

DXを実現する自然言語処理技術

サービスイノベーション部

 しらみず ゆうたろう おか けいすけ
 白水 優太郎 岡 慶介
 たつみ しゅうすけ
 辰巳 守祐

近年、AIを利用してDXの推進に資する技術開発がさかんである。特に、企業に蓄積される大量のテキストデータを高速かつ正確に処理するAI技術として、自然言語処理と呼ばれる分野に注目が集まっている。本稿では、ドコモが開発している自然言語処理AIとその技術的特徴、およびAIの利用をサポートするGUIツールを紹介するとともに、実際の導入事例やRPAツールとの相乗効果、今後の展望について解説する。

1. まえがき

近年、「働き方改革」が政府主導で進められており、RPA (Robotic Process Automation) ツール*1やチャットボット*2をはじめとするデジタル技術を活用した、ビジネスの加速、生産性の向上、稼働の削減などが強く求められている。特にRPAツールは、アンケートやユーザからの問合せをはじめとする大量のテキストデータの処理業務において改善効果が大きく、フロントオフィスやバックオフィスを中心に普及が進んでいる。さらに、テキストのカテゴリ分類やタグ付けのように、知的な判断を必要とする作業のデジタルトランスフォーメーション (DX: Digital Transformation)*3はRPAツールだけでは実

現が困難であることから、人間の言語理解を代替できる「自然言語処理」と呼ばれる分野のAIによる自動処理も需要が増大している。

しかし、実際の現場では利用できる計算資源や求められる性能が案件ごとに異なるためAIの多様性が必要であること、SaaS (Software as a Service)*4で社内のテキストデータを処理するのはセキュリティ面で抵抗感があること、システムエンジニアを抱えていない部門が独力でAIを導入して運用を担うのは負担が大きいことなどの諸問題が立ちはだかり、AIの導入は決して容易ではない。

そこでドコモは、計算資源を存分に活用した高性能な自然言語処理アルゴリズムを開発し、加えて、省リソースなPCでも動作する軽量高速なアルゴリ

©2022 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 RPAツール：パソコン操作をシナリオとして記録させ、パソコン操作を自動化するソフトウェア型ロボット。

*2 チャットボット：音声やテキストチャットを介して、人との会話を自動的に行うプログラム。

ズムも開発した。これにより、クラウドやローカルの処理環境、PCやサーバのスペック、利用者のセキュリティポリシーなど、ユーザの要件に合わせた柔軟な自然言語処理AIの選択・利用が可能となった。さらに運用者の負担を軽減するために、これらのAIを手軽に利用したりチューニングしたりできるGUI (Graphical User Interface)^{*5}ツールも併せて開発した。

本稿では、テキストのカテゴリ分類や個人情報の秘匿化など、現場の多様なニーズに広く対応した自然言語処理AIやGUIの技術開発、ならびに社内外での活用事例について解説する。

2. DXに推進に資する自然言語処理AI

2.1 概要

近年、自然言語処理の学術研究は世界中で活発に

行われており、学術研究における整備された実験設定下では、高い性能を誇るアルゴリズムが数多く提案されている。また、これらのアルゴリズムはOSS (Open Source Software)^{*6}として公開されることが多く、最新のアルゴリズムを誰でも手軽に利用できる。一方、ビジネス現場では、利用できる計算資源やデータに制約が多く、アルゴリズムをそのまま適用してもニーズを満たせない。そこでドコモでは、最新のOSSも活用しつつ、前述したような現場の多様なニーズに対応するため、追加機能の開発や展開方法の工夫に取り組んできた。

2.2 機能

(1) 文書分類

文書分類機能は、文書にラベルを自動付与する機能である (図1①)。ドコモでは、ニーズに応じて、軽量の分類器と高性能な分類器の2種類の分類器を

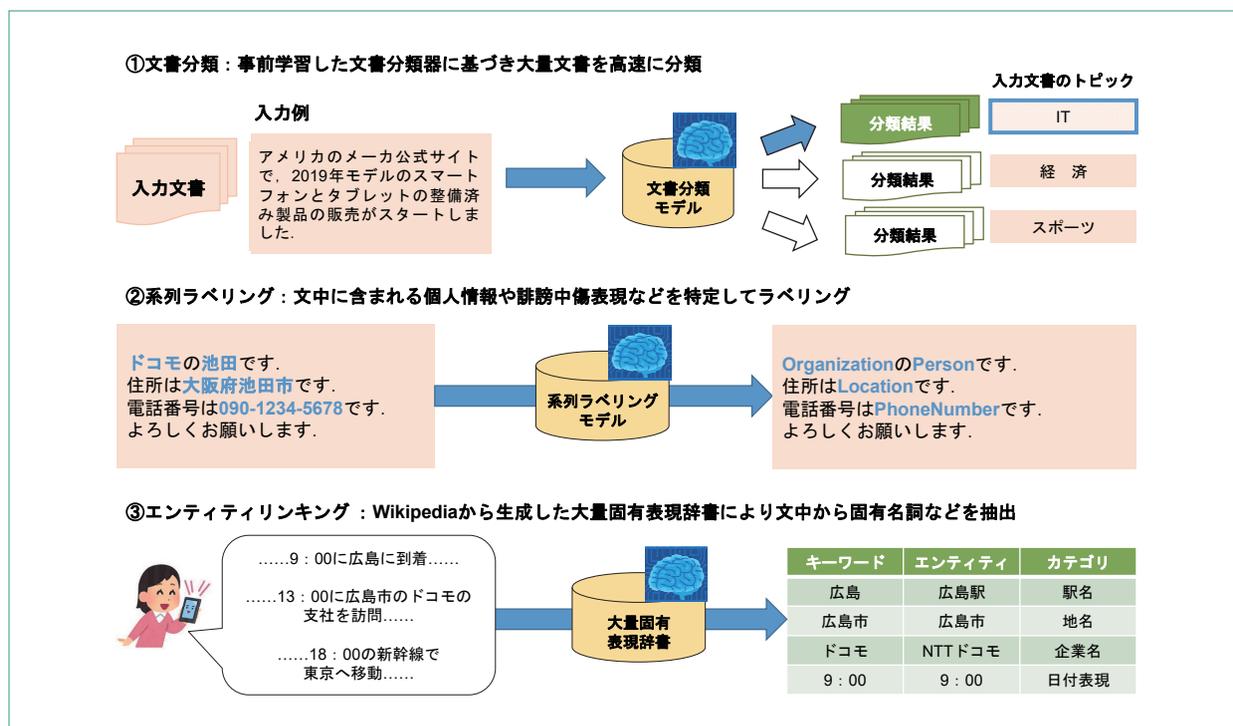


図1 自然言語処理AIの機能一覧

- *3 デジタルトランスフォーメーション (DX)：IT技術を活用してサービスやビジネスモデルを変革させ、事業を促進するとともに人々の生活をあらゆる面で良い方向に変化させること。
- *4 SaaS：インターネットなどのネットワーク経由で利用するソフトウェアのこと。
- *5 GUI：ボタンやアイコンの組合せからなり、直感的な操作性や

- 優れた視認性を特長とするインターフェース。
- *6 OSS：ソースコードが無償で公開されており、誰でも再利用や変更が行えるソフトウェア。

提供している。

(a)軽量な分類器

軽量な分類器は、コストやセキュリティの観点からクラウドが使えない場合など、通常業務用PCのローカル環境での利用ニーズに応えたものである。アルゴリズムには、実装が単純で動作が軽く、省リソースなPCでも実行可能な多層パーセプトロン^{*7}を利用している。また、特徴量^{*8}として、一般的に利用される形態素^{*9}に加えて、文字そのものや文字N-gram^{*10}も利用している。この工夫により、少量のコーパス^{*11}でも高い精度が得られ、さらには誤字脱字にも頑強な分類器を構築することができる。これらの特徴量が有効に機能する事例としてFAQ分類を図2に示す。FAQ分類では、入力された質問文を、用意されたFAQに自動分類する。この際、入力文に誤字脱字がある場合、形態素のみでは正しく分類できない。一方、文字や文字N-gramを利用すれば、部分文字列（図中のサンプルでは「b, l, u, e, t, o, …」など）をヒントに推論することで正しく分類できる。

(b)高性能な分類器

高性能な分類器は、GPU (Graphics Processing Unit)^{*12}が利用できるクラウドサーバかオンプレミス^{*13}サーバを保有している利用者の、潤沢な計算資源を最大限活用して高い性能の分類器を構築したいというニーズに応えたものである。アルゴリズムは、近年の言語処理のデファクトスタンダードであるBERT (Bidirectional Encoder Representations from Transformers) [1]を採用した。BERTは大量の学習データと計算資源を必要とするが、性能が非常に高い。BERTの事前学習モデル^{*14}には、NTT人間情報研究所の成果技術であるNTT版BERT^{*15}を利用している。

また、マルチラベル分類^{*16}への対応をドコモで独自で行った。NTT版BERTはシングルラベル分類^{*17}しか対応していなかったが、ビジネス現場ではマルチラベル分類のニーズが高い。そこで、ドコモでは、NTT版BERTのラベル出力部分の関数を工夫することで、出力をマルチラベル化し提供している。

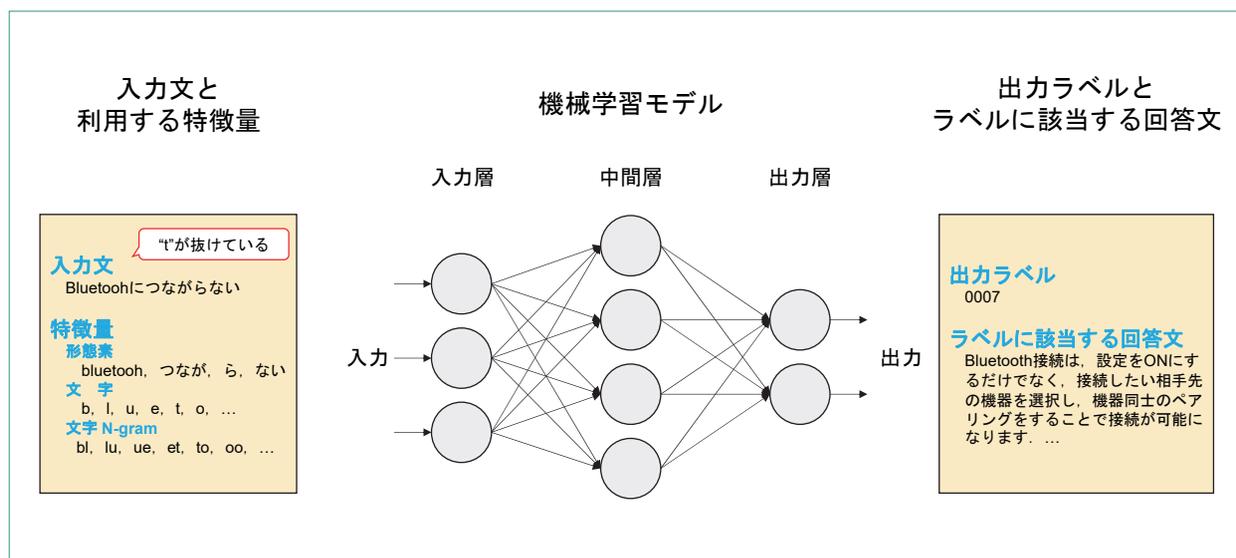


図2 軽量な分類器の特徴量が有効に機能する事例

*7 多層パーセプトロン：機械学習アルゴリズムであるニューラルネットワークの一種。
 *8 特徴量：データから抽出される、そのデータの特徴付ける量（数値）のこと。
 *9 形態素：言語表現の最小単位。
 *10 N-gram：任意の連続した n 個の要素。要素の例としては、単語

や文字などが挙げられる。
 *11 コーパス：テキストを大量に収集してデータベース化した言語資料。
 *12 GPU：並列的な計算処理に優れたプロセッサ。並列計算を要する深層学習と相性が良い。

(2)系列ラベリング

系列ラベリング機能は、系列データ^{*18}にラベルを自動付与する機能である。ドコモは、この機能を個人情報や誹謗中傷表現をはじめとする、テキストのマスキングに利用している（図1②）。当機能の開発においてドコモは、軽量なアルゴリズムを扱えて、かつ開発者コミュニティの大きいOSSとして知られるFlair [2]を採用した。OSSの選定において、アルゴリズムの軽量を重視した理由は、機密性の高いデータの処理では、セキュリティの観点からクラウドの利用を避けたいというニーズが多く、ローカル環境でも動作する、軽量なものが求められるためである。また、開発コミュニティの大きさを重視した理由は、多様な利用ユーザのフィードバックに基づくインターフェースの改善や継続的なバグフィックスが期待できるためである。コミュニティの小さいOSSでは、利用マニュアルやノウハウの貯蓄が少なく、運用が属人化してしまう傾向があるため、これの利用を避けた。

(3)エンティティリンキング

エンティティリンキング機能は、文中からキーワードを抽出し、さらにそのキーワードをエンティティと紐づける機能である。例えば、「ドコモ」「エヌティーイー・ドコモ」「DOCOMO」といった語は、一般的な単語（例えば、「机」や「電車」など）とは違い、固有名詞である。さらに、これらの語は文字列こそ異なるものの、いずれも「NTTドコモ」という同一の概念（エンティティ）を指している。このように、「文中の何がキーワードか（固有の意味をもっているか）」「そのキーワードがどのエンティティを指すか」を機械的に抽出・推定する機能がエンティティリンキングである（図1③）。

(a)Wikipediaからの抽出データに基づく辞書の構築

ドコモでは、特にWikipedia [3] から抽出したデータをエンティティリンキングに活用している。抽出された記事本文、アンカーテキスト、

ページ閲覧数などの各種データは統計的に処理され、キーワードの重み付けやエンティティーキーワード辞書の構築に利用される。Wikipediaを利用する利点として、①記事の新規作成・更新が頻繁であり新語や流行語などへの対応が早いこと、②毎日新たにデータベース・ダンプ^{*19}が公開されておりデータを利用しやすいこと、③データ抽出から辞書構築までを一気通貫で自動処理できるため、辞書のメンテナンスコストが低いことが挙げられる。

(b)カテゴリの付与

さらにドコモでは、「拡張固有表現階層^{*20}」[4]に基づくカテゴリを上位概念としてエンティティに付与しており、抽出した結果を200種類程度のカテゴリに分類した上で、それらを利用する。例えば、対話システムでの活用を考えた場合、「ラーメンを食べたい」「りんごを食べたい」のような、想定されるすべてのユーザ発話に対して1つひとつシステム応答を用意するのは現実的ではない。しかし、「ラーメン」「りんご」に付与されるカテゴリが「食べ物名」であることを利用すれば、「[食べ物名]を食べたい」だけに対してシステム応答を用意すればよいので、応答シナリオ作成の負担を大きく低減することができる。

(c)活用事例

エンティティリンキングの想定活用事例として、ニュースからのキーワード抽出が挙げられる。Wikipediaは時勢を反映して記事内容が更新されたり特定のページ閲覧数が伸びたりするため、Wikipediaの統計データを利用したエンティティリンキングは、ニュース本文からのキーワード抽出と相性がよい。加えて、単純な文字列抽出のみではなく「表記ゆれの吸収」も同時に行うため、抽出・リンキングされたエンティティあるいはエンティティに付与されたカ

*13 オンプレミス：企業がシステムを構成するハードウェアを自社で保有し、自社で保守運用すること。

*14 事前学習モデル：目的タスクの教師あり学習実施前に、大量のコーパスで教師なし学習したモデル。

*15 NTT版BERT：NTT人間情報研究所が独自で学習データを収集し、そのデータで事前学習させたBERTモデル。

*16 マルチラベル分類：1つのデータに複数のラベルを付与する分類方式。

*17 シングルラベル分類：1つのデータに1つのラベルを付与する分類方式。

*18 系列データ：文字列や音声波形、購入履歴などのように、要素が直列に並んでいるデータのこと。

テゴリを記事のタグとして活用しやすいというメリットがある。

3. 自然言語処理AIの展開を加速するGUIツールの開発

3.1 概要

自然言語処理AIの導入や活用には、AIの精度を維持・向上していくためのチューニングが欠かせない。チューニングとは、AIの精度検証、アノテーション*21作業による学習データの修正など、AIモデルの構築と継続的な更新に発生する一連の作業を指し、運用者による目視の作業なども伴うことから、一定の人的な運用コストが発生する。

このような作業にかかる運用コストを軽減するために、ドコモはAIのチューニングをサポートするGUIツールも併せて開発している。このGUIツールは、構築およびチューニングのような繰り返し作業が必要となるAIの学習・評価を、シンプルな画面操作で手早く実行することができる。また、ユーザが学習させたAIの予測結果をアノテーションに活用して、人手の作業量を調整したり作業の優先度を決めたりできるので、効率的な作業が可能である。さらに、構築したAIとデータはリビジョンという単位でシステムにより管理されるので、どのデータを利用してどのAIをいつ構築し、その精度がどれくらいだったかといった煩瑣な情報管理をユーザが意識する必要がなくなる。導入の観点からは、ローカル/クラウドサーバといった作業環境やホスト側のOSに依存しないよう、ツールのバックエンド*22エンジンをコンテナ仮想化技術*23の上で動作させている。また、GUIツールも併せてコンテナ化することで、バックエンドからフロントエンド*24までを簡単に導入できる。

このように、容易に導入できるGUIツールを利用することで、画面に沿った操作のみで手軽にAI構

築、チューニング、管理を行うことが可能となり、自然言語処理AIの導入・運用のハードルを軽減させる効果が期待できる。

3.2 GUIツールの開発とその機能

本ツールでは、AIの構築に必要なデータセットのアップロード、アノテーションデータ作成、学習、評価など一連の機能をGUIとして提供しており、AIの初期構築からメンテナンスに至るまでのすべてをツール画面上で実行可能である(図3)。

主な機能詳細を以下に示す。

(1)学習・評価

利用ユーザが、あらかじめ用意したアノテーション済みのテキストデータをツール経由でバックエンドサーバへアップロードし、簡易な画面操作を行うことで学習が行われ、独自のAIを作成できる。また作成したAIが、入力に対して自動付与するラベルの精度を判定するため、正解として用意したアノテーション済みデータをGUIツールへアップロードすると、本ツールが正答率を自動で算出し、精度評価することも可能である。

このように、利用ユーザはアノテーション済みのテキストデータを準備するだけで、容易にAIの構築と精度評価が可能となる。

(2)チューニング

学習済みのAIを用いて、テキストデータに自動でラベルを付与し、ユーザがラベルの確認・修正を行うことが可能な機能を提供している。一般にAIは、人によってラベル付けがなされたテキストを学習することで、人の判断を再現することが可能となり、学習テキストを増やしていくことで精度が向上する。ある程度学習を繰り返したAIにテキストのアノテーションを任せ、間違った部分を人間の判断で修正することでアノテーションの作業の効率化が可能となる。そして、アノテーションしたテキストをさらにAIに学習させることによる精度向上も可

*19 データベース・ダンプ：データベースの内容をそのままファイルに出力したもの。

*20 拡張固有表現階層：語を「人名」「市区町村名」「国名」のような200種類程度のカテゴリに分類したもの。カテゴリは「地名>天体名>惑星名」のように、最大3階層の構造になっている。

*21 アノテーション：テキストや画像などのデータに対して人手で

ラベルを付与すること。

*22 バックエンド：GUIを動作させるためのシステム部分。主にエンジンおよびエンジンとGUIの動作をつなぐシステム部分のこと。

*23 コンテナ仮想化技術：アプリケーション本体や、アプリケーションに必要なファイル群を「コンテナ」としてパッケージングし、コンテナエンジンというプロセス上で動かす技術の1つ。



図3 チューニングフローとGUIツールの画面イメージ

能である。

このように、本機能を利用することで、ラベル付きテキストデータの作成コストを軽減しつつ、AIの精度向上も実現できる。

また、学習時のハイパーパラメータ^{*25}の変更、学習実行、精度評価などが実行できる画面も提供しており、チューニング時のパラメータ調整^{*26}が容易である。

(3)リビジョン管理

チューニングを行う際には、過去のデータセットや評価結果の管理の煩雑さが往々にして運用上の課題となる。本ツールでは、学習したAIや学習に利用したデータセット、評価サマリなどを、「リビジョン」という作業単位ごとに管理できる機能を提供しており、前述したような煩雑な管理をユーザが意識することなく、過去のデータセットや評価サマ

りをシステム側で個別に管理し、各データセットのチューニングの効果やデータの差分を分析することが可能である。

4. 活用事例の紹介

4.1 概要

ドコモでは、自然言語処理AIやGUIツールの開発のみならず、社内外で自然言語処理AIによるDXを推進している。

社内においては、ユーザや従業員からの意見・要望がテキストデータとして集約されるCS（Customer Satisfaction）^{*27}部門と相互に協力し、自然言語処理AIの構築や導入サポート、蓄積してきた人手による分類結果の共有などを通して、ユーザ満足度向上のためのオペレーション改善に努めている。

*24 フロントエンド：ユーザが目にして操作することが可能なシステム部分。主にGUI部分のこと。

*25 ハイパーパラメータ：学習時の設定値のこと。設定値により性能が変化するため、最も性能が良くなるように最適化することが必要。

*26 パラメータ調整：性能が最も良くなるように関連する設定値を

設定すること。

*27 CS：顧客満足のこと。

また社外においても、法人営業部門と共同でRPA - AI連携ソリューション展開を進めており、自然言語処理のみならずOCR (Optical Character Recognition)*²⁸や音声認識*²⁹技術をも活用した実証実験を広く実施している。

4.2 社内CS部門との取組み

ドコモでは、ドコモショップやコールセンターなどでユーザや従業員からいただいた意見・要望を、特定の個人を識別することができないように加工した上で分析し、ユーザ満足度・従業員満足度の向上に繋がるようなサービスの開発や業務の改善に活用している。例えば、ユーザより受領したアンケートは専門のスタッフが目を通し、「要望」「賞賛の声」などの単位でまとめられて関連部署へ共有される。しかし、日々届くアンケートは大量であることから精読には時間がかかる。また、人の目で確認する以上、見落としや分類ミスなどによる精度の低下などは避けられない。さらに、分類作業はスタッフの知識や経験に依存するため、作業内容の質の均一化が難しいという問題がある。

そこでドコモは、ユーザ要望の速やかかつ適切な実現やオペレーションの効率化を目指し、自然言語処理AIの導入によるアンケート自動分類化を社内CS部門と進めてきた。収集したアンケートは社内閉じたネットワークの中で利用・分析されており、外部の高スペックなサーバによる自動処理が難しかったことから、前述した軽量な分類アルゴリズムをアンケート分類のAIとして開発・提供した。また、CS部門にはこれまで人手で分類してきたデータの蓄積があったことから、これらをAIの学習データとして活用した。

自然言語処理AIを作業に導入することで、これまで長時間人手をかけて実施していた内容の大半が自動化され、大きな稼働時間の削減効果があった。今後はアンケートの自動分類のみならず、ユーザが

誤って入力してしまった氏名や住所などの個人情報を、人の手を介さずに自動で秘匿するAIを導入することで、ユーザにより一層満足いただける価値の提供を目指す。

4.3 RPAツールとAIの連携ソリューション

近年、自然言語処理やOCR用のAIを「頭脳」に見立ててRPAツールと組み合わせることで、より複雑な業務を自動化するソリューション (Cognitive Automation*³⁰) が模索されている。ドコモにおいても、RPAツール「WinActor」の導入・拡大およびRPA - AI連携ソリューションの一環として、「既存システムへの電子カルテ自動投入」の実証実験を、奈良県総合医療センターと共同で実施した。

医療現場には、診察や検査といった「人にしかできない業務」のほかに、患者ごとのカルテ作成や電子カルテシステムへのデータ入力など、ITによる効率化が可能な業務も多くある。しかし、特にシステムへのデータ投入にあたっては、データ内のテキストを「理解」した上で適切な項目に転記する必要があるため、RPAツールのみでの対応は容易ではない。また、個人情報を扱うことから、外部ネットワークとのデータ通信が発生するSaaSの利用よりも、作業用PC間のローカルネットワークに閉じた処理が比較的好まれる。加えて、院内ではデータベースやシステムがすでに導入済み・稼働中である場合が多く、これらを改修しないでそのまま活用できることが求められる。

以上の点を踏まえ、ローカルでも動作する軽量な自然言語処理AIと外部サーバの不要なRPAツールを、院内のローカルネットワークに接続しているPC上だけに構築し、それらを連携させて、これまで人間が実施していたシステムへの転記作業をそのまま自動化することで、後続のフローに大きな変更を加えることなく医療従事者の稼働削減による超過勤務解消を図った。さらに音声認識ソフトを新たに

*28 OCR：文字認識技術のこと。

*29 音声認識：人間の発話音声を解析し、テキスト化したり感情を推定したりする技術のこと。

*30 Cognitive Automation：自然言語処理や画像認識、音声認識などのAIとRPAツールとを組み合わせることで、人間が行っている「判断に基づく作業」を自動化すること。

導入し、これまで紙に控えていた患者情報を音声発話から直接電子テキスト化することで、データ作成の効率化も目指した。

実証実験の結果、電子カルテシステムへの投入時間削減といった定量的な効果のみならず、「記入ミスの削減に繋がった」「患者と向き合う時間が増えた」という担当者からの定性的な評価もあった。このように、RPA-AI連携ソリューションの導入によって本来の業務に集中して取り組めるようになったこと、ひいてはお客様の新たな価値創造に繋がる結果が得られたことは、DXが単純なコスト削減だけでなく、業務そのものの質的向上にも寄与できる可能性を強く示唆している。

5. あとがき

本稿では、DXを実現する自然言語処理アルゴリズムの開発、操作性・利便性を向上させるためのGUIツールの開発、および社内外への展開と今後の展望について解説した。複数種類の手法やアルゴリズムをサポートした多様な自然言語処理AIを開発することによって、利用者は、用途や要件、実行環境などに合わせた適切なAIの選択が可能になった。

並行して、運用部門からの機能要望を盛り込んだGUIツールを開発することで、現場でのAI運用（学習、性能評価、精度改善）の負担軽減を目指した。さらに、開発した自然言語処理AIを導入・活用することで、稼働を定量的に削減できるだけでなく、RPAツールとの親和性の高さや、業務フロー改善に伴って作業の質が向上するといった効果も示せた。

今後は、新規アルゴリズムの追加などバックエンドの機能拡張と並行して、社内外での活用事例創出や利用組織拡大に向けてフロントエンドのGUIツールを訴求していきたい。

文 献

- [1] J. Devlin, M.-W. Chang, K. L. and K. Toutanova : “BERT : Pre-training of Deep Bidirectional Transformers for Language Understanding,” Proc. of NAACL, May 2019.
- [2] A. Akbik, T. Bergmann, D. Blythe, K. Rasul, S. Schweter and R. Vollgraf : “FLAIR : An easy-to-use framework for state-of-the-art NLP,” Proc. of NAACL, Jun. 2019.
- [3] ウィキペディアホームページ。
<https://ja.wikipedia.org/wiki/メインページ>
- [4] 森羅プロジェクト：“拡張固有表現。”
<http://ene-project.info>