

教員養成課程の学生を対象とした 生成AIリテラシー・ワークショップの開発と評価

齋木 匠^{1)*}・石井雄隆²⁾

¹⁾千葉大学大学院・融合理工学府・博士課程

²⁾千葉大学・教育学部

Development and Evaluation of a Generative AI Literacy Workshop for Pre-service Teacher Education Students

SAIKI Takumi^{1)*} and ISHII Yutaka²⁾

¹⁾Doctoral Student, Graduate School of Engineering, Chiba University

²⁾Faculty of Education, Chiba University, Japan

近年、ChatGPTをはじめとした対話型AIが急速に広がっており、教育分野への応用も期待されている。しかしながら、教師のAIリテラシー涵養に関する研究は十分に進んでおらず、児童・生徒のAIリテラシーを養う上での課題となっている。本研究では、ChatGPTの解説スライドと実際にChatGPTを利用するワークショップを開発し、教員養成課程の学生に対する教育効果を測定した。ワークショップ前後のアンケート分析より、開発したワークショップの有効性が示された。

In recent years, conversational AI, including ChatGPT, has been spreading rapidly, and its application to the field of education is also expected. However, research on improving teachers' AI literacy is still ongoing, and this is an issue for students in developing AI literacy. In this study, we created explanatory slides for ChatGPT and held a workshop using ChatGPT. Through the workshop, we measured the educational effect on students in the teacher education program. The results demonstrated the effectiveness of the workshop.

キーワード：生成AI (Generative AI), AIリテラシー (AI Literacy), ワークショップ (Workshop)

1. はじめに

近年のArtificial Intelligence (人工知能) の急速な進歩は様々な分野に影響を与えており、教育分野においても、AIを活用した教材開発や学習分析が行われている(緒方・江口, 2023; Piech, 2015)。中でも、2022年に登場した生成AIの一つであるChatGPT (<https://openai.com/chatgpt/>) は、人間と遜色ない対話を行なえるだけでなく、各種学力テストで高得点を出す知識の量と論理を兼ね備えたことから普及し、様々なサービスも登場した。GIGAスクール構想の進展に伴って、ChatGPT等の生成AIを活用した教育実践が蓄積される中で、児童・生徒のAIリテラシー教育の重要度は今後とも高まると考えられる。

文部科学省(2023)は、「初等中等教育段階における生成AIの利用に関する暫定的なガイドライン」を策定し、その利活用について、ガイドラインを整備しているが、教師のAIリテラシー向上についての研究は十分になされていない。児童・生徒がAIを安全に利用するためには、教師がAIを理解し、児童・生徒に適切な指導を行う必要がある。生成AI普及以前に開発されてきた情報リテ

ラシー教材では、生成AIの特徴に対応できない側面があるため、新たな教材の開発が必要である。

本研究では、最も普及している生成AIサービスの一つとしてChatGPTを取り上げ、教員養成課程の学生を対象に、ChatGPTの仕組みについて解説するスライドを作成し、実際にChatGPTを用いるワークショップを開発した。ChatGPTを用いた理由としては、教育分野での応用を考えた際、画像や音声ではなくテキストを出力するAIサービスが適しているという点とChatGPTの基礎となっているモデルについては、詳細な論文が出ており、仕組みが把握しやすい点が挙げられる。

2. 先行研究

2.1 AIリテラシーの定義

Casal-Otero et al. (2023) は179の文献をレビューし、AIリテラシー・アプローチの2つの大まかなグループ、すなわちAIを用いた学習経験とAI技術の理論的理解に分類した。最初のグループは、特定の関心領域における技術的、概念的、応用的スキルの学習経験を対象としていた。第二のグループは、AIリテラシーの提案の骨格となるモデルを設計するために大きな努力が払われていることを明らかにした。Ng et al. (2021) は、AIリテラ

*連絡先著者：齋木 匠 takum.saiki@gmail.com

シーを育成するための4つの側面「知る・理解する」、「使う・応用する」、「評価する・創造する」、「倫理的問題」を提案している。またKoch et al. (2024) は、AIリテラシーの4つの相関する潜在的な次元を発見した。これらの潜在的な次元は、(1)AIを使用し、AIと相互作用する能力、(2)AIを設計・プログラムする能力(深い技術的知識を含む)、(3)AIに関する複雑な認知操作を行う能力(倫理的配慮など)、(4)AIを検出あるいはAIと人間の区別を行い、AIの説得的な影響を管理する能力(AIの説得力に対するリテラシー)に共通する因子として解釈可能であった。Chiu et al. (2024) は、AI4futureという名称の香港の7年生から9年生を対象としたAIカリキュラムの開発プロジェクトにおいて、小学生向けのAIリテラシー・テストを開発した。このテストは、AI4futureプロジェクトのカリキュラムに基づいている。このカリキュラムは12のトピックで構成されており、AI知識、AIプロセス、AIインパクトの三つの学習分野に分かれている。AI知識では、AI技術、データからの学習方法、現実世界での仕組みについて学ぶ。AIプロセスでは、機械がどのように感覚情報を知覚し解釈するのか、また、どのように表現と推論を用いて実世界の状況において適切な判断を下すのかを学ぶ。AIインパクトでは、AI技術が私たちの社会にどのような影響を与えるのか、社会善、倫理、仕事の未来に焦点を当てて学ぶ。各トピックには、認識(基礎知識)、知識(技術的知識)、倫理(具体的な倫理問題)、相互作用(コーディング)、エンパワーメント(高度な知識)の五つのモジュールが存在する。最初の三つのモジュールは全生徒の必修科目となっており、最後の二つのモジュールは学校の選択科目となっている。上述のように、AIリテラシーについては様々な研究が行われており、様々な定義が提案されている状況である。

2.2 介入研究

AIリテラシーを涵養するための介入研究も様々な文脈において実施されている。Moorhouse et al. (2024) は、香港の大学の語学教師教育プログラムで実施された研究である。この介入研究は、言語教育に生成AIツールを使用するための明示的なトレーニングが、語学教師の専門的生成AI能力(P-GenAI-C)を向上させるかを検討した。Kajiwara and Kawabata (2024) は、大規模言語モデル(LLM: Large Language Model)の倫理的利用を学ぶカリキュラムを提案し、12歳から24歳の学生を対象に教育効果や技術受容性への影響を検証した。Ngo and Hastie (2025) は、AI for Academic Purposes (AIAP) と呼ばれる新しいフレームワークに基づいて、スコットランドのカレッジで10週間のAIを統合した学術目的のための英語教育モジュールを開発・実施することで、AIリテラシーを養うための実践上の課題に対処した。Chadha et al. (2024) は、医師とソーシャルワーカーを対象に、AIリテラシー向上を目的としたFDワークショップを開催した。このワークショップでは、ChatGPTに病気の診断をさせ、与えたプロンプトによってAIの出力がどのように変化するかを学ばせた。ワークショップの後ではAIツールを使いたいと考える人が増えただけでなく、それらのAIツールが有用であると

考える人も増えたと報告されている。

生成AIは、学習させた内容を元に毎回異なる結果を生成する。さらに、間違った内容を生成するハルシネーション(Hallucination; 幻覚)という現象があるため、利用者側が生成AIの出力を適切に判断する必要がある(Ji, 2023)。ChatGPTでは、内部モデルがGPT3.5からGPT4になった際、ハルシネーションが削減されたと報告されたが、依然ハルシネーションは発生している。したがって、特に初等中等教育においては、教師が生成AIを適切に利用するためのAIリテラシーが求められる。しかし、児童・生徒のAIリテラシー向上と比較して、教員のAIリテラシー向上に注目した研究はまだ十分に行われていない。生徒のAIリテラシーを向上させる教材の価値を高めるためには、教員側にも高いAIリテラシーが求められるため、学習者を対象としたものよりも先に検討する必要があると考えられる。

3. ChatGPT

ChatGPTは、2022年にOpenAIによって発表されたLLMである。OpenAIは、2018年に発表したGPT (Generative Pre-trained Transformer) をはじめとして、アップデートを重ね、Web上の対話型AIとしてChatGPTをリリースした。ChatGPTリリース以降、様々な企業、大学が独自のLLMを発表しており、研究も加速している。本節では、はじめに、ChatGPTをはじめとするLLMに共通する要素である深層学習、自己教師あり学習、人間によるフィードバックの三点、及び2024年5月に発表されたGPT4oで導入されたマルチモーダルモデルについて説明する。次に、LLMを利用する際の性能を大きく左右するプロンプトエンジニアリングについて説明する。

3.1 深層学習

人工知能の開発において、人間の脳構造に着目した計算構造としてニューラルネットワークがある。人間の脳では、多数の脳細胞が他の脳細胞からの信号に反応して、信号を伝えるネットワークになっている。ニューラルネットワークでは、この構造を計算機上に再現している。深層学習は、このニューラルネットワークを層状に重ねたモデルである。人工ニューロンの結合方法によって様々な処理が可能であり、自然言語処理、画像認識、音声認識など、様々な分野で利用されている。

3.2 自己教師あり学習

深層学習をはじめとする教師あり学習では、モデルの学習にあたって膨大な量の入力と正解のペアデータが必要となる。自然言語処理の分野では、入力と正解のペアになっている文章データを大量に集めることは困難であり、学習上の課題であった。

そこでGPTモデルでは、あらかじめ一部を欠損させた文章から元の文章の復元を学習させる、自己教師あり学習を用いた事前学習を導入した(Radford, 2018)。この手法では集めた文章データに対する正解、不正解のタグ付けが不要で、文章をそのまま用いることができるため、大規模データの収集コストが大幅に削減された。

3.3 人間によるフィードバック

GPT3.5では、収集したテキストを用いた学習だけでなく、人間からのフィードバックによる学習も行なわれている。これは、ChatGPTの出力に対して、人間が返答の好ましさをランク付けし、その結果を元にモデルを学習させることで、より自然な返答を生成するためである。人間からのフィードバックは返答の精度には大きく寄与しないが、人間が対話時により自然と感じる返答を生成する上で役立つことが示されている（OpenAI, 2023）。

3.4 マルチモーダルモデル

2024年9月1日時点のChatGPTの最新版、GPT-4oでは、それまで主流だった、テキスト入力を用いた命令だけでなく、アップロードした画像をテキストの命令によって処理させることなどもできるようになった。こうした複数の方法で情報を入力し、出力するモデルをマルチモーダルモデルと呼ぶ。マルチモーダルモデルの使い方の例として、画像からデータを読み取り表にすることや、音声データを文字起こしして要約することなどが考えられる。画像生成AIはStable DiffusionやDall-Eなどが既に存在するが、これらは最終生成物が画像に限られるためマルチモーダルモデルとは異なる。

3.5 プロンプトエンジニアリング

ChatGPTは与えた文章の形式、表現によって回答の質が異なることが報告されており、様々な入力上の工夫（プロンプトエンジニアリング）が研究されてきた。

1. Few-Shotプロンプティング：実施してほしいことに対して、解答例をプロンプトで与えることで、回答精度が向上する（Brown, 2020）。
2. Zero-Shotプロンプティング：解答例なしに、解答のルールだけをプロンプトで与えても、推論できる（Wei, 2021）。
3. Chain-of-Thoughtsプロンプティング：解答例に思考の流れを含めることで、複雑な推論、計算問題も回答可能になる（Wei, 2022）。

教育現場における生成AI利用では、教材として使用できる精密な回答が要求されるためプロンプトエンジニアリングの知識は重要となる。

4. 方 法

本研究では、ChatGPTの特徴とその仕組みを解説するスライドとワークショップを開発し、その効果を測定する。本節では、ワークショップに用いたスライド、実験対象、ワークショップの内容、アンケートについて説明する。

4.1 スライド

解説スライドのアウトラインを以下に示す。

- ・ AIとChatGPTの歴史
- ・ ChatGPTの会話例
- ・ ChatGPTのできるまで
- ・ 最近の動向と注意点

- ・ ChatGPTの教育応用
- ・ まとめ

解説スライドの作成にあたっては、(1)参加者はAIについて初心者であること、(2)短時間で全ての技術を詳細に理解してもらうのは困難であること、(3)参加者が生成AIを使う際に注意すべき点を知ってもらうこと、の三点を考慮した上で解説内容を選択した。実際のスライドは附録Aに掲載している。

4.2 実験対象

千葉大学で実施された集中講義「教育AI論」を受講した学生を対象にワークショップを実施し、講義前後での理解度と、生成AIに対する考え方の変化を比較した。講義には12人が参加し、内11人からアンケート結果の利用許諾を得た。

4.3 ワークショップの内容

ワークショップは講義1コマ（90分）の中で実施した。はじめに、生成AIに対するアンケートを実施し、事前知識を確認した。続いて、Steiss et al. (2024)で行われた人間のフィードバックと生成AIのフィードバックの比較研究に基づき、人間のフィードバックと生成AIのフィードバックを検討するタスクを行った。その後、本研究で作成したChatGPT解説スライドを説明した。そして、実際にChatGPTを用いて、スライド内で紹介した英語教員のためのChatGPTプロンプト集（<https://sites.google.com/office-shunsuke.com/prompts-for-english-teachers/>）などを用いて、プロンプトエンジニアリングを実際に体験するワークショップを行い、事前知識との違いや、ChatGPT解説スライドで述べた注意点などを確認するように進行した。また、ワークショップ後にアンケートを実施し、参加者のAIリテラシーの変化を確認した。

4.4 アンケート

AIリテラシーを測定する方法はいくつか提案されてきた。本研究では、Moorhouse et al. (2024)の質問紙を用いて教員養成課程の学生を対象に介入研究を行なった。Moorhouse et al. (2024)は、Professional GenAI Competence (P-GenAI-C) と呼ばれる指標を提案した。この指標では、教師の生成AIリテラシーを以下の五つに分類している。

- TP (GenAI technological proficiency; 生成AIの道具、機能に関する意識)
- PC (Pedagogical compatibility of GenAI in English language teaching; 生成AIを生徒の学習に役立てる能力(補助教材開発など))
- PW (Teachers' professional work; 生成AIの授業外利用)
- EU (Risk, well-being and the ethical use of GenAI; 生成AIのリスクに関する意識)
- PS (Preparation of students for a GenAI worlds; 生徒が生成AIを学習、余暇、将来に使いこなすために指導する能力)

表1. 事前事後アンケートの項目

	質問項目
事前事後	あなたは、テクノロジーを教育に活用する能力がどの程度あると思いますか。(5-item Likert Scale) 生成AIツール (ChatGPT, New Bing, Midjourneyなど) をどの程度知っていますか。(5-item Likert Scale) 生成AIツールは良い英語教師になるのに役立つと思いますか。 もしそうなら、それらはどのようにあなたを助けることができますか。そうでない場合、なぜ助けてもらえないと思いますか。 英語教育における生成AIツールの使用について、懸念や心配はありますか。 はいの場合、どのようなことが心配ですか。そうでない場合、不安や心配がない理由を教えてください。
事前のみ	生成AIのツールを使ったことはありますか。 もしそうなら、ツールの名前と使い方を教えてください。使っていない場合は、なぜ使わないのかを教えてください。
事後のみ	生成AIの活用と英語教育について知りたいことはありますか。(自由記述) このワークショップで学んだことは、役に立つと思いますか。(5-item Likert Scale) 上記の回答について、具体的な理由をお書きください。 このワークショップはわかりやすかったと思いますか。(5-Item Likert Scale) 上記の回答について、具体的な理由をお書きください。 このワークショップの内容に満足しましたか。(5-item Likert Scale) 上記の回答について、具体的な理由をお書きください。 生成AIの活用と英語教育について、このワークショップで何を学びましたか。

今回の実験は、講義1コマを対象としているため、長期的なAIリテラシーの計測は困難である。そこでMoorhouse et al. (2024) の提案した五つの指標のうち、生徒の将来に対する能力育成であるPSを除く、TP, PC, PW, EUの四つの指標に関するアンケートを実施し、能力を計測することとした(表1)。

5. 結果と考察

5.1 ワークショップ事前事後共通アンケート

ワークショップ事前事後で共通して、テクノロジーの教育活用能力と、生成AIツールに対する知識レベルを調査した(表2)。結果から、ワークショップ事前事後でどちらの能力、知識も高まったと参加者が考えていることがわかった。したがって、ワークショップを通してChatGPTの仕組みや強み・弱みを参加者が理解できたと考えられる。

5段階解答としたアンケート以外に表3に示す、はい、

いいえの二択から選択するアンケートも実施した。生成AIが良い英語教師になるのに役立つと考える人はワークショップ前から大半を占めていたが、特にワークショップ後では全員が役に立つと考えている結果となった。その一方で、生成AIに対して教育上の懸念を聞いた項目では、多くの参加者が懸念、心配を表明した。理由として、事前アンケートでは個人情報などのプライバシー、事後アンケートでは、スライドでも示したハルシネーションを気にするものが多かった。生成AIに対し不安を感じていない理由は「不安を感じるほど生成AIを使っていないから」という回答であった。

5.2 ワークショップ前にも実施したアンケート

ワークショップ前のアンケート項目では、全ての参加者(11人, 100%)が生成AIのツールを使用していると回答した。利用ツールについては、多くの参加者がChatGPTを利用(10人, 90.9%)していたが、一部はCopilotを利用(4人, 36.4%)しており、併用している

表2. 事前事後共通で実施し、5段階解答としたアンケートの解答結果

	Pre		Post	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
あなたは、テクノロジーを教育に活用する能力がどの程度あると思いますか。	2.18	0.63	2.45	0.69
生成AIツール (ChatGPT, New Bing, Midjourneyなど) をどの程度知っていますか。	2.18	0.63	2.81	0.40

表3. 事前事後共通で実施し、2段階解答としたアンケートの解答結果

	Pre		Post	
	はい	いいえ	はい	いいえ
生成AIは良い英語教師になるのに役立つと思いますか。	10 (90.9%)	1 (9.1%)	11 (100%)	0 (0%)
英語教育における生成AIツールの使用について懸念や心配はありますか。	10 (90.9%)	1 (9.1%)	10 (90.9%)	1 (9.1%)

参加者もいた。

生成AIの活用と英語教育について知りたいことについては、(1)テスト作成や採点におけるAI活用、(2)授業における現在の活用例、(3)協働的な学びへの活用などを知りたいといった意見があった。

5.3 ワークショップ後のアンケート

本節では、ワークショップ後に実施した5段階解答のアンケート結果について紹介する(表4)。ワークショップが役に立つか、わかりやすかったかという設問について、ほとんどの回答者が4以上の選択をした。回答の理由については、具体例が挙げられていたことや、元々知りたかったハルシネーションについて解説していたこと、理論と実践の両面の解説があったことなどが挙げられた。以上の結果から、今回のワークショップの有効性が示されたと言える。生成AIの今後の利用に関しては目立った変化は見られず、事前事後アンケートいずれにおいても多くの参加者が前向きな意見であった。このワークショップで学んだこととしては、生成AIの利用上注意すべきハルシネーションなどの問題についてと、人間と生成AIの出力結果が似ていることに対する驚きなどが見られた。

表4. ワークショップ後に実施した5段階解答のアンケート

	M	SD
このワークショップで学んだことは役に立つと思いますか。	4.64	0.50
このワークショップはわかりやすかったと思いますか。	4.45	0.69
このワークショップの内容に満足しましたか。	4.91	0.30

6. おわりに

ChatGPTをはじめとした生成AIの発展は目覚ましく、活用能力が今後情報リテラシーの一部となることが予想される。一方で、児童・生徒の情報リテラシー教育において、その基盤となる教員の生成AIリテラシーの涵養についてはまだ十分に検討されていない。本研究では生成AIの仕組みを説明するスライドを作成するとともに、講義において生成AI活用のワークショップを開催し、事前事後アンケートによって参加者のAIリテラシーの変化を調査した。アンケートの結果より、ワークショップを通して参加者の生成AIに対する意識が、より生成AIを活用する方向へと変化したことから、AIリテラシー向上に寄与したと考えられる。今後の方向性としては、より長期的な授業における教材・ワークショップの開発と評価が挙げられる。

参考文献

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger,

G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., ... Amodei, D. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems*, 33, 1877-1901. <https://doi.org/10.48550/arXiv.2005.14165>

Casal-Otero, L., Catala, A., Fernández-Morante, C., Taboada, M., Cebreiro, B., & Barro, S. (2023). AI literacy in K-12: a systematic literature review. *International Journal of STEM Education*, 10, 29.

Chadha, N., Popil, E., Gregory, J., Armstrong-Davies, L., & Justin, G. (2024). How do we teach generative artificial intelligence to medical educators? Pilot of a faculty development workshop using ChatGPT. *Medical Teacher*, 1-3. <https://doi.org/10.1080/0142159X.2024.2341806>

Chiu, T.K., Chen, Y., Yau, K.W., Chai, C.S., Meng, H., King, I., ... & Yam, Y. (2024). Developing and validating measures for AI literacy tests: From self-reported to objective measures. *Computers and Education: Artificial Intelligence*, 7, 100282. <https://doi.org/10.1016/j.caeai.2024.100282>

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y.J., Madotto, A., & Fung, P. (2023). Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys*, 55(12), 1-38. <https://doi.org/10.1145/3571730>

Kajiwar, Y., & Kawabata, K. (2024). AI Literacy for Ethical use of Chatbot: Will Students accept AI Ethics? *Computers and Education: Artificial Intelligence*, 6, 100251. <https://doi.org/10.1016/j.caeai.2024.100251>

Koch, M.J., Wienrich, C., Straka, S., Latoschik, M.E., & Carolus, A. (2024). Overview and Confirmatory and Exploratory Factor Analysis of AI Literacy Scale. *Computers and Education: Artificial Intelligence*, 7, 100310. <https://doi.org/10.1016/j.caeai.2024.100310>

Moorhouse, B.L., Wan, Y., Wu, C., Kohnke, L., Ho, T.Y., & Kwong, T. (2024). Developing language teachers' professional generative AI competence: An intervention study in an initial language teacher education course. *System*, 125, 103399. <https://doi.org/10.1016/j.system.2024.103399>

Ng, D.T.K., Leung, J.K.L., Chu, S.K.W., & Qiao, M.S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041. <https://doi.org/10.1016/j.caeai.2021.100041>

Ngo, T.N., & Hastie, D. (2025). Artificial Intelligence for Academic Purposes (AIAP): Integrating AI literacy into an EAP module. *English for Specific Purposes*, 77, 20-38. <https://doi.org/10.1016/j.esp.2024.09.001>

OpenAI, Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., Avila, R., Babuschkin, I., Balaji, S., Balcom, V., Baltescu, P., Bao, H., Bavarian, M., Belgum, J., ... Zoph, B. (2023). *GPT-4 Technical*

- Report*. <https://doi.org/10.48550/arXiv.2303.08774>
- Piech, C., Bassen, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L.J., & Sohl-Dickstein, J. (2015). Deep Knowledge Tracing. *Advances in Neural Information Processing Systems*, 28. <https://doi.org/10.48550/arXiv.1506.05908>
- Radford, A., & Narasimhan, K. (2018). *Improving Language Understanding by Generative Pre-Training*. https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf (最終アクセス日2024年10月14日)
- Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., ... & Olson, C.B. (2024). Comparing the quality of human and ChatGPT feedback of students' writing. *Learning and Instruction*, 91, 101894. <https://doi.org/10.1016/j.learninstruc.2024.101894>
- Wei, J., Bosma, M., Zhao, V., Guu, K., Yu, A. W., Lester, B., Du, N., Dai, A.M., & Le, Q.V. (2021). *Finetuned Language Models are Zero-Shot Learners*. International Conference on Learning Representations. <https://doi.org/10.48550/arXiv.2109.01652>
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q. V., & Zhou, D. (2022). Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. *Advances in Neural Information Processing Systems*, 35, 24824-24837. <https://doi.org/10.48550/arXiv.2201.11903>
- 文部科学省 (2023) 「初等中等教育段階における生成AIの利用に関する暫定的なガイドライン」 https://www.mext.go.jp/content/20230718-mtx_syoto02-000031167_011.pdf (最終アクセス日2024年10月14日)
- 緒方広明・江口悦弘 (2023) 「学びを変えるラーニングアナリティクス」 日経ビジネス・パブリケーション.

附録 A

ChatGPT概説
ー作成方法と特徴、使い方ー

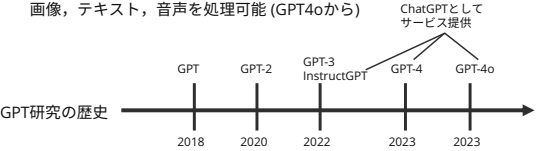
石井雄隆 齋木匠

ChatGPTとは

人工知能開発会社 OpenAIが2022年に発表したGPT技術を用いた対話型AI
最初のGPTは2018年に発表、定期的に改良版が発表
現在の最新は2024年発表のGPT-4o

特徴

- 自然な文章
- 多言語対応
- 画像、テキスト、音声进行处理可能 (GPT4oから)



会話例: よく見ると...

間違いが散見
(Hallucination)

• 伝統的に農学や生物学の分野で強みを持ち、農学部や生物資源科学研究科などが著名です。また、工学部や理学部、医学部なども含めて幅広い学部・研究科を備えています。

• 千葉大学は、産学連携や国際交流にも力を入れており、産業界との共同研究やインターンシッププログラム、留学プログラムなどを提供しています。

学部・研究分野:

- 農学部
- 法政経学部
- 園芸学部

• 千葉大学には、以下の学部があります: 農学部、教育学部、理学部、工学部、医学部、薬学部、法学部、経済学部、文学部、地域創造学部、情報科学研究科、生物資源科学研究科、先端融合科学研究科、総合人間科学研究科、国際総合科学研究科、大学院医学研究科、大学院薬学研究科、大学院工学研究科、大学院農学研究科、大学院理学研究科など。

• 研究分野としては、農学、生物学、環境科学、工学、情報科学、医学、薬学、人文社会科学など、幅広くカバーしています。

他の会話型AIとの違い

ChatGPT以降、多くの会話型生成AIツールが発表

- Google: Gemini
- Microsoft: Bing AI (内部ではGPT利用)
- Perplexity: Perplexity AI

返答のくせは異なるが、内部構造は多くが共通
本発表の内容はChatGPTに限らず最新の会話型AI全般に共通

このパワポの目的と流れ

教員に知ってほしい最低限のChatGPT知識 (内部構造)
使う上で知っておきたい、プロンプトエンジニアリング

目次

- ChatGPTの歴史
- ChatGPTの会話例
- ChatGPTができるまで
- ChatGPTの注意点 (Hallucination)
- プロンプトエンジニアリング

会話例

千葉大について聞いてみた

千葉大学について教えてください

千葉大学 (ちばだいがく) は、日本の国立大学の一つです。千葉県千葉市に位置し、首都圏における重要な教育研究拠点として知られています。以下に、千葉大学の概要と学部・研究分野、キャンパスなどについて詳しく説明します。

特徴:

- 千葉大学は1949年に設立され、現在は総合大学として幅広い学問領域をカバーしています。そのため、多様な学問分野で高い教育・研究成果を上げています。
- 伝統的に農学や生物学の分野で強みを持ち、農学部や生物資源科学研究科などが著名です。また、工学部や理学部、医学部なども含めて幅広い学部・研究科を備えています。
- 千葉大学は、産学連携や国際交流にも力を入れており、産業界との共同研究やインターンシッププログラム、留学プログラムなどを提供しています。

学部・研究分野:

- 千葉大学には、以下の学部があります: 農学部、教育学部、理学部、工学部、医学部、薬学部、法学部、経済学部、文学部、地域創造学部、情報科学研究科、生物資源科学研究科、先端融合科学研究科、総合人間科学研究科、国際総合科学研究科、大学院医学研究科、大学院薬学研究科、大学院工学研究科、大学院農学研究科、大学院理学研究科など。

会話例2: GPT4-oの場合

5月配信の最新版
精度が大きく改善
いまだに間違いあり

千葉大学について教えてください

千葉大学 (ちばだいがく, Chiba University) は、日本の千葉県千葉市に位置する国立大学です。1949年に設立され、複数の学部・大学院を持つ総合大学として知られています。以下に、千葉大学の概要を説明します。

流暢さ = 知識の正確さ
ではない

学部と学科 正解

千葉大学には以下の主要な学部があります:

- 文学部 - 人文学、社会学、心理学などの分野をカバー。
- 教育学部 - 教育学、幼児教育、小学校・中学校・高等学校教育など。
- 法政経学部 - 法学、政治学、経済学、経営学。
- 理学部 - 物理学、化学、生物学、地球科学などの自然科学分野。
- 工学部 - 機械工学、電気電子工学、情報工学、建築工学など。
- 園芸学部 - 農業、園芸、バイオサイエンスなど。
- 医学部 - 医学、看護学、保健学などの医療関連分野。
- 薬学部 - 薬学、薬科学など。

看護学部

ChatGPTの内部構造を知ることで正しい使い方を学ぶ

ChatGPTができるまで

- 人間の書いた文章 (手本) を収集しAIに学ばせる (言語モデル)
- 会話形式の応答例を人間に作らせる
- 会話形式のデータを学習させ、返答結果を人間が順位づけ
- 回答の順位を点数化 (報酬モデル)
- 回答の点数が高くなるよう言語モデルを再調整

ChatGPTの学習は全自動ではなく人手の作業が大きな役割

手本の収集と学習 (言語モデル)

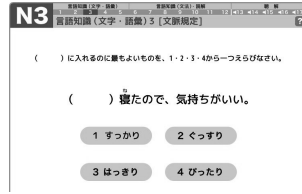
人間が右の問題を解けるのはなぜか？

1. 沢山の文章を経験
2. 親や教師からの指導

AIも同様のプロセスで学習

1. 文章データを収集
2. 問題を解かせる

PCはそのまま文章を扱えないため
言語モデルを用いる



日本語能力試験問題例より
https://www.jipt.jp/samples/forlearners.html

言語モデル

言語モデルでは各単語を数字の列 (ベクトル) として表現

例: [動物, 家畜, 植物, 野菜, 果物] の場合

犬 = [1, 0, 0, 0, 0]

牛 = [1, 1, 0, 0, 0]

レタス = [0, 0, 1, 1, 0]

りんご = [0, 0, 1, 0, 1]

類似単語は
ベクトルも類似

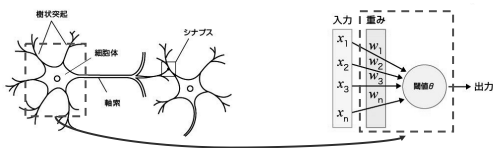
実際には人手で意味設定せず深層学習が変換を学習

深層学習 (Deep Learning; ディープラーニング)

教育学における「深い学び」とは別物

AIにおける学習: 与えられた問題に正解するよう自らのプログラムを改良

深層学習: 人間の脳構造 (ニューロン) を模したネットワークによる学習方法

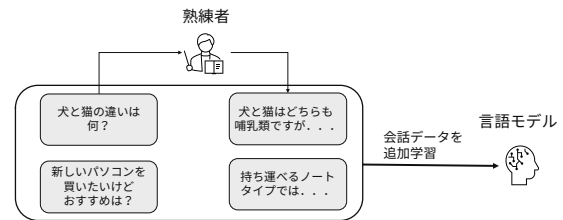


https://ipsj-catalog.jp/glossary/parlance/26.html

会話形式の学習

会話形式のデータは通常の文章に比べて少ない

人間に作らせた会話データを使って言語モデルを追加学習



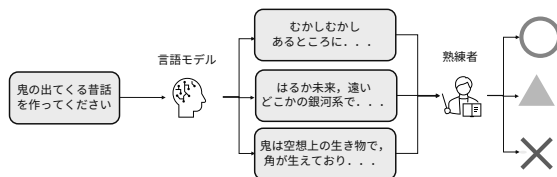
人間の基準を学習

言語モデルでは文法的に合っていないでも不自然な文章を出すことがある

そこで、言語モデルに人間の基準を学ばせる

人間の弱み: どう不自然かの説明は苦手

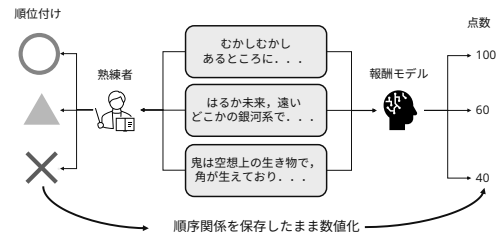
人間の強み: 一目で不自然な文章がわかる



報酬モデル

順位のままでは学習は困難

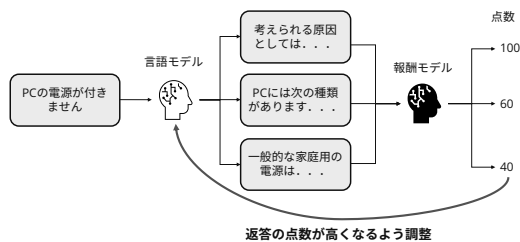
熟練者の順位を点数へ変換



人間の基準を基にした再調整

人間の比較によって学習した報酬モデルを使って

より多くの報酬を獲得できるよう再度言語モデルを学習させる



マルチモーダル学習

マルチモーダル: テキスト, 画像, 音声など複数の情報形式を扱えること

GPT-4oで導入 (詳細技術は非公表)

GPT-4oでは画像からExcelファイル作成ができる等、話題に



https://medium.com/nextbeat-engineering/gpt-4oを使って手書きの図からhtmlを出してもらった件-0814564f915b

Hallucination (幻覚)

Hallucination (幻覚): 言語モデルの生成するおかしな返答

Hallucinationの種類

- (1) 事実性 Hallucination: 事実と異なる内容を生成
1. 間違った発言
 2. 事実の捏造
- (2) 忠実性 Hallucination: 指示に即さない内容を生成
1. 指示に答えない
 2. コンテキストの読解間違い
 3. 間違った推論

GPTのバージョンアップごとにHallucinationは減っているが0ではない
元情報を提示するAI (Bingなど) の利用や、**独自チェックが必要**

国立情報学研究所 大規模言語モデルにおけるハルネーションとその対策：
https://www.nii.ac.jp/event/upload/20240717_05_thuta.pdf

プロンプトエンジニアリング

ChatGPTでは質問の工夫によって返答精度が向上
様々な工夫を総称して「プロンプトエンジニアリング」と呼ばれる

プロンプトエンジニアリングの種類:

役職指定: 「あなたは〜です」と特定の立場で考えさせる

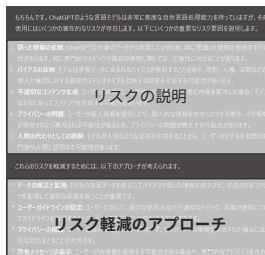
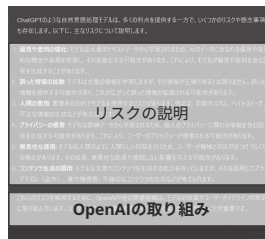
Few-Shot Prompting: 指令に加えて例を与える

Chain-of-thought: 指令に加えて、例とその解答に至った思考の流れを含める

プロンプトによる変化の例

ChatGPTのリスクについて教えてください

あなたは弁護士です。
ChatGPTのリスクについて教えてください



役職指定の例

あなたは大学のFAQの応答担当者です。以下のように質問に対し回答しました。
これらの内容の特徴づけるタグを付与してください。

質問: ~~~

回答: ~~~

Few-shot Promptingの例

I want you to create 5 fill-in-the-blank multiple-choice questions like the example below to assess students' understanding of English vocabulary. I need 4 answer options for each question, and only one of them is the correct answer. (省略) Correct answers include "単語," "単語," "単語," "単語," and "単語."

#example
His attitude is always _____, which makes everyone feel uplifted.
A) negative
B) neutral
C) melancholy
D) positive (Correct Answer)

指示

事例

<https://sites.google.com/office-shunsuke.com/prompts-for-english-teachers/#h.76ivdbz0cuus>
英語教員のための ChatGPT プロンプト集より抜粋

プロンプト資料集

- 英語教育の哲学的資料集3 (柳瀬陽介 京都大学)
<https://yanase-yosuke.blogspot.com/2023/05/chatgptbing.html>
- 英語教員のための ChatGPT プロンプト集 (高木俊輔 聖光学院中学校高等学校)
<https://sites.google.com/office-shunsuke.com/prompts-for-english-teachers>
- ChatGPT翻訳術 新AI時代の超英語スキルブック (山田優 立教大学)
<https://book.alc.co.jp/book/b10038198.html>

Chain-of-thoughtsの例

問題: ハチには6本の足があります。2匹のミツバチの足は何本でしょう？

解答: ハチには6本の足があります。2匹のミツバチがいるので、 $6 \times 2 = 12$ 、合計で12本の足があります。よって答えは12です。

問題: 現在公園には22本のクルミの木があります。(省略) 作業を終えると公園には55本のクルミの木があることになります。今日、公園作業員が植えるクルミの木は何本ですか？

解答:

今、公園には22本のクルミの木があります。(省略) $55 - 22 = 33$ 、あと33本のクルミの木を植える必要があります。よって答えは33本です。

サンプル問題と解答 (思考の流れ)

解かせたい問題

ChatGPTの出力

堀尾海斗 et al., "日本語における Chain-of-Thought プロンプトの検証," 人工知能学会全国大会論文集, vol. JSAI2023, p. 371GS602-371GS602, 2023, doi: 10.11517/jpsai.jsai2023.0_371GS602.

まとめ

ChatGPTはOpenAIの作ったGPT技術を用いた**対話型AI**

深層学習を用いた言語モデルの学習によって**高性能な言語処理を実現**

人間の基準を学ばせることで会話能力を更に向上

現在でも**間違った解答 (Hallucination)** には注意が必要

活用時にはプロンプトが返答精度に大きく影響