

生成系 AI の教育および教育研究へのインパクトについて

武田俊之^{*1}

^{*1} 関西学院大学

The Impact of Generative AI on Education and Educational Research

Toshiyuki Takeda^{*1}

^{*1} Kwansei Gakuin University

対話型の大規模言語モデル (Large Language Model) である ChatGPT は、大量のテキストデータを学習することによって、文章生成、要約、QA、翻訳などのさまざまなタスクを実行可能な生成系 AI (Generative Artificial Intelligence) である。ChatGPT は従来の AI と比較して人間らしく知的水準の高い応答が可能であり、教育および教育研究へのインパクトは大きい。この報告では生成系 AI が適用可能な教育の場面と今後の教育研究にあたえる影響を探究する。

キーワード: 生成系 AI, 大規模言語モデル, ChatGPT

1. はじめに

2022 年 12 月に発表された対話型の大規模言語モデル (Large Language Model: LLM) である OpenAI 社の ChatGPT[1]は、従来の AI よりはるかに人間らしく知的水準の高い応答が可能で、自然言語処理分野だけではなく社会的なインパクトをあたえた。その後 Microsoft 社の Bing Search との連携や、より大規模な学習をおこなった GPT-4[2], OpenAI API など、ChatGPT の関連技術が公開されている。

ChatGPT のように、新しい情報やコンテンツを生成する人工知能を生成系人工知能 (生成系 AI, Generative Artificial Intelligence: GAI) と呼ぶ。GAI の研究開発はテキスト、画像、音楽、動画などで活発におこなわれており、人間の専門家が制作するレベルのアウトプットが得られる製品が増えている。なかでも Transformer アーキテクチャの発表以降、大規模言語モデルの性能向上と適用範囲の拡大はめざましい。

ChatGPT は、(1) 誰でも使えるインターフェース (チャット) 上で、(2) システムと対話することによって広く深い知識を引き出すことが可能であり、(3) 他

のシステムと比較してバイアスが少ないため安心して利用することができることが特徴である。このような生成系 AI の利用によって、人間の知的活動の生産性は向上する可能性がある。

一方で生成系 AI は社会活動に脅威となる可能性がある。ChatGPT のような LLM は入力テキストからそのつづきを予測して出力したテキストは実社会で正確であるとは限らない。また学習で使用されたテキストに含まれる差別的表現等のバイアスが LLM でも表現されていて出力に含まれることがある。入力に個人情報や企業の機密情報がふくまれた場合に漏洩する可能性がある。その他、生成系 AI を悪用する方法は多数あり、より社会的な問題としては「AI が仕事を奪う」などの議論もある。リスクの評価をおこなうためには、ChatGPT の技術的詳細を公表するなど、基盤技術の透明性確保も必要である。

このような生成系 AI のインパクトが大きい領域の一つが教育である。AI が教員のさまざまなタスク (教材作成など) のアシスタントとなるなどの益がある一方で、AI で生成したレポートの提出や AI による試験

代行など大きなリスクに直面している。さらに「生成系 AI が日常的な知的ツールとなる時代になにをどのように教えるか」のようなより大きなインパクトのある課題がある。コンピューターやインターネットの普及が教育を変えたように、実践上のタスクの内容と方法の見直しだけではなく、上位の目的や目標の変更はさらに大きな変化が必要となる。生成系 AI のインパクトは人間の知的活動の多くを代替する可能性があり、過去の技術進歩よりも大きいと思われる。

このような教育の内容と方法に大きな変化が必要なら、それは教育研究の過去の研究に基づくことが望ましい。それは新技術の適用が効果的、効率的におこなうこと、そして、ステークホルダーのリスク低減のためである。特に学校教育において児童生徒学生等は新技術の実験台とすべきではなく、リスクを最小化しなければならない。

以上の点からこの報告では、(1) 生成系 AI の中でも大規模言語モデルをベースとした ChatGPT の仕組みを説明した上で、(2) 教育場面における生成系 AI の利用とリスク、今後の研究開発の可能性について記述する。新規性が高く技術進歩も速いため、内容は不十分であるが、なるべく体系的な記述を試みる。

2. ChatGPT とは

2.1 大規模言語モデル

ChatGPT は大規模言語モデルの一種である。言語モデル (Language Model) とは、自然言語の単語や文章が生成される確率をモデル化したもので、テキストのデータセットを用いて作成される。言語モデルの中でも非常に大きなデータセットで訓練した言語モデルを大規模言語モデル (LLM) と呼ぶ。

言語モデルの主な機能は単語や文章などのテキストを生成する確率の計算であるが、ニューラルネットワークに基づくニューラル言語モデルは、さまざまな自然言語処理 (NLP) タスクに対応できるように設計されている。たとえば、機械翻訳、質問応答、固有表現抽出、文書要約、単語や文章の生成、文章の補完、文法誤り訂正、文章の内容判定、感情分析などを、言語モデルを用いておこなうことができる。

Transformer アーキテクチャ[3] に基づく大規模言語モデルである BERT[4] が高い性能と汎用性を示して以降、さまざまな派生モデルや他の Transformer ベースのモデル (GPT[5] 等) の大規模化と性能の向上が進んだ。

2.2 ChatGPT

2022 年 12 月に公開された OpenAI の ChatGPT は、自然で「知的」な文章を生成する性能の高い LLM ベースのチャットボットである。専門家以外にも使いやすい対話型のインターフェース (チャット) を備えており、社会に大きなインパクトをあたえた。ChatGPT 以前のチャットボットの多くは、あらかじめ定形の回答を用意しておき、入力文に近い回答を返す。しかし、一方で、ChatGPT との対話は自然で、ほとんどの質問に自然な形でもっともらしい答えを返す。

チャットの中の ChatGPT への指示はプロンプトとして機能する。プロンプトとは、特定のタスクを実行させるための LLM に対する指示で、質問と回答の例示や処理するデータを含む。LLM が生成するテキストの品質や適切性はプロンプトによって変わる。望ましいテキストが得られるようなプロンプトの設計をプロンプト・エンジニアリングと呼ぶことがある。また、LLM を特定のタスクやドメインに適応させるために、そのタスクを解くための入力と出力のペアなどをあたえることをファインチューニングと呼ぶ。

ChatGPT は、大規模言語モデル (GPT-3.5) を、対話テキストをデータとして、人間によるフィードバックを用いた強化学習 (Reinforcement Learning from Human Feedback: RLHF) によってファインチューニング (微調整) をおこない、人間の指示、嗜好に合った応答を返せるようにしている。

さらに ChatGPT は安全な利用にも配慮されている。悪意のある指示や言及すべきではない内容のフィルターや、さらに社会的、文化的な問題 (差別などのバイアス等) のある応答の監視と、頻繁な改善とリリース (Iterative deployment) をおこなっている。ただしプライバシー上の懸念は残っており、次節で述べる。

これらの高度なエンジニアリングによって、

ChatGPT は膨大な知識を持つだけでなく、安全に利用できることのできる Web サービスとして、広範囲な利用を獲得した。ChatGPT は Web インターフェースだけではなく、OpenAI API を通じて利用することが可能である。API によって外部と連携したアプリケーションを作成することが可能であり、2021 年以降のデータやドメイン知識の少ない分野でも ChatGPT の機能を強化することができる。

ChatGPT は GPT 3.5 がベースの LLM である。その前のバージョン (GPT-3) (2020) は 570GB のデータセット (元となるテキストデータは 45TB) を用いて作成した 1750 億パラメータの言語モデルである。最初の GPT の 1.17 億パラメータと比較して極めて大きくなっている。

その後発表された GPT-4 はさらに性能が高く、米国の司法試験の内容で上位 10% のスコアを獲得するなど人間の専門家に匹敵する。教育分野では、Duolingo, Khan Academy などが GPT-4 と連携している。また、GPT-4 は画像の出力を含めたマルチモーダルなモデルであるが、2023 年 3 月現在公開はされていない。

2.3 ChatGPT への批判

実用性が高く、リスクへの対処よりも普及が急速であったこともあり、さまざまな点において ChatGPT は批判を受けている。

(1) 不正確な出力

LLM の性質上、出力は知識を反映しているのではなく、入力から予測したものである。ChatGPT は (おそらく差別的な内容は別として) 不正確な内容であってもそのまま出力する。ChatGPT は 2021 年までのテキストで学習しているためそれ以降のことは正確に答えられない。場合によっては平然と嘘をつくかのように答えることもある (幻覚: hallucination)。

また人種、性別、宗教などに関するバイアスがあるかもしれない。これらは LLM 構築時のテキストの正確性、即時性の向上や学習時間の増加によって改善が期待できるが、本質的に解消できないため、別の方法で出力の検証が必要である。

(2) プライバシー、著作権、機密情報の問題

ChatGPT が学習した内容には個人情報が含まれている。また API 版を除くデータは LLM の追加学習に使われる可能性がある。これらは歴史上の人物を含めた著名人以外は個人名の削除等によってプライバシー侵害にならないよう処理されているということであるが処理方法は公開されていない。この点への懸念からイタリアの個人データ保護局は ChatGPT へのアクセスを一時的に禁止した。EU の他国も同様の懸念を持ち調査をおこなっている。

また安全保障上のリスクへの対策も諸国で検討されている。企業の機密情報の漏洩やソースコードの流出を防止するための利用ガイドラインの策定などを進める組織も少なくない。

(3) 寡占

ChatGPT 発表以降同等性能の LLM がいないため、サービスの寡占状態となっている。競争上の問題の解消は重要であり、また、データソースが異なる他の LLM があれば不正確な出力の検証やバイアスの検出に役立つ可能性がある。

2017 年にアシロマ AI 23 原則 [6] を発表した Future of Life Institute は、OpenAI の生成系 AI 開発を 6 ヶ月停止して、安全な AI を開発のプロトコルと監査の策定することを求める "Pause Giant AI Experiments" [7] を公表した。批判を受けて Open AI は "Our approach to AI safety" [8] を発表した。

2.4 教育機関等の反応

ニューヨーク市教育局は学校から ChatGPT へのアクセスをブロックした [9]。理由は ChatGPT の使用が批判的思考などのスキル育成の妨げとなることと、提供されるコンテンツの安全性・正確性に対する懸念であった。初等教育においてはこのような利用制限をかける自治体は少なくない。

大学の場合は、生成系 AI の日常化を前提として、その特徴とリスクに対する心構えを述べる [10] か、一定の制約とともに教員の責任における利用の許諾 [11] か、が示されることが海外を含めて多いようである。いずれにしても技術の進歩を見極めながら、教育の価値を高めるよう施策を更新することになろう。

3. 教育へのインパクト

教育における ChatGPT の登場のインパクトは大きい。社会的に AI の利用が進む中で各個人の AI の支援を受けた知的活動をおこなう以上、その教育への反映は必須である。

このような AI が知的リソースと日常化した社会において、教育はどのように変化するであろうか。以下のようないくつかの問いが考えられる。

(1) 個人より優秀な AI が日常化したときに、個人が習得すべき能力はどのようなものか。そしてその能力を獲得する方法はどのようなものか。

ChatGPT と GPT-4 でも性能や特徴が異なっており、今後はさらに特徴の異なる AI が利用可能になる。さらに複数の AI を連携させながら知的生産活動に従事するならば、個人に必要な能力は何で、それを学校教育ではどのように学ぶのであろうか。この問いは簡単に答えられるものではなく、また技術の変化の影響も大きい。現時点の生成系 AI を対象として答えを出せるものでもない。

しかし、この問いの検討は緊急性があり、教育研究だけではなく、他の社会科学、人文学、自然科学、工学など、人間と社会に関連する学問の領域横断的で精緻な議論が必要であろう。

生成系 AI の道具としての AI の位置づけをインターネットや検索エンジンの利用（さらに電卓、コンピューターなどまで遡ることも）と同様とする意見もある。しかし、個人の知的生産のかなりの部分を代替する AI はこれらと同等ではないと思われる。詳細な比較、検討が必要であろう。

(2) 教育の公平性、公正性をいかに維持するか。

「ChatGPT の利用を教えて当然」という主張は、OpenAI 社が戦略として無料にしていることを無視している。上位版の値段は 2023 年 4 月現在で 20 ドル/月であり、今後さらに差がつくともあり得る。個人の能力が AI を含めて判断されるならば、経済格差が能力格差であり、学校入学時から格差は広がることになる。このような格差と学校教育の公平性、公正性のコンフリクトをいかに解消するか（または容認するか）。

以上は教育研究を越えた問いであるが、教育の科学的、技術的問いをいくつか挙げる。

(3) AI の使用を前提としたときの個人の能力の定義とその評価方法

これは問いの(1)のサブタスクであり、新しい能力（観）とその評価方法の研究開発である。能力の定義と評価を含めた妥当性の確認は大きな問いへの貢献となる。実際には現代の能力観で問題ないかもしれないが検証は必要である。

(4) AI の特徴と人間との違いはなにか

GPT-3 で認知心理学のタスク実行と人間との比較 [12] のように知的主体としての AI の特徴の研究が進んでいる。教育や学習の支援において、AI の特徴を知することは必須である。一方で、ある生成系 AI の特徴が普遍的であるとはかぎらない。LLM の数理的性質の研究の進展も取り入れながら、教育利用に必要な項目を整理する必要がある。

4. 生成系 AI を用いる教育の場面

この節では、今後の教育の研究と開発の手がかりとなるよう、授業関連のプロセスで生成系 AI を用いる場面、内容、方法について記述する。

4.1 教員の授業のプロセス

教員の授業のプロセスは以下のステップは、授業目標の設定、目標と連結した評価方法の設定、これらを元にした授業の構造・内容の設計と学習活動・教材の選択の3段階で進む。そして、授業中、評価、そして授業全般に関わる場面で生成系 AI を用いることができる（表 1）。

4.2 学生の受講のプロセス

学生の授業受講プロセスを、授業の選択、授業中および授業外学習、評価とフィードバック、成績評価、そして全般に分類した。表 2 にこれらの場目において生成系 AI の利用法を挙げた。

4.3 教育における生成系 AI の研究開発

以上比較的単純な例として項目を示した。これらは

方法の効果に関して研究と開発は課題である。

今後の教育における生成系 AI の研究開発において特徴的と考えられる事項について述べる。

- AIを用いて生成したテキストを直接課題レポートにすることを防ぐことやその検出は困難である。ゆえに個人の能力を測定するための新しい方法の開発が必要であり、オンライン試験も同様である。
- ChatGPT の普及により、言語モデルから知識を引き出すために対話（プロンプトの入力と出力の利用）の重要性が一般に認識された。AI との対話は、個別学習、適応的学習、アクティブ・ラーニング、グループ学習におけるファシリテーションにおいても重要と考えられる。対話による知識構築の理論的、実践的研究とそのアプリケーション開発は重要な課題となるであろう。
- ファシリテーターとしての生成系 AI を考える場合に生成系 AI の知性の特徴研究は重要である。
- 以上は既存の ITS (Intelligent Tutoring System) と関連する点も多い。特定の領域を対象にすることが多い ITS の知見で、汎用性のある LLM ベースの AI で適用されることも多いであろう。

5. まとめと今後の課題

生成系 AI が教育に革命をもたらすのか、さまざまなリスクが発現するのかわからないが、多くの新技術がそうであるように課題を解決しながら適応していくのであろう。短期的にはレポートや試験での不正利用が課題であり、教育評価やレポート課題などの方法の見直し、不正検出や利用のルールづくりが喫緊の課題である。一方で教材開発や授業スクリプト作成の支援は教員にとって大きな力であり、適切な方法の確立と普及が望まれる。

現状は個人（教員、研究者、その他）の経験に基づく思いつきで ChatGPT の活用が試されている時期である。その中には学校の学習者を実験台とするような倫理上問題のある実践もあるかもしれない。今後は生成系 AI のインパクトと進歩の速さの負の側面を軽減するために、教育への生成系 AI についてシステマティックな検討が必要であろう。この報告は、今後の研

究開発の可能性とリスク評価のために、教育のプロセスにおける生成系 AI が支援、代替する要素を列挙する手がかりとなることを意図している。教育研究にはさまざまな価値やパラダイムのもとでおこなわれている。そのため、アドホックなアプローチに陥らずに生成系 AI のような革新的な技術の適用、評価を行うことは容易ではないが、既存の理論やエビデンスと生成系 AI のある教育環境デザインに教育研究者の果たす役割は大きい。

謝辞

本研究は科研費 JP21K02759 の助成を受けている。

参 考 文 献

- (1) OpenAI: "Introducing ChatGPT"
<https://openai.com/blog/chatgpt>
- (2) OpenAI: "GPT-4 Technical Report",
arXiv:2303.08774[cs.CL]
- (3) A Vaswani, N Shazeer, N Parmar, J Uszkoreit, L Jones, A N. Gomez, L Kaiser, I Polosukhin. 2017. Attention is All You Need. In NIPS, pp. 5998-6008.
- (4) Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. _arXiv preprint arXiv:1810.04805_.
- (5) Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training.
- (6) Future of Life Institute: "Pause Giant AI Experiments", <https://futureoflife.org/open-letter/ai-principles/>
- (7) Future of Life Institute <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>
- (8) OpenAI: "Our approach to AI safety",
<https://openai.com/blog/our-approach-to-ai-safety>
- (9) Michael Elsen-Rooney: "NYC education department blocks ChatGPT on school devices, networks",
<https://ny.chalkbeat.org/2023/1/3/23537987/nyc-schools-ban-chatgpt-writing-artificial-intelligence>
- (10) 太田邦史: “生成系 AI(ChatGPT, BingAI, Bard, Midjourney, Stable Diffusion 等)について”,

<https://utelecon.adm.u-tokyo.ac.jp/docs/20230403-generative-ai>

(11) 上智大学: “ChatGPT 等の AI チャットボット (生成 AI) への対応について”,

<https://piloti.sophia.ac.jp/assets/uploads/2023/03/27162222/23f430e7f216cbe188652f8a6855c493.pdf>

(12) Binz, M., & Schulz, E. (2023). Using cognitive psychology to understand GPT-3. _Proceedings of the National Academy of Sciences_, _120_(6), e2218523120.

	オン補助 授業内容に関する学習者個別の対話と補足説明 学習者への課題やプロジェクトのアイデアの提案
評価	レポート課題の提案 レポート課題の模範解答の作成支援 試験問題作成支援 学習者へのフィードバック支援

表 1 生成系 AI を用いる場面 (教員)

ステップ	AI の指導, 支援内容
授業目標の設定	学校の教育目標, カリキュラムの理解支援 他科目の理解の支援 カリキュラムやポリシーに関連した知識やスキルの列挙 学習者の事前知識の把握支援 知識, スキル, 学習者レベルに応じた授業目標の提案
評価方法の設定	授業目標に合致した達成度の評価方法の提案 評価方法の解説 定量的・定性的な指標の理解の促進 形式的・総括的評価の理解の支援
授業の構造・内容の設計 学習活動・教材の選択	授業各回のトピックの例示, 提案 学習目標に沿った活動の列挙, 提案 学習目標に沿った教材の列挙, 提案 学習者の事前知識に沿った教材の提案 教材作成の支援 インストラクショナル・デザイン 授業進行のシナリオの提案
授業中	進行, 教材に基づく授業スクリプトの作成 学習者個別の課題の提示, 応答へのフィードバック 学習者との対話支援 グループ学習でのファシリテーシ

表 2 生成系 AI を用いる場面 (学生)

ステップ	AI の指導, 支援内容
授業選択	学生の興味や専攻, 卒業要件に適した授業を提案 シラバスの内容の分析 教員の評判の分析 履修の推薦 スケジュールの最適化
授業内/ 授業外学習	適応的な課題の提示 回答への個別フィードバック 個別化された対話 AI への出題 課題トピックの探索 授業内容に関する対話 追加の資料や説明の提示 課題やプロジェクトのアイデアの提案
評価とフィードバック	練習問題の提供 弱点の提案
成績	成績と授業内容の関連付け 授業の総括 (対話) 達成状況の分析 次の学習トピック, 履修の提案
全般	目標設定 時間管理 学習方法の提案 自己評価と成長に関する対話

