期待できる。また、三菱電機の高度な音声分離技術は、複数話者の音声混在するWeb会議や店頭窓口の会話を話者ごとに分離できるため、自動議事録作成とテキスト分析によって業務プロセスの効率改善が期待される。MDISは2000年代初頭の大手金融機関向けの通話録音システム導入以来20年間にわたって培った音声系システムのノウハウと“人を理解するAI技術”を活用した音声ソリューションで、コンタクトセンターや金融機関などでのデジタルトランスフォーメーション(DX)を支援し、市場競争力の向上と安心・安全・快適な社会実現に貢献する。

## 1. ま え が き

近年、コンタクトセンターや金融機関等の対面営業では、音声データの効果的な利活用が求められている。AI技術の進化によって、音声データのテキスト分析や感情分析が可能になり、顧客の苦情やニーズを正確に把握できつつある。これは商品やサービスの最適化につながり、企業の市場競争力を向上させる。また、コンタクトセンターでのオペレーター離職率の高さは深刻な課題で、感情分析でオペレーターのストレス度合いを推定して、早期対策につなげられることが期待される。

さらに、Web会議と音声認識AIの普及によって、Web会議や店頭窓口での会話の自動議事録作成が可能になり、バックエンド業務の合理化が期待されている。

MDISは、このようなトレンドに適応するために、音声ソリューションの整備を進めて、コンタクトセンターや金融機関での音声データをビジネスに結び付けて、導入企業のDXに対応していく(図1)。



(注1) nokos(ノコス)は、通話・会議を自動で録音/管理するMDISの音声ソリューションである。  
(注2) Maisart(マイサート)は、三菱電機AI技術ブランドの名称である。

図1-MDIS音声ソリューション

## 2. 市場動向と課題

### 2.1 市場動向

コロナ禍を契機にDX化が求められる時代になった。それに加えて顧客体験(カスタマーエクスペリエンス、以下“CX”という。)向上を目指す企業が増加し、顧客の生の声を拾うことができるコンタクトセンターの重要性が増してきている。コンタクトセンターのサービス品質を現場で支えているのはオペレーターである。従業員体験(エンployeeエクスペリエンス、以下“EX”という。)の向上がCX向上につながる好循環を生み出すため、人の力を最大化するテクノロジーへの投資を強化すべきと考えられている<sup>(1)</sup>。

また金融機関では電話、スマートフォン、Web会議ツール、コンタクトセンター、窓口販売、ディーリングフォン等、複数のコミュニケーションの手段が利用されており、コンプライアンスを目的とした録音は金融機関にとって重要な位置付けになっている。金融機関のうち、特に証券、銀行及び保険業界ではコンプライアンスリスクもあり、対面営業での会話の録音のニーズは更に強まることが想定されている。

### 2.2 課題

コンタクトセンターにはオペレーターの高い離職率やトレーニングの課題があり、これらの課題の解決はEX向上に必要不可欠である。オペレーターのパフォーマンスを継続的に評価し向上させる体制を整える必要があるが、膨大な対応履歴を全て人手で管理し、モニタリングすることは困難である。

また、AIの活用でこれらの課題を解決する場合、複数のコミュニケーション手段の音声テキストデータを統合的に管理し、活用することが求められる。金融窓口販売や訪問型営業に関しては、一つのマイク(以下“シングルマイク”という。)から複数話者の音声を取得することが多く、話者が混在した音声データが生成されてしまい、音声テキスト化が困難になっている。

## 3. 人を理解するAI技術

コンタクトセンターや金融機関等の対面営業でのCX/EX向上のために、当社は“人を理解するAI技術”の開発に注力している。3章では、そのうち音声分離技術と感情分析技術について述べる。音声分離技術によって、シングルマイクで録音した複数話者の音声を個々の発話に分離してテキスト化できる。また感情分析技術によって、オペレーターのスキルやストレス度合い、顧客の満足度を効率的に推定できる。

### 3.1 音声分離技術

音声分離はシングルマイクで録音された複数話者の音声を話者ごとに分離する技術である。既存パブリッククラウドサービスなどで音声分離の機能は提供されているが、これらは話者が切り替わるタイミングを特定して音声を分離する仕組みのため、同時に発話した場合の音声に対する音声分離には対応できないケースがある。一方で、三菱電機が開発した最新のAI技術“Maisart”を活用した音声分離技術は、複数話者が同時に発話している場合でも音声分離が可能である。図2に、話者の切り替わりを検知して分離する方法と、三菱電機の音声分離技術の音声分離処理の違いを示す。図2の左側は話者Aと話者Bが会話している音声データを示している。話者AとBが同時に会話すると両者の音声波形が重なる部分がある。この状態の音声データに対して、図2のように音声分離を行うと、図2上段に示す話者の切り替わりを検知する方法では同時に会話している部分で分離ができないケースがあった。図2下段の三菱電機が提供する音声分離技術で分離すると、波形が重なる部分も分離できている。

従来、金融店頭窓口・訪問型営業といったシングルマイクで録音している業務、及びモノラルの録音システムが採用されている環境下の業務については、音声データのテキスト化やデータ分析が困難であった。しかし、この技術によって音声分離が可能になるため、それらの業務でのデータ利活用が期待される。また、この技術はGPU(Graphics Processing Unit)などの高価な設備を必要とせず、CPUだけで動作する利点もある。導入のハードルが低いいため、既存の録音システムに低コストでアドオンするという利用方法も可能である。

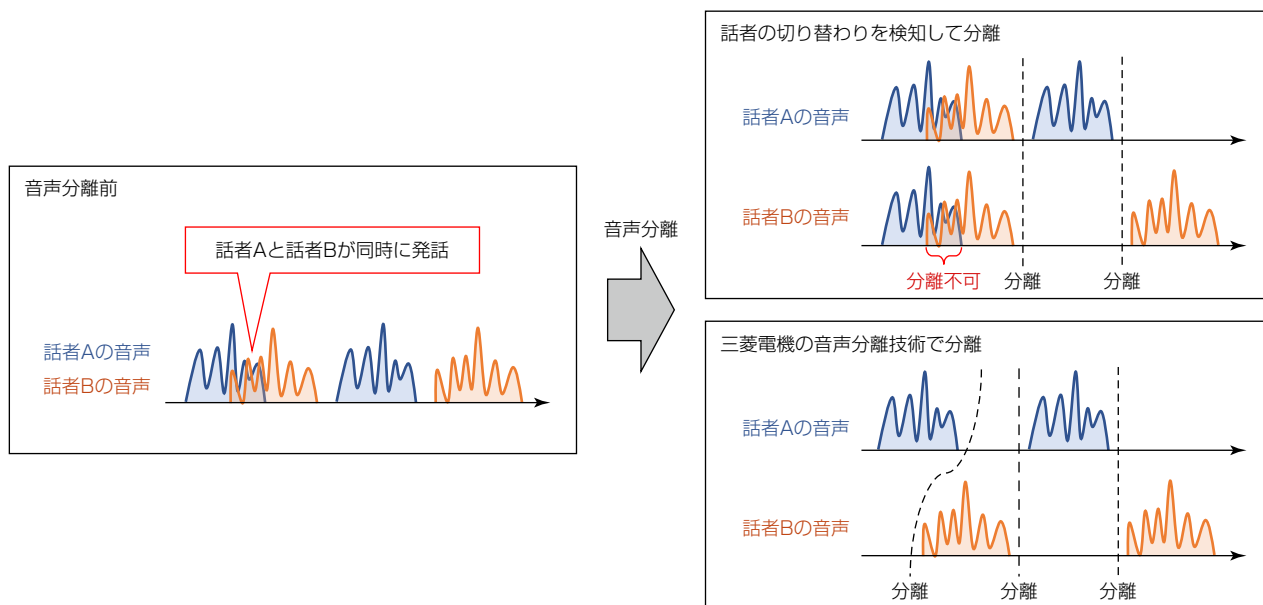


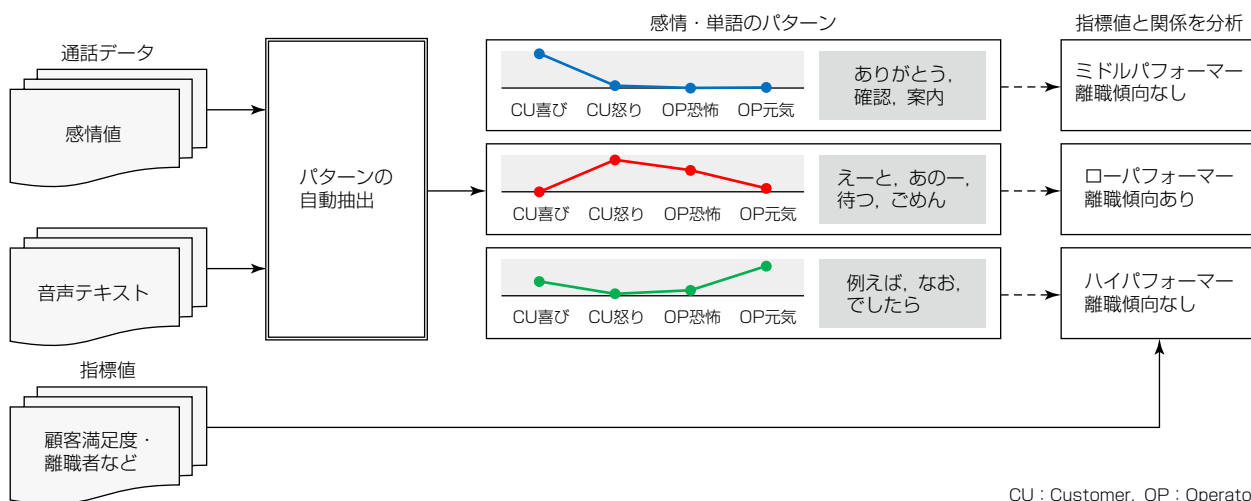
図2-音声分離処理の違い

### 3.2 感情分析技術

音声に対する感情分析とは、人間の声の大きさや速さ、抑揚などの物理的な特徴を解析することで、喜び、怒り、恐怖などの感情状態を推定する技術である。感情分析をコンタクトセンターでの通話音声に適用することで、顧客とオペレーターの感情を把握できるようになり、顧客満足度や、オペレーターの対応品質向上・離職防止への活用が期待できる。

しかしながら、従来の活用では、感情値のリアルタイムでの監視や、感情値だけの傾向分析や集計などにとどまっており、業務改善まで至ったケースは少ない。これは、感情値だけでは具体的な改善施策の立案が難しいためである。通常、コンタクトセンターでの対応業務では、会話内容や顧客に応じて声の抑揚や言葉遣いなどの話し方をオペレーターが適切に切り替える必要があり、会話内容や文脈を把握した上で感情を分析することが重要になる。

MDISでは、音声から認識した感情値と音声のテキストを同時に分析する感情分析技術を提供している（特許出願中）。図3に感情分析の概要を示す。音声データから認識した感情値と音声テキストを基に、複数の通話に共通する感情と単語のパターンを自動で抽出する。各パターンには、感情だけでなく音声テキスト内の特徴的な単語も抽出できるため、より具体的な対応内容を把握できる。さらに、抽出したパターンと顧客満足度や離職者などの指標値との関係を分析することで、パターンごとのオペレーターの対応スキルレベルや離職傾向を推定可能になる。



CU : Customer, OP : Operator

図3-感情値と音声テキスト同時分析の概要

## 4. 期待される効果

“人を理解するAI技術”によって、MDISが提供する音声ソリューションは付加価値の高いデータ分析機能を提供する。オペレーターや営業担当者を支援することで、オペレーターの応対品質向上や業務プロセスの効率改善が見込める。また、オペレーターの高い離職率を抑える効果も期待できる。

この技術は学習済みのモデルであるため、一般的な特徴や知識を獲得しており、これを新しいタスクに適用することで、データ量や計算資源が少なくても良好な性能を達成することが期待できる。また、蓄積されたデータやコンタクトセンターの応対マニュアルなどを生成AIと組み合わせれば、電話応対などの顧客対応も支援できるため、MDIS音声ソリューションによるDX推進の適用業務領域は更に拡大が可能である。

## 5. む す び

生成AIの実用化が進んで、サービス業での接客対応や保守員の教育、営業支援など、人間の行動主体に関わる分野に広がりを見せている。しかし、ハルシネーションや再現性の問題などから、リスクのある業務については今後も人が担うことが想定される。そのため、人を中心としたシステムの設計・構築がますます重要になる。

MDISは、人手不足が懸念されているコンタクトセンターのような知識・労働集約型の業務に対して、導入企業やエンドユーザーが“安心・安全・快適な社会”を実現するため、20年間培った音声系システムのノウハウを活用しつつ、顧客との共創を通じて、AI技術を活用した人を理解する音声ソリューションの提供を進めていく。

## 参 考 文 献

- (1) デロイト トーマツ グループ：デロイト トーマツ調査、日本企業のコンタクトセンター AI導入率が約50%に伸びるも、成果創出には苦戦(2023)  
<https://www2.deloitte.com/jp/ja/pages/about-deloitte/articles/news-releases/nr20230808.html>