



生成系AI(ChatGPT, BingAI, Bard, Midjourney, Stable Diffusion等)について

目次

[何ができるか、「検索」ではなく「相談」するシステム](#)

[仕組み上、書かれている内容の信憑性には注意が必要](#)

[機密情報や個人情報などを安易にChatGPTに送信することは危険](#)

[将来著作権や文書を用いた試験・評価に問題が発生する可能性がある](#)

[社会に対する影響](#)

[本学の学生や教職員はどう対応したらよいか](#)

2023年4月3日

東京大学理事・副学長（教育・情報担当）

太田 邦史

この半年ほどの期間で、**生成系人工知能(Generative AI)**が複数発表され、社会的に大きな注目を集めています。基本的には、インターネット上などに存在する既存の文章や画像イメージを大量に機械学習し、これに強化学習を組み合わせなどして、一定レベルの品質の文章や画像を生成するシステムです。とくに、2022年11月に公開され、話題になった大規模言語モデル**ChatGPT**はバージョンが更新され、最新の**GPT-4**では生成される文章などの質や正確性が著しく向上しています¹。

これらの生成系AIは、平和的かつ上手に制御して利用すれば、人類の言語活動や知的創作活動の一部を補完し、私たちのwell-being向上に大きく寄与することができるでしょう。一方で、生成系AIには技術的な課題も存在しており、今後の社会への悪影響も懸念されています。実際、2023年3月末にはテスラ創業者のイーロン・マスク氏やアップル創業メンバーのステーブ・ウォズニアク氏、歴史学者のユヴァル・ノア・ハラリ氏らが、半年間人工知能の開発を凍結し、管理しながら開発すべきであるとの公開書簡を発表しています。多くの分野の学者が社会の大変革を予感しており、原子力やコンピューターの登場ぐらいのインパクトがあるだろうと語っています。私は分子生物学や合成生物学が専門ですが、この分野でいうと「組換えDNA技術」の登場に匹敵する変革ではないでしょうか。この時も、研究者が一度立ち止まり、カリフォルニアのアシロマで会議を行い、研究のガイドラインを策定し

ました。AI分野でも2017年に**アシロMAI23原則**が発表されましたが²、生成系AIの最近の進歩を踏まえ、再度ガイドラインが見直されるかもしれません。

このお知らせでは、東京大学の学生や教職員が生成系AIに対してどのように向き合うべきか、現在の状況（急速に変化している最中ですが）を踏まえ、第一報として情報を共有したいと思います。生成系AIには複数のシステムが存在しますが、以下では話を単純化するためにChatGPTを中心に議論を進めます。

何ができるか、「検索」ではなく「相談」するシステム

ChatGPTは、メールアドレスを登録すれば無料で（現在GPT3.5までのレベルまで）利用できます。何ができるかですが、ここでいろいろ説明するよりも、まず皆さん自身で試してみるのが良いと思います。ChatGPTに何か質問を書き込むと、いろいろな情報を交えて文章を返してくれます。**質問・指示文章（プロンプト**とか呪文と呼ばれます）のやり方を工夫すると、文書作成や文書アウトラインの生成、翻訳、文章の修正・トーン変更、文章要約、コンピューター・プログラムの作成や修正、試験問題作成、調査・分析、ディスカッション・ブレインストーミングの壁打ち相手など、様々なことに使えます。どのような長さやどのような文体（です・ます調、硬い文章、子供向けなど）で回答するか、また箇条書きや表形式で回答を要求することもできます。ネットで使える私設コンサルタントのような使い方ができます。

既存のGoogleなどの検索エンジンでは、調べたことに関連する内容を記したサイトがリストになって出てきます。私たちは、リストの中からリンクをクリックしてサイトを開き、情報を読み取る作業を行います。それらのサイトを書いてある内容を自分で咀嚼し、知識を総合して理解する必要があります。ところが、ChatGPTに質問すると、**質問に対してかなりの確かな内容**（時々間違いはありますが）を自動的にとりまとめて回答してくれます。

API(アプリケーション・プログラミング・インターフェース)というソフトウェア間の情報のやりとりをする仕組みも提供されていますので、**表計算ソフトや音声認識・発声ソフト、画像生成が可能な生成系AI等と連動**させることで、かなりの**作業を自動化**することができます。たとえば、表計算ソフトにGPTのAPIを連動させると、いろいろな国の人口や首都、主要産業、GDPなどを自動的に一瞬で調べることができます。スマートフォンの音声認識機能と連動させると、ChatGPTに料理のレシピ等いろいろな質問を話しかけるだけで、音声で答えてくれるようになります。Pythonなどのプログラミング言語を用いて指示を行うと、さらに複雑なことが自動化できます。たとえば、章立てした文書も自動的に作成できます（なお、この文書作成にはChatGPTは用いていません）。現在多くの人々が、ChatGPTを用いた新しい使い方を模索しており、もうすぐ多数の新規ビジネスが生まれるでしょう。パソコンやインターネット、スマートフォンの登場時と同等、あるいはそれ以上の社会的な影響があると思います。

仕組み上、書かれている内容の信憑性には注意が必要

ChatGPTの原理は、大量の既存の文章やコンテンツの機械学習と強化学習を通じ、確率的にもっともらしい文章を作成していくものです。したがって、**書かれている内容には嘘が含**

まれている可能性があります。非常に話し上手な「知ったかぶりの人物」と話しをしているような感じです。（もっとも、最新バージョンGPT4ではかなり正確性が増しており、相談相手としても相当優秀な存在です。）そのため、ChatGPTを使いこなすには、相当の**専門的な知識が必要**であり、**回答を批判的に確認し、適宜修正することが必要**です。また、既存の情報にない新しい知見に関する分析や記述はできません。つまり、ChatGPTが出たからといって、**人間自身が勉強や研究を怠ることはできない**ということです。ただ、教養や専門知識を有する人物が回答内容を批判的に分析し、上手に使いこなせば、定型的な業務の効率を格段に向上させることが可能でしょう。

機密情報や個人情報などを安易にChatGPTに送信することは危険

ChatGPTに送った質問の文章も、ChatGPTのシステムに蓄積され、学習される可能性があります。原理的には送信した内容を他者が引き出すことも可能だと思った方が良いでしょう。したがって基本的に、**業務で知り得た機密情報、未公開の発明内容、研究費などの申請内容、入学試験問題の原稿、個人情報などは質問に含めない**ことが求められます。また、設定が**オプション**（申請すれば送信情報が取り込まれる）か**オプトアウト**（申請すれば送信情報が取り込まれない）になっているかを確認した上で、利用することも必要です。

将来著作権や文書を用いた試験・評価に問題が発生する可能性がある

画像生成系のAIでは、インターネット上のコンテンツを取り込んで学習し、画像を生成しています。現在、これらの元画像を作成した著者が、**知らないうちに著作権を侵害されたとして問題提起**を行っています。将来的に生成系AIが生み出したコンテンツが訴訟の対象になる可能性があるかもしれません。また、学位論文やレポートを提出する際や、学術誌などに論文投稿する際に、生成系AIをだけを用いた文書を提出してはならないという規定が出始めています。**本学では学位やレポートについては、学生本人が作成することを前提としておりますので、生成系AIのみを用いてこれらを作成することはできません**。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す。したがって、**教員はレポートや提出論文の審査に関しては、十分そのことを認識した上で評価を行う必要**があります。つまり、論文やレポートなどの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**が出てきます。（後日、「AIツールの授業における利用について」を公表しますので、そちらもご確認下さい。）

社会に対する影響

今後生成系AIの利用が普及すると、**教育・研究活動など大学の活動にも大きな影響**が生じると考えられます。私たちを取り巻く**社会の構造や産業構造にも大きな変化**が生じると考えられます。現在の社会は法律や制度面においても、今回の生成系AIの登場が織り込まれていません。下手をすると失業者の増大、産業構造の変化、社会の階層化の進行など、様々な悪影響

が生じる可能性があります。現在議論されているのは、定型的な業務に関与するホワイトカラーに大きな影響が生じる可能性があるとのこと。最悪の事態は、映画『ターミネーター』のスカイネットや、『マトリックス』のようにAIに人間が支配される世界が来るかもしれないと恐れる人もいます。

本学の学生や教職員はどう対応したらよいか

しかしながら、生成系AIを有害な存在であるとして利用禁止する（イタリア、米国の事例など）だけでは問題は解決しません。秘密裏に開発が進行する可能性や、地下で悪用されることも懸念されます。人類はこの数ヶ月でもうすでにルビコン川を渡ってしまったのかもしれないのです。むしろ、**どのようにしたら問題を生じないようにできるのか、その方向性を見出すべく行動することが重要**であると思います。何にせよ、大きな変革の時期に来ていると考えられますので、本学構成員の皆様は、この変化を傍観するだけでなく、**大規模言語モデルに「創発」（能力が突然飛躍的に向上すること）が起きた³原因を考察**したり、生成系AIがもたらす様々な社会の変化を先取りし、積極的に**良い利用法や新技術、新しい法制度や社会・経済システムなどを見出し**ていくべきではないでしょうか。今後、生成系AIの活用法や問題点、改善策などについて、**学内で**（本部・教育システム整備検討WG、情報システム本部などを介して）**議論の機会を設ける予定**です。このような対話を通じて、是非社会とともにより良い世界の構築に貢献して頂きたいと願っています。

2023年4月3日訂正：GPT-4のパラメーター数については公式に開示されていないとの情報を得ましたので、削除いたしました。

本ポータルサイト (utelecon) の生成AI（ChatGPT等）関連情報ページはこちら

1. “GPT4 technical report.” arXiv:2303.08774 [cs.CL] (2023) OpenAI
<https://doi.org/10.48550/arXiv.2303.08774> ↑
2. “The Asilomar AI Principles” <https://futureoflife.org/open-letter/ai-principles/> ↑
3. “Characterizing Emergent Phenomena in Large Language Models.” Google Research. Blog (2022) Wei J. Tay Y., <https://ai.googleblog.com/2022/11/characterizing-emergent-phenomena-in.html> ↑



utelecon: オンライン授業・Web会議ポータルサイト

まずはここから

大学生活に必要な情報システムの準備について（新入生向け）
東京大学における情報システムの準備について（教員向け）
オンラインを活用するために
サポート窓口

東京大学のシステム

UTokyo Account
UTAS
ITC-LMS
Zoom
Webex
ECCSクラウドメール (Google Workspace)
UTokyo Microsoft License (Microsoft 365)
UTokyo Wi-Fi
ECCS端末
その他

オンラインの活用

オンラインを活用するために
使えるツールから探す
やりたいことから探す
よりよい授業の実施方法を探す
東大の実践を知る
資料作成における著作権

お知らせ・イベント

お知らせ
オンライン授業情報交換会
説明会等

サポート

サポート窓口
よくある質問 (FAQ)
システム運用状況（障害情報等）
授業入室トラブル