

生成系 AI と共生した文章表現基盤教育の実践：次世代医療人育成を目指して

二瓶裕之・北海道医療大学薬学部

中山章・北海道医療大学薬学部、木村治・北海道医療大学薬学部、西牧可織・北海道医療大学心理科学部

〒061-0293 北海道石狩郡当別町金沢 1757・0133-23-1265・nihei@hoku-iryu-u.ac.jp

【概要】 生成系 AI を組み込んで開発した統合型知的学修支援システムを活用して、生成系 AI が演じる学生、教員、相談役と学生が共に学ぶなど、AI と共生した文章表現基盤教育を実践した。学生自身が生成系 AI を検証しながら AI の仕組みや革新性を体験できるようにし、そのプロセスを段階的に深めた。これにより、批判的視点を持つことが AI を使いこなすスキル修得につながるとの認識を醸成でき、AI が日常となる次世代医療の現場で活躍する医療人育成を目指した基盤教育ができたものとする。

キーワード： 生成系 AI、統合型知的学修支援システム、医療人教育、文章指導、基盤教育

1. 教育改善の目的・目標

医療分野でも実用化が進む人工知能 (AI) であるが、近年注目を集めているのが大規模言語モデル (LLM) に基づく生成系 AI である。LLM が学習しているテキストデータは多岐にわたり、あらゆる視点、意見、表現方法を含む。LLM のなかでも、Open AI 社の GPT-4 モデルは 100 兆個のパラメータを持ち、回答精度も高い。GPT-4 を基にした対話型の生成系 AI である Chat GPT や Bing AI は人間が気付きにくい視点や発想にもとづいて自然な会話を生成する。GPT-4 は API (Application Programming Interface) を提供している。API を使うことで、独自に開発したアプリケーションに GPT-4 を組み込むことができ、費用も開発側が負担する。個人情報等を学習させないためのオプトアウト設定が可能である。さらに、開発側が指定したシナリオをプロンプトとして GPT-4 にロールプレイさせて、意図した役割を GPT-4 に演じさせることもできる。

教育の文脈では、生成系 AI は知的学修支援システム (ITS: Intelligent Tutoring Systems) ^[1] と強いつながりを持つ。API 経由で ITS に生成系 AI を組み込んだうえで、例えば、AI に教員を演じさせることで、学生個人のニーズに応じた学修支援などの機能を強化できると考えられる。しかし、大きな期待の一方、急速な普及のため、医療人教育の現場でも生成系 AI とどのように向き合うべきかが喫緊の課題となっている。

このような中、北海道医療大学 (以下、本学) では、DX 推進計画^[2]を策定し、様々な教育支援システムを内製化しながら個別最適化教育を進めてきた。最近では、自然言語処理の機械学習や AI 技術を使った ITS の開発や活用も始めた^[3]。さらに、数理データサイエンス AI 教育プログラム (MDASH) も全学的に実施している。特に、薬学部の MDASH では、初年次基盤教育科目である「文章指導」も取り入れて、ITS により講義ノートを自然言語処理して学生の学びを可視化するなどの DX を進めてきた。このように、中小規模ながらも多様な AI がある学びの実践を重ねており、本学では、ICT や AI を教育に活用することの理解が醸成されてきた。

そこで、今回、内製化してきた ITS に API 経由で生成系 AI を組み込んで統合型 ITS を開発し、生成系 AI が演じる学生、教員、相談役と学生が共に学ぶなど、AI と共生した文章表現基盤教育を実践することとした。鍵は、生成系 AI の留意点や欠点を単に教員が伝えるだけではなく、学生自身が生成系 AI を体験したうえで、批判的視点を持って生成系 AI の検証もするようにした点である。体験と検証のプロセスについては「気づく、培う、醸成する」の 3 つの段階で深め、段階ごとの教育改善の目的・目標を下記のように定めた：

- ① 気づく : AI の回答に対する信頼性を疑うきっかけを掴み、批判的視点を持つ大切さに気付くこと
- ② 培う : AI の回答に対する適度な信頼と健全な懐疑心を持って、批判的視点を培うこと
- ③ 醸成する : 批判的視点を持つことが AI 活用スキルの修得につながるとの認識を醸成すること

これらの段階を経て、AI の仕組みや革新性に加えて欠点に触れることで、生成系 AI との向き合い方を探求するとともに AI を使いこなすスキルを身に付けられる次世代医療人教育を目指す。

2. 授業概要と教育改善の内容

教育改善の対象とした「文章指導」は薬学部 1 年生前期の必須科目であり、2018 年度から開講した。履修者概数は 150 名、単位数は 2 単位である。授業回数は 15 回であるが、毎週水曜日の午後に 3 回連続で開講し、入学してからすぐの 5 週程で授業を終える。入学して間もない短期間での開講としているのは、「文章指導」において、アカデミックスキルと ICT・AI の活用スキルを入学当初に修得できるようにするためである。

修得するアカデミックスキルは、薬学教育に必要なレポートや論文の作成スキルと定めている。また、ICT・AI の活用スキルの修得も目指す背景には、DX 推進計画に沿った MDASH の実践がある。自然言語処理を用いた機械学習などを「文章指導」に取り入れながら、教育の DX 化を図っている。今回、開発した生成系 AI を

組み込んだ統合型 ITS は前処理、中間処理、後処理の3つのプロセスからなる。

前処理は、内製化してきた自然言語処理の機械学習システムをベースとしている。開発環境は Google colab、使用言語は Python 3 である。目指したことは、統合型 ITS への入力データとなる学生のレポートや質問などのテキストデータに対して、クラスタリングにより全体的な内容を事前に確認したり、異常値検出により不適切な内容を事前に確認したりすることである。文書間の一致度 (cos 類似度) を計算する機能も実装し、例えば、学生が提出した文章がどの程度正答に近いのかを定量化できるようにもした。

中間処理では、API 経由で生成系 AI へのデータ送信と回答受信を行う。使用した LLM は GPT-4 モデルである。目指したことは、教員が指定したシナリオに沿って生成系 AI にロールプレイさせたり、オプトアウト機能によって情報を保護したり、生成系 AI の利用コストを公費負担して学生への公平性を担保することである。生成系 AI からの回答をデータベース化することで、レポートへの剽窃行為にも対応できるようにした。

後処理は、内製化してきた教育支援システムをベースとしている。使用言語は VB Script であり、学生が統合型 ITS へログインすることで、学生の要求に応じて生成系 AI の回答を提供できるようにしている。

統合型 ITS により、AI と共生した文章表現基盤教育を実践したのは、2023 年度前期である。表 1 は、授業内容、授業形態、教育改善で定めた 3 段階の目標である。各段階では、学生が、AI を体験するだけでなく、体験の後に AI を検証するプロセスを段階的に深めながら、AI の回答への批判的視点を養えるようにした。

4 回目授業では AI の仕組みや留意点も講義したが、それを踏まえ、学生に AI を検証させる各プロセスにおいて、批判的視点を持ちながら AI の利点や欠点を自由記述のアンケートとして文章で入力させた。その検証のため、統合型 ITS には、自由記述されたテキストの集計機能を実装した。

表 1 「文章指導」の授業内容(テーマ)と改善の授業形態

授業回	授業内容 (テーマ)	授業形態	目標
1~3	事実と意見	グループワーク	気づく
4~6	ノートの取り方	課題解決型学修	
7~9	ループリック評価とピアレビュー	ライティング学修	培う
10~12	文章の読解と要約	ライティング学修	
13~15	文章指導のまとめ	探求型学修	醸成する

3. 教育実践による教育効果とその分析

3. 1. 「気づく」に向けた教育実践による教育効果：学生を演じる AI (GPT-4) との共生

表 1 の「気づく」に該当するテーマである「事実と意見」のグループワークでは、中学校で制服が良いのか私服が良いのかというテーマでディベートをさせた。ポイントは、統合型 ITS で GPT-4 を学生としてロールプレイさせ、学生を演じる AI が発想した意見も学生へ提示した点である。GPT-4 の意見は 29 件あり、そのうち 14 件は半数以上の学生が「気付かなかった意見」とした。その例の 1 つが「制服の着用によって反抗的な態度をとる生徒がいることがある」である。129 名 (87%) が「気付かなかった意見」としたが、半数以上の 69 名が「参考にならない」とした。生成系 AI の発想が学生個人の経験に基づいた発想に反することを体験したことにより、学生が AI の回答に対する信頼性を疑うきっかけを得たと考える。

同じく、「ノートの取り方」の課題解決型学修では、医療分野における AI 活用の調査というテーマで情報検索をさせた。ポイントは、グループワークで検索結果を共有し、さらに、AI の仕組みや留意点を学んだうえで、検索に適した生成系 AI である Bing AI と一緒に 2 度目の情報検索をさせた点である。はじめて生成系 AI を利用した学生が半数近くであり、興味深く操作していた。グループワークと生成系 AI の有効性の度合いを 4 件法で比較した結果、新しい知識の修得に関しては生成系 AI の利用が有意に効果的であった一方 (カイ二乗検定 $p < 0.01$)、学修意欲や理解度には差を認めなかった。このことは、AI の効果的な利用場面を想定するきっかけになったと考える。

3. 2. 「培う」に向けた教育実践による教育効果：教員を演じる AI との共生

「培う」に該当する「ループリック評価とピアレビュー」のライティング学修では、医療分野における AI 活用の調査結果をレポートにまとめさせた。ポイントは、統合型 ITS を通して、教員を演じる AI がレポートを添削した点である。図 1 (a) は教員としてロールプレイさせるためのシナリオとなるプロンプトの一部である。レポートの 4 つの章、目的、方法、結果、考察を評価項目としたループリック評価表に加え、統合型 ITS により把握したレポート全体の内容や異常値検出の結果を踏まえて作った指導方針を明記した。今回は、記述不十分な箇所も見受けられたことから、学生毎に補完箇所を最適化して指導する回答が得られるように、生成系 AI からの回答を繰り返し確認するなどの試行錯誤を経てシナリオとなるプロンプトを作った。図 1 (b) が添削の例である。学生には、AI の添削を踏まえてレポートを修正して再度提出させたが、添削を参考にする際に、AI の回答であるからこそ信頼性を検証し、批判的視点を持つことの重要性も伝えている。

```
[ "role": "system", "content": "レポートは、目的、方法、結果、考察から" ]
[ "role": "system", "content": "レポートでは、学生は、自身が設定した医" ]
[ "role": "system", "content": "目的の採点項目は、様々な「医療分野にお" ]
[ "role": "system", "content": "方法の採点項目は、インターネットを検索" ]
[ "role": "system", "content": "結果の採点項目は、複数の検索結果につい" ]
[ "role": "system", "content": "考察の採点項目は、何を感じ、何を思った" ]
```

あなたのレポートでは、医療分野における AI 活用に関心を持ち、特にスマホで診断できる医療アプリに焦点を当てていますね。目的の部分では、自身が選んだテーマの背景も説明されており、読者に興味を持たせることができていると思います。方法の部分では、検索の経緯が記載されていますが、もう少し具体的なキーワードや検索方法を追加することで、読者が理解しやすくなります。

(a) シナリオ (プロンプト)

(b) 添削の例

図 1 教員を演じる AI によるレポートの添削

図2は、4つの章ごとに、添削前後での単語数の変化を表したヒストグラムである。単語数が最も増えたのが「考察」であり、「結果」は変化なかった。「結果」には十分な記載がされていたが、「考察」はAIにより欠けていた点が補完されたことが示唆され、個別最適化した添削が効果的であったと考える。修正箇所を問うたアンケートでも、149名中54名が特に考察を深めたとし、他の学生も考察を含め内容の充実や説得力を高めたとした。添削に対するアンケートからも、7割近くが「有効性が高い、今後もあるとよい」とした一方「正確性に欠ける、褒められすぎて不安」との意見もあった。

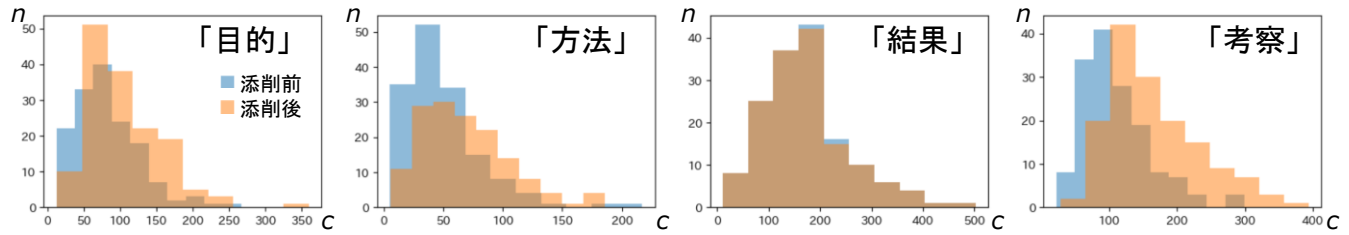


図2 単語数の変化を表すヒストグラム：縦軸が学生数(n)、横軸が文字数(c)で、青が添削前、橙が添削後
「文章の読解と要約」では、要約のコツを徐々に伝えながら、同じ小説の要約を3回提出させた。ポイントは、要約の提出のたびに、教員を演じるAIが採点した点である。要約には正答を用意しており、統合型ITSを通して正答との一致度を採点の点数として学生へ伝えた。一致度は統合型ITSにより算出したcos類似度(過去の1回目要約の平均値と標準偏差から偏差値SSに換算)であり、仕組みも学生へ伝えている。

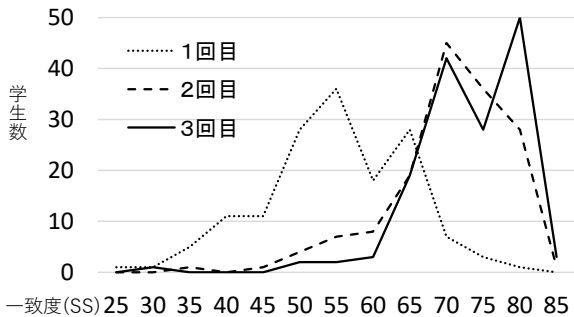


図3 要約の一致度(横軸)と学生数(縦軸)

図3は3回提出した要約の一致度の変化を表すヒストグラムである。要約を繰り返すたびに、全体的に一致度が高まった。アンケートでも、要約スキルの向上が一致度の高まりで伝わってうれしいなどの意見が多数あった。さらに、一致度算出の欠点もアウトプットさせた。回答数は152名、総単語数は約5,000個である。統合型ITSの集計機能を使って学生の意見を5つのクラスターに分けた結果が表2である。上位3件では、単語の網羅性、文のつながり、単語の言い換えにより一致度が変わってしまうことが指摘されており、単語の重要度であるTF-IDF値に基づいて算出されるcos類似度の挙動と同じ傾向を示していると考えられる。なお、統合型ITSによる添削と採点の目的は学生へのフィードバックであり、成績には用いていない。

3. 3. 「醸成する」に向けた教育実践による教育効果：相談役を演じるAIとの共生

「醸成する」に該当する「文章指導のまとめ」の探求学修では、実務家教員が、かかりつけ薬剤師などの薬剤師を取り巻く現状を講義したうえで、AIが日常となる自身の将来像を学生に探求させ、その結果を最終レポートとして提出させた。ポイントは、相談役を演じるAIが学生からの質問を受け付けた点である。

そこで検証をしたのがAIへの質問力である。AIから適切な回答を引き出すには、自身の意見や考えを明確に表現し、AIが理解できる具体的な指示を出す必要がある。今回は、統合型ITSを通して学生がGPT-4へ投げかけた質問をQ1からQ3で検証した。Q1：質問するに至った自身の考えを質問とは別の文章として記載したか、Q2：質問に対する自分の意見を述べて、それに対してAIがどう思うのかといった対比記述があるか、Q3：GPT-4の回答を制御するコマンド(例えば、詳細に、箇条書きに、会話風になどに)があるかである。表3には、各項目に該当した学生数の比率を示した。対照実験として、統制群のX学部に対して、「人工知能に聞いてみたいこと」とした質問に対する結果を示した。授業の組み立て上、「文章指導」では質問の提出時刻に制限を設けてしまいn=110名の回答となったが、全ての項目について統制群と十分有意な差がみられた。

最終レポートの記述を統合型ITSに実装したNetworkXの共起ネットワーク分析を用いて可視化した。最終レポートは2019年度から同一のテーマであり、4年分の分析(2020年度はコロナ禍の影響で中止)で抽出した単語数は約11万個である。図4には、出現頻度上位10個の名詞、年度の見出し、ノード間の重み(%)を記載した。結果、2023年度は薬剤師の次にAIとの結びつきが強く、AIから、かかりつけ薬剤師、患者、医療、薬局、薬、私、人、知識などの単語を経由して他年度とつながった。薬局など医療の現場におけるAI

表2 一致度算出の欠点とされた意見

1. 計算方法や評価基準が不明確・不十分であると感じる意見：58件
例：句読点の有無で点数が変動する、単語の網羅性だけでは不十分等
2. 文法や文章の流れ、内容の理解についても評価されるべき意見：23件
例：単語の網羅性が正答率に反映されない、文のつながりが不自然でも一致度が高い等
3. 単語の出現頻度による一致度計算に疑問を持つ意見：29件
例：言い換えた単語で正答率が下がる、重要でない単語で点数が変わる等
4. 具体的な改善点やアドバイスが欲しい意見：22件
例：どこを直せばいいかわからない、どの点で良かった・悪かったのかわからない等
5. 不正確であると感じなかった・良かったと思う意見：20件
例：正答率の算出は正確であったと思う、結構正確に採点されていた等

表3 AIへの質問力の検証

	学生数	Q1	Q2	Q3
文章指導	110	96%	28%	42%
統制群	160	24%	2%	9%

と人のつながりを考察しているものと読み取れる。

最終レポート提出時にはアンケートも実施した。回答数は149名である。設問は「Q1:将来、薬剤師として働くときに、AIとの共生で注意すべき点があれば、自由にお答えください、Q2:医療系学部である薬学部における文章指導に、生成系AIを取り入れたことについてどのように感じるか、自由にお答えください」である。各々、回答の総単語数は約8,000個と約9,000個であった。統合型ITSの集計機能を使って意見を6個のクラスターに分け、クラスターサイズの大きな3つを要約した結果が表4である。

Q1の結果から、過信しない、頼りすぎない、人の判断が大切など批判的視点を持ってAIを利用する姿勢が見受けられる。Q2でも新しい視点や効率的な学びが提供されたとの意見が多く、AIと共生した学びから、将来の薬剤師像を具体的に描くことができたとの意見も多かった。

最後に、今後の生成系AIとの向き合い方を探求する観点から、希望する学生に、AIアバターが最終レポートの講評を語り掛ける映像を提供した。図5は映像から切り出した画像である。アバターはOpen AI社のDALL E2モデル、発話映像はD-ID社のサービス(<https://studio.d-id.com/>)により生成した。学生からは、講評の内容やAIアバターの表情の変化も適切であり、「AIの今後の発展が楽しみ」などの感想を得た。

表4 最終レポート提出時のアンケート

Q1	AIを過信せず、人間のチェックと共存が重要であり、情報管理にも注意する。: 49件
	AIと共存する薬剤師は、患者とのコミュニケーションや個人情報保護に注意し、AIに頼りすぎず自己研鑽を続けるべきである。: 38件
	AIは薬剤師のサポート役であり、コミュニケーション能力や人間の判断は重要: 30件
Q2	AIの活用により、考えの幅が広がり、新しい視点や発見が得られた。: 63件
	生成系AIを使った添削や質問応答が効率的で新規性があり、学習に役立った。: 39件
	AIとの関わりを学び、医療現場での活用や将来の薬剤師像を描くことができた。: 33件



図5 AIアバターの語り掛け

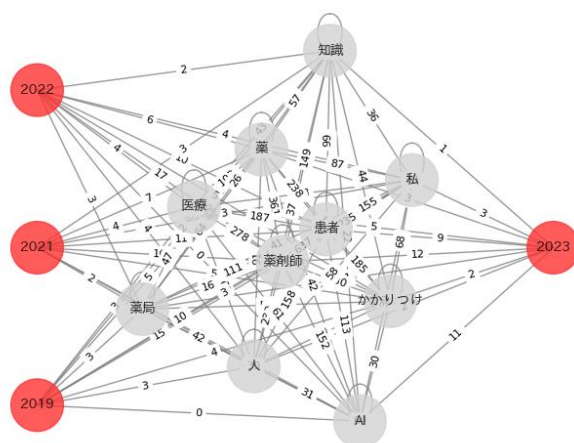


図4 共起ネットワーク分析の結果

4. 結果の考察

次世代医療人育成を目指して開発した統合型ITSを用いて、学生自身がAIの体験と検証を繰り返しながら生成系AIと共生する文章表現基盤教育を実践した。学生を演じるAIと共生したグループワークでは、学生は自分の経験と生成系AIの学習データとの違いに気づき、AIの回答に対する信頼性を疑うきっかけを掴めたと考える。教員を演じるAIと共生したライティング学修では、自分の意見がAIに添削や採点されることで、AIに対する反論や自分の見解を再評価し、TF-IDF値など自然言語処理の仕組みも理解し、AIの回答に対する適度な信頼と健全な懐疑心を培ったと考える。相談役を演じるAIと共生した探求型学修では、将来像の探求といった主観的で多面的な課題であるからこそ、相談役であるAIから、より適切な回答を引き出せるような質問の仕方を学んだ。最終レポートにおいても、薬局など医療の現場におけるAIと人のつながりの考察を深めた。最終アンケートでは、AIと共存しつつも批判的視点を持ちながらAIへの過信を避けて人の判断を大切にする姿勢が見受けられた。過信を避けながらもAIから適切な回答を引き出すなどの過程を経たことで、批判的視点を持つことがAI活用スキルの修得につながるとの認識を醸成できたのではないかと考える。しかし、AI活用スキルは1つの授業科目で醸成されるものではなく、今後、MDASHにおける情報系科目やより高度なAIに関わる専門教育科目においても取り組みを続けたい。

最後に、APIにより生成AIにロールプレイをさせる手法は、医療人教育における生成系AIの有効な利用方法の1つになると考える。また、AIアバターなどの直感的なインターフェースを介すことで、例えば、グループワークに、様々な学部にも所属する学生を演じるAIアバターを参加させれば、AIの発言にリアリティーを持たせながら学部の枠を超えた議論を展開できる可能性もある。一方、学修効果の検証や倫理的問題への対応など検討すべき事項も多い。そこで、今回、API技術も含めて、統合型ITSのpythonコードを本学DX推進計画サイト<https://dx.hoku-iryu-u.ac.jp/dx/OSS/>にCC BY 4.0で順次公開することとし、AIと共生した学びの今後の検証や教育の質向上へ寄与したいと考える。

参考文献および関連URL

- [1]林勇吾:私のブックマーク「知的学習支援システム」、人工知能学会, 33(4), 527-530, 2018
- [2]北海道医療大学DX推進計画: <https://dx.hoku-iryu-u.ac.jp/> (2023年7月23日参照)
- [3]二瓶裕之他:教育システム情報学会(JSiSE)2023年度第2回研究会, 1-1-2, 2023