

## 事例紹介1 14:10~14:40

生成AI発明の権利化ポイント、知財実務における生成AI利活用の最前線

パテント・インテグレーション株式会社 代表取締役CEO 弁理士 大瀬 佳之 氏

### 1. 生成 AI とは

生成 AI は、与えられた入力に基づいて新しいコンテンツを生成する人工知能モデルです。この技術は、テキスト、画像、音声など、さまざまな形式のデータから新たなコンテンツを作り出す能力を持ちます。例えば、テキストから画像を生成するモデル(例: Stable Diffusion、Midjourney、DALL-E )は、テキストによる説明や指示に基づいて、視覚的に表現される画像を生成します。また、テキストから新しいテキストを出力するモデル(例: ChatGPT、Gemini、Claude)は、翻訳、要約、コンテンツ生成など、多様なタスクに利用できます。

生成 AI の核心は、**大量のデータと深層学習アルゴリズムを活用して、入力データに含まれるパターンや構造を学習し、それを基に新しいデータを生成すること**にあります。従来の AI が既存のデータを分析・分類するのに対し、生成 AI は創造的な側面を持ち、全く新しいコンテンツを生み出すことが可能です。

### 2. AI の歴史

AI の歴史は長く、いくつかのブームを経て現在に至っています。この文書では、2012 年からの深層学習技術の進展に焦点を当てています。**深層学習の登場により、画像認識、機械翻訳、画像生成、自動運転といった、従来は困難であった複雑な問題が解決可能になりました。**深層学習は、**多層のニューラルネットワークを用いてデータを解析する技術**であり、これにより、従来の機械学習では捉えきれなかった複雑なパターンを学習できるようになりました。その結果、AI の性能は飛躍的に向上し、実用的な応用範囲が大きく広がりました。

### 3. 大規模言語モデル(LLM)

**大規模言語モデル(LLM)は、大量のテキストデータを学習し、人間が使う自然言語を理解し、生成する能力を持つ AI モデル**です。この文書では、LLM が革新的な対話型 AI(チャットボット)として登場し、多様な情報処理タスクを実行できるようになった点を強調しています。

LLM は、質問応答、要約、プログラミング、文書翻訳など、多岐にわたるタスクに対応できます。3。例えば、質問に対して適切な回答を生成したり、長い文章を要約したり、プログラムのコードを生成したり、文章を異なる言語に翻訳したりすることが可能です。

ただし、LLM は入力文に連なる文章を確率的に推論する文章生成器に過ぎないという点に注意が必要です。LLM は、文法的な正しさや表面的な意味の一貫性を重視して文章を生成しますが、**必ずしも論理的な整合性や事実に基づいた正確性を持っているとは限りません。**

LLM は、**大量の学習データに基づいています**が、事前学習済みであり、ゼロまたは少数の入力から自然な回答を出力できます。これは、LLM が特定のタスクやドメインに特化していなくても、一般的な知識やパターンに基づいて、ある程度の品質の回答を生成できることを意味します。

#### 4. プロンプトの種類

LLM の性能を最大限に引き出すためには、適切なプロンプト(指示文)を与えることが重要です。この文書では、**Zero-Shot プロンプティング**と**Few-Shot プロンプティング**という2つの主要なプロンプトの種類を紹介しています。

•**Zero-Shot プロンプティング**:直接的な指示でモデルが実行するタスクを指示します。例えば、「〇〇について説明してください」といった具体的な質問や指示を与えることで、LLM は学習データに基づいて回答を生成します。

•**Few-Shot プロンプティング**:少量の入出力例によりタスクを指示します。これは、LLM に対して、タスクの例をいくつか示すことで、LLM がそのパターンを学習し、同様のタスクを実行できるようにする方法です。指示内容を具体化できない場合に有効です。

プロンプトの種類を適切に使い分けることで、タスクの精度と効率を向上させることができます。

#### 5. LLM の仕組み

LLM は、自然法則や原理を理解しているわけではありません。LLM は、入力された質問に基づいて、過去のデータから適切な回答を選択します。つまり、LLM は大量のテキストデータを統計的に解析し、入力されたプロンプトに最も関連性の高い情報を組み合わせて回答を生成します。

LLM は、プログラムが明示的に内包している自然法則やピタゴラスの定理のような演繹的な知識を理解しているわけではありません。LLM は、過去の学習データに基づいて、最も可能性の高い回答を生成するにとどまります。

#### 6. 生成 AI の課題

生成 AI は多くの可能性を秘めている一方で、いくつかの重要な課題も抱えています。この文書では、以下の5つの課題を挙げています:

•**ハルシネーション問題**:存在しない情報や事実を生成する可能性があります。LLM は、学習データに存在しない情報をあたかも事実であるかのように生成することがあります。

•LLM は単なる言語モデルなので、「対象物の構造」や「科学原理」などを理解して回答しているわけではありません<sup>4</sup>。LLM は、表面的な言語パターンに基づいて回答を生成するため、物理的な制約や科学的な法則を無視した回答を生成することがあります。

•**入力データの学習内容によっては、不適切な内容が含まれるおそれがあります**<sup>4</sup>。LLM は、学習データに含まれる偏見や差別的な表現をそのまま学習してしまう可能性があり、その結果、不適切な回答を生成することがあります。

•LLM に対する入力に対して、**どのような LLM の内部構造により出力結果が出力されるかは不明**です<sup>4</sup>。LLM の内部動作は複雑であり、ブラックボックス化されているため、なぜ特定の回答が生成されたのかを正確に理解することが困難です。

•**生成物が何らかの著作物に類似している場合、著作権を侵害する可能性があります**<sup>4</sup>。LLM が生成したコンテンツが既存の著作物と類似している場合、著作権侵害の問題が生じる可能性があります。

これらの課題を認識し、適切に対応することが、生成 AI の責任ある開発と利用のために不可欠です。

## 7. 業務利用時の注意点

生成 AI を業務に利用する際には、いくつかの注意点があります。特に **ChatGPT の再学習と個人情報・セキュリティ情報**の取り扱いについて注意を促しています。

•**ChatGPT の再学習に利用されることを防ぐことができます。**ChatGPT は、ユーザーとの対話内容を学習データとして利用する可能性があります。そのため、業務上の機密情報や顧客情報などが ChatGPT に学習されないように、設定を変更する必要があります。

•**個人情報、セキュリティ情報等はできるだけ入力しないようにしてください。**やむを得ず個人情報やセキュリティ情報を入力する場合には、匿名化や暗号化などの対策を講じることが重要です。

これらの注意点を守ることで、情報漏洩のリスクを低減し、安全に生成 AI を利用することができます。

## 8. DeepSeek

DeepSeek は、**中国の AI スタートアップ**であり、**OpenAI の GPT と比較してコストが 1/10 のモデルを発表**しました 7。このモデルは、**蒸溜と呼ばれる技術**や人間が出力したコンテンツによる**自主学習**により、高品質かつ低コストな AI モデルを実現しています。

**DeepSeek のモデル自体は、オープンソースで公開**されています。これにより、誰もが DeepSeek のモデルを自由に利用し、カスタマイズすることができます。

DeepSeek は、低コストで高性能な AI モデルを提供することで、AI の民主化を推進し、より多くの企業や個人が AI 技術を利用できるようにすることを目指しています。

## 9. OpenAI とソフトバンクによる企業 AI 提供

OpenAI とソフトバンクグループは、**企業向けに最先端 AI を開発・販売**することに合意しました 7。この提携により、ソフトバンクは OpenAI の技術を活用し、**企業のデジタルトランスフォーメーションを支援**することを目指しています。

OpenAI とソフトバンクの提携は、企業が自社のニーズに合わせた AI ソリューションを導入し、ビジネスの効率化や新たな価値創造につなげることを可能にします。

## 10. AI エージェント

**AI エージェントは、環境と対話し、データを収集し、そのデータを使用して自己決定タスクを実行**します 8。AI エージェントは、**事前に決められた目標を達成するためのソフトウェアプログラム**であり、自律的にタスクを実行し、目標を達成するために最適なアクションを選択します。

AI エージェントは、**複雑なタスクを個別の細かいタスクに分解し、足りない情報を AI が自律的に収集し、与えられた目標を達成するように自律的に動作**します。例えば、顧客対応、在庫管理、品質管理など、様々な業務プロセスを自動化することができます。

## 11. GPT-4o の出現とエージェント機能への注目

GPT-4o は、OpenAI が新たにリリースした AI モデルであり、単一の入力に対して、高品質な出力結果が得られます。GPT-4o は、テキスト、画像、音声など、様々な種類のデータを組み合わせることで処理することができ、より自然で人間らしい対話を可能にします。

GPT-4o の登場により、AI エージェントの性能は飛躍的に向上し、より複雑なタスクをより効率的に実行できるようになると期待されています。

## 12. LLM 関連発明の特許出願

LLM 関連技術の進展に伴い、プロンプトエンジニアリングや RAG (Retrieval-Augmented Generation) といった技術が注目されています。これらの技術は、LLM の性能を向上させ、より高品質な回答を生成するために重要です。

•**プロンプトエンジニアリング**: LLM に与えるプロンプトを最適化することで、LLM の回答の質を向上させる技術です。

•**RAG (Retrieval-Augmented Generation)**: LLM が外部の知識ソースを参照し、その情報を基に回答を生成する技術です。

これらの技術は、LLM の知識の限界を克服し、より正確で信頼性の高い回答を生成するために役立ちます。

## 13. 公報読解/拒絶対応/侵害予防調査プロセス整理

特許情報を扱う業務においては、公報読解、拒絶対応、侵害予防調査といったプロセスが不可欠です。これらのプロセスは、非常に時間と労力を要するため、AI を活用して効率化することが望まれています。

AI は、大量の特許情報を解析し、必要な情報を迅速に抽出することができます。例えば、AI は特定の技術分野に関する特許文献を自動的に検索し、重要な情報を抽出したり、特許請求の範囲を分析し、侵害の可能性を評価したりすることができます。

## 14. サマリア

サマリアは、特許の読解支援アシスタントとして、弁理士が開発した実践的なツールです。サマリアは、明細書の読み込み作業を支援し、ポイント抽出や観点抽出などを行います。

サマリアは、AI を活用して特許文献を解析し、重要な情報を自動的に抽出することができます。これにより、弁理士は特許文献の読解にかかる時間を大幅に削減し、より創造的な業務に集中することができます。

## 15. 拒絶理由通知対応における利活用

特許出願が拒絶された場合、拒絶理由通知に対応する必要があります。このプロセスは、拒絶理由を理解し、反論を準備し、必要な証拠を収集するといった複雑な作業を伴います。

AI は、拒絶理由通知の解析を支援し、関連する特許文献や技術情報を迅速に提供することができます。また、AI は反論の構成を支援し、効果的な主張を提案することも可能です。

## 16. 専用サービス/汎用サービスの比較

生成 AI を活用したサービスには、専用サービスと汎用サービスの 2 種類があります。専用サービスは、特定のタスク向けに用意された機能を提供します。一方、汎用サービスは、様々なタスクに対応できる汎用的な機能を提供します。

専用サービスは、汎用サービスと比較して、高品質な出力結果が得られ、「ハイライト機能」「機械翻訳」「OCR」など実務上必要な周辺機能も提供されます。そのため、特定の業務においては、専用サービスの方がより効率的かつ効果的な場合があります。

#### 17. 特許実務における利活用

生成 AI は、特許実務における様々なタスクを支援することができます。例えば、拒絶対応を例として、AI エージェント機能を活用することで、拒絶理由の概要把握、関連文書の精読、審査経過の確認といった作業を効率化することができます。

AI エージェントは、特許弁理士の業務をサポートし、より高度な知的財産戦略の策定に貢献することが期待されています。

#### まとめ

この文書は、生成 AI の基本概念から最新の技術動向、そして特許業務への応用までを網羅的に解説しています。生成 AI は、多様な分野で活用されており、その応用範囲は今後ますます拡大していくと考えられます。

生成 AI を理解し、適切に活用することで、ビジネスの効率化や新たな価値創造につなげることができるでしょう。