

## 社内業務への生成 AI 適用事例

### Example of Applying Generative AI to Internal Operations

杉山 義治

**要約** 近年の AI 技術の発展は目覚ましく、特に自然言語処理や画像処理技術を活用した生成 AI は文章や画像などを作成できるため、業務の効率化や創造性の向上が期待されている。深層学習で有用性が認知されたことで AI の研究が盛んになり、遂には生成 AI が誕生するに至るといった技術的背景がある。生成 AI は大量データからパターンを検出してモデルを作成する。このモデルへ新しい入力データを与えると過去のパターンを基に尤もらしいデータを新しく生成する。業務に適用する場合は、このような生成 AI の特徴を把握して、利点や注意点も考慮した上で導入していくことが求められる。ユニアデックスでは生成 AI を利用した概念実証用に複数のアプリ (Chat, RAG, CoT) を作成している。各アプリは ChatGPT の特性を理解するのに役立つ。AI は、その高度な機能で生活やビジネスに大きな影響を与えるが、一方で多数の課題も存在する。そのため実証実験を重ねてから段階的に実用化していくことが望ましい。

**Abstract** The development of AI technology in recent years has been remarkable, and in particular, generative AI that utilizes natural language processing and image processing technology can create sentences, images, etc., and is expected to improve work efficiency and creativity. The technological background is that the recognition of the usefulness of deep learning caused AI research to increase, which eventually led to the birth of generative AI. Generative AI detects patterns from large amounts of data and creates models. When new input data is given to this model, it generates new plausible data based on past patterns. When applying it to business, it is necessary to understand the characteristics of such generative AI and consider its advantages and precautions before introducing it. Uniadex has created multiple apps (Chat, RAG, CoT) for proof of concept using generative AI. Each app helps you understand the characteristics of ChatGPT. AI has a great impact on life and business with its advanced functions, but there are also many challenges. Therefore, it is desirable to carry out repeated demonstration experiments and then gradually put it into practical use.

#### 1. はじめに

近年の AI (Artificial Intelligence : 人工知能) 技術の発展は目覚ましく、特に自然言語処理や画像処理技術を活用した生成 AI は文章や画像などを作成できるため、業務の効率化や創造性の向上が期待されている。生成 AI は大量データからパターンを検出してモデルを作成する。このモデルへ新しい入力データを与えると過去のパターンを基に尤もらしいデータを新しく生成する。業務に適用する場合は、このような生成 AI の特徴を把握して、利点や注意点も考慮した上で導入していくことが求められる。ユニアデックス株式会社 (以降、ユニアデックス) では生成 AI を利用した概念実証用に複数のアプリ (Chat, RAG, CoT) を作成して、社内業務に適用している。

本稿では生成 AI の仕組みや課題に触れた後、ユニアデックスでの社内適用事例を紹介する。まず 2 章で生成 AI の構成要素を紹介した後、3 章で文章生成 AI、4 章で画像生成 AI の仕組みや課題を述べる。5 章で社内業務への適用事例から生成 AI を上手く使う方法を紹介する。

## 2. 生成 AI とは

AI は、人間の知能を模倣して学習・推論・認識・自然言語理解などのタスクを自動化するシステムのことである。近年急速に発展して私たちの生活やビジネスに大きな影響を与えている。その中でも特に注目されているのが「生成 AI」である。

生成 AI とは人間の創造性を模倣して新しいコンテンツを生成する AI のことを指す。現在は文章や画像を中心としたブームが起こっており、生成 AI の応用範囲は音声や音楽など多岐にわたる。本章では生成 AI の誕生までの歴史を簡単に振り返り、生成 AI を構成する技術について説明する。

### 2.1 AI の歴史

AI の歴史は 1950 年代に始まった。初期の AI ブームではコンピューターが「推論」や「探索」を行うことで特定の問題に対して解答できるようになった。しかし現実世界で発生する問題は一定のルール内で解決できるものではないため、現実の課題への応用ができず、非常に貧弱な仕組みであるとして AI ブームは沈静化した。

1980 年代の第 2 次 AI ブームでは、専門分野の「知識」をコンピューターに取り込み推論を行う「エキスパートシステム」が注目された。しかしその作業は専用辞書の作成に近いものであり、全ての事例に対応することが難しいなどの理由により限定的な利用に留まった。

2000 年代に始まり現在進行中のブームが第 3 次 AI ブームで、「ビッグデータ」と呼ばれる大量のデータから AI 自身が知識を獲得する「機械学習」と、それをさらに発展させた「深層学習」(ディープラーニング) という技術が登場した。深層学習では特定領域で人間並みの判断を高速で行えることが判明したため、AI が現実社会へ適用されるようになった。

### 2.2 生成 AI の誕生

今話題となっている「生成 AI」も機械学習や深層学習といった既存技術の延長線上にある。生成 AI の例として、大量の文章データを学習した結果を基に新たな文章を生成する文章生成 AI や、大量の画像データから学習して新たな画像を生成する画像生成 AI が挙げられる。

### 2.3 生成 AI の構成技術

生成 AI の仕組みを理解するには、まずはその背後にある構成技術の理解が不可欠である。生成 AI の基盤となる技術は、機械学習や深層学習そして自然言語処理といった技術である。これらの技術は人間が行うような複雑な判断や創造的な作業をコンピューターが自律的に行えるようにするためのものである。生成 AI を構成する技術の名称と要約を表 1 に挙げる。

生成 AI はこれらの技術を用いて、大量のデータからパターンを学習し、その学習結果を基に新たなデータを生成する。これらの技術は、生成 AI の可能性を引き出すための重要な要素である。ただし、生成 AI を活用する上では必須の知識ではないため、参考程度に捉えていただきたい。

表 1 生成 AI の構成技術

項番	名称	技術要約
1	機械学習	データから学習しモデルを作り新たな予測や決定を行う。スパムメールフィルタリングやウェブページのランキングなどに利用される。
2	深層学習 (ディープラーニング)	機械学習を発展させたもので、ニューラルネットワークを用いることで機械学習よりも複雑な関係性を導ける。今日の AI ブームの火付け役である。
3	ニューラルネットワーク	人間の脳の神経細胞の接続を模倣したコンピューターの学習モデルを指す。深層学習に用いられるモデルであり、複雑な法則を導出できる。
4	自然言語処理 (NLP : Natural Language Processing)	人間が日常的に使う言語をコンピューターが理解し、また生成するための技術を指す。 文章生成 AI で利用される大規模言語モデル (LLM : Large Language Model) も自然言語処理の深層学習モデルの一種である。
5	単語のベクトル化	Word2Vec で発見された、単語をベクトル (数値の配列) として表現することで単語の関係性をモデル化する技術のこと。関係性を捉えると例えば「王様 - 男 + 女 = 女王様」という導出ができるようになる。 「王様」から「男」を引くと、性別を除いた「王位」の概念が得られ、その「王位」に性別の「女」を足すことで「女性の王位」である「女王様」が得られる。
6	Transformer	大規模言語モデルの基礎となるモデル。 Attention 機構を用いて入力文中の各単語が他の単語とどれだけ関連しているかを計算し、それに基づいて新たな文章を生成する。単語と同様に文章もベクトル化して関係性を示す。
7	GANs (Generative Adversarial Networks : 敵対的生成ネットワーク)	二つのモデルを使用して競争させることで品質向上を図る仕組み。「生成モデル」が新たなデータを生成し、「識別モデル」がそのデータが本物か偽物かを判断する。
8	コンピュータービジョン	コンピューターから捉えたら画像はピクセルの集合であり、そこに描かれているものがりんごなのかバナナなのかもわからない。 特定の円形範囲に赤が多く含まれるパターンはりんごであるというように、特徴を見つけて分析することを指す。

### 3. 文章生成 AI

生成 AI の中でもビジネスシーンで特に期待されているのが文章生成 AI である。著名な文章生成 AI の Chat サービスとしては OpenAI 社の ChatGPT、Google 社の Gemini などが挙げられる。

本章ではこれらのサービスで使用されている文章生成 AI の仕組みや使い方、現時点で判明している課題や実行例について紹介する。

### 3.1 文章生成 AI の仕組み

文章生成 AI は、大量のテキストデータから学習し、その学習結果を基に新たな文章を生成する AI の一種である。このプロセスは機械学習や自然言語処理などの技術を利用している。最初は大量のテキストデータを読み込んでそのパターンを理解することから始まり、その後は学習したパターンを基に新たな文章を生成する。ここでいうパターンとは、人間が書く文章の特徴や文法、表現の傾向などを捉えることである。パターンの例を挙げると「昔々あるところに」と書き出した際に「おじいさんとおばあさんが住んでいました」と続くということで、文意同士を紐づける処理でもある。

単語のベクトル化（2.3 節の表 1）で紹介したように言語にも概念や関係性があり、文章も同様に論理モデルにすることができる。こういったパターンを小説やニュース記事などの様々なテキストデータから学習することで、それらのジャンルに特有の表現や文体を学び取り、それに基づいた新たな文章を生成することができるようになる。これにより法律文書のような堅い書き方から友人との会話のような砕けた書き方まで模倣することができる。

### 3.2 文章生成 AI の応用例

文章生成 AI の技術は、ニュース記事の自動生成、SNS の投稿内容の生成、カスタマーサポートの自動応答、商品レビューやブログ記事の生成、自動翻訳など、様々な分野での応用が期待されている。

例えばニュース記事の自動生成では、最新のニュース情報を基にした記事をリアルタイムで生成することができるだろう。また SNS の投稿内容の生成では、ユーザーの投稿傾向や好みに合わせた投稿内容を自動的に生成することができるだろう。さらにカスタマーサポートでは、顧客からの問い合わせ内容に対して適切な回答を自動的に生成して迅速に対応することができるだろう。

### 3.3 文章生成 AI の課題

文章生成 AI にはまだ解決すべき課題が存在している。生成された文章の品質管理や、生成過程の透明性、そして倫理的な問題などが挙げられる。これらの課題を解決するためには、技術の進化だけでなく、社会全体でのルール作りや教育が不可欠である。

特に今後は学習データについても透明性が求められていくことになる。2023 年 12 月には、New York Times 社と OpenAI 社との間で知的財産権侵害に関する訴訟が開始された。New York Times 社は OpenAI 社が自社コンテンツを無断で使用し、AI チャットボット ChatGPT のトレーニングに利用していると主張している。この問題はまだ係争中であり、今後の動向が注目されている。今後は、文章生成 AI が使用するデータセットがどこから来たものか、どういうプロセスを経て集められたものかという透明性がより重要となってくる。

倫理問題の一例として、自動生成された文章によるスパムメールや SNS へのスパム投稿などが挙げられる。利用者の悪意を判断して生成を拒絶するような振る舞いが求められる。

他には電力問題が挙げられる。現時点では Transformer 技術には大量の CPU、GPU リソースを消費する。当然それらを作成して実行するにも大量の電力を消費することになる。生成 AI のサービス価格が高止まりする要因にもなり得るため、電力消費を改善していかなければならない。

また、文章生成 AI の弱点としてハルシネーション（幻想）と呼ばれる現象が知られている。ハルシネーションとは、端的に言えば AI が誤った回答をすることである。文章生成 AI はあくまでも過去の学習済みパターンから推測したデータを生成するのである。新しいデータに関するものは予測から外れてしまうが、それを生成 AI 自身が意識することはない。他の回答と同じように尤もらしい<sup>\*1</sup> 回答であるかのように振る舞うのである。AI には一足飛びに回答を求めのではなく、順を追って理解を深めていくなど、ある程度の対処方法も知られているが、根本的な理解としては「生成 AI はハルシネーションが起きうるもの」と理解して過信しないことが肝要である。

### 3.4 文章生成 AI の実行例

本節では、5章で述べる業務適用の前段階として、文章生成 AI を実際に動かした実行例を示す。2023年現在で ChatGPT は一番著名な文章生成 AI であるが、他の選択肢も存在する。特に企業で使用する場合には機微なデータを使用しなくてはならないケースがあるだろう。そういった場合には社内環境で実行できる LLM を利用することも考えられる。

例えば、サイバーエージェント社の CALM2-7B-Chat や国立情報学研究所の LLM-JP 13B は、一般的なワークステーションで実行できるモデルとして提供されている。本稿では LLM-JP 13B の実行結果を紹介する。なお実行環境にはそれなりのサイズの VRAM が搭載された GPU が要求されるため、今回は Jetson AGX Orin 64GB というコンピューターを用意した。

まずサンプルプロンプトを実行した結果を以下の表 2 に掲示した。自然言語処理について尋ねたが書籍の紹介が行われたので、そのものズバリの回答とはいかないようである。

表 2 自然言語処理プロンプト

プロンプト	自然言語処理とは何か
実行結果	自然言語処理とは何か 第 2 版」が出版されていました。本書は、自然言語処理とは何かについて、わかりやすく解説した入門書です。自然言語処理という言葉初めて聞いたという方でも、本書を読めば、その基本概念を理解することができます。(後略)

次に業務を想定したタスクとして「氏名の抽出」を与えてみたがうまく動作しなかった。表 3 に示した通り、オウム返しの後に意図しない文面が生成されていた。

表 3 氏名の抽出プロンプト

プロンプト	個人情報を抽出したいです。文章を「」で囲んで渡すので文章の中から名前を抜き取って JSON 形式 { "name": <name> } で回答してください。 「私は杉山と申します。エンジニアとして勤務しています。」
実行結果	個人情報を抽出したいです。文章を「」で囲んで渡すので文章の中から名前を抜き取って JSON 形式 { "name": <name> } で回答してください。 「私は杉山と申します。エンジニアとして勤務しています。」 <name> には「私は杉山と申します。エンジニアとして勤務しています。」の中から抽出したい名前を入れてください。 このような文章を渡します。 「私は杉山と申します。エンジニアとして勤務しています。」

OpenAI 社の ChatGPT であればこういったタスクも標準で処理できるが、LLM-JP 13B では意図した結果を受け取ることができなかった。提供されているモデルはまだ研究段階ということで複雑なタスクは標準では難しいのであろう。今後は改善が進むものと思われる。

また、大規模なデータセットで訓練した AI モデルにおいては、特定のタスクに対する追加学習を行うことでそのタスクの性能を向上させる手法（以降、ファインチューニング）が存在する。提供されているモデルだけでも簡単な受け答えはできているため、ファインチューニングを行うことで特定のタスクをこなすことができるのではないかと期待している。

## 4. 画像生成 AI

今までは人間以外には難しかった絵を生み出すという行為が、画像生成 AI の登場によりコンピュータにも行えるようになった。本章では画像生成 AI の仕組みや応用例と課題、そして実行例を紹介する。

### 4.1 画像生成 AI の仕組み

画像生成 AI は、大量の画像データから学習し、その学習結果を基に新たな画像を生成する AI の一種である。このプロセスは、深層学習や GANs などの技術を用いて行われる。

文章生成 AI と同様に、最初は大量の画像データを読み取ってそのパターンを理解することから始まる。その後、学習したパターンを基に新たな画像を生成する。ここでいうパターンとは人間が描く絵や写真の特徴や色彩、構図の傾向などを捉えることを指す。この学習パターンを基に新たな画像を生成するのである。

例えば、風景や人物などの画像データから学習することで、それらのジャンルに特有の表現や色彩を学び取り、それに基づいて新たな風景画像や人物画像を生成することができる。

また学習の際にアノテーションと呼ばれる単語や文章と紐づけることで、プロンプトと呼ばれる文章を基に画像を生成できるようになる。生成指示の入力情報としては、文章だけでなく画像を用いることができる。画像を入力した上でプロンプトにて追加指示することで、入力画像を基にプロンプトに沿って改変した画像を生成することができる。論理モデルが表示されている物体の意味を理解して単語のベクトル化と同じように別の物体に差し替えることになる。

例えば、かごに入ったりんごの写真に対してりんごをバナナに差し替えるように指示すると、かごやかごが乗っているテーブルなどはそのままにして、りんごの部分だけをバナナに変える。ただし、一旦すべて論理データに置き換えられるため、りんご以外のかごやテーブルはそのまま出てくるのではなく、細部には違いが見られることが多々ある。

### 4.2 画像生成 AI の応用例

画像生成 AI は、アート作品の生成やゲームや映画の背景画像の生成、バーチャルリアリティのコンテンツ生成など、様々な分野での応用が期待されている。

例えばアート作品の生成では、既存のアート作品からインスピレーションを得て新たなアート作品を生成することができるだろう。またゲームや映画の背景画像においても、制作者が意図した世界観や雰囲気に合わせた背景画像を自動的に生成することができるようになるはずだ。さらにバーチャルリアリティでは、ユーザーが没入感を得られるようリアルな 3D 空間を自動的に生成することができるだろう。

### 4.3 画像生成 AI の課題

画像生成 AI は文章生成 AI と同様にいくつかの課題を抱えている。それらは生成された画像の品質管理、生成過程の透明性、そして倫理的な問題などである。

画像品質は登場以来どんどん向上してはいるが、未だに指や手足が正しく生成できないケースなどがある。パターンにおいて数の認識が難しいため、他にも PC の画像を描かせたときに 109 キーボードが正しく生成されるのはまだ先と考えられる。特定のオブジェクトは専用モデルで生成したものを組み合わせるなど、何らかの手当てをする場合もある。

透明性については著作権侵害が主な課題となる。例えば特定の個人や団体が著作権を持つキャラクターが画像生成 AI によって生成されてしまう可能性がある。学習データの透明性が担保されなければならない。そうした課題に対して、Adobe 社は Firefly というプロダクトにおいて権利問題をクリアしていることを公表し、利用企業が権利侵害で訴訟を起こされた場合には法的に補償することを表明している。今後は、学習データを把握しているサービス提供側が著作権侵害に対して責任を負うようになり、課題が解消していくことが期待される。

また、倫理的な問題の一例として「ディープフェイク」が挙げられる。これは、人物の画像や動画・音声を人工的に合成し、元とは異なるデータを生成することができる技術であり、その結果として人々を騙す可能性のある高精度の画像が生成されることがある。既に大統領経験者や総理大臣のような有名人がディープフェイクの標的にされ、ニュースにもなっている。このような非倫理的な指示を画像生成 AI がユーザーから受けた場合は、それを拒否するような改善が不可欠である。悪意のあるユーザーが非倫理的な利用法を新たに考えるため、継続性のある対応が求められる。

### 4.4 画像生成 AI の実行例

画像生成 AI を利用できるプロダクトやサービスは、複数の企業から提供されている。OpenAI 社の DALL-E3、Stability AI 社の Stable Diffusion、Midjourney 社の Midjourney などである。本節では、DALL-E3 を簡単に利用できる Microsoft Bing Copilot から画像生成の指示を出した例を紹介する。

今回はユニアデックスのコーポレートメッセージである「同じ未来を想うことから」を基軸に複数の画像を生成した。合計で 50 枚程度の生成指示を出し、図 1 のような画像を受け取ることに成功した。

画像生成 AI を利用する際は、複数回の試行からより自分のイメージに近いものを拾い上げることが肝要である。現状では選別するまでの試行に労力を要する。細かい指示通りに生成できるようになるにはもう少し時間がかかるものと思われる。



図1 DALL-E3 生成画像

## 5. 社内業務への生成 AI 適用事例

ユニアデックスでは、生成 AI によるデータ分析の PoC (Proof of Concept: 概念実証) を行っている。PoC ではプロトタイプとして 3 種類のアプリを開発し、社内業務で欠かせない機能を追加開発する形で PoC を進めている。本章ではプロトタイプのアプリを通じて PoC の事例を紹介する。

一つ目は、ChatGPT の API の基本機能である Chat による会話型のアプリである。初めて文書生成 AI に触る人向けのデモアプリであり、特出した機能はなく履歴保存機能があるだけの簡素なものである。主に API の利用方法を確認する開発サンプルとしての性質が強いが、生成 AI を説明するのに重宝している。

二つ目は RAG (Retrieval Augmented Generation: 検索拡張生成) アプリである。RAG アプリによって ChatGPT 単体では難しい企業独自の回答を実現できるようになり、業務での活用につながる。

三つ目は、CoT (Chain of Thought: 思考の連鎖) アプリである。CoT アプリによって

ChatGPT と 1 回のやり取りでは難しい処理を定常的に実施することができる。

以降の節では、RAG アプリと CoT アプリを用いて、生成 AI を利用する際の問題を解決する方法について紹介し、検証環境での確認結果を述べる。なお、Chat アプリは通常の会話アプリであるため詳細な紹介は割愛する。

### 5.1 COE サービス for ITSM とは

RAG アプリを紹介する前に、社内業務にどのように AI を組み込むのかを理解するため、RAG アプリの PoC を行っている COE サービス for ITSM について説明する。ユニアデックスでは、CoE (Center of Excellence) チームが、顧客のシステム運用における IT サービスマネジメント (ITSM) を支援している。その活動を支える仕組みが COE サービス for ITSM (図 2) である。

この仕組みを利用して、ユニアデックスでサポートしているソフトウェア/ハードウェアベンダーの不具合情報を顧客向けにサービスポータル上に提供することで、顧客によるセルフサポートが実現できるようにしている。具体的な機能としては、管理対象機器のレポート提供、障害発生時の復旧支援や問い合わせ対応がある。さらに、設定変更のような情報提供後の後工程の作業を他サービスと組み合わせることもできる。顧客との契約によってラインナップやサービス範囲を選択できるようになっている。

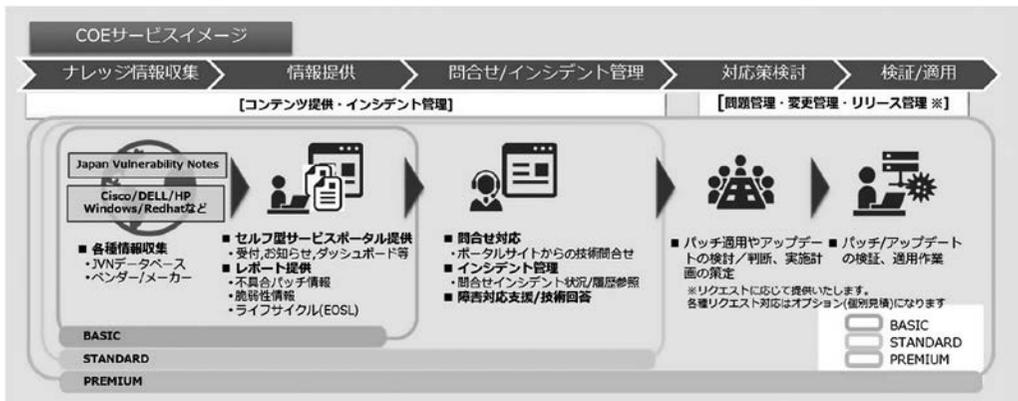


図 2 COE サービス for ITSM イメージ

### 5.2 RAG アプリを活用した PoC

ユニアデックスでは広範な製品の顧客サポートを提供しているため、社内で蓄積している過去事例も多い。そこで得られた知見をナレッジベースとして取りまとめている。また、製品ベンダーが公開する不具合情報も能動的に収集して蓄積している。PoC ではそれらのナレッジベースを有効活用することを目的に実証を進めている。この PoC では、顧客の新規の問い合わせに対して過去事例があれば回答例を作成して提供するオペレーター Bot を作成し検証している。

3.3 節で述べた通り、文章生成 AI にはハルシネーション (幻想) という問題がある一方、顧客に誤った回答をすることはできない。そこで、RAG および Grounding というテクニックを利用することでハルシネーションを回避する手法を採った。

RAGとは、生成 AI に検索機能を組み合わせることで、生成 AI の拡張を図ることである。Grounding は、生成 AI が学習したモデルが固定されているため現在に掛かる情報を知りえないという性質を補完するために、追加情報を提示することである。

例えば、ある製品で先月発見された不具合とそのパッチに関する情報があったとする。その不具合情報は新しすぎて生成 AI モデルは学習していないため、生成 AI はその情報を回答することができない。結果的に、過去の別の不具合だけを基にして誤った回答が生成される。その回避策として、社内に蓄積しているデータを先に検索して、該当不具合情報を見つけた後、生成 AI モデルにその不具合情報を渡して回答を得るという方式を採る。この際、Grounding とは先月という指示から 2023 年 11 月などの現実時間を付与することであり、RAG とは社内システムの検索により生成 AI へ情報を追加することである。

このように追加のデータを生成 AI に提供することで、生成 AI モデルが学習していない内容についての回答を生成させることもできるようになる。

AI では一般的に出力を改善する際に、3.4 節で述べたファインチューニングと呼ばれるモデルの改善処理を行うことが多いが、生成 AI の場合にはファインチューニングそのものの難易度が少し高いため、手軽に実装できる RAG や Grounding が多用されている。

### 5.3 RAG アプリをコンセプトにした検証環境での結果

本稿で紹介する検証用のデモ環境を図 3 に示す。デモ環境では、生成 AI に Azure OpenAI Service の ChatGPT 3.5 Turbo を利用する。2023 年 12 月現在、ChatGPT で学習されているデータは 2021 年 9 月までのものを使用していると発表されている<sup>\*2</sup>。そのため、RAG の例として 2022 年以降の情報を使用してハルシネーションが回避できることを検証する。この検証では ChatGPT が学習していない情報としてニュース記事を利用することとする。

まず検証に利用する News データを、News API というニュース記事を取得できる API を使って取得した。これを MySQL データベース (DB) に登録する。この DB は業務データが登録されているデータベースを想定したものである。更にユニアデックスのニュースリリースから本文を抽出して CSV 化して DB へ追加した。この作業はデモで使用するクエリーを意識した特別データである。

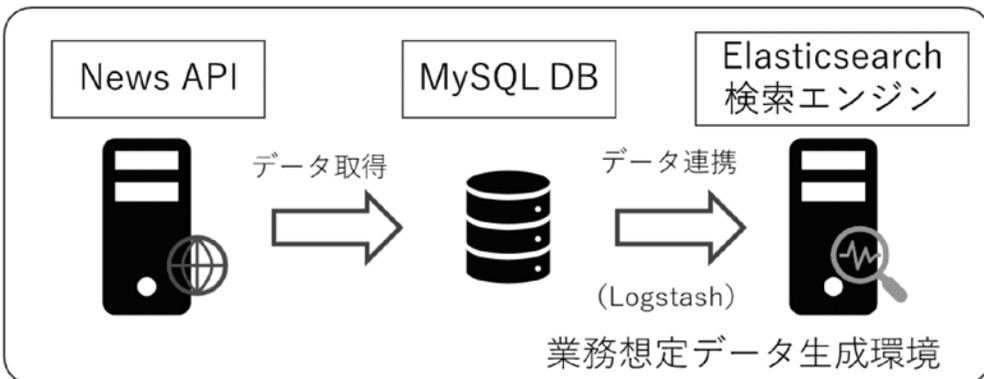


図 3 業務想定データの作成環境

続いてDBをLogStashというツールを介してElasticsearchへ登録する。Elasticsearchは検索エンジンであり、データベースのような製品ではあるが、検索に特化しているためチャットアプリの応答速度の向上が期待できる。また業務影響を避けるために既存DBとは独立させるという目的がある。

そして今回の検証用アプリは次の仕様で作成した(図4)。

- 1) ChatGPTに依頼し、ユーザー質問から検索キーワードを生成する
- 2) 検索キーワードを利用して検索エンジンにクエリーを実行する
- 3) 実行結果の文章をChatGPTに渡して、その情報を基に質問に対する回答を生成する

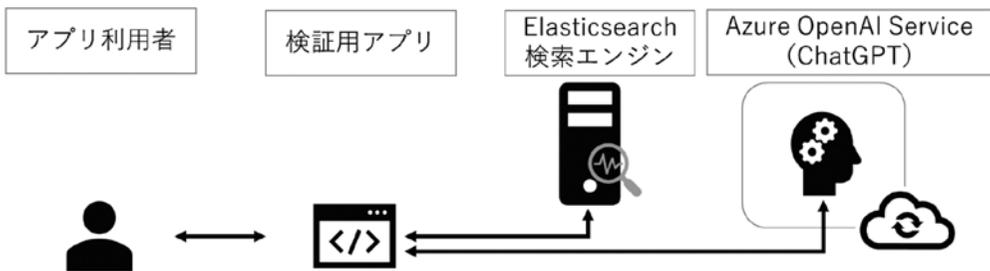


図4 検証用アプリのイメージ

検証用アプリの画面イメージは図5の通りで、アプリ内の中間処理を可視化しつつ結果を受け取れるようにしている。

問い合わせ内容を提示してください

Azure OpenAI ServiceスターターセットPlusとは何ですか。

送信

検索ワード抽出 (クリックで展開)

```
["search_query":["Azure OpenAI ServiceスターターセットPlus"]]
```

社内Web検索 (クリックで展開)

```
[
  {
    "name": "企業向け「ChatGPT」利用環境構築サービス「Azure OpenAI ServiceスターターセットPlus」の提供を開始",
    "url": "https://www.uniadex.co.jp/news/2023/20230809-openai-service.html",
    "summary": "BIPROGYとユニアデックスのBIPROGYグループは、情報管理統制を高めることにより、お客様の負担を低減し効率的にChatGPTの業務利用が行える環境を構築するサービス「Azure OpenAI ServiceスターターセットPlus」の提供を8月から開始します。本サービスは、Microsoft Azure OpenAI Serviceの活用でセキュアな生成AI環境構築と、お客様の生成AI活用シナリオの作成を伴走型で支援します。BIPROGYグループは、生成AI技術を業務シス
```

回答:

Azure OpenAI ServiceスターターセットPlusは、BIPROGYとユニアデックスのBIPROGYグループが提供するサービスです。このサービスは、Microsoft Azure OpenAI Serviceを活用してセキュアな生成AI環境を構築し、お客様の生成AI活用シナリオの作成を伴走型で支援します。BIPROGYグループは、生成AI技術を業務システムに組み込むための最善策を提案し、企業のシステム化構想の段階から生成AI技術を活用したDXの推進を支援します。具体的なサービス内容については、情報管理統制や利用履歴分析、利用促進のための拡張機能を提供し、エンタープライズシステムと連携した運用を可能にします。また、基礎教育や活用シナリオの作成なども伴走型で支援します。

図5 検証用アプリ実行結果

以下の表4のようにChatGPTが知らないことを質問して回答を得た。

表4 ニュース質問プロンプト

プロンプト	Azure OpenAI Service スターターセット Plus とは何ですか。
実行結果	Azure OpenAI Service スターターセット Plus は、BIPROGY とユニアデックスの BIPROGY グループが提供するサービスです。このサービスは、Microsoft Azure OpenAI Service を活用してセキュアな生成 AI 環境を構築し、お客様の生成 AI 活用シナリオの作成を伴走型で支援します。BIPROGY グループは、生成 AI 技術を業務システムに組み込むための最善策を提案し、企業のシステム化構想の段階から生成 AI 技術を活用した DX の推進を支援します。具体的なサービス内容については、情報管理統制や利用履歴分析、利用促進のための拡張機能を提供し、エンタープライズシステムと連携した運用を可能にします。また、基礎教育や活用シナリオの作成なども伴走型で支援します。

Azure OpenAI Service スターターセット Plus は 2023 年に提供を開始したサービスであり、ChatGPT が持つ 2021 年 9 月までの学習データには含まれていない。RAG によって、2022 年以降の情報を追加することで、検索エンジンの結果を基に正しい回答を得ることができた。このような仕組みを利用することで社内ナレッジを有効活用できるようになる。

なお、詳細は割愛するが、取得したニュース情報に対する回答も意図した通りに動作したことを補足しておく。

#### 5.4 CoT アプリを活用した PoC

ChatGPT は非常にパワフルではあるが、複雑な依頼に対する回答精度は実は高くない。そこで CoT (Chain of Thought : 思考の連鎖) という段階的に問い合わせを行う手法が編み出された。人間が複雑な計算に途中式を用いるように、ChatGPT への問い合わせも段階を追って整理するのである。

問題を解決するための一連の流れを明確にして、それを複数のステップに分解する。そしてそれぞれのステップで ChatGPT に具体的な問い合わせを行うことで最終回答へ導くのである。CoT により ChatGPT が各ステップで生成した回答を基に次の問い合わせを行うことによって、より複雑な問題に対する解答を得ることができる。

CoT アプリは、一連のステップを予め定義しておき複数の入力に対して同じ処理を実行するような、バッチ処理を行うアプリである。複雑な変換処理を色々なデータに対して実行したい場合に利用することができる。

#### 5.5 CoT アプリをコンセプトにした検証環境での結果

RAG アプリと同様に、CoT アプリについても本稿用のデモシナリオを用意した。業務との関連性は少し薄いですが、ChatGPT が苦手としていることの一つである「計算」を例に挙げる。採用するシナリオは標準偏差の算出である。標準偏差を得る際に単発の質問と CoT アプリによる順序立てた質問でどのような違いが見られるかを確認する。

まず、Chat アプリに標準偏差を質問する。数列 [10, 20, 30, 40, 50] の標準偏差について質問したところ、分散を求めることに失敗して「3.464」という誤った回答が導かれた。アプリ上の表示を図 6 に示す。



図 6 CoT を使用しない場合

続いて CoT アプリで標準偏差を質問する。まず途中式のように段階的に値を導くようにプロンプトを検討する。Chat アプリでは分散をうまく導くことができなかったので、分散を求める処理を複数の工程に分けることで正答を導ける可能性が高い。それを踏まえて今回は以下のプロンプトを作成した。

- 1) system : ユーザーが与えるデータを順番に計算していきます。
- 2) 総和 : 与えられた数列を足し合わせてください。
- 3) 平均値 : 総和と数列の数から平均値を算出してください。
- 4) 偏差 : 平均値を使用して、数列の偏差を求めてください。
- 5) 偏差の二乗 : それぞれの偏差を二乗します。
- 6) 分散 : 偏差を二乗した値を足し合わせて、数列の数で割って分散を求めます。
- 7) 標準偏差 : 分散の平方根を算出して標準偏差を求めます。

CoT アプリのプロンプト設定内容を図 7 に示す。

# プロンプト設定

	role_name	role_content
0	system	ユーザーが与えるデータを順番に計算していきます。
1	総和	与えられた数列を足し合わせてください
2	平均値	総和と数列の数から平均値を算出してください。
3	偏差	平均値を使用して、数列の偏差を求めてください。
4	偏差の二乗	それぞれの偏差を二乗します。
5	分散	偏差を二乗した値を足し合わせて、数列の数で割って分散を求めます。
6	標準偏差	分散の平方根を算出して標準偏差を求めます。

図7 プロンプト設定画面

CoT アプリではプロンプトを順次実行して以前の回答を記憶したまま次の質問を処理するように実装されている。そのため、人間が計算する際に途中式を書いてその結果を踏まえて次の計算をするのと同等の結果を得ることができる。

CoT アプリに先程と同じく数列 [10, 20, 30, 40, 50] の標準偏差について質問すると、結果として約 14.142 という正答を得ることができた。実行結果を図 8 に示す。

## CoTアプリ

標準偏差を求める数列を記載してください:

[10, 20, 30, 40, 50]

回答:

分散の平方根を算出すると、標準偏差は $\sqrt{200} \approx 14.142$ です。

総和実行中 (クリックで展開) ▼

平均値実行中 (クリックで展開) ▲

与えられた数列の要素の総和は150です。数列の要素数は5ですので、平均値は $150/5 = 30$ です。

偏差実行中 (クリックで展開) ▼

偏差の二乗実行中 (クリックで展開) ▲

各偏差を二乗すると、[400, 100, 0, 100, 400]となります。

分散実行中 (クリックで展開) ▲

偏差を二乗した値を足し合わせると、 $400 + 100 + 0 + 100 + 400 = 1000$ となります。数列の要素数は5ですので、分散は $1000/5 = 200$ です。

標準偏差実行中 (クリックで展開) ▼

図8 CoT アプリを使用した場合

このように、CoT を利用してバッチ処理のように順次実行することで、ChatGPT はより複雑なタスクもこなすことができるようになる。

## 6. おわりに

AI には、その高度な学習能力と予測能力を活用することで、私たちの生活やビジネスに大きな影響を与えることが期待されている。今後、より高度な自然言語理解や、リアルタイムでの高品質な生成ができるようになるだろう。一方で AI には課題も存在するため、まずは社内利用などで実証実験を重ねて、影響範囲が小さい業務から段階的に実用化していくことが望ましい。

また、課題自体も技術的發展に伴って徐々に解決していくと考えられる。AI の進歩とともに、倫理的に正しく、透明性を担保した学習データによる AI サービスがどんどん増えていくだろう。AI の利点を最大限に活用しつつ、倫理的な問題や社会的な影響を最小限に抑えることで、より良い社会を実現することを目指すべきである。これらの視点を踏まえ、私たち一人一人が AI について理解し、適切に利用していくことで AI と人間社会との共生が実現し、これからの未来がより良いものになっていくことだろう。

本稿では、生成 AI の現状と、業務への適用事例を通して生成 AI の使い方の一例を示した。業務への生成 AI 適用を検討中の読者にとって参考になれば幸いである。今後も、汎用人工知能 (AGI) の実現を目指して、より良い未来を AI と共に歩んでいきたい。

- 
- \* 1 一般的な「もっともらしい」は「いかにも道理にあっているように見える」というニュアンスだが、ここでの「尤もらしい」は統計用語の「結果から原因を推測するときの尤もらしさ」に根ざしている。
  - \* 2 「Azure OpenAI Service に関してよく寄せられる質問」の中で、「私が GPT-4 に実行中のモデルを尋ねると、GPT-3 を実行していると答えます。なぜこのようになるのですか？」の質問に対して、その理由の中に「GPT-4 の場合、2023 年 8 月の時点で、基になるトレーニング データは 2021 年 9 月までしかありません。」と記載あり。  
<https://learn.microsoft.com/ja-jp/azure/ai-services/openai/faq>

- 参考文献** [1] Ashish Vaswani / Noam Shazeer / Niki Parmar / Jakob Uszkoreit / Llion Jones / Aidan N. Gomez / Lukasz Kaiser / Illia Polosukhin, “Attention Is All You Need”, 2017 年 6 月 13 日, <https://arxiv.org/abs/1706.03762>
- [2] Jacob Devlin / Ming-Wei Chang / Kenton Lee / Kristina Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”, 2018 年 10 月 11 日, <https://arxiv.org/abs/1810.04805>
- [3] Robin Rombach / Andreas Blattmann / Dominik Lorenz / Patrick Esser / Björn Ommer, “High-Resolution Image Synthesis with Latent Diffusion Models”, 2021 年 12 月 21 日, <https://arxiv.org/abs/2112.10752>
- ※ 上記注釈および参考文献に含まれる URL のリンク先は、2024 年 1 月 30 日時点での存在を確認。

**執筆者紹介** 杉山 義治 (Yoshiharu Sugiyama)

2018年ユニアデックス(株)入社。開発部門であるテクノロジーセンターに所属し、AIを中心としたプログラム開発に従事する。前職で培ったインフラ知識を活用して、SRE (Site Reliability Engineering) チームとしても活動している。

