

MIT Technology Review

Published by KADOKAWA / ASCII

Vol.

61

2024.05

検証「生成AI革命」 破壊と創造は進んだか

News&Trend

「味の素」の牙城を狙う半導体ベンチャー
遺伝子編集ブタの腎臓を人間に移植

Interview

伊藤昌平 (フルデブス)

003

特集

検証「生成AI革命」 破壊と創造は進んだか

004

生成AI革命で残された
6つの重要課題

012

「Sora」登場で激震、
生成AIが変える映像制作の未来

018

誰もがオープンと言い出したAI業界、
オープンソースの定義を巡り攻防

023

生成AIアシスタントは
プログラマーの仕事をこう変えた

029

コンピューター科学者は
LLMをなぜ「科学」するのか

034

U35 イノベーターの軌跡 #13

伊藤昌平 (フルデプス)

社会課題に挑む水中ドローン開発者は「深海」を夢見る探求者

037

News&Trends

知られざる半導体材料の巨人「味の素」その牙城を狙う米スタートアップの勝算
ブタの腎臓を人間に移植、1カ所の遺伝子改変で拒絶反応を抑制
「おいしさ」をリバース・エンジニアリング、AI考案の植物性チーズ
低コストで即効性あり、飛行機のルート見直しが温暖化対策になる理由
羊水からオルガノイド作製に成功、胎児発達の謎解明へ

●本PDFに収録した記事の情報は原則として、初出時の情報です。記事中の初出日をご確認ください。

●WebサイトのURLやソフトウェアのバージョン等は予告なく変更されている場合があります。

●本PDFは情報の提供のみを目的としています。本PDFを運用した結果について、著者およびMIT Technology Review Japan/株式会社角川アスキー総合研究所は一切の責任を負いません。

●本PDFに登場する会社名、商品名は該当する各社の商標または登録商標です。本PDFでは®マークおよびTMマークの表示を省略しています。

検証「生成AI革命」 破壊と創造は進んだか

2022年末に登場した「チャットGPT (ChatGPT)」のブームは、世界に大きな衝撃をもたらした。大手テック企業は生成AI技術を組み込んだ製品を矢継ぎ早に市場投入。ホワイトカラーの生産性向上やクリエイティブ産業での活用が叫ばれ、停滞していたAIの社会実装は一気に進みつつあるように見える。一方で、訓練データをめぐる倫理的な問題やデマ・フェイクの問題、新たなセキュリティ・リスクなど、解決すべき重要な課題も山積している。生成AI革命の現在を多角的に捉え、その光と影、未来への展望を検証する。

Selman Design



Story

1

検証「生成AI革命」 破壊と創造は進んだか

生成AI革命で残された 6つの重要課題

2023年、生成AIは世界を席卷したが、未解決の問題も残っている。生成AIの未来、そして私たちの未来は、私たちが次にどうするかで決まってくるだろう。

by Will Douglas Heaven (米国版AI担当上級編集者)

2 023年のトレンドの変化の大きさを私に初めて教えてくれたのは、見ず知らずの人であった。3月のことだ。なかなか来ないエレベーターを待っていたとき、たまたま居合わせたある女性は、チャットGPT (ChatGPT) を使って、マーケティングの仕事の報告書を書く手伝いをしてもらったと話してくれた。彼女は、報告書の作成があまり得意でないと思っており、書くことが嫌いだ。だが、今回はマネージャーが褒めてくれたのだそうだ。「ズルをしている気分でしたか？」と私が尋ねると、「とんでもない」と彼女は答えた。「仕事をこなすためなら、使えるものは使います」。

この女性の生成AI (ジェネレーティブAI) の経験は、数多くある例の1つである。いま、街中 (そしてエレベーターの中) にいる人たちは、この斬新なテクノロジーが何のためにあるのかを理解しようとし、どんなことに使えるかを考えている。現在、生成AIをめぐる喧伝は、いろんな意味でインターネットの黎明期を彷彿とさせる。そこには興奮や期待感、そして前に進みながら作り上げているという感覚がある。

言うなれば、私たちはいま、2000年頃のドットコム・バブルの最中に

いるのと同じなのだ。この先、多くの企業が破綻するだろう。この時代のフェイスブック (現メタ)、ツイッター (現「X」)、ティックトック (TikTok) が出現するまで、何年もかかるかもしれない。「人々は10年後の未来を想像しがりません。誰も愚か者と思われたくないからです」。テクノロジー・コンサルティング会社ブーズ・アレン・ハミルトン (Booz Allen Hamilton) で生成AI部門の責任者を務めるアリソン・スミスは言う。「でも、それは私たちの予想を大きく超えるものになると思います」。

インターネットは、私たちの仕事のやり方や遊び方、友人や家族との過ごし方、学び方、消費の仕方、恋愛のやり方など、あらゆることを変えた。その一方で、ネットいじめやリベンジポルノ、トロール・ファクトリー (デマ発信組織) が出現した。インターネットは大量虐殺を促し、メンタルヘルス危機を煽った。さらに、中毒性のあるアルゴリズムと略奪的な広告を駆使した監視資本主義を誕生させ、これは現代の支配的な市場勢力となっている。こうしたデメリットが明らかになったのは、膨大な数の人々がインターネットを使い始め、ソーシャルメディアのようなキラアプリが登場してからだ。

生成AIも同じ道をたどる可能性が高い。オープンAI (OpenAI)、グーグル、メタなどが、いくつかの基盤生成モデルというインフラを整備し、その構築に携わっていない人たちが、作成者が夢にも思わなかった形で、このインフラの利用や悪用を始めることになる。「個々のユーザーに実際に試してもらわないと、その可能性とリスクを完全に把握する

「問題は、テクノロジーが使われるまで、その悪用方法をすべて知ることとは不可能だということです」

ことはできません」(スミス責任者)。

生成AIはインターネット上のコンテンツで訓練されているため、バイアス、誤った情報、著作権侵害、人権侵害、あらゆる面での経済的混乱に関するものなど、未解決の問題の多くを引き継いでいる。とはいえ、闇雲に進もうとしているわけではない。

生成AI革命の展開を観察するうえで、心に留めておくべき6つの未解決の疑問を見ていこう。インターネット革命の失敗を踏まえて、今回はもっと上手くやるチャンスが私たちにはある。

1. バイアスの問題がなくなる日は来るのか?

バイアスが人工知能 (AI) 関連の害悪の代名詞となっているのには、相応の理由がある。現実世界のデータ、特にインターネットから収集したテキストや画像は、性別をめぐる固定観念から人種差別に至るまで、バイアスに満ちあふれている。それらデータに基づいて訓練されたモデルは、そうしたバイアスがエンコードされており、どこで使われようとさまざまなバイアスを強化する方向に働く。

チャットボットや画像生成AIは、エンジニアを白人男性、看護師を白人女性として表現する傾向がある。黒人は、警察の顔認識プログラムによって誤認され、不当逮捕に見舞われるリスクを抱えている。人材採用アルゴリズムは、女性よりも男性を好む傾向があり、しばしばその対処のために取り入れられるアルゴリズムによって、バイアスが強化されている。

新しいデータセットやモデルの新しい訓練方法(どちらもその準備に



SELMAN DESIGN

何年もかかる可能性がある)がなければ、バイアス問題の根本原因は依然として残ることになる。しかし、これが注目の研究テーマであることに変わりはない。オープンAIは、人間のフィードバックによる強化学習 (RLHF: Reinforcement Learning from Human Feedback) などの手法を使って、大規模言語モデルのバイアスを減らすことに取り組んでいる。これにより、人間のテスターが好ましいと思うようなテキストを出力するようにモデルが方向付けられる。

他の手法としては、合成データセットの使用が挙げられる。例えば、映像制作用の生成モデルを構築するスタートアップであるランウェイ (Runway) は、民族、性別、職業、年齢が異なる人々のAI生成画像など、合成データを使って、人気の画像生成モデル「Stable Diffusion (ステーブル・ディフュージョン)」のあるバージョンを訓練している。同

社の報告によると、このデータセットで訓練されたモデルでは、肌の色が濃い人の画像や女性の画像がより多く生成される。実業家の画像をリクエストすると、スカーフをかぶった女性の画像が出力されるようになった。また、医師の画像では、肌の色や性別が多様な人たちが描かれるといった具合だ。

評論家らはこれらの対策を、壊れた基本モデルの対処療法に過ぎず、問題の解決ではなく隠蔽であるとして否定している。しかし、ブーズ・アレン・ハミルトンの「責任あるAI」部門のスミス責任者の同僚、ジェフ・シェーファーは、このようなアルゴリズムにおけるバイアスは、長期的には有益な形で社会のバイアスを暴露する可能性がある」と主張する。

一例として、人種に関する明示的な情報がデータセットから削除されたとしても、人の住所から人種を推測できるため、人種バイアスが依然としてデータに基づく意思決定を歪

める可能性がある。これによって人種差別や住宅差別のパターンが明らかになるとシェーファーは指摘する。「大量のデータを1カ所に集めたことで、その相関関係が非常に明確になりました」。

シェーファーは、同様のことがこの世代のAIにも起こる可能性があると考えており、「社会全体に存在するこうしたバイアスが明らかになるでしょう」と言う。そして、それがよりの絞った政策立案につながるだろうと語る。

しかし、多くの人はそのような楽観主義に洗い顔をするだろう。問題が掛けになったからといって、それが解決される保証はない。政策立案者たちは、住宅、雇用、融資、治安維持などの面で、数年前に露呈した社会的バイアスに今も対処しようとしている。その間、人はそれを受け入れて生活することになる。

予測：バイアスは、ほとんどの生成AIモデルに内在する特性であり続けるだろう。しかし、次善策と認識の広がり、政策立案者が最も明白なバイアスに対処するうえでの助けとなるかもしれない。

2.AIで著作権の扱いがどう変わるのか？

テック企業が同意なしに作品から利益を得ることに憤慨したアーティストや作家（およびプログラマー）らは、著作権侵害を主張して、オープンAIやマイクロソフトなどを相手取って集団訴訟を起こしている。Gettyは、画像生成モデルのStable Diffusionを提供するスタビリティAI (Stability AI) を訴えている。

これらは大きな問題となっており、サラ・シルバーマンやジョージ・R・R・マーティンなどの著名人が原告となり、メディアの関心を集めている。こうした訴訟により、少なくとも米国では、他者の作品の何がフェアユースで、何がそうでないかに関する規則が書き換えられることになりそうだ。

しかし、期待は禁物だ。280社以上のAI企業の代理人を務めるガンダーソン・デットマー (Gunderson Dettmer) 法律事務所、知的財産ライセンス専門パートナーのケイティ・ガードナー弁護士は、裁判所が最終決定を下すまでには何年もかかるだろうと話す。最終決定が下るころには、「このテクノロジーは経済にしっかりと定着し、元に戻すことはできなくなっているでしょう」と同弁護士は言う。

その一方で、テック業界はこのような侵害疑惑を猛スピードで積み上げていっている。「企業が様子見をするとは思えません」とガードナー弁護士は言う。

「ある程度の法的リスクはあるかもしれませんが。しかし、それ以外に、追従しないことによるリスクも、大変多く存在するのです」。

侵害の可能性を制限する対策を講じている企業もある。オープンAIとメタは、クリエイターが将来のデータセットから自分の作品を削除できる方法を導入したと主張している。オープンAIは現在、「DALL-E (ダリー)」のユーザーが、生存中のアーティストのスタイルで画像をリクエストできないようにしている。しかし、「これらはすべて、訴訟で自分たちの主張をサポートするための措置です」とガードナー弁護士は話す。現在、グーグル、マイクロソフト、

オープンAIは、自社モデルのユーザーを潜在的な訴訟から守る姿勢を示している。マイクロソフトは、訓練に使われたコードを作ったソフトウェア開発者を代表とする集団訴訟の対象となっている生成コーディングアシスタント「ギットハブ・コパイロット (GitHub Copilot)」の免責ポリシーで、裁判所が問題を解決する間、原則としてそれを使用する人たちを保護することになる。「当社製品のユーザーが心配せずに済むよう、責任は私たちが負います」と、マイクロソフトのサティア・ナデラ最高経営責任者 (CEO) はMITテクノロジーレビューに語った。

一方で、新しい種類のライセンス契約も登場している。シャッターストック (Shutterstock) は、オープンAIと画像の使用に関する6年間の契約を締結した。また、アドビは、「ファイアフライ (Firefly)」という独自の画像作成モデルを、許諾された画像、アドビ・ストック (Adobe Stock) のデータセットの画像、または著作権が消滅した画像のみで訓練したと主張している。ただ、アドビ・ストックへの投稿者の中には、相談を受けておらず、不満を抱いているという人もいる。

憤りは激しい。アーティストたちはいま、独自のテクノロジーを使って反撃している。「ナイトシェイド (Nightshade)」というツールを使うと、人間には気づかれない形で画像に手を加えることができる。しかし、その変更は機械学習モデルにとっては壊滅的な影響をもたらす、訓練の際に画像を誤って分類してしまうという。こうした技術により、ネット上でのメディアの共有と再利用に関する規範が大幅に再整理されることが期待される。

予測：注目の訴訟は引き続き関心を集めるが、だからといって企業が生成モデルの強化を止める可能性は低いだろう。倫理に則ったデータセットを中心に、新たなマーケットプレイスが誕生し、企業とクリエイターとの間のいたちごっこが展開されるだろう。

3.AIで仕事はどう変わるのか？

私たちは長い間、AIが私たちの仕事を奪ってしまうと聞かされてきた。今回、1つ違うのはデータアナリスト、医師、弁護士、(妄信的な)ジャーナリストなどのホワイトカラーも危機に直面しているように思われることだ。チャットボットは高校の試験、医師免許試験、司法試験で楽々と合格点を取ることができる。会議を要約したり、基本的なニュース記事を書いたりさえできる。人間の仕事には何が残っているのだろうか？しかし、現実はその単純ではない。

研究者の多くは、大規模言語モデルのパフォーマンスが真の知能の証示ではないとしている。しかし、たとえ大規模言語モデルが知能を持っているとしても、ほとんどの専門職には、大規模言語モデルにはこなせないタスクが大量に存在する。

2023年夏、ペンシルベニア大学ウォートン・スクールでイノベーションについて研究しているイーサン・モリック准教授は、ボストン・コンサルティング・グループと協力して、コンサルタントに与えるChatGPTの影響を調べる実験をした。実験チームは、数百人のコンサルタントに、「未開拓の市場やスポーツをターゲットにした新しい靴のアイデアを10個以上提案する」「ユ

ーザーに基づいて履物産業市場をセグメント化する」など、架空の靴会社に関連するタスクを18個与えた。そして、いくつかのグループはChatGPTの助けを借りることができるようにした。

その結果は驚くべきものであった。「GPT-4を利用したコンサルタントは、あらゆる面で、パフォーマンスを測定したあらゆる形で、利用しないコンサルタントよりもはるかに高いパフォーマンスを示しました」とモリック准教授はこの研究についてのブログ投稿で書いた。

ベンチャーキャピタル企業エア・ストリート・キャピタル(Air Street Capital)の創業者で、AI研究とその業界動向の包括的な年次概要である「AIの現状報告書(State of AI Report)」を発行するチームのリーダーを務めるネイサン・ベナイッチは、多くの企業がすでに情報の検索や取得に大規模言語モデルを使っていると話す。彼はこれを歓迎し、「AIモデルがアナリストになることを願っています。こういったことは、大抵とても面倒ですから」と言う。

ベナイッチ創業者が言いたいのは、単調な仕事を機械に任せることで、人は仕事の達成に特化した部分に集中できるということだ。また、AIによって従業員全体のスキルを平準化できるようだ。モリック准教授のコンサルタントや他のプログラマーに関する初期の研究では、経験の浅い人がAIを使うことで、より大きな後押しを受ける可能性が示されている。ただし、注意点もある。同准教授は、GPT-4に依存しすぎると不注意になり、モデルが間違いをしたときに、それに気づく可能性が低くなることを発見した。

生成AIで変わるのはデスクワー

クだけではない。画像および映像作成モデルを使えば、人間のイラストレーター、撮影スタッフ、俳優なしで、写真や映画のストリームを無限に作成できるようになる。2023年に米国で作家や俳優がストライキを起こしたことで、このことが今後何年にもわたって争いの火種になるであろうことが明らかになった。

それでも、多くの研究者は、このテクノロジーが労働者全体に取って代わるのではなく、力を与えるものであると見ている。結局のところ、産業革命以降、テクノロジーは人の仕事を奪ってきた。古い仕事が無くなれば、新しい仕事生まれる。「生成AIは正味でプラスに働くと強く感じています」とスミス責任者は言う。

しかし、変化には常に痛みが伴い、正味のプラスによって個人の損失が隠されてしまう可能性がある。テクノロジーの激変は、富と権力を集中させ、不平等を助長する傾向もある。

「私の考えでは、問題はもはやAIが仕事を奪えるかどうかではなく、私たちがそれをどうしたいかということです」とモリック准教授は書いている。

予測：大量の雇用喪失の懸念は誇張されていたことが判明するだろう。しかし、生成ツールの使用は、職場で今後も急増していくだろう。人の役割が変わる可能性があり、新しいスキルを学ぶ必要があるかもしれない。

4.生成AIでどのような誤った情報の発信が可能なのか？

2023年に最も拡散された3つの画像は、バレンシアガのダウンコートを着たローマ教皇、警官に組み伏せられるドナルド・トランプ、国防

Insider Online限定

eムックはMITテクノロジーレビュー[日本版]の
有料会員限定サービスです。
有料会員はすべてのページ、バックナンバーを
ダウンロードできます。

ご購入はこちら



<https://www.technologyreview.jp/insider/pricing/>

No part of this issue may be produced by any mechanical, photographic or electronic process, or in the form of a phonographic recording, nor may it be stored in a retrieval system, transmitted or otherwise copied for public or private use without written permission of KADOKAWA ASCII Research Laboratories, Inc.

本書のいかなる部分も、法令または利用規約に定めのある場合あるいは株式会社角川アスキー総合研究所の書面による許可がある場合を除いて、電子的、光学的、機械的処理によって、あるいは口述記録の形態によっても、製品にしたり、公衆向けか個人用かに関わらず送信したり複製したりすることはできません。