

2023年5月19日

金融機関は ChatGPT にどう対処すべきか

—対話型 AI の技術概説、金融機関の大規模言語モデル(LLM)の活用動向、将来展望まで—

先端技術ラボ 間瀬英之、大沼俊輔

《要 点》

- ◆ ChatGPT は、2022年11月に米 OpenAI が発表した大規模言語モデル、または対話型 AI サービスのこと。生成 AI の一種で、従来よりも自然な言語理解や言語生成を行い、対話形式での指示出しや質問応答ができる。発表から、わずか 2 カ月ほどで月間アクティブユーザ数が 1 億を超え、大きな注目を浴びた。後継モデルでは、専門分野やアカデミックのベンチマークで人間を上回る性能を示した。
- ◆ 爆発的に広がった背景としては、①自然な対話が可能インタフェース、②多言語対応、③マルチモーダル AI の進展、の 3 点が考えられる。一方で、ChatGPT にはリスクや課題が存在し、目下、研究開発や対策が進められている。
- ◆ ChatGPT や大規模言語モデルによって、文章の添削や校正、要約、アイデアの提案、ハイライト抽出といった業務の効率化、精度向上が期待されている。金融機関においては、情報漏洩などの側面から ChatGPT 自体を直接、サービス利用するのではなく、IT 企業が提供する専用サービスの利用や自社専用の大規模言語モデル構築、およびそのモデルによる生成 AI の活用を模索している。
- ◆ ChatGPT を取り巻く将来展望としては、①大規模言語モデルの拡張、②人間社会への影響に関する研究の進展、③AGI (汎用人工知能) を目指した研究の進展、が想定される。今後は、複数の機能を組み合わせた高度な機能を持つ AI の社会実装が進んでいくであろう。
- ◆ 金融業務は高い公共性や社会的責任の大きさから、取引の透明性や公平性、説明責任などが求められる。そのため、①技術的な特徴の把握、②現業務の見直し、③AI 利用ガイドラインの策定・アップデート、といった点に留意しながら、ChatGPT や大規模言語モデルの活用によるユースケース探索や利活用を進めていくことが肝要である。

本件に関するご照会は、先端技術ラボ 間瀬英之、大沼俊輔 宛にお願いいたします。

Mail : 101360-advanced_tech@ml.jri.co.jp

日本総研・先端技術ラボについては、以下をご覧ください。

<https://www.jri.co.jp/company/business/system/advtechlab/>

本資料は、情報提供を目的に作成されたものであり、何らかの取引を誘引することを目的としたものではありません。本資料は、作成日時時点で弊社が一般に信頼出来ると思われる資料に基づいて作成されたものですが、情報の正確性・完全性を保証するものではありません。また、情報の内容は、経済情勢等の変化により変更されることがあります。本資料の情報に基づき起因してご閲覧者様及び第三者に損害が発生したとしても執筆者、執筆にあたっての取材先及び弊社は一切責任を負わないものとします。

1. ChatGPT とは何か。トレンドの背景と技術的課題

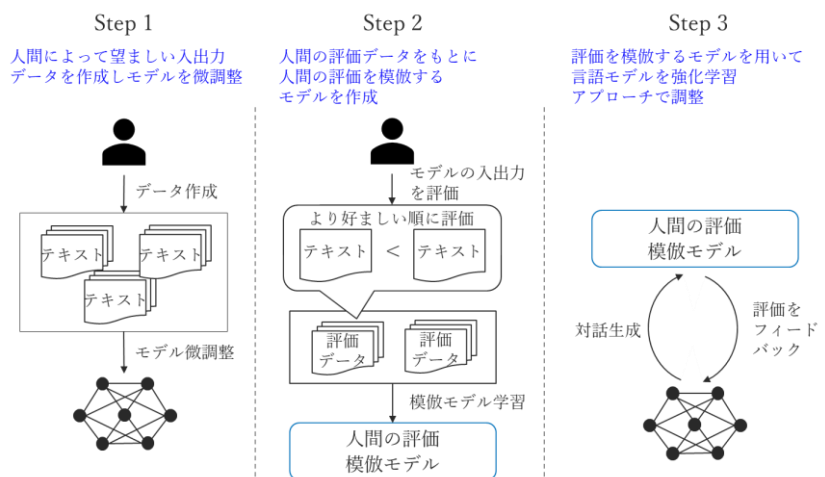
(1) ChatGPT とは何か

「ChatGPT (Generative Pre-trained Transformer)」は、2022 年 11 月に米 OpenAI が発表した大規模言語モデル (Large Language Model; LLM、人間が用いる言語を確率的に計算するモデルで、特にパラメータ数が膨大な深層学習モデルを指す)、または、そのモデルを用いた対話型 AI サービスを指す¹。生成 AI (画像、テキスト、音声などのコンテンツを生成する AI) の一種で、従来よりも自然な言語理解や言語生成を行い、対話形式での指示出しや質問応答ができる。OpenAI の発表から、わずか 2 カ月ほどで月間アクティブユーザー数が 1 億を超えたと推定されている²。2023 年 3 月には後継モデルとなる「GPT-4」が発表された。GPT-4 は、さまざまな専門分野やアカデミックなベンチマークで人間を上回る性能を示し、例えば、米国司法試験の模擬試験では全受験者の上位 10%以内の成績を達成した、と報告されている (一方、現実的な設定のタスクにおいて、まだ人間に勝るものではない点は留意)³。

ChatGPT に関する技術的な詳細は公開されていないが、その前身は OpenAI が 2020 年 7 月に発表した「GPT-3」というモデルである。GPT-3 は、ウェブ上の大量のデータを基に学習された 1,750 億パラメータを持つ大規模言語モデルで、さまざまな言語処理タスク (質問応答、翻訳など) を自然言語による指示や少数のサンプルデータを参考にして、ある程度の正確さで解けることが挙げられる⁴。GPT-3 (他の大規模言語モデルも同様) の仕組みは非常にシンプルで、簡単に言えば、ある文を途中まで受け取り、その次の単語を予測するような学習をしている。GPT-3 は、ある程度の正確さでタスクを解ける一方で、不正確な表現や有害な表現などの、人間の入力意図に沿わない出力生成が問題視されていた。OpenAI はこの点を改善するために、2022 年 1 月に「InstructGPT」と呼ばれるモデルを開発した。

InstructGPT は、GPT-3 で見られた有害な出力や人間の意図に沿わない出力を抑制するために、従来行われていた正解データを基にした教師あり学習⁵に加え、「RLHF (Reinforcement Learning from Human Feedback)」と呼ばれる人間の評価情報を活用してモデルを調整する強化学習⁶アプローチを取り入れた。データ作成者が人間にとって望ましい関係の入出力データを作成し、人間の評価に合わせてモデルを調整⁷することで、モデルは人間にとって好ましい出力や、事実性の向上、有害な出力の抑制が可能となった。ChatGPT は、InstructGPT の学習プロセスにおいて、対話データを追加して発展させた対話ベースのモデルである (図表 1)。

(図表 1) ChatGPT の学習概要



¹ ChatGPT の学習では、学習データ (上記のテキスト、評価データ) に対話データが追加されている。

² Step 2 と Step 3 の学習方法を RLHF (Reinforcement Learning from Human Feedback) と呼ぶ。

(出所) OpenAI Introducing ChatGPT を参考に日本総研作成

ChatGPT の登場以前から、大規模言語モデルやそれを利用した対話形式のアプリケーション開発は行われてきたが、ChatGPT の爆発的な普及を契機に、世界中で製品・サービスの発表が相次ぐようになった。例えば、2023 年 2 月には、米 Google が対話型 AI サービス「Bard」^{8,9}を、米 Meta (旧 Facebook) は「LLaMA」という比較的軽量のモデルを発表した¹⁰。また、中国では Baidu が 2023 年 3 月に対話型 AI サービス「文心一言 (ERNIE Bot)」を¹¹、Alibaba は 2023 年 4 月に対話型 AI サービス「通義千問 (Tongyi Qianwen)」を発表した¹²。なお、いずれのサービスも詳細な仕様は基本的に公開されておらず、言語モデルの原理自体も深層学習であるため、現時点で各モデル間の技術的な違いや性能比較を示すことは難しい。

(2) ChatGPT が爆発的に普及した背景

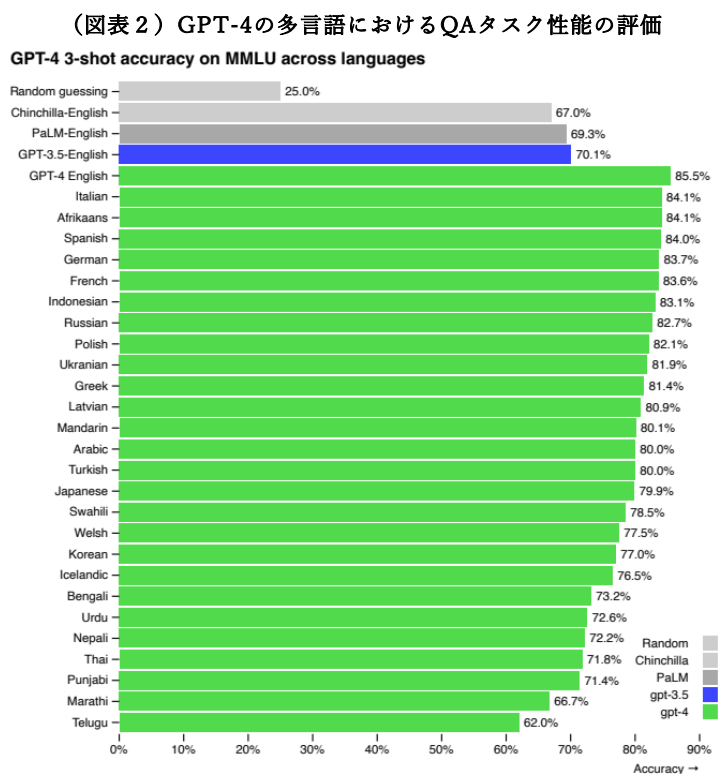
では、ChatGPT の利用は、なぜこれほどまでに爆発的に広がったのであろうか。その背景としては、①自然な対話が可能なインタフェース、②多言語対応、③マルチモーダル AI の進展、の 3 点があったと考えられる。

① 自然な対話が可能なインタフェース

ChatGPT 登場以前の GPT-3 や InstructGPT は、言語処理タスク（質問応答、翻訳など）をある程度の正確さで解くことができたが、主に開発者向けのモデルであり、一般ユーザが直接利用しやすいものではなかった。一方で、ChatGPT は対話型のモデルとして公開され、その対話が人間同士の会話のような自然なものであり、かつ多様な処理を自然言語で行えることから注目が集まり、爆発的に利用が増加したと考えられる。

② 多言語対応

ChatGPT の学習データの大半は英語であるが、英語以外の言語も含まれている。英語以外の言語でも、ある程度の正確な会話ができただことは、ChatGPT の利用者増加の一因と考えられる。日本語も使うことができるので、この点は日本国内においても同様である¹³。OpenAI の GPT-4 に関する報告¹⁴では、各言語における一部タスクの評価がされている（図表 2）。図表 2 における青いバーは、英語における GPT-3.5¹⁵の性能を示しており、緑バーは各言語に翻訳したデータセットにおける GPT-4 の性能が示されている。学習データ量が少ない言語（ラトビア語、ウェールズ語など）においても、GPT-4 は比較的高い性能を示している。



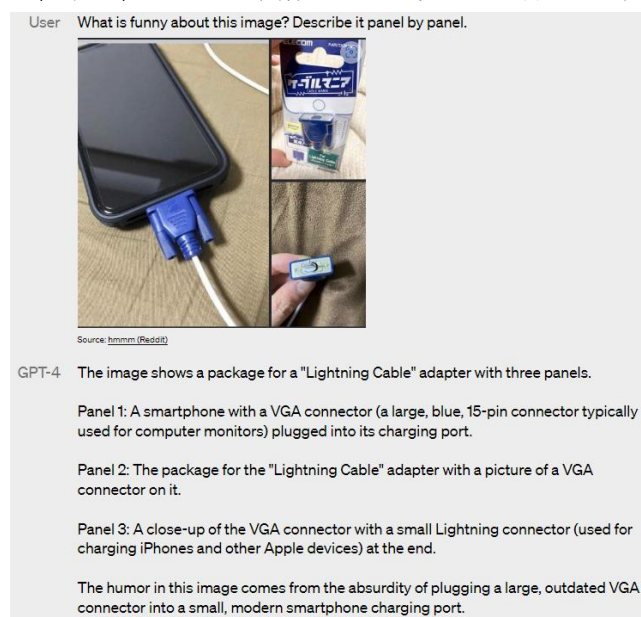
(図出所) OpenAI, GPT-4 Technical Report (<https://cdn.openai.com/papers/gpt-4.pdf>)

※閲覧日：2023年5月17日

③ マルチモーダル AI の進展

近年のマルチモーダル AI（画像、テキスト、音声など複数種類のデータを組み合わせて処理をする AI）の進展の影響も考えられる。2022年に、Stable Diffusion¹⁶などのテキストを入力として画像を生成するモデルが話題となった。これら画像を生成する AI の進展と ChatGPT の話題性が相まって生成 AI ブームにつながったことも、ChatGPT が爆発的に広がった一因と考えられる。2023年3月に発表された ChatGPT の後継モデルである GPT-4 では、画像入力が可能となっている（現時点で一般公開はされていない）³。例えば、図表3にあるように、GPT-4 に「なぜこの写真が面白いのか」を説明させることができるかとされている。

（図表3）GPT-4に画像入力して状況を説明させる例



（図出所）OpenAI, GPT-4 (<https://openai.com/research/gpt-4>)

※閲覧日：2023年5月17日

（3） ChatGPT が抱えるリスクと課題

現在の AI 技術全般と同様に、ChatGPT は万能な AI ではない。以下に、ChatGPT や関連技術におけるリスクや課題の代表例を示す。

① 入力内容の漏洩

ChatGPT にユーザが入力した内容は、OpenAI が学習データとして利用する可能性があるため、個人情報や機密性の高い内容を入力すると、不特定多数にその内容が流出する恐れがある。（2023年4月に、OpenAI は ChatGPT のチャット履歴を無効にする機能を追加。無効状態で開始された会話は、モデルのトレーニングや改善には用いられない）¹⁷

② 事実と異なる、不正確な出力

この現象は「幻覚 (Hallucination)」と呼ばれ、ChatGPT を含む生成用途の言語モデルの課題である。加えて、ChatGPT や GPT-4 の場合は、事実に基づかないような内容であっても説得力があるように出力をする。そのため、ユーザを過度に信頼させることがあり、結果として利用者の不利益につながる可能性がある¹⁴。また、言語モデルを学習した時点のデータより新しい情報は含まれていないため、利用時点では正確な情報が得られないことがある。

③ ユーザの悪意ある入力により有害な内容を出力

ユーザの入力に対して返答をする性質から、ユーザの悪意ある入力に従って有害な内容を出力してしまうことがある。例えば、入力を工夫することで、マルウェアの作成を支援するような内容の出力をさせるなど、サイバー犯罪者に悪用されている懸念がある¹⁸。

④ バイアスやステレオタイプが再生産

言語モデルは、大量のテキストデータから学習されているため、ある時点までの学習データに含まれるバイアスやステレオタイプが再生産されてしまう懸念がある。例えば、テキストデータに含まれるジェンダー・人種的なバイアスが学習データとして利用されると、そのバイアスが解消されないまま固定されてしまう可能性がある。ある時点の常識や文化なども同様である。そのため、政治や医療、法律関係など、社会的に重要な意思決定については慎重な利用検討が必要である¹⁴。

2. 金融機関の取り組み動向

(1) 金融機関におけるユースケース

ChatGPT を含む大規模言語モデルによって、主にテキストを入力とした文章の添削や校正、要約、アイデアの提案、ハイライト抽出といったさまざまなタスクの効率化、精度向上が期待されている。そのため、金融業務だけでなく、通常の業務にも利用できると考えられる（図表4）。

（図表4）金融機関の大規模言語モデル活用のユースケース例

	ユースケース	概要
金融業務	顧客サービス	顧客の問い合わせ、苦情などに迅速・効率的に対応
	ローン組成	顧客データの収集、申請処理などのステップ自動化
	ウェルスマネジメント	顧客データ分析、パーソナライズされた投資推奨
	ファイナンシャルプランニング	予算編成、退職計画など、パーソナライズされた財務計画
	顧客のオンボーディング	新規口座の開設などの記入支援、プロセスの自動化
	リスクマネジメント	マーケットレポート等の分析、経済・政治的リスクの予測
	バーチャルアシスタント	24時間365日利用できる仮想アシスタントの提供
共通業務	マーケティング／営業	商品広告の自動生成、顧客フィードバックの分析や要約
	オペレーション	社内照会に対する回答、文章校正、プロセス自動化
	IT／エンジニアリング	効率的なソースコードへの変換、プログラミング言語変換
	リスク／リーガル	リーガル文書の草稿作成、文書要約、照会回答
	HR	面接の質問リスト作成、採用者のオンボーディング効率化
	社内業務	メール文の作成、プレゼンテーション資料作成、文書要約

（出所）日本総研作成

前節で示したリスクや課題が解消することを前提とするが、具体的なユースケースとしては、例えば、顧客からの問い合わせ対応や、口座開設時の記入支援の活用が想定される。他にも、ファイナンシャルプランニングにおいて、顧客のヒアリングデータなどを大規模言語モデルで分析し、分析結果を顧客の財務データなどと組み合わせることで、顧客に最適なソリューションの提案ができる可能性がある。金融業務でなくても、マーケティングでの広告キャッチコピーの作成や、IT 部門でのソースコードの作成、社内業務での従業員の返信メールやプレゼンテーション資料の作成などにおいて、作業の効率化・精度向上が見込める。

米マイクロソフトが2023年3月に開催したイベント「The Future of Work with AI」¹⁹では、大規模言語モデルをベースに、「Microsoft Graph」や「Microsoft 365」アプリ内のデータを組み合わせた「Microsoft 365 Copilot」のデモが行われた。例えば、OneNoteの顧客メモと別の社内文書から顧客へのメールの草稿を作成する、売上データから主要な傾向を示すようにグラフを作成する、といった作業が、ボタンのクリックとAIとの対話だけで完了する様子が公開されている。現在、レビュー公開は一部企業のみとなっているが²⁰、今後の本格展開が始まれば、金融機関を含むさまざまな企業で活用が広がると予想される。

(2) 金融機関の取り組み事例

では、ChatGPT を含む大規模言語モデルに対する金融機関の取り組みはどうか。

ChatGPT に限定すれば、金融機関は顧客情報や機密情報の漏洩、出力結果の正確性や説明性に対

する懸念等から従業員の ChatGPT 利用を制限している。実際、国内3メガバンクや欧米大手行（米 JP モルガン・チェース、米シティグループ、米バンク・オブ・アメリカ、米ゴールドマン・サックス、米ウェルズ・ファー

(図表5) 大手金融機関の主な取り組み

名称	国	概要
JPモルガン・チェース	米国	FRBの見解25年分をLLMを利用して、タカ派ハト派で数値化
モルガン・スタンレー		ChatGPTと連携して、財務アドバイザーに必要な回答を提示
ゴールドマン・サックス		コード作成支援の生成AIツールを内部で試験中
EQT ¹⁾	欧州	ChatGPTを利用して、社内データベース検索を効率化
中国工商銀行	中国	LLMを顧客の電話応答に適用
中国郵政儲蓄銀行		百度のLLMを用いて、マーケティングや資産管理等へ適用検討
中国農業銀行		金融AI大規模言語モデルを自社開発し、社内で試験中
三井住友フィナンシャルグループ	日本	自社専用環境で動作する対話型AIを行内業務に導入予定
三菱UFJフィナンシャル・グループ		自社専用環境で動作する対話型AIを行内業務に導入予定
オーストラリア・コモンウェルス	豪州	対話型AIのユースケースを模索
オーストラリア・ニュージーランド		対話型AIのユースケースを模索
ナショナルオーストラリア		生成AIによるコード作成支援を模索
ウエストバック		生成AIによる顧客へ手紙執筆やコード作成支援の利用を検討

¹⁾ スウェーデン・ストックホルムに本社を置く北欧最大のプライベート・エクイティ・ファンド
(出所) 公開情報を基に日本総研作成

ゴ、ドイツ銀行)などは従業員の ChatGPT 利用制限を発表している。そのため、金融機関は ChatGPT 自体を直接、サービス利用するのではなく、IT 企業が提供する専用サービスの利用や自社専用の大規模言語モデル構築、およびそのモデルによる生成 AI の活用を模索している (図表5)。

例えば、三井住友フィナンシャルグループは、日本マイクロソフトのクラウドサービス「Azure OpenAI Service²¹⁾」を用いて、SMBC グループ専用環境上で動作する AI アシスタントツールの開発を計画している²²⁾。SMBC グループの従業員のみが利用でき、文章の作成、要約、翻訳、ソースコードの生成などをチャット形式で手軽に行うことができるとされる。米モルガン・スタンレーでは、財務アドバイザー向けの利用を検討している²³⁾。同社は投資戦略、市場調査、アナリストの洞察などのコンテンツを含む調査レポートを数十万ページ保持しているが、主に PDF 形式で社内ポータルサイトに格納されているため、財務アドバイザーが特定の質問に対する回答を見つけるには多くの時間を要する²⁴⁾。同社は発行した調査レポートのうち、精査された約 10 万件を大規模言語モデルで学習させて、その質問に即座に回答できるようにすることで、財務アドバイザーの業務効率化、洞察の強化を図る計画である²⁵⁾。また、米ゴールドマン・サックスは、社内エンジニア向けに、ソースコード作成支援の生成 AI ツールを社内でテストしている。同社 CIO はケースによっては、ソースコードの 40%程度が生成 AI によって自動生成できたと述べている²⁶⁾。他にも、米 JP モルガン・チェースは、過去 25 年分の FOMC 声明や当局者の講演内容について、対話型 AI の「ChatGPT」ベースの大規模言語モデルを利用して政策シグナルを検出し、ハト派・タカ派スコアで緩和的か引き締め的かを効果的に評価する仕組みを考案した。なお、米連邦準備銀行²⁷⁾やフロリダ大学²⁸⁾の研究者も似たような研究を実施している。このような ChatGPT や大規模言語モデルを利用する動きは、中国（中国工商銀行、中国郵政儲蓄銀行、興業銀行²⁹⁾、中国農業銀行³⁰⁾）やオーストラリア（オーストラリア・コモンウェルス、オーストラリア・ニュージーランド³¹⁾、ナショナルオーストラリア銀行³²⁾、ウエストバック³³⁾）の金融機関でも見受けられる。

ChatGPT に関する取り組みは、大手金融機関に限らない。ネオバンク³⁴⁾最大手の米 Chime は、ChatGPT や Google Bard を採用せず、自社のソースコードを学習データとして、独自バージョンの生成 AI の開発を進めている³⁵⁾。スウェーデンの BNPL フィンテック企業 Klarna は、一部地域を対象に、ChatGPT のプラグイン機能（サードパーティーのアプリケーションとの連携が可能）³⁶⁾を用いた、商品レコメンデーション機能の提供を開始した³⁷⁾。例えば、「150 ドルありますが、どのヘッドフォンを購入できますか？」と質問すると、Klarna の小売りパートナー50 万店舗から、リクエ

ストにあった商品のレコメンデーションがされる(図表6)。その他、「ユニコーンが好きな姪の誕生日に何をかうべきですか」と質問すれば、ユニコーンをテーマにしたプレゼントアイデアが表示される。これにより、従来のキーワード単位の検索方法から、より自然で直感的な商品検索、レコメンデーションが可能とされる。

また、金融機関ではないが、経済メディアの米ブルームバーグは、米ジョンズ・ホプキンス大学の研究者

らと「BloombergGPT: A Large Language Model for Finance」³⁸という論文を発表している。幅広い金融データで学習させた500億のパラメータの大規模言語モデルで、例えば「アップルの直近の株価と時価総額を教えてください」と入力すると、簡潔に回答を提示してくれるという。

3. 今後の展望

(1) ChatGPT を取り巻く将来展望

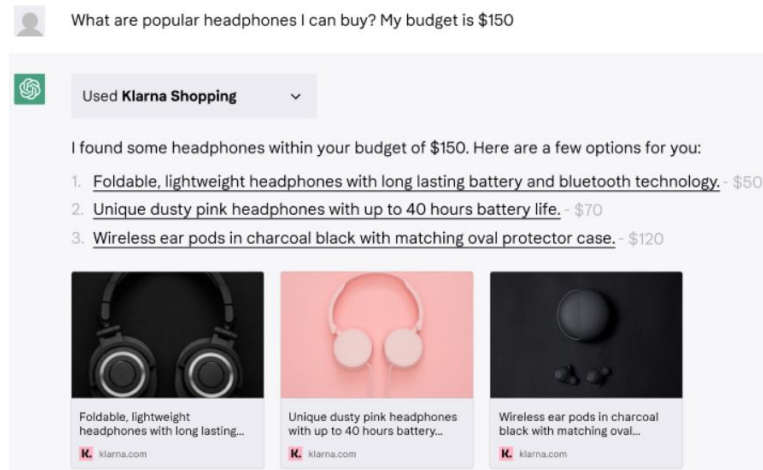
ChatGPT に関する世の中の動きは想像以上に速く、現時点で、将来を見通すことは難しい。さまざまな将来展望が考えられるが、以下、我々の想定する将来展望を3つ述べる。

① 大規模言語モデルの拡張

現在の ChatGPT を含めた大規模言語モデルでは、入力できるテキスト長の制限やモデルの動作原理から、入力情報を長期間にわたって記憶することや、大量の文書を一度に処理することが難しい。そのため、利用者が ChatGPT などの大規模言語モデルを単体で用いるのではなく、他のツールやシステムと組み合わせて利用する手法の開発が進むと予想される。例えば、長期間の記録は、既存のデータベース機能と組み合わせることにより実現ができる。大量の文書に対する検索ならば、検索エンジンと組み合わせることで実現が可能である。

このような周辺ツールとの接続によって、より柔軟に ChatGPT などの大規模言語モデルが IT システムへ組み込まれ、結果、新たなサービス創造などにつながると考えられる(図表7)。

(図表6) 予算\$150内のヘッドフォン製品のレコメンド

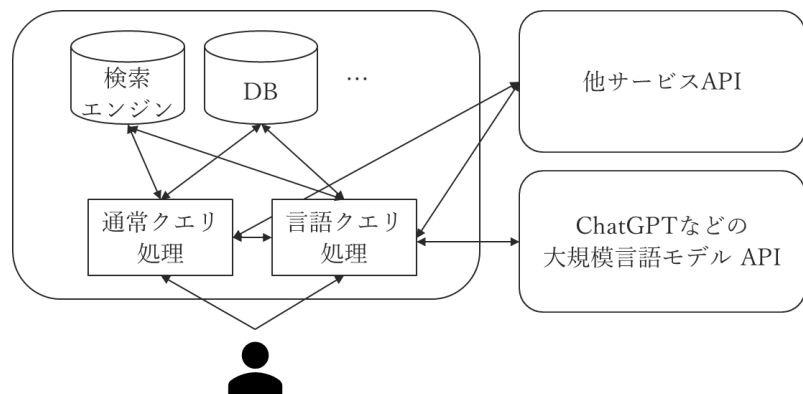


(図出所) Klarna, Klarna brings smooth shopping to ChatGPT

(<https://www.klarna.com/international/press/klarna-brings-smooth-shopping-to-chatgpt/>)

※閲覧日: 2023年5月17日

(図表7) 大規模言語モデルをシステムに組み込む概念図



(出所) 日本総研作成

② 人間社会への影響に関する研究の進展

ChatGPT が抱えるリスクと課題で述べた通り、ChatGPT は幻覚 (Hallucination) と呼ばれる課題や、事実に基づかないような内容であっても説得力があるように出力する可能性がある。そのため、利用者が誤った情報やバイアスのある出力を信用してしまう悪影響が考えられる¹⁴。

また、機械学習モデルは学習した内容を基に動作するため、バイアスやステレオタイプの再生産の問題がある。学習したモデルに含まれる特定の知識だけを修正することは現状では難しいが、今後は技術的な研究だけでなく、人間や社会に対する影響・緩和策についての研究も進展していくと予想される。

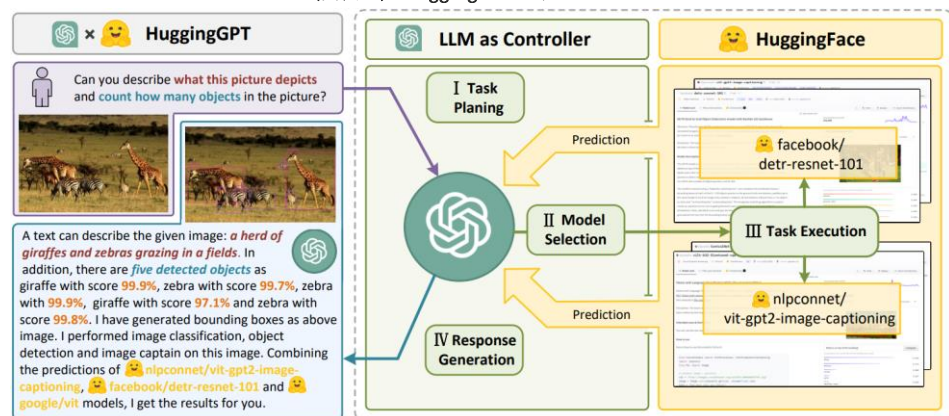
③ AGI (汎用人工知能) を目指した研究の進展

米マイクロソフトは 2023 年 4 月に、「Sparks of Artificial General Intelligence: Early experiments with GPT-4」という論文³⁹を発表した。Artificial General Intelligence (汎用人工知能、以下 AGI という) は、人間のようにさまざまな状況や問題に対応できる汎用性を持つ AI のことで、これまで実現されてきたような特化型の AI の対極にある概念である。この論文では、GPT-4 の初期バージョン⁴⁰について、知的能力 (人間と同等かそれ以上の推論や計画策定、経験から学ぶ能力を含む幅広い能力) に焦点を当てて検証を行っている。大規模言語モデルなどに共通する課題は未だにあるものの、GPT-4 の能力は、AGI の実現に向けた進歩であると主張している。

この論文で検証の焦点となっている計画策定が可能となると、複雑なタスクについてモデルが自動で解決方針を立てて処理を実行することが期待できる。例えば、モデルが入力された目標をタスクに分解して、ユーザに確認を求めつつ半自動でタスクをこなしていく「Auto-GPT」というアプリケーションが登場している⁴¹。また「HuggingGPT」という、与えたタスクを複数の機械学習モデルを組み合わせて自動で解くためのシステムも提案されている (図表 8)⁴²。図表 8 (左) では、ユーザが「画像の内容を説明し、映っている物体の数を数えて」という指示を画像と共に入力している。その指示に HuggingGPT が「野原で草を食べているキリンとシマウマの群れ」と答え、続けて 5 つの物体を検知したことを出力している。このシステムは、ユーザの指示入力を受け取り、ChatGPT を用いて実行できるタ

スクに分解する。分解したタスクを実行するため、Hugging Face Hub⁴³上で提供されている機械学習モデルを選択して実行する。最後に各タスクの実行結果を統合して回答を生成している。

(図表 8) HuggingGPT の処理イメージ



(図出所) Yongliang Shen, et al., HuggingGPT: Solving AI Tasks with ChatGPT and its Friends in Hugging Face (<https://arxiv.org/pdf/2303.17580.pdf>)

※閲覧日: 2023年5月17日

他の AGI の研究課題として、人間と同様に現実世界での動作や学習がある⁴⁴。その実現には言語処理以外に、画像処理や映像処理技術が重要であり、それらを組み合わせたマルチモーダル AI 研究が期待される。例えば Google Research と Everyday Robots が大規模言語モデルを実装したヘルパ

ーロボット「PaLM-SayCan」の記事を公開している⁴⁵。このヘルパーロボットは限られたシチュエーションにおいて、自然言語による指示を受け、可能な行動を組み合わせることで現実世界の物体を操作することが可能である。例えば、「飲み物と軽食が欲しい」という指示を受けて内容を解釈し、飲み物を利用者に届けるため、「水の入っているボトルを見つける」、「ボトルを持ち上げる」、といった動作を実行する。また、次に、軽食のために「リンゴを見つける」、「リンゴを持ち上げる」といった動作を実行する。

AGI 実現に向けたモデルが統一的なモデルになるか、それとも個別のモデルが協調する形式になるか、現時点では見通すことは難しいが、既に開発されている複数の機能を連携・統合させることで、より高度な機能が実装されていくと考えられる。

(2) 金融機関は ChatGPT にどう対処すべきか

金融業務は高い公共性や社会的責任の大きさから、取引の透明性や公平性、説明責任などが求められる。そのため、ChatGPT や類似サービスが抱えるリスクや課題を正しく認識すること、ChatGPT だけでなく他の大規模言語モデルや生成 AI の活用も含めたユースケース探索や利活用を進めていくことが肝要である。また、現業務の見直し、AI 規制の対応といった観点での対応も必要となるであろう。以下、対処法を3つ述べる。

① 技術的な特徴の把握

従業員は、ChatGPT や類似サービスの利用にあたり、技術的な特徴を把握しておくことが必要である。例えば、ChatGPT は事実と異なる内容を出力することがあるため、ある調査に ChatGPT を利用した際には、別の資料を用いて裏付けをとる必要がある。文書作成の作業が容易になるための活用を検討することも大事であるが、作成された文書の信頼性担保の仕組みも合わせて検討することが望ましい。また、機械学習の原理上、専門性の高い領域ほどデータ量が少ないことから、金融などの領域ではモデルの出力は正確でない傾向があると考えられる。そのため、ChatGPT などの大規模言語モデルをそのまま利用するほか、対象ドメインのデータを用いて微調整することも考えられる。ChatGPT に入力する際に、指示を詳細に書くことで精度が上がる傾向があるとも言われており、ChatGPT にタスクを行わせる前提で業務知識を文書化していくことも重要である。

その他にも、出力結果のバイアスやステレオタイプの問題では、OpenAI などのベンダー企業がモデルの出力からバイアス等を除くようにさまざまな処理を検討しているが、サービス提供者やモデル作成者が認識しているバイアスと利用者の考えるバイアスが異なる可能性がある。そのため、利用に際しては、自社のユースケースに沿って、適切な処理を検討する必要がある。なお、言語モデルのバイアスの確認や修正は、技術的にも困難な課題であり、現状では注意深く利用する必要がある。

利用サイドとしては、ChatGPT や類似サービスをまず使ってみることも重要であるが、前提として、抱えるリスクや課題を正しく理解したうえで、ユースケース探索や実証を進めていく必要がある。現時点では、技術的に解決していない課題もあるので、まずは社内に閉じた形での実証実験を推奨する。

② 現業務の見直し

2023年3月、OpenAI と米ペンシルバニア大学は「GPTs are GPTs: An early look at the labor market impact potential of large language models」と題した論文を公開した⁴⁶。この調査では、「米国

の労働力の約 80%が GPT の導入によって、作業の少なくとも 10%が影響を受ける可能性があり、約 19%の労働者においては少なくとも 50%のタスクが GPT 導入による影響を受ける可能性がある」と報告されている。1,016 の職業を対象に調査⁴⁷したところ、プログラミングと執筆のスキルは大規模言語モデルの影響を受けやすく、科学的、批判的思考スキルを必要とする職業は影響を受けにくいという。このように多分野で一部の仕事が AI によって影響を受けると予想されている。

金融機関も例外ではない。ChatGPT などの大規模言語モデルを活用することで、業務の効率化や精度向上などが期待できるが、現在の仕事の仕方、仕組みにも影響を与えるであろう。例えば、リスクマネジメント業務では、マーケット分析結果のレポート草案の作成や文書要約といった場面で、ChatGPT などの大規模言語モデルを活用が想定される。その際には、ChatGPT が他のツールやシステムと連携をとる、ChatGPT にマルチモーダルな入力（図、数値、表テーブルなど）をする、といった前提で業務の見直しが必要であろう。また、ChatGPT を有効活用するには、DX 推進の業務改善のように、業務をタスク単位に分解し、AI に向いているタスクの処理内容を言語化・実行、その出力内容の正誤や妥当性を確認する、というステップが必要となるであろう。また、金融業務の場合は、その業務の特性上、人間が AI の出力内容を確認し、業務上の判断を下すことの重要性や機会が増す可能性がある。基本的な業務が AI に置き換わってしまう可能性がある中、どのように業務知識や経験を積ませ、重要な判断を的確に下せるようにするのか、人材育成の観点でも検討が必要であろう。

さらに、中長期の目線でみれば、ChatGPT の登場は「人間にしかできない業務とは本質的に何か？」ということを人間に問う機会にもなっている。例えば、ウェブや社内文書等の関連文書から単に文書要約するレポート作成であれば、将来的には、AI に完全に置き換わってしまう可能性があるだろう。その際には、当事者が経験した生の体験や独特な着眼点、発想などがレポートの価値になる（もちろん、これは現時点でも重要である）。これを機に、自業務にとって本当に価値のある業務とは何かを、関係メンバーなどで考え、準備・検討しておくことも重要となるであろう。

③ AI 利用ガイドラインの策定・アップデート

AI については、2016 年以降、IT 全般の課題（セキュリティ・プライバシー等）や AI 固有のリスク（透明性、公平性等）を中心に、国際的に AI 倫理原則の議論がなされてきた。そして、今回の ChatGPT の急速な普及を受け、透明性の促進や偽情報への対処、著作権の保護などの生成 AI の利活用に関する規制・ガイドライン策定の動きが強まっている（図表 9）。

特に、EU は新たに法律による厳格な規制案の導入を検討している。欧州委員会のベステアール上級副委員長は、AI が作成した文章や画像に「メード・ウィズ AI」といったラベルを表示させる規制案を提起し、生成 AI の提供企業に説明義務を課す考え

(図表 9) 生成AI関連の各国・地域の対応状況

国・地域		対応状況
欧州	EU	・EUのデータ保護委員会がChatGPT対策タスクフォースを立ち上げ ・生成AIに関する規制新法を検討し、年内の合意を目指す
	イタリア	・データ保護の観点からChatGPTが一時使用禁止（4/29解除）
	フランス	・個人情報の無断収集、使用等の苦情を受け、データ保護当局がOpenAIを調査
	イギリス	・データ保護の当局がChatGPTを含む生成AIの使用や開発に関する留意点を公表
北米	米国	・商務省がAIの説明責任に関するパブリックコメント募集を開始 ・バイデン大統領はAIリスクを巡り、マイクロソフトやグーグル首脳らと会合
	カナダ	・個人情報保護委員会が、個人情報の無断収集、使用、開示の苦情を受け、OpenAIを調査
アジア	中国	・生成系AIに対する規制案の草案を発表。社会主義体制の転覆などを扇動する内容を禁止し、当局による事前審査を義務付け
	日本	・利用に前向き。開発状況や国際的な動向を見極め、規制の必要性を判断 ・国際的なルール作りに関して、G7広島サミットの議題に

(出所) 公開情報を基に日本総研作成

を明らかにしている⁴⁸。一方で、米国や日本のように、厳格な規制ではなく、柔軟かつ慎重にガイドラインや規制を定めていく方向性もあり、各国で規制の足並みは揃っていない。

生成 AI に限った話ではないが、AI に関する法規制はまだ実運用に至っておらず、今後も、各国の標準・ガイドライン等が乱立する可能性は排除できない。では、ただ静観すればよいかというと、そうではない。AI を社会で安全安心に利用してもらうという本来の目的に振り返れば、先進企業・機関にならって、現時点から、技術面の実態把握、事例調査等の取り組みを開始しておくことが必要である。金融機関の場合は、金融商品の販売時などに顧客への説明責任といった AI の公平性・説明可能性⁴⁹が求められる。また、AI 技術は本格導入や運用開始と同時に、入力データの傾向や外部環境変化により陳腐化（運用時のモデル精度の低下）が始まる。そのため、AI 全般に関するフレームワークや原則⁵⁰に加えて、ChatGPT や生成 AI における法的な留意点（著作権侵害、名誉棄損等）⁵¹を参考に、自社に適した生成 AI（ChatGPT 含む）利用ガイドラインの策定や、自社の AI 関連のポリシーや手順書の継続的なアップデートが重要となるであろう。

執筆者

間瀬英之（Hideyuki Mase）

2014 年(株)日本総合研究所入社。三井住友銀行のシステム企画・開発業務等を経て、18 年から現職。IT 全般および量子コンピュータの動向調査に従事。本レポートでは 2 章、3 章を担当。

大沼俊輔（Onuma Shunsuke）

2016 年(株)日本総合研究所入社。三井住友カードの基盤システム開発業務を経て、18 年から現職。自然言語処理の技術調査・検証に従事。本レポートでは 1 章、3 章を担当。

参考文献・補足

¹ OpenAI, Introducing ChatGPT (<https://openai.com/blog/chatgpt>)では ChatGPT は Model であると表現されているが、API を通じたサービスとも見なせる。

² Yahoo!finance, ChatGPT on track to surpass 100 million users faster than TikTok or Instagram: UBS (<https://finance.yahoo.com/news/chatgpt-on-track-to-surpass-100-million-users-faster-than-tiktok-or-instagram-ubs-214423357.html>)

³ OpenAI, GPT-4 (<https://openai.com/research/gpt-4>)

⁴ 自然言語による指示や少数のサンプルデータを参考にして解くことを、zero-shot 学習や few-shot 学習と呼ぶ。

⁵ 機械学習の一種。学習データに正解を与えて学習させる手法。

⁶ 機械学習の一種。試行錯誤をしながら価値を最大化するような選択を学習する手法。

⁷ OpenAI, Aligning language models to follow instructions (<https://openai.com/research/instruction-following>)

⁸ Google The Keyword, An important next step on our AI journey (<https://blog.google/technology/ai/bard-google-ai-search-updates/>)

⁹ 2023 年 5 月 11 日に新バージョンが出て、日本語対応が発表された。(Google Japan Blog, Bard が日本語に対応 (<https://japan.googleblog.com/2023/05/bard.html>))

¹⁰ Meta AI Blog, Introducing LLaMA: A foundational, 65-billion-parameter large language model (<https://ai.facebook.com/blog/large-language-model-llama-meta-ai/>)

¹¹ <https://jp.reuters.com/article/baidu-ai-ernie-idJPKBN2VI0KO>

¹² <https://jp.reuters.com/article/alibaba-ai-idJPKBN2W805M>

-
- ¹³ similarweb , ChatGPT のユーザ数が多い国別ランキングのトップ5は上から、アメリカ、インド、日本、コロンビア、カナダの順（2023年4月結果）（<https://www.similarweb.com/website/chat.openai.com/#geography>）
- ¹⁴ OpenAI, GPT-4 Technical Report（<https://cdn.openai.com/papers/gpt-4.pdf>）
- ¹⁵ GPT-3.5はGPT-3に対してデータの更新や学習方法が変更された更新版モデルを指す。（OpenAI, Model index for researchers（<https://platform.openai.com/docs/model-index-for-researchers>））
- ¹⁶ 英 Stability AI から公開されている text-to-image モデル。オープンソースソフトウェアとしてリリースされ、画像生成 AI における実験やサービス化の敷居を大きく下げた。
- ¹⁷ OpenAI, New ways to manage your data in ChatGPT（<https://openai.com/blog/new-ways-to-manage-your-data-in-chatgpt>）
- ¹⁸ Check Point Research, OPWNAI : CYBERCRIMINALS STARTING TO USE CHATGPT（<https://research.checkpoint.com/2023/opwnai-cybercriminals-starting-to-use-chatgpt/>）
- ¹⁹ Microsoft YouTube, The Future of Work With AI - Microsoft March 2023 Event（<https://www.youtube.com/watch?v=Bf-dbS9CcRU>）
- ²⁰ 2023年5月9日時点で、全世界で600社にプレビュー公開されている。（<https://blogs.microsoft.com/blog/2023/05/09/introducing-the-microsoft-365-copilot-early-access-program-and-2023-microsoft-work-trend-index/>）
- ²¹ OpenAI と同じモデルを実行しながら、Microsoft Azure のセキュリティ機能が利用可能。
- ²² 三井住友フィナンシャルグループ, SMBC グループの専用環境における AI アシスタントツール「SMBC-GPT」の実証実験の開始について（https://www.smfg.co.jp/news/j110416_01.html）
- ²³ Morgan Stanley, Morgan Stanley Wealth Management Announces Key Milestone in Innovation Journey with OpenAI（<https://www.morganstanley.com/press-releases/key-milestone-in-innovation-journey-with-openai>）
- ²⁴ OpenAI, Morgan Stanley wealth management deploys GPT-4 to organize its vast knowledge base.（<https://openai.com/customer-stories/morgan-stanley>）
- ²⁵ CNBC, Morgan Stanley is testing an OpenAI-powered chatbot for its 16,000 financial advisors（<https://www.cnbc.com/2023/03/14/morgan-stanley-testing-openai-powered-chatbot-for-its-financial-advisors.html>）
- ²⁶ CNBC, Goldman Sachs is using ChatGPT-style A.I. in house to assist developers with writing code（<https://www.cnbc.com/2023/03/22/goldman-sachs-experiments-with-chatgpt-like-ai-to-help-devs-write-code.html>）
- ²⁷ Anne Lundgaard Hansen and Sophia Kazinnik, Can ChatGPT Decipher FedSpeak?（https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4399406）
- ²⁸ Alejandro Lopez-Lira and Yuehua Tang, Can ChatGPT Forecast Stock Price Movements? Return Predictability and Large Language Models（https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4412788）
- ²⁹ 中国城乡金融报, 人工智能赋能银行业迈入新时代（https://h5.newaircloud.com/detailArticle/21475837_59709_cxjrb.html）
- ³⁰ Sohu , 发布 | 农业银行金融 AI 大模型产品 ChatABC 成功发布（https://www.sohu.com/a/661390379_672569）
- ³¹ itnews, CBA and ANZ explore conversational AI use cases（<https://www.itnews.com.au/news/cba-and-anz-explore-conversational-ai-use-cases-591286>）
- ³² The Australian Financial Review, How ChatGPT will change these four industries（<https://www.afr.com/work-and-careers/workplace/how-chatgpt-will-change-these-four-industries-20230328-p5cvzx>）
- ³³ The Australian Financial Review, Westpac exec warns AI bots may 'hallucinate'（<https://www.afr.com/companies/financial-services/westpac-and-aws-swim-deeper-into-generative-ai-for-banking-20230224-p5cne4>）
- ³⁴ 銀行免許を持たず、既存銀行との提携を通じて金融サービスを提供する企業のこと。
- ³⁵ Business Insider, Chime's top tech exec details how the fintech is building an AI chatbot to speed up how quickly it launches new products and features（<https://www.businessinsider.com/fintech-chime-to-launch-internal-chatgpt-ai-tool-this-year-2023-2>）
- ³⁶ OpenAI, ChatGPT plugins（<https://openai.com/blog/chatgpt-plugins>）
- ³⁷ Klarna, Klarna brings smooth shopping to ChatGPT（<https://www.klarna.com/international/press/klarna-brings-smooth-shopping-to-chatgpt/>）
- ³⁸ Shijie Wu et al., BloombergGPT: A Large Language Model for Finance（<https://arxiv.org/abs/2303.17564>）
- ³⁹ Sébastien Bubeck, et al., Sparks of Artificial General Intelligence: Early experiments with GPT-4（<https://arxiv.org/pdf/2303.12712.pdf>）
- ⁴⁰ 言語情報のみを扱う初期バージョンを対象としている。
- ⁴¹ Auto-GPT（<https://github.com/Significant-Gravitas/Auto-GPT>）
- ⁴² Yongliang Shen, et al., HuggingGPT: Solving AI Tasks with ChatGPT and its Friends in Hugging Face（<https://arxiv.org/pdf/2303.17580.pdf>）

-
- ⁴³ Hugging Face Hub は機械学習関連のモデルやツールを開発している Hugging Face 社が運営するプラットフォーム (Hugging Face (<https://huggingface.co/>))
- ⁴⁴ 荒川直哉ら, 汎用人工知能の研究動向 (<https://www.ai-gakkai.or.jp/jsai2014/webprogram/2014/pdf/174.pdf>)
- ⁴⁵ Google Research, Using language to better interact with helper robots. (<https://sites.research.google/palm-saycan>)
- ⁴⁶ OpenAI, GPTs are GPTs: An early look at the labor market impact potential of large language models (<https://openai.com/research/gpts-are-gpts>)
- ⁴⁷ 人間が特定のタスクを実行するのに必要な時間を、GPT-4 を使うことで少なくとも 50%短縮できるかどうか調査した。
- ⁴⁸ 日本経済新聞, EU、生成 AI に統一規制論 「メイド・ウィズ AI」表示案 (<https://www.nikkei.com/article/DGXZQOGR234M20T20C23A4000000/>)
- ⁴⁹ 日本総研, 【先端技術リサーチ】AI 公平性・説明可能 AI (XAI) の概説と動向 (<https://www.jri.co.jp/page.jsp?id=104040>)
- ⁵⁰ 参考例;米国標準技術研究所 (NIST) 「AI 管理フレームワーク」(<https://www.nist.gov/itl/ai-risk-management-framework>)、シンガポール金融管理局 (MAS) が主導する Veritas コンソーシアム「FEAT (Fairness, Ethics, Accountability, Transparency) の原則」
- ⁵¹ 参考例;JDLA (一般社団法人日本ディープラーニング協会), JDLA が、『生成 AI の利用ガイドライン』を公開 (<https://www.jdla.org/news/20230501001/>)