

QuantumBlack, AI by McKinsey, Build by McKinsey

生成AIについて CEOが知っておくべきこと

CEOたちは、生成AIがビジネスにもたらす価値とリスクを見極めようとしている。一方で、その間にも生成AIは驚異的なスピードで進化し続けている。本記事では、生成AIについてCEOが押さえておくべきポイントについて解説する。

本記事は、マッキンゼー・デジタルのMcKinsey Technology CouncilおよびQuantumBlack、AI by McKinseyの見解として、Michael Chui、Roger Roberts、Tanya Rodchenko、Alex Singla、Alex Sukharevsky、Lareina Yee、Delphine Zurkiyaが共同で執筆したものである。



ChatGPT、Bard、Claude、Midjourneyなどのコンテンツ作成ツールがリリースされて以来、生成AIをめぐる熱狂が広がるなか、CEOたちは、これを「過度な期待のピーク期」として捉えるべきか、あるいは競争原理を変えてしまうほどの機会として捉えるべきか、もし後者であるとしたら、自社の事業にどのような価値がもたらされるのかと思索している。

ChatGPTは、一般公開を開始してからわずか2カ月で利用者が1億人に到達した。それは画期的な機能を通じて「AIの民主化」を実現し、史上最も急速に成長しているアプリケーションとなっている。生成AIは、誰でも気軽に利用できる点で、従来のAIと異なる。利用者は機械学習に関する専門知識がなくとも、対話型AIに質問を投げかけるだけで回答を得て、価値を引き出すことができる。そして、パソコンやiPhoneといった他の革新的な技術と同様に、一つの生成AIプラットフォームを起点として、あらゆる年齢層および教育レベルの人々が、インターネットに接続できる場所であればどこでも、新たな用途での活用方法を探求することができる。

これは、生成AIチャットボットが基盤モデルの上に構築されていることで実現している。基盤モデルとは、大量かつ多様な形式(文書や音声など)でラベルなしの非構造化データで訓練された、拡張性のあるニューラルネットワークのことを指す。従来のAIモデルは用途が「狭い」ことが多く、顧客の離脱率の予測など、単一タスクにのみ対応している一方で、基盤モデルは幅広いタスクに適用できる。例えば、ある基盤モデルは、量子コンピューティングに関する2万字にのぼる報告書のエグゼクティブサマリーや、伐採事業の市場参入戦略案を作成できるだけでなく、冷蔵庫にある10種類の食材を使用して、5種類のレシピを紹介することもできる。ただし、生成AIは多様な用途に対応する一方で、現時点では必ずしも正確な情報を提示するわけではないため、AIのリスク管理に関する懸念も生じている。

そのため、「ガードレール」を設置し、リスクに適切に対応する必要があるが、生成AIによって新たなユースケースが創出されるだけでなく、既存のユースケースの導入の加速化や規模の拡大、高度化が実現する。例えば営業の場合、これまでは、事前に収集した人口統計学的な属性や購買パターンといった静的な顧客データに基づいて販促活動を行っていたが、今後はAIモデルに特別な訓練を行うことにより、アップセル(顧客単価の向上)の機会を特定したり、営業活動の提案を受けたりすることができるようになる可能性がある。また、最近の技術の進歩により、

これまでコールドスタート¹が難しいとされた、スパース(低密度)な大量データを保有する総合小売業界において、Transformerの実装により、データの密度に限定されず、より精度の高い購買パターンを特定し、対象SKUを推奨することが可能となった。生成AIツールは、営業担当が顧客と実際に行ったやり取り、社内の顧客データ、市場トレンドなどの外的要素、SNSのインフルエンサーに関するデータなどに基づき、リアルタイムに営業担当にアップセルの機会を提示するだけでなく、提案資料の素案も作成できるようになることが期待される。

上記の例では、生成AIが特定の職種に与え得るメリットを示しているが、ほぼすべての知識労働者が、生成AIから恩恵を享受できるようになるとみられる。生成AIによって、一部のタスクが自動化される可能性があるが、知識労働者が日常的に活用するツール(Eメールや文書作成ソフトウェアなど)にこの技術を組み込むことで、大きな価値が生まれる。このようにアップグレードされたツールによって、生産性は大幅に向上する。

CEOたちは、今すぐ行動すべきか、もしそうであれば、どこから着手すべきかと思索している。生成AIを競合の一步先に行く好機として捉え、業務にどう生かせるかを検討するCEOもいれば、慎重な姿勢をとり、大規模な投資を行う前に効果を確認すべく、いくつかのユースケースで実証実験を行うCEOもいるであろう。また、生成AIを導入するとすれば、それに伴い必要となる技術的専門知識、技術とデータのアーキテクチャ、オペレーティングモデル、リスク管理プロセスなどが自社に備わっているかを評価する必要もある。

本記事は、CEOおよび経営陣に対して、生成AIの価値創造シナリオや業務変革の進め方の検討を支援することを目的としている。まず、「生成AI入門」のセクションでは、急速なAIの進化と利用可能な技術的選択肢の理解を深める。次のセクション「生成AIをビジネスに生かす」では、組織の有効性を高めることを目的とした4つの導入事例を通して、企業がどのように生成AIを活用できるかを検討する。ここでは、アーリーアダプター(早期導入者)の事例を取り上げており、技術、コスト、オペレーティングモデルの観点から様々な選択肢を提示している。そして最後のセクション「生成AIを導入する際に考慮すべきこと」では、生成AI導入を成功させるにあたり、CEOに求められる役割について取り上げている。

¹ コールドスタート問題: 新規ユーザーがシステムを使用開始する時や新規コンテンツが追加された時など、学習に必要なデータが十分に蓄積されていないために機械学習による予測が困難となり、レコメンデーションの精度が落ちてしまうこと

世界が生成AIの熱狂に包まれる中、経営陣は慎重に、かつ意図したスピード感をもって導入を進めたいと考えているのではないであろうか。本記事が、将来有望な生成AIについてバランスの取れた見解を提示するものとして、ビジネスリーダーの理解を深める一助となれば幸いである。

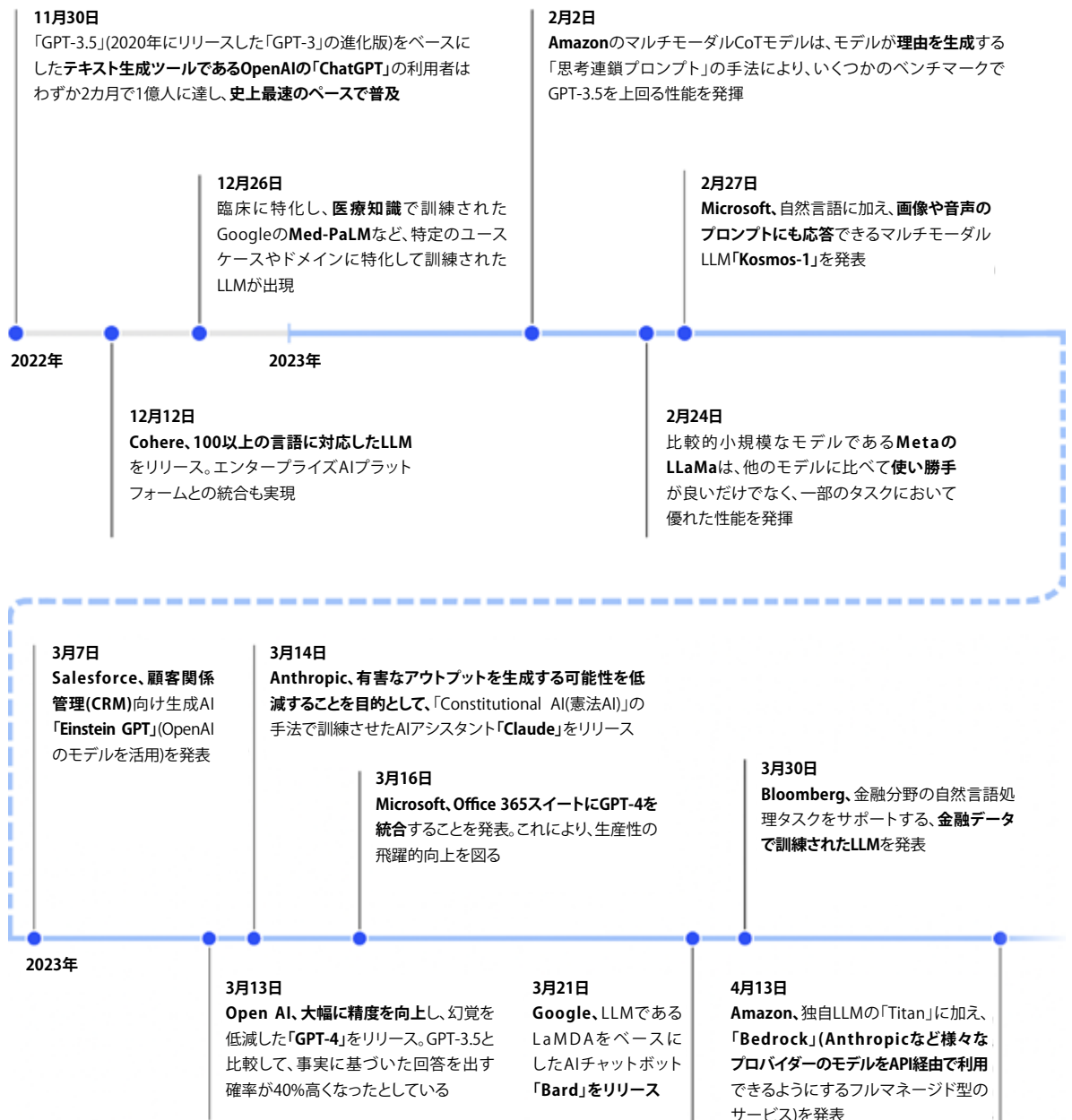
生成AI入門

生成AIの技術は飛躍的に進歩している(図表1)。リリースサイクル、スタートアップの数、既存のソフトウェアアプリケーションとの迅速な統合など、その勢いには目を見張るものがある。ここでは、生成AIの応用範囲の

図表1

生成AIは急速に進化している

ChatGPTのリリース後、数カ月間の主要な大規模言語モデル(LLM)の動向



広さと技術について、従来のAIとの違いも含めて概説する。

生成AIは単なるチャットボットではない

生成AIによって、業務の自動化や高速化、また生産性向上が可能となる。なお、ここでは、生成AIによって代替される人間のタスクではなく、生成AIによってどのように業務の効率化を図れるかについて取り上げていく。

現在、ChatGPTといったテキスト生成チャットボットに注目が集まっているが、生成AIは、画像、動画、音声、ソースコードなど、幅広いジャンルのコンテンツで能力を発揮することができ、分類、編集、要約、質問への回答、新しいコンテンツの素案の作成など、組織内で様々な機能を果たすことができる。これらの機能は、あらゆるビジネス部門やワークフローにおけるタスクの進め方を改善し、価値を生み出す可能性を持っている。例として、以下のようなものが挙げられる。

分類

- ― 不正検知専門のアナリスト: 取引や顧客に関する情報を生成AIツールに入力することで、不正取引を特定
- ― 顧客サービスマネージャー: 生成AIを使用して、顧客との音声通話を録音したデータを顧客の満足度に基づき分類

編集

- ― コピーライター: 生成AIを使用してコンテンツの文法ミスを修正し、クライアントのブランドのスタイルやトーンに合わせて調整
- ― グラフィックデザイナー: 画像から特定の要素(古いロゴなど)を削除

要約

- ― 制作アシスタント: 数時間分のイベント映像からハイライト動画を作成
- ― ビジネスアナリスト: 経営幹部のプレゼン資料の要点をまとめたベン図を作成

質問への回答

- ― テクニカルサポート: 製造現場の社員からのオペレーションに関する技術的な質問に、生成AIベースの「バーチャルエキスパート」を活用して回答
- ― 顧客サポート: 家具の購入者からの組み立て方に関する問い合わせに対して、チャットボットと分担して対応

素案の作成

- ― ソフトウェアエンジニア: 生成AIがソースコードを生成、または部分的に記述されたコードを完成させるためのコードを提案
- ― マーケティングマネージャー: キャンペーンのキャッチフレーズを考案する際に、生成AIを使用して様々なパターンの素案を作成

生成AIが進化・成熟するにつれて社内システムへの統合が進み、ミーティング終了時に議事録を自動送信するなど、タスクの自動化や特定のアクションを即座に実行できるようになる。事実、すでに様々なツールが生み出されている。

生成AIと他のAIとの違い

生成AIは、その名の通り、手書きの文書や画像など、表形式で表せない非構造化データの形式であっても新しいコンテンツを生成できるという点で、従来のAIやアナリティクスと大きく異なる(生成AIに関する用語については、コラム「用語集」を参照)。

生成AIを支える基盤モデルは、ニューラルネットワークの一つである。ニューラルネットワークは、何十億ものニューロンが相互接続された人間の脳の仕組みに着想を得ており、ディープラーニング(層の数が非常に多いニューラルネットワークを用いた機械学習の手法)を使用して訓練される。ディープラーニングは、近年のAIの進化に大きな影響を与えている。

基盤モデルは、前世代のディープラーニングモデルとは異なる、独自の特徴もいくつか持ち合わせている。まず、基盤モデルは、極めて多様かつ大規模な非構造化データセットで訓練することができる。例えば、基盤モデルの一つである大規模言語モデルは、インターネット上で公開されている多種多様なトピックが含まれる膨大な量のテキストで訓練することができる。他のディープラーニングモデルも大量の非構造化データを取り込むことができるが、通常、より具体的なデータセットで訓練される(特定の画像データセットでモデルを訓練することによって、写真の中の特定のオブジェクトを認識できるようにするなど)。他のディープラーニングモデルが実行できるタスクは、写真の中のオブジェクトの分類あるいは予測というように一つに限られていることが多いが、基盤モデルでは、これら両方の機能を実行できるだけでなく、コンテンツを生成することもできる。基盤モデルは、大量の訓練データからパターンと相関関係

用語集

API (Application Programming Interface: アプリケーションプログラミングインターフェース): プログラムを介して、(一般的に外部の)モデル、データセットや他のソフトウェアの機能にアクセスする手段

AI (Artificial Intelligence: 人工知能): 通常人間の知能を必要とするタスクを実行する能力を備えたソフトウェア

ディープラーニング: 機械学習の一部で、深層ニューラルネットワークを使用し、相互接続された「ニューロン」の層で構成され、訓練可能な重みパラメータがある。特に、画像、テキスト、音声といった非構造化データを学習データとして活用する場合に効果的

ファインチューニング: 特定のタスクでモデルのパフォーマンスを向上させるために、あらかじめ訓練された基盤モデルを微調整すること。事前訓練に使用した膨大な量のデータセットとは異なり、少量のラベル付きデータセットを用いて短期間で学習させる。これにより、少量のデータセットに隠されたニュアンスや専門用語、特定のパターンをモデルが学習し、適応することができる

基盤モデル (FM: Foundation Model): 大量のラベルなし・非構造化データで訓練されたディープラーニングモデルで、そのまま幅広いタスクに活用できる。また、ファインチューニングにより、特定のタスクにも適応させることができる。例としては、GPT-4、PaLM、DALL-E 2、Stable Diffusionなどが挙げられる

生成AI: 基盤モデルを活用して構築されており、コンテンツ生成など、従来のAIにはなかった能力を持つAI。また、基盤モデルは非生成的な用途にも使用でき(例: 通話記録に基づいて利用者の感情をネガティブまたはポジティブに分類)、従来のモデルのパフォーマンスを大幅に向上させることもできる。単純化のため、本記事で生成AIに言及する際には、基盤モデルのすべてのユースケースを対象としている

GPU (Graphics Processing Unit: グラフィックスプロセッシングユニット): 元々は、ビデオゲームなどのコンピューターグラフィックス(CG)を作成するために開発されたコンピューターチップで、ディープラーニングにも活用できる。一方、伝統的な機械学習やその他の分析は、通常、コンピューターにおける中心的な処理装置とされるCPU(中央処理装置)上で実行される

大規模言語モデル (LLM: Large Language Model): 基盤モデルの種別の一つ。非構造化テキストデータを大量に処理し、単語や単語の一部(トークン)の間の関係を学習することができる。これにより、LLMは自然言語テキストを生成し、情報の要約や抽出といったタスクを実行できる。例としては、GPT-4 (ChatGPTの技術基盤)やLaMDA (Bardが採用しているモデル)が挙げられる

機械学習 (ML: Machine Learning): AIのサブセットで、数多くのデータポイントを学習することでモデルの能力を構築。機械学習アルゴリズムは、プログラムによる指示ではなく、データ処理を通じて経験を積むことでパターンを検出し、予測や提言を行う。これらのアルゴリズムは適応することも可能で、新しいデータや経験にも効果的に対応できる

MLOps (Machine Learning Operations): AIとMLをスケーリングし、継続的にモデルを開発していくための取り組みを指す。これには、MLのライフサイクル全体(データ管理、モデルの開発、デプロイ、運用)をカバーする一連の取り組みが含まれる。現在、これらの多くは、サポートソフトウェア(タスクの標準化、効率化または自動化を支援するツール)によって実現あるいは最適化されている

プロンプトエンジニアリング: 生成AIモデルが期待通りの(正確な)アウトプットを生成するように、プロンプト(生成AIへの指示、入力文)を設計、改良、最適化するためのプロセス

構造化データ: テーブル、データベース、スプレッドシートなどの表形式のデータで、一部の機械学習モデルを効果的に訓練するために使用できる

非構造化データ: テキスト、画像、音声ファイルなど、形式や構造に一貫性がなく、洞察を導出するには一般的に高度な手法を必要とするデータ

を学習することで、このような能力を構築・強化していく。これによって、文章の次の単語を予測することなどができる。この仕組みにより、ChatGPTが様々なトピックに関する質問に答えたり、DALL-E 2 (ダリ・ツー)やStable Diffusion (ステーブルディフュージョン)が文章から画像を生成したりすることが可能となる。

企業は、基盤モデルの多機能性を生かし、単一モデルを複数のビジネスユースケースに活用できる。これは従来 of ディープラーニングモデルではほぼ不可能であった。基盤モデルに製品情報を取り込むことで、顧客からの問い合わせに対応できるだけでなく、アップデート版の開発を行うエンジニアをサポートすることもできる。このように、企業は基盤モデルを様々な用途に適用することで、早期に便益を得ることができる。

ただし、現在の基盤モデルはその性質上、すべての用途に適しているわけではない。例えば、特定の個別分類タスク(例: 発話者を逐次特定していくなどの推論)において、個別の分類の正確性を保ちつつ、処理速度のパフォーマンスも合わせて訴求する場合は、LSTM (Long Short-Term Memory) のようなニューラルネットワークの方が高速かつ正確に分類できる。このため、領域によっては、コスト効率性の観点からも、異なるアルゴリズムと併用することが望ましい結果を得られる可能性もある。また、大規模言語モデルは、時に「幻覚 (ハルシネーション)」を起こし、質問に対して一見妥当なように見えても誤った回答をすることがある(コラム「生成AIの責任ある使用」参照)。更に、推論結果の根拠や情報源は必ずしも提示されるわけではない。つまり、ミスによって被害が生じたり、説明可能性が求められたりするような用途で、人間による監視なしに生成AIを統合する場合は、細心の注意を払う必要がある。また、現時点では、生成AIは膨大な量の表形式データを直接解析したり、高度な数値最適化問題を解いたりする用途には適していない。研究者はこれらの制約を取り払うべく、日々取り組みを重ねている。

新たな生成AIエコシステムの出現

基盤モデルは生成AIの「脳」として機能しているが、その訓練や活用を支えるバリューチェーンが出現している(図表2)²。大規模な処理能力を持つ専用ハードウェアが各種モデルの訓練を行い、クラウドプラットフォームを通

じてこれらのハードウェアにアクセスする。MLOpsおよびモデルハブのプロバイダーは、組織が基盤モデルを適応させてエンドユーザーアプリケーション内に展開するために必要となるツールや技術、手法を提供している。生成AI市場には多くの企業が参入しており、基盤モデルを活用してアプリケーションを構築し、顧客のサービスに関する課題解決など、特定のタスクの実行を支援している。

初期の基盤モデルを開発する際には、モデルの訓練に多大な計算リソースが必要で、またそれらの精緻化に人間の労力も求められたことから、莫大な投資が不可欠であった。そのため、これらのモデルの開発は、一部の大手テック企業や大規模な投資を受けたスタートアップ、またはBigScienceのようなオープンソースの共同研究プロジェクトチームを中心に進められていた。しかし、現在は一部のタスクに特化し、より効率的に訓練できる小規模モデルの開発が進められている。これによって、市場参入のハードルが下がる可能性がある。実際、モデルの独自開発を進めるスタートアップが現れ始めており、Cohere、Anthropic、AI21 Labsなどは独自の大規模言語モデルを構築し、訓練している。

生成AIをビジネスに生かす

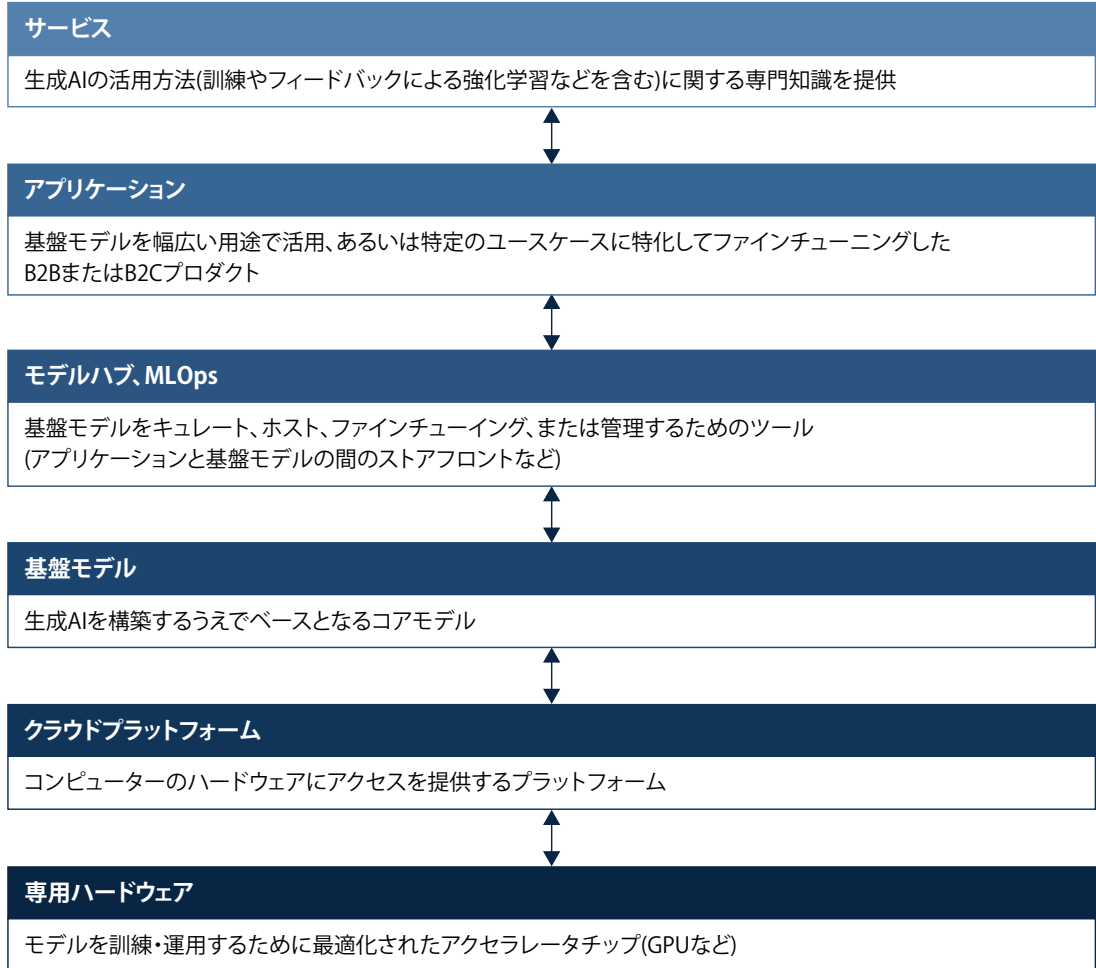
CEOは、生成AIの探求を「可能性の一つ」としてではなく「必須のもの」として捉える必要がある。生成AIを取り入れることで、幅広いユースケースを通じて価値を生み出すことができる。生成AI導入の経済的・技術的ハードルはそれほど高くないに関わらず、行動に移さないままでは、競合企業との差が急速に広がるおそれがある。CEOは経営陣と共に、どこでどのように勝負するかを検討する必要がある。企業によっては、生成AIの導入を研究開発からマーケティング、営業、顧客対応に至るまで、あらゆる業務を再考する変革の機会として捉えるところもあれば、最初は小規模に展開し、順次規模を拡大していくことを選択するところもあるであろう。いずれにせよ、方向性が定まれば、AI専門家はユースケースに応じて、開発方針に沿って戦略を実行する。

ほとんどの場合で、生成AIの機能を既存のソフトウェアに組み込むことで、生成AIを活用することになるとみられる(ただし、生成AIの価値はこれに限定されない)。

² 詳細は、マッキンゼーの記事“Exploring opportunities in the generative AI value chain”(2023年4月26日)を参照、<https://www.mckinsey.com/capabilities/quantumblack/our-insights/exploring-opportunities-in-the-generative-ai-value-chain> (英文のみ)

図表2
生成AIシステムを支えるバリューチェーンが急速に進展している

生成AIのバリューチェーン



電子メールシステムには、メッセージの下書き作成機能が加わり、業務効率化アプリケーションでは、必要な情報を入力するだけでプレゼンテーション資料の素案を作成できるようになるであろう。また、財務ソフトでは、財務報告書の重要項目に関する説明文を作成できるようになり、CRM(Customer Relationship Management: 顧客関係管理)システムでは、顧客とのコミュニケーション方法を提案してもらえるようになるであろう。これらの機能によって、知識労働者の生産性が飛躍的に向上すると考えられる。

ただし、生成AIは、特定のユースケースで更に大きな変革をもたらす可能性がある。以下に4つの事例を取り上げ、様々な業界の企業が現時点でどのように生成AIを活用して、業務改革を実現しているのかについて検討していく³。事例は、最小限のリソースで実施できるものから、膨大なリソースが必要なものまで多岐にわたる(事例の比較および技術的な詳細については図表3を参照)。

³ これらの事例は、ある特定の企業における事象をそのまま反映したものではなく、マッキンゼーがこれまでに支援したプロジェクトや一般的な事例を抜粋し、統合したものである

コラム: 生成AIの責任ある使用

生成AIは、様々なリスクを招き得る。CEOたちは、急速に進化し続ける規制要件に対応するだけでなく、ビジネスを守り、消費者からの信用を得るために、チームやプロセスを組み立て、初期の段階からこれらのリスクを回避・軽減する必要がある¹。

公正性のリスク: 不完全な訓練データや、モデルを開発するエンジニアの判断によって、アルゴリズムにバイアスが入る可能性がある

知的財産権(IP)に関するリスク: 訓練データやモデルのアウトプットによって、著作権、商標権、特許権、その他、法的保護の対象となっているあらゆる権利の侵害を含め、知財リスクが高まるおそれがある。サービスプロバイダーの生成AIツールを使用する場合でも、どのようなデータで訓練されているのか、そしてそのデータがアウトプットでどのように利用されているのかを把握しておく必要がある

プライバシーに関するリスク: ユーザーが入力した情報がモデルに取り込まれ、その個人を特定できるような形でアウトプットが生成された場合、プライバシーに関する懸念が生じる可能性がある。また、生成AIを悪用して、ディスインフォメーション(偽情報)、ディープフェイク(画像、動画、音声などの一部を結合させて作成する合成メディア)、ヘイトスピーチなどの悪意あるコンテンツを生成・拡散されるおそれもある

セキュリティに関するリスク: 生成AIが悪意のあるアクターによって悪用され、サイバー攻撃の洗練度とスピードが増すおそれがある。また、攻撃者が学習データを改ざんすることで、悪意あるアウトプットが生成される可能性もある。例えば、プロンプトインジェクションと呼ばれる攻撃手法では、第三者がモデルに新しい指示を与え、モデルの作成者やエンドユーザーが意図しないアウトプットを生成させる

説明可能性に関するリスク: 生成AIは、何十億ものパラメータを持つニューラルネットワークに依存しているため、どのようにして特定の回答が導き出されたのか、その根拠を説明することが困難となる

信頼性に関するリスク: モデルに対して同じプロンプトを繰り返しても同じ回答が得られないことがあるため、アウトプットの正確性と信頼性を評価することが困難である

組織への影響: 生成AIは労働者に大きな影響を与え、特定のグループや地域コミュニティに不均衡な影響を及ぼすおそれがある

社会のおよび環境的影響: 基盤モデルの開発と訓練を行うことで、炭素排出量が増加するなど、社会および環境に悪影響を及ぼす可能性がある(例: 一つの大規模言語モデルを訓練することで、約315トンの二酸化炭素が排出される)²

¹ Jim Boehm, Liz Grennan, Alex Singla, and Kate Smaje, "Why digital trust truly matters" マッキンゼー (2022年9月12日)

² Ananya Ganesh, Andrew McCallum, and Emma Strubell, "Energy and policy considerations for deep learning in NLP," Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (2019年6月5日)

図表3

組織が生成AIを導入・活用する際のハードルの高さはユースケースによって異なる



Low ————— High

ユースケース	開発方針	コスト	テック人材	独自のデータ	プロセスの見直し
ソフトウェアエンジニアリングの業務改革	SaaSツールを活用	多くのSaaSツールは、定額制のサブスクリプションで1ユーザー当たりの月額料金は10~30ドルとなっているが、一部、従量課金制のものもある	適切なソリューションの選定や簡単な統合に関わる作業のために必要となる可能性はあるが、テック人材はほとんど必要ない	外部のモデルをそのまま使用するため、独自のデータは不要	プロセスはほとんど変わらないが、作業者がモデルのアウトプットの正確性や正当性をチェックする仕組みが必要となる
リレーションシップマネージャーに対する情報・データの収集支援	モデルAPI上にソフトウェアレイヤーを構築	ユーザーインターフェースの開発、ソリューションの統合、後処理レイヤーの構築などへの先行投資が必要となる ランニングコストとして、APIの利用やソフトウェア保守に関わるコストが発生する	ソフトウェア開発、プロダクトマネジメントおよびデータベース統合に関するケイパビリティが必要となるため、データサイエンティスト、機械学習エンジニア、データエンジニア、デザイナー、フロントエンド開発者をそれぞれ少なくとも1人は確保しなければならない	外部のモデルをそのまま使用するため、独自のデータは不要	プロンプトやアウトプットを蓄積するためのプロセスを整備し、リスクやコスト面の理由から利用を制限するためにガードレールを設ける必要が生じる可能性がある
カスタマーサポート担当者の負担軽減および高付加価値業務へのシフト	自社でオープンソースモデルのファインチューニングを実施	データのクリーニングやラベリング、モデルのファインチューニングのための人件費がかかるため、API上に構築する場合に比べて2倍以上のイニシャルコストがかかる ランニングコストとして、モデルの保守やクラウドコンピューティングに関わるコストが発生する	経験豊富で機械学習オペレーション(MLOps)に関する知識とリソースを持ち、必要に応じてラベル付きデータを作成・チェックできるメンバーから成るデータサイエンス&エンジニアリングチームを設ける必要がある	モデルのファインチューニングのために、独自のラベル付きデータセットが必要になるが、比較的小規模で済む場合もある	問題をトリアージし、人間にエスカレーションするプロセスを整備するとともに、モデルの安全性を定期的に評価する必要がある
創業プロセスの加速	基盤モデルをゼロから訓練	人材や技術インフラについて先行投資的な支出が発生するため、API上に構築する場合に比べて10~20倍のイニシャルコストがかかる ランニングコストとして、モデルの保守やクラウドコンピューティングに関わるコストが発生する	トピックに関する博士号レベルの知識、MLOpsのベストプラクティス、データやインフラ管理に関するスキルを有するメンバーから成る大規模なデータサイエンス&エンジニアリングチームを設ける必要がある	一般公開されている大量のデータを活用して基盤モデルを訓練することができるが、長期的には、独自のラベル付きまたは比較的収集しやすいラベルなしデータをモデルに加えることで差別化を図ることができる	上記を含め、外部データを活用してモデルを訓練する場合、知的財産権の侵害を防ぐため、リーガルチェックを徹底して行う必要がある



ソフトウェアエンジニアリングの業務改革

最初に、複雑度が比較的 low、既製の生成AIソリューションをカスタマイズせずに活用することで、早期に業務効率化を実現した事例を紹介する。

ソフトウェアエンジニアの主要な業務の一つはコードを書くことで、何度も試行錯誤を繰り返し、公開、非公開を問わず様々な文献を調べる必要があるため、多くの時間と労力を要する。事例の企業では、熟練のソフトウェアエンジニアが不足していたことから、機能改善要求やバグ修正に関する未処理タスクが増え続けていた。そこで同社では、エンジニアの生産性を向上させるため、開発環境にAIベースのコード補完ツールを統合することを決定した。これにより、今ではエンジニアがコードの説明を自然言語で入力するだけで、AIから条件を満たす複数のコードブロックの候補が提案される。エンジニアはその中から適切なものを選び、必要に応じて修正を加え、クリックするだけでコードに挿入できるようになった。

マッキンゼーの研究では、このようなツールを導入することで、開発者がコードの作成に費やす時間を最大50%削減できることが明らかになっている。また、デバッグにも活用することで、開発済製品の品質向上を図ることもできる。しかし、現時点では、生成AIが熟練ソフトウェアエンジニアに取って代わることはない。実際、経験豊富なエンジニアは、このようなツールの便益を最大限享受し、生産性を上げている。一方、経験の浅い開発者では、目を見張る成果は得られておらず、時にはマイナスの影響が及んでいる。また、既知のリスクとして、AIが生成するコードには脆弱性や他のバグが含まれる可能性があるため、ソフトウェアエンジニアがコードの品質とセキュリティを担保する必要がある(リスク低減策については、本記事の最終セクションを参照)。

既製の生成AIコーディングツールのコストは比較的 low、またすぐに入手可能で自社内にて大規模な開発を行う必要はないため、導入までの期間が短い。事業者によってコストは異なるものの、定額制のサブスクリプションサービスでは、ユーザー1人当たりの月額料金は10~30ドルとなっている。なお、ツール選定時には、生成されたコードによって権利侵害が発生しないよう、プロバイダーと使用許諾契約や知的財産権などの課題に関する議論・意見交換を行うことが重要となる。

事業者の選定や新しいツールのパフォーマンスの監視は、少人数の部門横断チームが担当し、知的財産権やセキュリティに関する問題の確認もこのチームのタスクに含まれる。ツールは、ワークフローやポリシーを変更するだけで導入できる。また、純粋なSaaS(サービスとしてのソフトウェア)であるため、追加のコンピューティングコストやストレージコストはまったく発生しないか、最小限に収まる。



リレーションシップマネージャーに対する情報やデータの収集・活用支援

企業によっては、既製のツールを使わず、(APIやオープンモデルを介して)基盤モデルを活用し、独自の生成AIアプリケーションの構築を選択することも考えられる。その場合、前述の例と比べて投じるリソースは増えるが、自社固有の状況やニーズに合わせてカスタマイズできるというメリットがある。

ここでは、リレーションシップマネージャー(以下、RM)の生産性向上を図るために生成AIを導入した大手コーポレートバンクの事例を取り上げる。同行のRMは、各顧客の状況や優先事項を把握し、それぞれの特定のニーズに応じたサービスを提供するために、年次報告書や決算資料など、膨大な量の文書の確認に多くの時間を費やしていた。そこで、同行はAPIを介して基盤モデルにアクセスするソリューションを構築した。

このソリューションは、様々な文書をスキャンしてその情報をまとめ、RMからの質問に回答する機能を備えている。また、基盤モデルにレイヤーを追加し、ユーザー体験の合理化、社内システムへのツールの統合、リスクやコンプライアンスの管理などを行っている。ただし、大規模言語モデルは幻覚を起こすことがあるため、若手アナリストのアウトプットをチェックするのと同様に、モデルのアウトプットを検証する必要がある。更に、RMに対しては、ソリューションから高精度な回答を得られるような質問をするためのトレーニングが行われ(プロンプトエンジニアリング)、またツールのアウトプットと情報源の検証作業を効率化するためのプロセスも整備している。これにより、RMは分析業務に費やす時間を数日から数時間に削減させることができ、仕事に対する満足度が高まり、RMが見落としていた可能性のある洞察も捉えられるようになった。

開発コストの大半は、ユーザーインターフェースの構築とツールの統合作業に伴うもので、その主な費用はデータサイエンティスト、機械学習エンジニア、データエンジニア、デザイナー、フロントエンド開発者の人件費となっている。継続的に発生する費用としては、ソフトウェア運用保守費とAPIの利用料が挙げられる。なお、コストは、選択するモデルやサードパーティベンダーの費用、チームの規模、MVP開発に要する時間によって左右される。



顧客サポート担当者の負担軽減および高付加価値業務へのシフト

次に、基盤モデルをファインチューニングした事例を紹介する。事例の企業では、会話に特化して最適化された基盤モデルを、自社独自の質の高い顧客サポート用チャットボットや業界別Q&Aに合わせてファインチューニングを行った。同社が事業を展開する業界は専門用語が多く(法律、医療、不動産、金融など)、顧客からの問い合わせにスピーディに対応することが差別化要素の一つとして捉えられている。

同社の顧客サポート担当者が1日に対応する問い合わせ件数は何百件にもなると、時には、対応にかかる時間が長くなり、これが利用者の不満の原因となっていた。そこで、顧客からのほとんどの問い合わせに対応できる、生成AIを搭載した顧客サービスボット(以下、生成AIボット)の導入を決定した。そして、自社のブランドや顧客の好みに合ったトーンで、問い合わせに迅速に対応できる生成AIボットの構築を目指した。基盤モデルのファインチューニングとテストの一環として、回答が業界の専門用語、ブランドの提供価値、自社のブランドイメージを反映したものとなっているかの確認を行った。また、継続的なモニタリングを通じて、顧客満足度を含め、様々な観点からシステムのパフォーマンスを検証した。

加えて、同社ではモデルの潜在的エラーを最小限に抑えるために、複数のフェーズで構成したプロダクトロードマップを作成した。最初のフェーズでは、生成AIボットを社内ですべてのユーザーに導入した。社員がモデルからの提案を「良い」「悪い」の2択で評価を行い、モデルはこれらのインプットを通じて学習し、精度を上げていった。次に、モデルに顧客サポートの会話を「聞き取らせ」、提案を表示させるようにした。このようなテストを重ねた後に次のフェーズに進み、人間の監視のもと、モデルを顧客対応業務のユースケースに展開していった。リーダーたちが技術の信頼性を十分に検証した上で、最終的には顧客サポートの大半の業務を自動化することができた。

その結果、顧客サポート担当者に余剰のキャパシティが生まれ、それを付加価値や複雑度の高い問い合わせに充てられるようになったことで、業務を効率化するだけでなく、社員の仕事に対する満足度、サービス水準、顧客満足度も向上した。この生成AIボットは、社内のすべての顧客データにアクセスでき、電話を含む会話履歴を参照できる点で、従来型のチャットボットとは大きく異なっている。

この便益を享受するためには、ソフトウェア、クラウドインフラ、テック人材に多額の投資を行い、リスクやオペレーションに関して部門間連携も強化する必要があった。一般的に、基盤モデルのファインチューニングには、API上に複数のソフトウェアレイヤーを構築する場合と比べて2~3倍のコストがかかる。コスト増加の要因として、クラウドコンピューティング(セルフホスティングモデルをファインチューニングする場合)やAPI(サードパーティーAPIを介してファインチューニングする場合)に要する人件費およびサードパーティーコストなどが挙げられる。また、ソリューションの導入に際しては、DataOpsおよびMLOpsの専門家によるサポートに加え、プロダクトマネジement、デザイン、法務、顧客サービスといった各部門からのインプットも必要となった。



創薬プロセスの加速

最も複雑で高度なカスタマイズが必要となる生成AIのユースケースは、適切な基盤モデルがなく、企業がゼロからモデルを構築しなければならない場合である。このような状況は、以下の製薬会社の事例のように、専門性の高い領域や、既存の基盤モデルの学習に使用されるデータと大きく異なる独自のデータセットを扱う場合に生じる。基盤モデルをゼロから学習させる場合、テクノロジー、エンジニアリング、リソースの面で様々な課題に直面する。しかし、より高性能なモデルに投資することで、財務費用や人件費を上回る利益を得ることができる。

事例の製薬会社では、創薬を専門とするリサーチサイエンティストが、顕微鏡画像に基づいて次にどのような実験を行うかを判断していた。数百万枚もの顕微鏡画像のデータセットには、創薬に関連する細胞の特徴に関する多種多様な視覚情報が含まれていたが、人間がこれらの情報を解釈することは困難であった。これらの画像は、治療薬候補の評価に使用されていた。

そこで同社は、研究開発を加速させるために、創薬化学と顕微鏡による観察結果との関係を把握するのに役立つツールを構築することにした。このようなマルチモーダルモデルは発展途上であるため、同社は独自の学習モデルを構築することを選択した。モデルの構築にあたって、チームメンバーは実臨床画像(画像ベースの基盤モデルの訓練に使用)と、自社の膨大な顕微鏡画像のデータセットの両方を活用した。

訓練されたモデルは、望ましい効果が期待できる治療薬候補の予測、および創薬に向けて細胞の特徴を高精度に把握する能力の向上という付加価値をもたらした。これにより、創薬プロセスの効率性と有効性が高まり、価値創造までの時間が短縮するだけでなく、不正確で誤解を招くような分析や、失敗に終わる分析の数を削減できるようになった。

一般的に、ゼロからモデルを訓練する場合、APIベースのソフトウェアを構築する場合に比べて10倍から20倍のコストがかかる。なお、コストは、チームの規模や構成(例えば、機械学習に関する博士号レベルの専門家がメンバーに含まれている場合など)、コンピューティングやストレージにかかるコストによっても変わってくる。基盤モデルの訓練に要するコストは、期待する性能やモデリングの複雑性によって大きく異なる。これらの要素は、必要となるデータセットの規模、チーム構成や計算リソースに影響を与える。このユースケースでは、エンジニアリングチームと、継続的に発生するクラウドサービスの費用が全コストの大半を占めていた。

この企業では、自社の技術インフラとプロセスを大幅に刷新する必要があると判断した。これには、モデルを訓練するための様々なGPUインスタンスへのアクセス、多数のシステムに学習させるためのツール、コスト削減およびプロジェクト期間の短縮に向けたMLOpsのベストプラクティス導入などが含まれていた。また、データ収集、統合(多種多様なデータセットのファイルの形式と解像度を統一)、クリーニング(低品質データの排除、重複データの削除、使用目的に沿ったデータ分配)のために、膨大なデータ解析作業が必要となった。更に、ゼロから基盤モデルを訓練したため、モデルの性能と安全性を厳密に検証する必要もあった。

これらの事例からCEOが得られる学び

CEOは、自社で生成AIの導入を検討するにあたり、ここで紹介したユースケースを参考事例として活用いただきたい:

- 仕事や職場に実用的なメリットをもたらす画期的なユースケースはすでに存在する。製薬、銀行、小売りなど、様々な業界の企業が、様々なユースケースを立ち上げ、価値創造の可能性を追求している。組織は、それぞれが目指すものによって、スモールスタートあるいはビッグスタートを切ることができる。
- 生成AIの導入コストは、ユースケースや、ソフトウェアで必要となるデータ、クラウドインフラ、求められる技術的専門知識やリスク対策によって大きく変わってくる。企業はどのようなユースケースにおいても、リスクに関わる課題を考慮することが必須となる。企業によっては、他社よりも多くのリソースを投じる必要がある場合もある。
- 生成AIを迅速に導入すれば、それだけ早くメリットを享受できるが、まずは基本的なビジネスケースを構築することで、効果的に生成AIを導入できる。

生成AIを導入する際に考慮すべきこと

CEOは、生成AIに重点を置いた取り組みを推進する上で、極めて重要な役割を担う。この最終セクションで、生成AIの導入を進めるにあたって、CEOが念頭に置くべき要素を提示する。その多くは、新たな技術に対するこれまでの対応と同様であるが、生成AIの技術は過去に例を見ないほどの圧倒的なスピードで進化していることから、そのような技術をいかにして管理していくのかなど、生成AI特有の課題も存在することに留意する必要がある。

生成AIの導入・活用を成功させるための体制を整える

多くの組織は、従来のAIの可能性を探求するために実験的な試みを行っているものの、組織内で個別に進められているケースがほとんどである。しかし、生成AIには固有のリスクがあり、基盤モデルは組織の多種多様なユースケースに対応できる能力を備えていることから、組織としてより慎重かつ協調的なアプローチをとる必要がある。例えば、自社のブランドアイデンティティを反映するために独自の資料を活用してファインチューニングしたモデルは、様々なユースケース(パーソナライズされたマーケティングキャンペーンや商品説明の生成など)やビジネス部門(製品開発、マーケティングなど)に展開できる。これを実現するには、各部門(データサイエンス、エンジニアリング、法務、サイバーセキュリティ、マーケティング、デザインなど)のリーダーからなる部門横断チームを立ち上げることが望ましい。このような体制を整えることで、最も付加価値の高いユースケースを特定し、優先順位づけを行えるだけでなく、組織全体で足並みを揃えて着実に取り組むことができる。

ユースケースに焦点を当てるのではなく、包括的視点から事業領域を再考する

生成AIは、ビジネスの在り方を根本から変える強力なツールであり、バリューチェーン内の特定の事業領域に多大な影響を及ぼす(小売業の場合はマーケティング、製造業の場合はオペレーションなど)。生成AIの中には容易に導入できるものもあることから、組織の各所で様々なユースケースを散発的に適用してしまうことが想定されるが、ビジネス機能をまたいで最も変革の実現可能性が高い領域において一連のユースケースを描くことが重要となる。事実、多くの組織が従来のAIアプリケーションと生成AIを組み合わせ、今まで実現不可能と思われていた組織の「あるべき姿」や「働き方」を追求している。

総合的な技術スタックを構築する

生成AIの導入を成功させるには、どのようなアプローチをとっても、ほぼすべてのケースでモダンデータと技術スタックが鍵となる。CEOは、自社が計算ソース、データシステム、ツール、モデルへのアクセス(モデルハブ経由のオープンソース入手、またはAPI経由での商用利用)の観点で、必要な技術力を備えているかをCTO(最高技術責任者)とともに確認する必要がある。

例えば、生成AIにとっての生命線は、ビジネスの特定のコンテキストや問題に関する高品質なデータへの柔軟なアクセスである。したがって、企業はデータを効果的に連携させ、容易にアクセス可能な方法を確立しない限り、生成AIをファインチューニングすることができず、必然的に革新的なユースケースも開発できない。また、データガバナンスやセキュリティ手順を含め、拡張性のあるデータアーキテクチャを構築することも重要となる。ユースケースによっては、既存のコンピューティングインフラやクラウド環境で利用するツールのアップグレードが必要となるケースもある(クラウドプロバイダー経由で調達、または自社で構築)。そのため、生成AIが創出するビジネス価値と競争優位性を軸に、データおよびインフラに関する明確な戦略を立てることが極めて重要となる。

指針となる先行事例を確立する

日々、新しいモデルやアプリケーションが急速に開発・リリースされているなか、CEOであれば誰でも、計画段階で頓挫することは避けたいと願っているのではないだろうか。例えば、GPT-3は2020年、ChatGPT (GPT-3.5)は2022年11月にリリースされたが、そのわずか数カ月後の2023年3月に後継のGPT-4がリリースされている。ビジネス界では、「スピードが命」という考え方が一般的で、生成AIの成長速度の速さを考慮すると、企業がそのメリットを享受するには迅速に行動する必要がある。一方で、安定したペースで取り組みを進めていく方法もいくつか存在する。

生成AIはまだ初期段階にあるが、それが自社のオペレーティングモデルにどのような影響を与えるのかを先行事例の実装を通じて示すことが重要である。一例として、自社固有のナレッジベースとして生成AIを搭載した「パー

シャルエキスパート」を構築し、現場の社員が活用して、顧客に最適なコンテンツを提示できるようにすることなどが挙げられる。このような取り組みにより、生産性および組織の士気が向上することが期待され、社内で生成AIを試験的に導入した上で顧客向けのアプリケーションに展開できるようになる。

他の技術革新の波と同様に、「PoC (概念実証)疲れ」に陥ってしまうリスクもある。実際、パイロットフェーズでの停滞状態に陥り、抜け出せない企業も多くみられる。しかし、隣接するユースケースに拡大展開する前に、PoCを通じて価値のあるビジネスケースを迅速に検証し、精緻化することが最善の手法であることには変わりはない。有意義な結果を早期にもたらすクイックウィンに重点を置くことで、取り組みに弾みをつけ、生成AIの多目的性を生かして、スケールアウトとスケールアップを図ることができる。このアプローチにより、AIをより広範に導入し、競争力の維持に不可欠なイノベーションを推進する組織文化を醸成することができる。前述の通り、PoCは各部門のリーダーたちが中心となって、相互に連携しながら計画的に進めることが望ましい。

リスクと価値創造のバランスを取る

前述の4つのユースケースが示しているように、ビジネスリーダーは、生成AIがもたらすリスクと価値創造の機会のバランスを取る必要がある。マッキンゼーが最近実施したグローバル調査によると、半数以上の組織がすでに従来のAIを導入しているものの、大半はこれに関わるリスク対策を行っていない⁴。生成AIは、訓練データの偏りを排除できないといった従来のAIと同様のリスクのみではなく、幻覚を生成する傾向など新たなリスクをもたらすため、一層の注意を払う必要がある。

そのため、リーダーたちは、生成AIの活用に関する包括的な倫理原則とガイドラインを定めるだけでなく、各ユースケースがもたらすリスクについても十分に理解することが求められる。そして、組織全体のリスク許容度に合致し、また派生的に生じるリスクを低減する仕組みが整ったユースケースを探求することが重要となる。例えば、小売業の場合、マーケティングコンテンツの素案作成や人が関与するタスクなど、比較的価値は低いものの、リスクも低いユースケースを優先させ、高度にパーソナライズ

⁴ "The state of AI in 2022—and a half decade in review" マッキンゼー (2022年12月6日)、
<https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review> (英文のみ)

された営業メールの自動作成・送信ツールといった価値もリスクも高いユースケースは先送りすることが考えられる。このようにリスクを先送りすることで、組織は生成AIを適切に管理し、コンプライアンスを確保するために必要な体制を整えることができる。

また、CEOを含む経営陣は、会社を責任問題から守るために、消費者データの保護や知的財産権に関連する規則を含み、生成AI関連の法規制について最新動向を把握しておく必要がある。従来のAIやデータと同様、生成AIに関する法規制へのアプローチは、国によって異なる。したがって、企業は急速に進化する生成AIの規制環境やリスクに全社規模で対応できるようなアプローチをとり、プロセス管理、組織文化および人材管理を適宜見直すことが求められる。

エコシステム中心型のアプローチでパートナーシップを締結する

ビジネスリーダーは、バランスの取れたアライアンス（業務提携）の構築と維持に注力する必要がある。買収・アライアンス戦略については、引き続き、様々なビジネスの状況に適応できるエコシステムの構築に重点を置き、技術スタックの全階層に生成AIを搭載することを想定すると同時に、ベンダーロックイン（システムの開発・保守を過度に特定のベンダーに依存している状態）に注意しなければならない。

企業は、適切なアライアンスパートナーを見つけることで実行を加速できる。すべてのアプリケーションや基盤モデルを自社で構築する必要はなく、生成AIのベンダーや外部の専門家と提携することで、素早く行動できるようになる。例えば、モデルプロバイダーと提携してモデルを特定の業界向けにカスタマイズしたり、スケーラブルなクラウドコンピューティングをサポートしているインフラプロバイダーと提携したりすることなどが考えられる。

他社の専門知識を活用し、迅速に行動することで最先端の生成AI技術の便益を享受できる。ただし、生成AIモデルは氷山の一角に過ぎず、価値創造には他の複数の要素を考慮する必要がある。

必要となる人材とスキルに重点を置く

企業が生成AIを効果的に活用してビジネス価値を高めるためには、技術力を構築し、既存社員のスキルアップに取り組む必要がある。そのためには、リーダーたちが連携し、ユースケースの優先度に基づいて必要となる能力を特定する必要があり、エンジニアリング、データ、デザイン、リスク管理、製品関連など、技術職のみならず、様々なビジネス機能の人材も考慮しなければならない。

前述のユースケースに見られるように、必要となる技術や人材は、既製のソリューションを活用する場合、またはゼロから基盤モデルを構築する場合など、生成AIの導入アプローチによって大きく変わってくる。例えば、生成モデルを構築するには、機械学習に関する博士号レベルの専門家が必要となることが予想される。一方、既存のモデルやSaaSを利用して生成AIツールを開発する場合には、データエンジニアとソフトウェアエンジニアで十分対応できる可能性がある。

企業は最適な人材を確保するだけでなく、既存社員の研修・教育にも力を入れる必要がある。プロンプトベースの会話型ユーザーインターフェースにより、生成AIアプリケーションを使いやすくなる可能性があるが、ユーザーはプロンプトを最適化し、技術的な限界があることを認識した上で、ワークフロー上のどこで、いつアプリケーションを組み込むことが適切なのかについて把握しておくことが求められる。そして、リーダーたちは、生成AIツールの活用に関するガイドラインを明示し、社員が様々なリスクを認識できるように教育や訓練を継続的に実施する必要がある。また、自発的に研究や実験に取り組む組織文化を醸成することで、プロセスや製品のイノベーションを後押しし、これらのツールを効果的に取り入れることができる。

各企業は何年も前からAIの可能性を追求しており、その大半が新たな収益源の開拓、製品の改良、業務効率化を実現している。成功事例の多くは、特定のタスクに最適なツールを提供しているAI技術が鍵を握っており、企業は今後も導入範囲を広げていくことが望まれる。

加えて、生成AIは更なる飛躍的進化と新たな可能性に満ちた世界をもたらす。生成AIについては、オペレーションやリスク管理に関する制度や仕組みはまだ十分に整備されていない状況ではあるものの、ビジネスリーダーは生成AIを導入する必要性を認識している。ただ、「どこから何を始めればいいのか」という問いに対する答えは、企業によって、そして企業内の組織によっても異なる。ビッグ

スタートを選択する企業もあれば、試験的にスモールスタートを選択する企業もあるであろう。最適解は、企業が目指すものやリスク選好度によって決まる。このような判断を下す際に、CEOがその手腕を発揮することが期待される。そして、どのようなゴールを目指すにしても、まずは行動に移し、実践的な学びを深めることが肝要である。

著者

Michael Chui (パートナー) マッキンゼー・グローバル・インスティテュートおよびマッキンゼー ベイエリアオフィス (サンフランシスコ)

Roger Roberts (パートナー) マッキンゼー ベイエリアオフィス

Tanya Rodchenko (アソシエイトパートナー) マッキンゼー ベイエリアオフィス

Lareina Yee (シニアパートナー、McKinsey Technology Councilの委員長) マッキンゼー ベイエリアオフィス

Alex Singla (シニアパートナー、QuantumBlack、AI by McKinseyのグローバルリーダー) シカゴオフィス

Alex Sukharevsky (シニアパートナー、QuantumBlack、AI by McKinseyのグローバルリーダー) ロンドンオフィス

Delphine Zurkiya (シニアパートナー) ボストンオフィス

日本語版著者

工藤 卓哉 (パートナー、QuantumBlack、AI by McKinseyの共同リーダー) マッキンゼー 関西オフィス

茶谷 公之 (ベンチャーリーダー、Build by McKinsey日本統括) マッキンゼー 東京オフィス

小笠原 尚徳 (シニアデータサイエンティスト) マッキンゼー 東京オフィス

川村 俊輔 (プロダクト担当ディレクター) マッキンゼー 東京オフィス

Copyright © 2023 McKinsey & Company. All rights reserved.