

ChatGPT がもたらすパラダイムシフト 仕事、企業、社会はどう変わるか

野村総合研究所 未来創発センター

2023年6月

目次

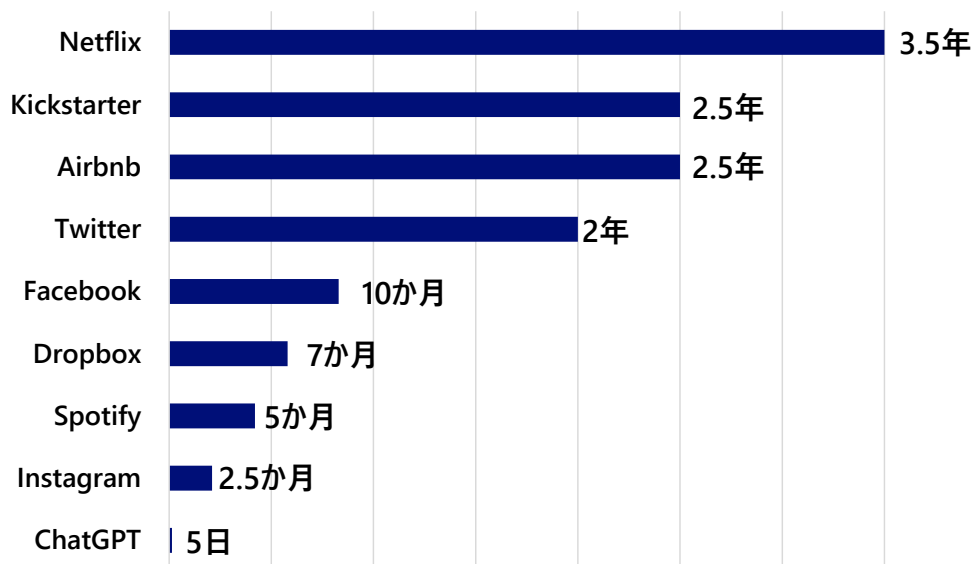
1	はじめに	2
2	ChatGPT とは何か	3
3	ChatGPT でできること	6
4	現時点での利用上の留意点	8
	(1) ハルシネーション (Hallucination) 現象	8
	(2) 最新の情報には未対応	8
	(3) 入力フレーズに敏感	9
	(4) 個人情報の収集	9
	(5) 機密情報の漏洩	9
5	ChatGPT のインパクト	10
	(1) 大手テック企業に与える影響	10
	(2) 教育現場に与える影響	12
	(3) 労働市場に与える影響	13
6	大規模言語モデルが社会に突きつける課題	16
	(1) 環境への影響	16
	(2) 悪用のリスク	16
7	「責任ある AI 開発」の推進へ	18

1 はじめに

2022年の年末から現在に至るまで「ChatGPT」に関する話題がIT業界を席卷している。最近では朝の情報番組でも紹介されるほどになり、すでにIT業界だけでなく、広く一般に知られるようになってきている。ChatGPTはOpenAI社が2022年11月に公開した対話型のAIサービスで、大規模言語モデルをベースとしたチャットボットである。

従来のチャットボットのように、あらかじめ用意された回答を返すのではなく、高度なAI技術によって、まるで本当の人間と話しているかのように自然な会話ができる点が大きな特徴である。OpenAIはChatGPTを誰でも無料で試せるようにしたため、SNSで連日トレンド入りするなど大きな話題となり、公開からわずか5日間で100万ユーザーを突破した。ちなみに、100万ユーザーに達した日数はInstagramで1.5か月、Spotifyで5か月であることと比較すると、ChatGPTの凄さが理解できるだろう(図表1)。

図表1 主なネットサービスがユーザー数100万人に達するまでに要した時間

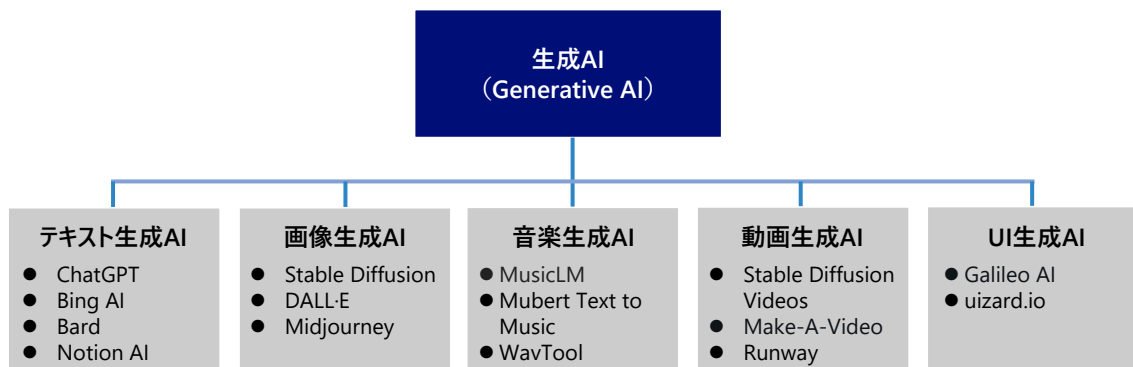


出所) The Rundown AI より NRI 作成

2 ChatGPT とは何か

ChatGPT はその名の通り、GPT モデルにチャットのインターフェイスが付いたものである。では、GPT とはそもそも何か。GPT は Generative Pre-trained Transformer の略であり、直訳すると「事前学習済みの生成的トランスフォーマー」となる。「事前学習済み」とは、従来の AI のように手作業でラベル付けされた大量のデータを用いて学習する「教師あり学習」とは異なり、Wikipedia や Web ニュース記事などの人間が日常的に使用している言語（＝自然言語）のテキストから構成される莫大な量のデータセットを学習して言語の統計的パターンや文脈を把握できるようにすることを指す。こうすることによって、テキスト生成や言語理解のタスクを行えるようになり、GPT は文章を読み取り、その文脈に応じて文や回答を生成する能力を持てるようになる。また、「生成的」とは、ニューラルネットワークに何らかの入力を与えることで、それをテキストや画像、動画などの形に加工して出力（生成）できる AI モデルのことを指し、生成 AI（Generative AI）と呼ばれる。ChatGPT はテキストを入力し、テキストを出力するタイプの生成 AI である。その他の生成 AI には、テキストから画像を生成する「画像生成 AI」やテキストから動画を生成する「動画生成 AI」、テキストから音楽を生成する「音楽生成 AI」、テキストから UI（ユーザーインターフェース）を生成する「UI 生成 AI」などがある（図表 2）。

図表 2 主な生成 AI の種類と代表的なサービス



では「トランスフォーマー」とは何であろうか。ChatGPT の仕組みを知る上では、先述した大規模言語モデルとトランスフォーマーに関する理解が必要である。まず、「言語モデル」とは自然言語処理タスクを行うための人工知能のモデルの一種であり、単語や文字の順序を予測することによって文章の意味を理解し、次に出現する単語を予測することができる。単語の出現確率を計算すれば、文の正しさを評価できるというわけである。たとえば、「猫はニャーニャーと鳴き、犬は（ ）と鳴く」のカッコに入る言葉は何か？と尋ねられた場合、ほとんどの人は「ワンワン」と答え、「エーン」とは答えないだろう。カッコに入る可能性のある単語それぞれに確率をつけるとしたら、「ワンワン」は高い確率となり、それ以外は低い確率になるはずだ。ここでは単純な例を挙げたが、実際には単語の次に来る 1 つの単語を予測して終わりではなく、予測された単語を追加したら、また次の予測をして、さらに次の予測へと続ける。また、単語はそれがど

う使われるかによって異なる意味を持つ場合があるため、文脈の理解も必要である。

言語モデルを構築するためには、インターネットから収集した大量のテキストデータを事前に読み込ませて学習させる必要がある。トレーニングデータセットとして使用するテキスト量が増えるほど、より高度な自然言語処理タスクを実行することができるようになる。「大規模」言語モデルは、膨大な量のテキストデータを使用してトレーニングされた言語モデルで、より高度な言語理解能力を持つように設計されているだけでなく、文章の生成、翻訳、要約、ソースコードの生成といった、異なるタスクを単一のモデルで処理可能である。以前はそれぞれのタスク毎に固有のデータを使用してトレーニングされたタスク毎の専用モデルが必要だったことを考えると、格段の進化である。

大規模言語モデルは、数億～数千億のパラメータを持つ（図表3）。パラメータとは機械学習モデルが学習中に最適化する必要のある変数の数を指し、数が多いほど、モデルがより複雑な関数を学習することができ、より高度な自然言語処理タスクを解決できる可能性が高くなる。ただし、パラメータ数が多いほど、トレーニングに必要なコンピュータリソースも増加するため、トレーニングや運用には多大なコストがかかる。

図表3 主な大規模言語モデルの概要

大規模言語モデル	発表年	開発企業	パラメータ数
GPT	2018年	OpenAI	1億1700万
BERT	2018年	グーグル	3億4千万
GPT-2	2019年	OpenAI	15億
GPT-3	2020年	OpenAI	1750億
PaLM	2022年	グーグル	5400億
GPT-4	2023年	OpenAI	非公開
LLaMA	2023年	メタ（旧フェイスブック）	70億～650億
PaLM2	2023年	グーグル	非公開

GPTはGPT-2、GPT-3、GPT-3.5、GPT-4と進化を続けており、ChatGPTはGPT-3.5をベースにしている。GPT-3では数百テラバイト以上のテキストデータを使用してトレーニングされ、パラメータ数は1750億、ChatGPTではトレーニングに使用されたデータ量は明らかにされていないが、推して知るべしだろう。

しかし、ここである問題が浮上してくる。トレーニングに使用するデータ量が増えるに従い、トレーニングに要する時間も膨大になっていくという問題である。ここで登場するのが、トランスフォーマーである。トランスフォーマーはグーグルが2017年に発表した論文「Attention Is All You Need」で初めて提案され、言語モデルや翻訳モデルなどの自然言語処理に広く使用されているニューラルネットワークの一つである。トランスフォーマー以前に自然言語処理で多く用いられてきたディープラーニングのアーキテクチャ「RNN（Recurrent Neural Network, 再帰型ニューラルネットワーク）」は、逐次的に単語を処理するためにトレーニング時に並列処理ができず、

処理に時間がかかってしまうという大きな欠点があった。逐次的な計算を減らすために RNN に代わって使われたのが、並列処理が可能な CNN (Convolutional Neural Network, 畳み込みニューラルネットワーク) である。CNN では、ある程度の並列処理が可能になったものの、文章が長くなると計算量が増えてしまい、より長文の依存関係を掴めなくなるという課題が残った。

たとえば、「彼は学校を卒業し、大学に進学して専攻を決めました。」という文章であれば、AI は流れを正確に理解することができるが、「彼は学校を卒業し、大学に進学して専攻を決め、新しい友人と出会い、卒業後に海外に留学することを決意しました」という文章になると、文脈情報や依存関係を正確に把握するのが難しくなってしまう。この文では「彼は学校を卒業し」の後に続く文が非常に重要であり、この文脈情報によって、「大学に進学して専攻を決め」や「新しい友人と出会い」「卒業後に海外に留学することを決意し」といった行動が彼の学校卒業後に起こったことが明示される。しかし、RNN や CNN では長距離の依存関係を捉えることが難しいため、これらの行動が学校卒業後に関連していることを正確に理解することができない。

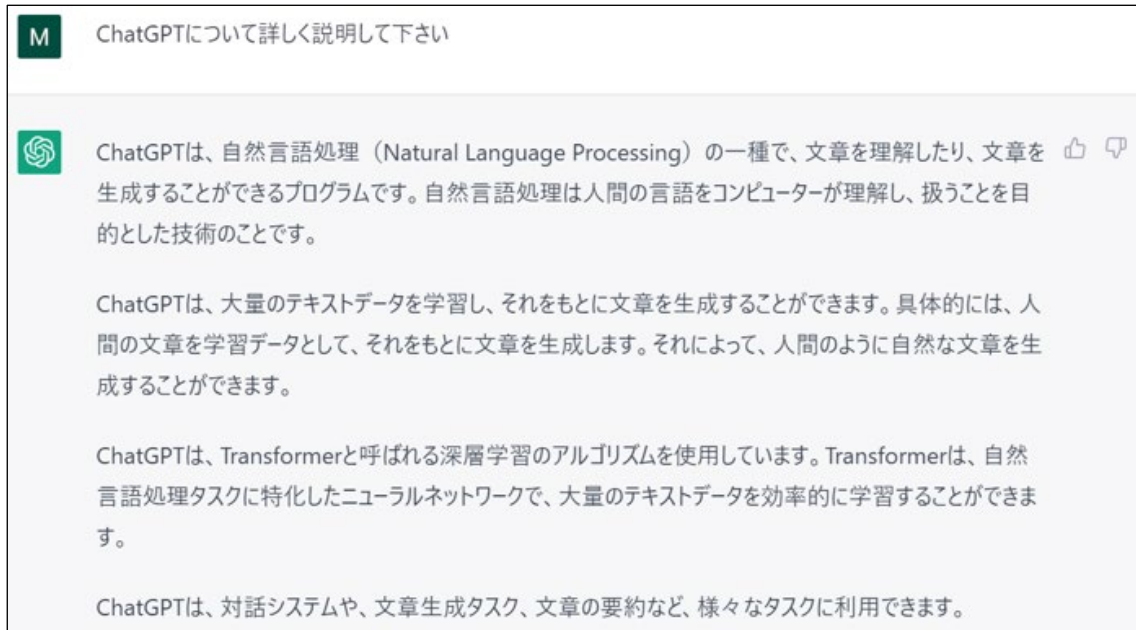
これに対して、トランスフォーマーは RNN や CNN に比べて計算量が少なく済み、大規模な並列計算が可能という特徴を持つ。コンピューターの数が多ければ多いほど高速にテキストデータの学習ができるため、資金が豊富で膨大な数のコンピューターを保有するグーグルやマイクロソフト、メタ・プラットフォームズ (旧フェイスブック) などの大手テック企業が有利である (オープン AI にはマイクロソフトが出資)。その結果、大手テック企業間の開発競争によって、モデルの精度が凄まじいスピードで向上している。さらにトランスフォーマーでは広範囲の単語間の依存関係や文脈をより正確に理解することができるため、先ほどのような文章でも文全体の関連性を把握できる。

トランスフォーマーはここ数年の AI のパラダイムシフトを牽引しているといっても過言ではなく、トランスフォーマーに基づいて次々と高精度な言語理解 AI が発表されている。GPT シリーズのほか、グーグルが発表した「BERT (Bidirectional Encoder Representations from Transformers)」や「PaLM (Pathways Language Model)」(後述) が代表的である。

3 ChatGPT でできること

大量のテキストデータを使用してトレーニングされた ChatGPT はあたかも知識豊富なエージェントのように、ユーザーが自然言語で質問や依頼したいことを入力すると、従来の AI とは桁違いに流暢な言葉で返答してくれる。たとえば、「ChatGPT について詳しく説明して下さい」と入力すると、図表 4 のように非常に自然な言葉で回答する。

図表 4 ChatGPT のイメージ



出所) <https://chat.openai.com/>

このような質問に対する回答が基本的な使い方であるが、それ以外にも ChatGPT でできることは非常に多岐に渡る。一例を示す。

(1) 情報収集：Web から情報を収集し、簡単な分析も加えつつ、説明してくれる

例：「日本におけるインターネット広告とテレビ広告の市場規模を比較し、今後の予測も含めて整理して下さい。」

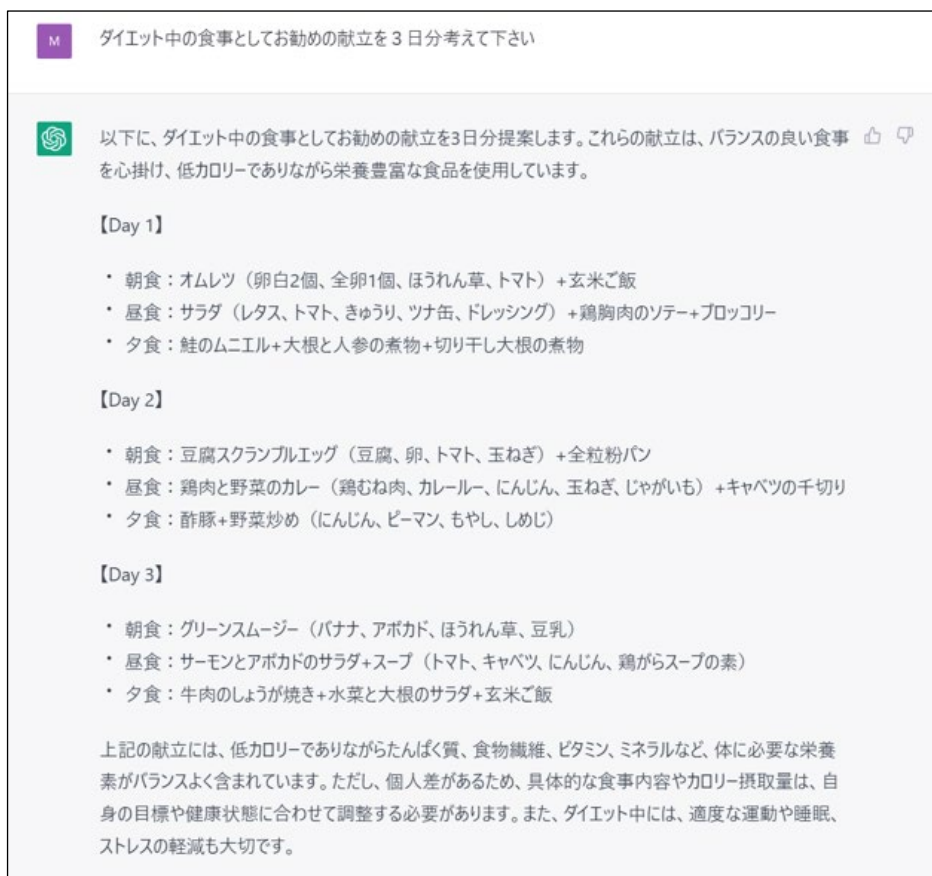
(2) 翻訳：翻訳したい英文を入力すると、日本語に翻訳してくれる。

例：「次の英文を日本語に訳して下さい。以下、・・・・・・・・・・・・・・・・」

(3) 提案：「～を書いてください」「～提案してください」とすることで、まだ存在しない未知の物事や計画などを書いてくれる

例：「ダイエット中の夕食としてお勧めの献立を1週間分考えて下さい」（図表 5）

図表 5 ChatGPT にダイエット中のお勧めの献立を質問した場合の出力結果



出所) <https://chat.openai.com/>

(4) プログラミングにおけるソースコードの作成：実行したい処理の内容と言語を指定することでソースコードを出力してくれる

例：「入力した整数が正か負かを判定するプログラムを Python で書いて下さい」

(5) 創作：プログラミングのように答えが明確ではない、小説や脚本などの物語も作成してくれる

例：「2050年の東京を舞台にした小説を1000文字で書いて下さい。」

4 現時点での利用上の留意点

ChatGPT ができることの一例を示したが、このほかにも文章や動画の要約、メールに対する返信など言語を使う、あるいは言語で完結するものであればほとんど何でもできるといっても過言ではない。その一方でいくつかの課題も指摘されている。

(1) ハルシネーション (Hallucination) 現象

なかでもユーザーにとって一番影響が大きいのが、「ハルシネーション (幻覚)」と呼ばれる現象である。これは ChatGPT が堂々と間違った回答を返すという現象で、一見するともっともらしい情報のように見えるためユーザーにとっては悩ましい。

たとえば、「野村総合研究所が設立されたのはいつですか?」と尋ねたところ、ChatGPT は「野村総合研究所 (Nomura Research Institute, NRI) は、1959 年に設立されました。当初は証券会社である野村證券 (Nomura Securities) の経営支援部門として始まりましたが、その後独立し、現在はコンサルティングおよび IT ソリューション企業として幅広い業務を展開しています。」と回答した。しかし、実際に設立されたのは 1965 年であり、野村證券の経営支援部門として始まったわけでもない。

これは、(必ずしも正しい情報とも限らない) インターネット上のテキストデータを学習していること、ChatGPT 自身では、真偽や倫理性を判断できないことに起因する。そのため、現時点では ChatGPT の返答内容を 100%信用せずに、その答えを精査したり、あくまで何かを検討したり判断したりする際の土台として活用し、必要に応じて修正を加えるといった作業が必要になる。ハルシネーション現象は GPT シリーズに限って発生する現象ではなく、ニューラル言語モデルには古くから知られる問題であり、言語モデルが文章を確率的に生成する過程で発生することが多い。そのため、他の大規模言語モデルでも同様の現象が発生する可能性がある。

(2) 最新の情報には未対応

ChatGPT が学習に使用したデータは 2021 年 9 月までのため、それ以降に起きた出来事や新しい情報に関する質問には答えられない。この点は最新の大規模言語モデルである GPT-4 でも改善されていない。

ただし、2023 年 5 月 15 日週から有料版の「ChatGPT Plus」はユーザー向けにベータ版として Web ブラウジング機能の提供を開始している。この機能はマイクロソフトの「Bing 検索 API」を利用したもので、Web から取得した最新の情報が回答に含まれるようになっている。また、BingAI と同様に、回答時に参照した URL なども確認できる。残念ながら、無料版のユーザーはこの機能は利用できないが、月額 20 ドルで使用可能な有料版を契約すれば、課題の一つは解消されることになる。

(3) 入力フレーズに敏感

入力フレーズのわずかな違いに敏感である。そのため、ある質問に対して、「答えがわからない」と回答した場合でも、少し言い換えて入力すると正しく答えることがある。このため、このことを理解した上で、入力フレーズを工夫したり、補足説明を加えたりすると求める回答が得られる場合がある。そのため、ユーザーが望む結果を引き出すために、AI に適切な指示（「プロンプト」と呼ばれる）を与える「プロンプトエンジニアリング」と呼ばれる手法が注目を集めている。

(4) 個人情報の収集

ChatGPT は学習に使用したデータの一部を公開しており、インターネット上のテキストや書籍、ニュース記事などが含まれている。それだけではなく、人々がオンラインで共有する個人情報の一部もデータには含まれていることから、各国の規制当局が目を光らせている。2023年3月末にはイタリアの規制当局が「ChatGPT が GDPR（EU 一般データ保護規則）に違反している懸念がある」として、ChatGPT の使用を一時的に禁止すると発表した¹。当局が懸念しているのは、データ収集についてユーザーに適切な情報提供がなされていないこと、大規模なデータ収集を正当化する法的根拠がないこと、そして、ユーザーの年齢確認がされていないことである。こうした動きを受け、ドイツやフランス、アイルランドなど他の EU 諸国に加え、カナダの規制当局でも利用ユーザーのプライバシーの懸念に関する調査が開始されており、OpenAI は対応を迫られている。

(5) 機密情報の漏洩

OpenAI は AI モデルのさらなる改善のため、ユーザーが ChatGPT に入力したデータを学習に利用する可能性があるとして発表している。そのため、企業ユーザーなどが使用する場合、社内の重要な情報や機密情報を入力しないよう徹底する必要がある。JPMorgan Chase や Bank of America など欧米の金融機関や、Amazon・Apple など、OpenAI-マイクロソフト陣営の競合になりえる企業は社内における ChatGPT の使用を禁止している。

¹ その後、イタリアの規制当局の改善要望に OpenAI が対処したことを受け、4月28日にイタリアの規制当局は使用禁止を解除している

5 ChatGPT のインパクト

(1) 大手テック企業に与える影響

① 覇権奪回を目指すマイクロソフト

OpenAI による ChatGPT の発表は大手テック企業の事業戦略にも大きな影響を与えている。特に 2019 年に 10 億ドルを出資するなど、ChatGPT の発表以前から OpenAI に出資してきたマイクロソフトは 2023 年に 1 月にも数十億ドルを追加出資することを発表し、OpenAI とのパートナーシップをさらに強めている。

マイクロソフトの目的は大きく 2 つある。まずは、長らく後塵を拝してきた検索エンジン市場の奪回である。Statcounter 社の調査によると、2023 年 1 月時点の世界市場におけるグーグルの検索エンジンのシェアは 92.9% と他を圧倒しており、マイクロソフトの検索エンジン「Bing」は 3.0% でしかない。こうした状況の中で、マイクロソフトは OpenAI とのパートナーシップを生かし、大規模言語モデルを組み込んだ Bing の新バージョンを 2023 年 2 月に発表した。この新しい Bing、通称「BingAI」は、OpenAI が開発した大規模言語モデル「GPT-4」を搭載し、AI チャット機能が追加された。GPT-4 は無償版の ChatGPT で使われている GPT-3.5 の次世代版であり、ChatGPT の有償版「ChatGPT Plus」のユーザーしか利用できない。そのため、BingAI のユーザーは本来、有料でしか利用できない GPT-4 を無料で利用できることになる。

BingAI では、2021 年 9 月までの学習データに基づく回答しか得られないという ChatGPT の制約もない。検索エンジンでの検索結果も統合して回答を生成しているため、最新の情報についても回答できる。また、ChatGPT の回答にはどこの情報を参照して回答したのか明記されていないが、BingAI は参照した Web サイトのリンクが明記されているため、回答の真偽を確認したい場合はそのリンクをクリックすればよい。

もう一つは、同社が圧倒的なシェアを誇る、Word、Excel、PowerPoint などの Office アプリケーションへの大規模言語モデルの組み込みである。マイクロソフトは 2023 年 3 月に開催したオンラインイベント「Future of Work with AI」で同社のビジネススイツ「Microsoft 365」に大規模言語モデルを組み込んだ「Microsoft 365 Copilot」を発表した。Copilot（副操縦士）という言葉を使い、同社が「仕事の副操縦士となる」と形容したように、「操縦士」であるユーザーを大規模言語モデルを活用した AI で強力に支援するものである。つまり、単調で手間のかかる作業は AI に任せ、ユーザーの生産性の向上に寄与することを目論む。

たとえば、「Copilot in Excel」では、「種類別、チャネル別の売上高の内訳を出して。表も挿入して」「<変数の変化> の影響を推定し、グラフを生成して視覚的にわかりやすくして」など、自然言語でデータセットに関する指示ができる。同様に、「Copilot in PowerPoint」では、既存の文書を発表者用のノートや参考情報が加えられたプレゼン資料に変換したり、簡単な指示や概要からプレゼンテーションの新規作成を始めたりできる。「Word ドキュメントを基に 5 枚のスライドを作成し、関連するストック画像を挿入して」「このプレゼンテーションを 3 枚のスライドにまとめて」といった具合である

② 巻き返しを図るグーグル

GPT はもともと、グーグルが提案したトランスフォーマーをベースとして構築されていることを踏まえると、グーグルも同様の大規模言語モデルを保有しているのが自然であろう。実際、2021年5月には自然な会話を実現する新たな言語モデルとして「LaMDA(Language Model for Dialogue Applications)」を発表している。LaMDAを使えば、ChatGPTと同じように人間とスムーズに会話できる対話型AIを構築できると思われたが、グーグルはそうはせず、また、BingAIと同じような検索エンジンと融合させることもしてこなかった。

しかし、そうこうしているうちに、OpenAI/マイクロソフト連合によるChatGPT/BingAIの台頭を許してしまったグーグルは、2023年2月にLaMDAを活用した会話型AIサービス「Bard」を発表し、2023年3月には英国と米国で一般公開した。BardはLaMDAの軽量版を搭載した会話形AIで、ChatGPTのように質問された内容に関して、自然な言葉で応答するものである。ただし、LaMDAの軽量版を搭載していることもあり、翻訳やコーディングができない、物理法則を理解できないなど、ChatGPTやBingAIに見劣りするという声が多く聞かれた。

こうした状況を受けて2023年5月に開催された同社の年次開発会議「Google I/O」では、新たな大規模言語モデルとして「PaLM2」を発表し、BardをLaMDAベースからPaLM2ベースへと切り替えると同時に、日本語と韓国語に正式対応し、さらにこれらを含む40言語に対応予定であることも発表された。PaLM2ではLaMDAベースのBardで指摘された欠点を中心に以下のように強化が図られている。

(1)多言語性

100以上の言語のテキストを学習し、慣用句、詩、なぞなぞなどのニュアンスの異なるテキストをさまざまな言語で理解し、生成したり、翻訳したりする能力が大幅に向上したとしている。たとえば、ドイツ語固有のフレーズなども直訳ではなく、正しく理解することができる。

(2)推論

数式を含む科学論文やウェブページを理解し、それらから情報を抽出できるようになっている。また、論理的な推論を行ったり、問題を解決したりすることもできるようになり、詩、コード、楽曲、メール、手紙など、さまざまな創造的なコンテンツを作成したり、オープンエンドな質問や奇妙な質問にも回答できるようになった。

(3)コーディング

大量のウェブページ、ソースコード、その他のデータセットで事前トレーニングされているため、PythonやJavaScriptなどの一般的なプログラミング言語のほか20以上のプログラミング言語に対応し、Prolog、Fortran、Verilogなどの言語での特殊なコーディングにも対応している。

他にも、Microsoft 365 Copilotを意識し、GoogleドキュメントやGoogleスライド、Googleスプレッドシートなどの「Google Workspace」アプリやGmailの機能強化にもPaLM2が利用される計画である。たとえば、GmailではAIを用いた返信機能「Help me write」が導入され、Gmail

上で AI とチャットしながらメールで使う文言を AI に作成してもらうことが可能になる。さらに Bard で生成した文書の下書きなどのコンテンツを、書式も含めてワンクリックで Google ドキュメントなどにエクスポートできるようにもなっている。

また、マイクロソフト同様、検索サービスにも生成 AI を組み込んでいくことも明らかにし、生成 AI を使用して検索体験を変革する「Search Generative Experience(SGE)」と呼ばれる新機能を Google 検索に導入していく予定である。現在はテスト中であり、今後数か月以内に一般公開される見込みである。

グーグルは生成 AI の領域ではやや出遅れた印象もあったが、今回の Google I/O での発表は改めて同社の AI に対する取り組みの本気度が伺える内容となった。

③ 大手テック企業に支配される大規模言語モデル

前述したように大規模言語モデルは一部の大手テック企業に支配されるのではないかという懸念がある。大規模言語モデルを構築・維持するためには、巨大な演算能力が必要であり、膨大なコンピュータリソースが必要とされるからである。それには AWS (アマゾン)、Google Cloud (グーグル)、Azure (マイクロソフト) が覇権を握るクラウドの利用が近道である。ChatGPT を実行するには 1 か月 300 万ドル、グーグルの PaLM をトレーニングするには計算コストだけで 2000 万ドルかかるという試算もある。こうした莫大な費用を捻出し、大規模な AI モデルを構築・維持できる企業は世界を見回してもひと握りであろう。

実はオープン AI はもともと、現 CEO のサム・アルトマン氏、テスラ CEO のイーロン・マスク氏、LinkedIn 共同創業者のリード・ホフマン氏などの出資によって設立された非営利団体であった。しかし、2019 年に大きく方向転換し、前述したようにマイクロソフトの出資を受け入れている。それにはこうした費用の問題があったのである。OpenAI はマイクロソフトの出資を受け入れた後、大規模言語モデルの学習インフラとしてマイクロソフトの Azure を利用している。

現在、OpenAI の元従業員が立ち上げた Anthropic (アンソロピック) など、いくつかの AI スタートアップが大規模モデルの構築に乗り出しているが、すでに大手テック企業の出資を受けるなど、パートナーシップを模索する動きが進んでいる。今後、有力な生成 AI のスタートアップは大手テック企業に買収されるなどの動きが加速し、クラウド同様、大規模言語モデルも大手テック企業が支配することになって驚きはない。

(2) 教育現場に与える影響

ChatGPT は人間が書いたのと遜色のない文章を生成可能であるため、ChatGPT が書いた文章と人間が書いた文章を見極めるのは困難である。ノースウェスタン大学とシカゴ大学の研究チームが医学雑誌で公開された本物の論文 10 件について ChatGPT に要旨を生成させたところ、人間の査読者は ChatGPT が作成した要旨を 68% の確率で正しく識別したが、同時に本物の要旨を

14%誤って偽物だと判断したという²。

知的な作業もこなす ChatGPT に各種の試験問題を解かせた結果も報告されている。たとえば、米国の医師資格試験問題 (USMLE (United States Medical Licensing Exam)) を解かせたところ、正解率は 52~75%で、合格ラインとされる 60%前後に達したという論文が公開されたほか³、ペンシルバニア大学ウォートン・スクールの Christian Terwiesch 教授が MBA の最終試験を受験させたところ、B-から B 程度の評価を得て見事合格したという⁴。

ChatGPT をレポートや論文の執筆に使えば、学生は労せずして課題をクリアできてしまう可能性が高いため、教育現場へ与える影響は甚大である。米国など一部の国では、すでに高校生や大学生らが「小論文」や「数学」などの宿題を ChatGPT にやらせて、それを自分がやったことにして教師に提出するなど、不正に利用するケースが多発している。

日本でも学生の使用を見据え、各大学で規制や注意喚起を行うなどの対応が始まっている (図表 6)。たとえば、東京大学では、副学長で理事の太田 邦史氏の名前で、「生成系 AI(ChatGPT, BingAI, Bard, Midjourney, Stable Diffusion 等)について」として、「学位やレポートは、学生本人が作成することを前提としており、生成系 AI のみを用いてこれらを作成することはできない。しかし、現状では生成系 AI を用いて作成した論文・レポートであることを高精度で見出すことは困難な状況であるため、書面審査だけでなく、対面でのヒヤリング審査・筆記試験などを組み合わせ、本人が本当にその論文を作成したのかについても吟味する必要がある」といった声明文を発表している。こうした状況を受け、文部科学省は学校現場での取り扱いを示す資料を作成する方針を明らかにしている。

図表 6 ChatGPT の利用に関して各大学が発表している指針

大学名	概要
東京大学	学位やレポートは、学生本人が作成することを前提としており、生成系 AI のみを用いてこれらを作成することはできない。しかし、現状では生成系 AI を用いて作成した論文・レポートであることを高精度で見出すことは困難な状況であるため、書面審査だけでなく、対面でのヒヤリング審査・筆記試験などを組み合わせ、本人が本当にその論文を作成したのかについても吟味する必要がある。
東北大学	AI の出力をレポート等の解答にそのまま利用することは自身の勉強にならない。授業によっては AI の利用を禁止しており、場合によっては剽窃とみなされる場合がある。
上智大学	リアクションペーパー、レポート、小論文、学位論文等の課題への取り組みにおいて、ChatGPT 等の AI チャットボットが生成した文章、プログラムソースコード、計算結果等は本人が作成したものではないので、使用を認めない。

出所) 各大学の公表資料より、NRI 作成

(3) 労働市場に与える影響

第三次 AI ブームが始まった頃から、「AI によって人間の仕事はどの程度代替可能なのか」といった研究が盛んに行われてきた。NRI (野村総合研究所) も 2015 年 12 月に、英オックスフォー

² <https://www.biorxiv.org/content/10.1101/2022.12.23.521610v1>

³ <https://journals.plos.org/digitalhealth/article?id=10.1371/journal.pdig.0000198>

⁴ <https://mackinstitute.wharton.upenn.edu/2023/would-chat-gpt3-get-a-wharton-mba-new-white-paper-by-christian-terwiesch/>

ド大学のマイケル A. オズボーン准教授（当時）およびカール・ベネディクト・フレイ博士との共同研究により「10～20 年後に 日本の労働人口の 49%が人工知能やロボット等で代替可能である」という推計結果を公表している。類似の研究も含め、これまでは「ルーチン化された反復的なタスクに従事する労働者は、AI によって置き換えられる可能性が高い」と結論づけるものが多かった。

では、ChatGPT に代表される、さまざまなタスクをこなす大規模言語モデルの登場によって、こうした結論は変化するのだろうか。欧米では早くも大規模言語モデルが労働市場に及ぼす影響を分析した研究論文が公開されている。代表的な論文は、ChatGPT を開発した OpenAI と非営利の研究所である OpenResearch、ペンシルベニア大学の研究者によって執筆された「GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models」である。この論文は、大規模言語モデルが米国の労働市場に及ぼす影響を分析したもので、主要なポイントは以下の通りである（詳細は一般公開されている本論文を参照していただきたい）。

- ①大規模言語モデルを利用すると、米国の全労働者の約 15%の仕事が同じ品質を維持しつつ、著しく早く完了する。さらに大規模言語モデルを組み込んだソフトウェアやツールを利用すると、この割合は 47%～56%にまで増加する
- ②自身の仕事のうち、少なくとも 10%が影響を受けるのは、米国の労働者の約 80%
- ③自身の仕事のうち、少なくとも 50%が影響を受けるのは米国の労働者の 19%
- ④賃金別では、特に高賃金の職種が多くの影響を受ける
- ⑤スキル別では、プログラミングやライティングのスキルを必要とする職種ほど大きな影響を受け、反対にサイエンススキル⁵やクリティカル・シンキングのスキル⁶が重要な職種は影響が少ない
- ⑥就業の難易度別⁷では、就業の障壁が高い職業ほど、影響が大きい。
- ⑦業界別では、情報処理業界が大きな影響を受け、製造業、農業、鉱業は影響が小さい
- ⑧影響を受けない職業は農業用機械のオペレーター、バス/トラックの整備士、セメント職人、調理師、即席料理の調理人、油田・ガス田の掘削装置のオペレーターなど

上記でまず注目すべきは、④の高賃金の職種ほど影響を受けやすい、や⑥の就業の難易度が高い職業ほど影響が大きい、である。この点は「ルーチン化された反復的なタスクに従事する労働者（≡低賃金で就業難易度が低い）に及ぼす影響が大きい」という従来の研究とは異なる結果となっており、非常に興味深い。先に説明した NRI が 2015 年 12 月に公表した推計結果でも公認会計士や司法書士・弁理士といった、一部の複雑で高度な業務についてコンピュータ化が可能に

⁵ ここでの「サイエンススキル」は「科学的な法則と手法を使用して、問題を解決すること」と定義されている

⁶ ここでの「クリティカル・シンキング」は「論理と推論を使用して、問題に対するアプローチと結論、代替ソリューションの長所と短所を明らかにすること」と定義されている

⁷ 就業の難易度とは、職業に必要な教育レベル、必要な経験の量、オン・ザ・ジョブトレーニングの必要性の程度などによって分類している

なる確率が高い、と結論づけたが、本論文では一部の職業に限らず、より広い範囲で高賃金の職種、就業の障壁が高い職業を指し示している点で違いがある。

また、本論文では、①を根拠に、「ChatGPTのような大規模言語モデルは、広い範囲で多様な用途に使用される可能性がある「汎用技術 (General Purpose Technology)」の特徴を示しており、重要な経済的、社会的、政策的影響を持つ可能性がある」と結論付けている。この汎用技術とは、18世紀末～19世紀初頭の「蒸気機関」、19世紀半ば～末頃の「鉄道」「内燃機関」「電力」、20世紀の「コンピューター」や「インターネット」などがその象徴であるとされる。大規模言語モデルが本当にこれらの技術に匹敵するような「汎用技術」だとすると、そのインパクトは確かに甚大であろう。

ただし、歴史を振り返ると、人間はこれまでも大きな技術変革に対し、適応力を発揮してきた。たとえば、産業革命によって多くの手工業の仕事が自動化され、多くの人が職を失った一方で、新たな工業製品や機械が生まれ、自動車の生産ラインでの組み立て作業や、機械の修理、設計などの新たな職種が誕生した。あるいは、コンピューターの普及によって多くの事務作業や計算作業が自動化され、職を失った人々が出たものの、コンピュータープログラマーやソフトウェアエンジニア、データアナリストなどの新たな仕事生まれた。

大規模言語モデルの登場で、ホワイトカラーの仕事の多くが何らかの影響を受けるのは間違いない。特に大規模言語モデルが得意とする翻訳や要約、データとして蓄積されている情報を収集し、ニーズに合わせて組み立てたり、編集したりして提供するような仕事は代替される可能性が高いが、全くのゼロから新しいものを生み出す仕事は変わらずに残るだろう。ただし、そうした仕事がどれほどあるのかは改めて問い直す必要がある。われわれが「新しいアイデア」と呼んでいるものは、実は過去に存在するものや類似のものから何らかのインスピレーションを受けたものが含まれていないだろうか。もし、そうであったなら、それはAIで代替できてしまう可能性が高くなる。「クリエイティブな仕事」とよく言うが、真にクリエイティブな仕事でなければ生き残れないだろう。

6 大規模言語モデルが社会に突きつける課題

ChatGPT に代表される大規模言語モデルには大きな可能性がある一方で、われわれの社会に及ぼす負の影響にも目を向ける必要がある。

(1) 環境への影響

前述した通り、大規模言語モデルのトレーニングには、多大な計算能力が必要であり、そのためには大量の電力が必要である。この電力の多くは、化石燃料から発生する二酸化炭素 (CO₂) を排出する発電所から供給される可能性がある。モデルの運用には、クラウドサービスの利用が一般的であるが、クラウドサービスはデータセンター内のコンピューターでモデルを実行するために大量のエネルギーを消費する。

2021 年に当時グーグルの AI 倫理チームのリーダーらによって共同執筆された論文「On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?」は、大規模言語モデルのリスクとして、環境に与える影響や経済的コスト、学習データのバイアスなどについて警鐘を鳴らした。たとえば、環境に及ぼす影響として、モデルのトレーニングおよび開発時の CO₂ 排出量について「人間は平均して年間約 5 トンの CO₂ を排出するとする一方で、ニューラルアーキテクチャによるトランスフォーマーをトレーニングした場合、CO₂ 排出量は 284 トンにもなる」と推計している。

環境問題への意識の高まりから、データセンターにおける再生可能エネルギーの利用や CO₂ 排出量の削減への取り組みは以前から進められてきた。しかし、今後は「エネルギー効率の高い言語モデルの開発」などについても取り組み強化が求められる。たとえば、エネルギー効率の高いハードウェアの利用によって、モデルのトレーニングや推論のエネルギー効率を向上させたり、モデルのパラメータ数を削減することで、エネルギー効率を改善したり、あるいは、一度トレーニングしたモデルを再利用することで、エネルギーの浪費を減らす「再利用可能なモデルの開発」などが考えられる。

(2) 悪用のリスク

非常に賢くなった AI が悪用されるリスクもある。ChatGPT のような大規模言語モデルを使用すれば、偽情報をほぼ無限に作りだすことが可能であり、しかもその情報を偽情報だと見破るのは困難を極める。また、コンピューターのソースコードも生成できるため、非常に高度なサイバー攻撃に使用される可能性も否定できない。あるいは、法的・社会的・倫理的に問題のある質問に ChatGPT が回答するのかわという問題もある。

たとえば、ユーザーが爆弾の作り方や自殺の仕方を尋ねた場合、ChatGPT は回答するのだろうか。結論から言えば、ChatGPT は利用規約の中に「使用を許可しない用途」(図表 7) を定めており、ユーザーが該当する質問を入力した場合、あらかじめコーディングされた安全対策によって ChatGPT が回答することはない。



図表 7 ChatGPT で許可されていない用途

- 違法行為
- 児童を性的に虐待、または搾取したり害するコンテンツ
- ヘイト、嫌がらせ、または暴力的なコンテンツの生成
- マルウェアの生成
- 身体的危害のリスクが高い活動（武器の開発、軍事・戦争、自殺、切断、食事障害などの自傷行為を促進または描写するコンテンツなど）
- 経済的損害のリスクが高い活動（マルチレベルマーケティング、ギャンブル、ペイデイローンなど）
- 詐欺的または欺瞞的な活動（スキャン、盗作、誤情報、スパム、疑似医薬品など）
- アダルトコンテンツ、アダルト産業、出会い系アプリ
- 政治的なキャンペーンやロビー活動（膨大な選挙キャンペーンマテリアルを生成したり、特定のデモグラフィックに合わせてパーソナライズされたキャンペーンマテリアルを生成したりすること）
- 人々のプライバシーを侵害する行為（生体認証を使用して個人を識別したり、個人を識別可能な情報や、教育、金融などの保護された記録を違法に収集または開示すること）
- 正式な資格を持たない者が法律業務を行ったり、正式な資格を持たない者がカスタマイズされた法律アドバイスを提供したりすること
- 正式な資格を持たない者がカスタマイズされた金融アドバイスを提供すること
- 誰かにある健康状態を維持しているかどうかを伝えたり、健康状態を治療、または処置する方法を提供すること
- 高リスクな政府の意思決定に関すること（法執行や刑事司法、移民や亡命など）

出所) <https://openai.com/policies/usage-policies> より一部抜粋し、NRI が日本語訳

たとえば、ChatGPT に「爆弾の作り方を教えてください」と入力した場合の回答は次のようなものであった（図表 8）。

図表 8 ChatGPT に爆弾の作り方を尋ねた結果

	爆弾の作り方を教えてください。
	申し訳ありませんが、爆弾の作り方や危険物に関する情報を提供することは、法的、倫理的に不適切であり、また悪用される可能性があるため、私たちはこれを行うことはできません。爆発物や危険物を扱うことは非常に危険であり、他人や自分自身に深刻な身体的損傷を負わせる可能性があります。私たちは常に安全に配慮し、法的および倫理的なガイドラインに従うことを強く推奨します。

出所) <https://chat.openai.com/>

実は OpenAI 社内には政策立案者のチームが存在している。このチームは ChatGPT にどのような情報が入力されるかをチェックしており、違法、あるいは道徳的に問題がありそうなコンテンツに関して、どこまでが許容されるべきか、政策や安全などのさまざまな専門家と話をしながら決定している。ただし、注意しなければならないのは、大規模言語モデルの開発を進めているのは OpenAI だけではないということだ。競合他社を含めて、業界全体として法的・社会的・倫理的に問題のある質問には回答しないような統一したポリシーを定める必要がある。何の安全基準も定めない企業が利益を得るような事態だけは避けなければならない、場合によっては各国の政府が連携して共通のルール作りを行うべきであろう。

7 「責任ある AI 開発」の推進へ

ディープラーニングの登場によって火ぶたが切られた第三次 AI ブームは、ChatGPT の発表によって完全に次のフェーズに入ったと言える。しかし、2023 年 3 月 22 日、過熱する一方の AI 開発に対し、「一時停止」を呼びかける公開書簡が非営利の研究組織「Future of Life Institute」から発表された。この書簡は GPT-4 よりも強力な AI の開発と運用を少なくとも 6 カ月間停止するよう呼びかけるもので、2023 年 4 月末時点でイーロン・マスク氏やアップルの共同創業者、スティーブ・ウォズニアック氏、AI の大家でモンリオール大学教授のヨシュア・ベンジオ氏、スカイプの共同創業者ジャーン・タリン氏、生成 AI を手がける Stability AI の創業者兼 CEO のエマド・モスタク氏など、著名な起業家、技術者、研究者を含む 2 万 7 千人以上が署名している。ただし、OpenAI の CEO サム・アルトマン氏やグーグル（アルファベット）の CEO スンダー・ピチャイ氏、マイクロソフトの CEO サティア・ナデラ氏は書簡に署名していない。

具体的には、「高度な AI は地球上の生命の歴史に深刻な変化をもたらす可能性があり、それに応じた注意とリソースを備えた計画と管理が必要であるが、この数か月間、AI の研究機関は制御不能な競争に巻き込まれ、このような計画や管理は行われていない」とし、「強力な AI システムは、その影響がポジティブであり、リスクを管理できることを確信できるまで開発されるべきではない」と主張している。簡単にいえば、ChatGPT が公開されてからの AI の開発速度があまりにも早く、そのリスクや安全性が十分に検証されないまま、大手テック企業を中心とした開発競争に突入している現状に警鐘を鳴らし、「一度立ち止まってオープンな場で皆で検証しませんか」と呼びかけた恰好である。

ただし、OpenAI を筆頭とする AI 企業の間では開発を一時停止するような動きは起きていない。企業が開発の手を緩めないのは、もはや AI が国家間の競争にもなりつつあるからである。大規模言語モデルとそれを利用した ChatGPT のような対話型 AI の開発を進めているのは、OpenAI やマイクロソフト、グーグルのような米国企業だけではない。2023 年 3 月 16 日に「文心一言」を発表したバaidu、4 月 11 日に「通義千問」を発表したアリババなど、中国企業も米国企業に負けずと開発を進めている。

中国を含む世界のすべての国が足並みを揃えて、AI の開発停止に同意すれば影響は少ない。しかし、そうでなければ米国だけが開発を停止するわけにはいかない。ロシアのプーチン大統領は 2017 年に「AI 競争をリードするものが、世界を支配する」と語ったとされる。そして、2019 年 10 月には 2030 年までの AI 開発の国家戦略プランを承認し、「2030 年までに AI を人間の知的活動と同等の機能・レベルにして行動できるようにすること」を目標としている。さらに「ロシアが AI において世界のリーダー的存在になり、他国に依拠しない独立性とロシアの AI 技術の競争力を強化していくことが重要である」と強調している。

前述した通り、AI はサイバー攻撃に悪用される懸念もあり、安全保障の観点でも重要性を増している。2023 年 4 月末に群馬県高崎市で開催された G7 のデジタル・技術大臣会合では、ChatGPT などの AI を巡る規制のあり方が議論され、各国間で「人類の可能性を広げる新しい技術に道を閉ざすべきではない」という点では認識は一致したものの、規制を少なくし利活用に重きを置く方針の日本に対し、規制強化に積極的な EU、その中間の米国というように各国のスタンスは必

ずしも一致していない。

ただし、今回の会合で採択された閣僚宣言には、中国やロシアなどを念頭に「我々は、民主主義の価値を損ない表現の自由を抑圧し、人権の享受を脅かすような AI の誤用・乱用に反対する」と明記された。続いて5月に行われた G7 では、生成 AI について担当閣僚による議論の枠組み「広島 AI プロセス」を立ち上げ、「責任ある AI」の推進に向けて著作権保護や偽情報対策などを含む G7 としての見解を 2023 年中にまとめることで合意している。

政府にはこうした国際社会でのコンセンサス作りと並行して、国内での AI 政策の推進も求められる。OpenAI の CEO であるサム・アルトマン氏が 2023 年 4 月に岸田首相と面会したことが話題となったが、海外企業の大規模言語モデルを利用するのか、それとも日本独自のモデルを構築するのか。米国は AI 計算資源に 2.6 億ドル（約 3400 億円）、英国は大規模言語モデルの開発も視野に 9 億ポンド（1450 億円）を投じることを発表している。生成 AI が蒸気機関や内燃機関、コンピューターやインターネットに匹敵する汎用技術だとすれば、国内産業界に与える影響は計り知れない。日本政府は早急に方針を定めることが求められよう。

執筆者



城田 真琴

株式会社野村総合研究所（NRI）

DX基盤事業本部付 兼

未来創発センター デジタル社会研究室 プリンシパル・アナリスト

2023年6月発行

NRI 未来創発センター

NRI 未来創発センターは、「未来志向型シンクタンク」を目指しています。各領域で高い専門性を有したメンバーが、日本・世界が直面する社会課題・経済課題を洞察し、科学的な判断に基づき、その処方箋を提言・発信していきます。

<https://www.nri.com/jp/service/souhatsu>

E-mail : miraisouhatsu-report@nri.co.jp

Envision the value,
Empower the change

株式会社野村総合研究所 未来創発センター

〒100-0004 東京都千代田区大手町 1-9-2 大手町フィナンシャルシティ グランキューブ