

対話の不確実性に鑑みた学習支援環境の設計初期検討 —対話型 AI の API 活用試行—

A Discussion for Surmounting AI Hallucination in Learning Support Systems -Trial Use of Generative AI-

松浦 健二^{*1}, 矢代 涼^{*2}, 竹内 寛典^{*1}
Kenji MATSUURA^{*1}, Ryo YASHIRO^{*2}, Hironori TAKEUCHI^{*1}
^{*1}徳島大学情報センター
^{*1}Center for AIT, Tokushima University
^{*2}徳島大学理工学部
^{*2}Faculty of Science and Technology, Tokushima University
Email: ma2@tokushima-u.ac.jp

あらまし：インターネットの膨大な情報に基づいて生成される対話型生成 AI の学習支援活用は、オープンエンドな環境とみなせる。ただし、虚偽や古い情報、矛盾等を含む場合もあり、その学習支援における活用方法の事例を早期集積と共有の必要がある。本稿では、生成された対話内容の不確実性に鑑みて、語句の置換やマスキング等の簡易手法を適用した汎用的学習支援のためのシステム試行段階での経験的な知見を速報することで今後の方向性の議論に繋げる。

キーワード：生成 AI, API, AI Hallucination

1. はじめに

2023 年は生成 AI によるイノベーションが注目され、OpenAI 社や Microsoft 社、Google 社等によるプロダクトは世界中で大きな潮流を生み出している。本稿執筆時点でも、日々新機能、新サービス、新計画が報道され、大学や自治体等の組織の方針策定、政策上の戦略決定と規制等がその開発に先んじることが難しいほどの加速で拡大傾向を示している。生成対話を信憑性の吟味なく人間の学習に使えるかという現時点での生成 AI の想定される仕組み上問題があるが、少なくともメタな層の学習でも、実務レベルでも、今後大きく影響を受ける分野として教育・学習分野があり、関係者が意識高く捉えていく必要がある。そこで、高等教育を含む学校教育や、技能訓練、社内教育、リカレント・リメディアル・リスキリング等で、その適性の検討と、活用の適否事例蓄積が急がれる。

AI の一般社会への実践投入という広義での捉え方では、例えば家電製品等にも導入されてきたが、2017 年の Transformer モデル⁽¹⁾以降 5 年ほどでの研究開発の進捗の間、生成 AI が契約約款に基づく一般サービス化、あるいは OS やブラウザへの組み込み機能化されるに至っている。極言すれば、プログラマの世界にあった高度な技術開発に対し、今や自然言語インタフェースや簡単な API を通じて、直接的に誰もが生成 AI を活用できる状況である。ただし、入力時の機密情報の取扱い、出力を利用する際の虚偽等を含む不確実性や著作権の課題についてサービス側の記載や利用組織での記載もあり、不確実性等に十分注意した適切な活用が求められる。

以上の議論から、体系的研究を今後展開としつつも、待ったなしの教育現場やその学習支援システム

研究に照らして、本稿では生成 AI 活用の知見を共有すべく初期試行・検討を報告する。

2. 生成 AI 活用検討

2.1 生成 AI と LLM

生成 AI では、その定義や特徴に関して他 AI との比較により齟齬の軽減が図れる。例えば Python 等言語環境での機械学習のコード実装や、表計算ソフトでのソルバー利用といったこれまでの AI 活用と、現在注目されている大規模言語モデル(LLM: Large Language Model)に基づく生成 AI 活用との比較議論はその特徴把握に一定貢献する。現時点では、生成 AI を Transformer モデルの発展であり LLM 適用とすると齟齬を生じないと思われる。また、DALL・E2 (<https://openai.com/product/dall-e-2>)のようにテキストではないメディア生成等の応用については現時点での対象外である。

個々のサービスでその利用規約、生成コンテンツの取扱い、研究での利用等に関して、約款の他にポリシーが公開されている場合には、それらを熟読して利用する必要がある。例えば、OpenAI 社ではコンテンツ利用ポリシーだけでなく、共有・公開に関するポリシーの中で研究面の方針も示され、研究活用時にはこれらを読んだ上での適切な活用が求められる (<https://openai.com/policies/sharing-publication-policy>)。

2.2 思考過程と出力

現・生成 AI の学習支援への貢献の可能性検討を試みる。ただし、人間の思考プロセスと生成 AI のレスポンス導出プロセスは異なると考えられる。事象毎に人間一人一人の内部評価関数が異なっている部分があることから、生成 AI の導出過程が人間内部過程の一般化と同様とは考えにくい。また、固有対

象のファインチューニング等を経た場合も、思考プロセスの異同の議論ではない。レスポンスの判別しにくさにおける AI と人間との差を明らかにするのは難しいが、その出力の特徴を捉えて応用した学習支援活用は可能と考えている。人間の内部プロセスのモデル化は、本学会等でも多数優れた研究があるが、反対にその差異を考慮しない道具としてのアプローチも多数ある。

2.3 一般的活用シナリオ

日本ディープラーニング協会が、「生成 AI の利用ガイドライン」⁽²⁾を公開した際、大別して3つの方向性が示されていた。下記(1)~(5)のうちの(1)(4)(5)に相当する。この公開後も、ローカル環境にセットアップし検証可能な Meta 社による LLaMA モデル⁽³⁾(<https://github.com/facebookresearch/llama>)や、OS への機能組込みが Microsoft 社から発表される⁽⁴⁾などもあり、本稿執筆時点では概ね下記のように大別できる。

- (1) 対話活用(検索統合含む)
- (2) OS・ブラウザ・文書ソフト等機能活用
- (3) ローカルインストール活用
- (4) API 活用の統合環境活用
- (5) 生成 AI 自体の研究開発

教育・学習での活用を捉える際には、(1)や(2)での対話的シナリオが典型的実践となるが、学習支援システム研究としては(4)により生成 AI を組込む設計への研究上の関心が高まると思われる。

3. 環境設計

3.1 フレームワーク

前節(4)については、学習者と直接の UI を介すフロントソフトウェアに対して、バックエンドの生成 AI があり、その API に対して中継する機構を設ける構成で一般化設計可能と考えている。テキスト対話を念頭にすると、同(1)(2)の実践的側面も有する。図1はそのシステム構成のモデル図である。

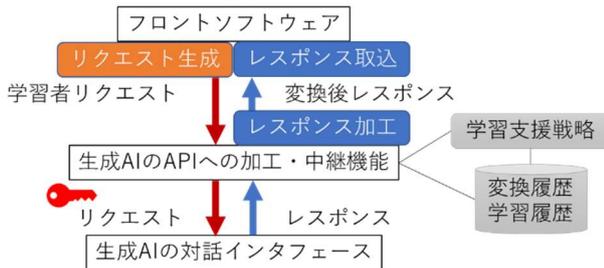


図1 システム構成図

3.2 フロントエンド

ユーザに対するフロントには、代表的なものではブラウザがあるが、Office 系文書、コード開発環境等も想定される。Copilot⁽⁴⁾の環境では直接的な組込みインターフェースが想定されるが、学習支援の環境設計上は、中継処理を設け、その入出力を制御する形での実現は柔軟性が高い。したがって、API 鍵はこの中継機能実装で利用する。中継機能は、フロントと一体化も分離実装も可能である。戻ってきたレ

スponsは、取り込む際に加工処理が必要になる。活用時には、生成 AI のコンテンツ利用明示が必要な場合はそれを明示するなどの適切な対応も行う。

3.3 API 利用中継機能

学習支援の中心的な制御機構である。この機能を Office 等では VBA で直接ファイル内に設けることも可能であるが、汎化のためプライベートなウェブ中継機構を設けることも有用である。この場合、学習戦略モジュールや履歴 DB の具備も容易になる。

3.4 ケーススタディ

(case1)スライド作成時の初期作業軽減と参考シナリオ利用を目的に、スライド編集時の支援実装を試みている。表紙スライド内のシェイプから他の生成ページへのテキストシェイプへの初期割り当ての検証であるが、文献⁽⁵⁾のような応用も検討可である。

(case2)いわゆる AI Hallucination の課題に対して、学習者とその不確かさに対するの吟味を学習に活用する。そこでは、学習支援戦略に沿った類型化の下で、例えば、特徴語を学習支援のために穴抜きや置換・位置変更等の手法が応用可能である。

(case3)問題アイテム集を事前定義し、生成 AI と学習者の出力と対比させるといった方向性も典型的応用である。例えば定積分問題などの数学分野でも、プロンプト次第で解導出過程が含まれるため、答えではなく答え方として、その導出過程を学習に活用できる。ただし、現時点ではドメインの適不適については経験的に判断していく段階である。

4. おわりに

本稿執筆時点では、学習の具体的なドメインを設定せずに環境評価の段階として汎用的なシステム設計を試みている。特に環境面での抽象的な設計指針と試行シナリオ3種を述べたが、このような枠組みでは戦略部分や履歴保存等は従来の学習支援システム設計の知見が活用できると考えている。

参考文献

- (1) Ashish, V., et al., "Attention is all you need." *Advances in neural information processing systems* 30 (2017).
- (2) 日本ディープラーニング協会(2023年5月参照)“生成 AI の利用ガイドライン”, <https://www.jdla.org/document/>
- (3) Touvron, H., et al., "LLaMA: Open and Efficient Foundation Language Models", arXiv preprint arXiv: 2302.13971. (2023) <https://arxiv.org/abs/2302.13971v1>
- (4) Microsoft blog (2023年5月参照)“Bringing the power of AI to Windows 11 – unlocking a new era of productivity for customers and developers with Windows Copilot and Dev Home”, <https://blogs.windows.com/windowsdeveloper/2023/05/23>
- (5) 松浦健二, 仁木啓司, 後藤田中, 金西計英, 矢野米雄: スライド教材の編集課題による教育・学習支援環境の試作, 電子情報通信学会論文誌(D), Vol.J91-D, No.2, pp.259-268, 2008年.