
人文科学系教養教育における生成系 AI の活用法の検討

研究代表者	藤原 崇人	(文学部)
共同研究者	渋谷 亮	(文学部)
	田口 律男	(経済学部)
	竹内 綱史	(経営学部)
	手嶋 英貴	(法学部)
	手嶋 泰伸	(文学部)
	東山 薫	(経済学部)
	中田 裕子	(農学部)
	入谷 秀一	(文学部)

はじめに

現在、Chat GPT をはじめとする所謂「生成系 AI」が急速に社会に浸透しつつある。大学教育の現場においても、その存在を無視することは許されず、現実問題として学生たちに対して生成系 AI の利用を禁ずることは困難な状況になっている。

この点、本学が「生成系 AI (Chat GPT 等) の活用について」(2023.4.23 付 大学 HP News) と題して、

生成系 AI は大学での学びをより効果的に進めていくうえで有意義なツールとなる可能性があります。適切に活用することで、自らの考えを検証したり、物事の理解を深めたりするようになれることなどが期待されるからです。

と述べているように、生成系 AI の利用が一定の学習効果をもたらすことも予測され、大学教育においては、教員と学生の双方がそのメリットとデメリットをしっかりと理解したうえで適切に活用することが現実的な対応となるだろう。

本 FD プロジェクトは、上記の事柄を踏まえ、大学教育、とりわけ本部会の扱う教養・人文科学系科目の授業における生成系 AI の活用法を検討するものである。まずは生成系 AI そのものについての理解を深め、そのうえで授業や試験(含レポート)における実際の活用法や対応を研究するとともに、そこに現れた課題を抽出し、その解決策を探るものである。この研究活動の一環として年度内に研究会を 2 回開催し、計 3 人が報告を行った。以下にそれぞれの報告と議論の概要を掲げる。

1. 第 1 回研究会

日 時：2023 年 10 月 13 日 (金) 17:00～18:30

場 所：深草学舎 22 号館 303 教室 (Zoom 併用)

報告者：田口律男 (経済学部教授)

「実践報告：生成系 AI に探りを入れてみる」

【報告概要】

本報告では、現在の生成系 AI をめぐる動向と問題点が確認された後、報告者が本学で担当している授業で生成系 AI を導入して行ったいくつかの実践が紹介された。

2022 年 11 月に OpenAI が ChatGPT を公開して以降、最近では新聞報道でも生成系 AI についての記事を見ない日がないくらい、私たちの日常や教育の現場に大きな影響を及ぼしつつある。報告者の授業に出ている学生からも、普段から ChatGPT で遊んでいるとか、自分の考えを深めるのに便利だという声が届いている。文科省は 2023 年 7 月 13 日付で「大学・高専における生成 AI の教学面の取り扱いについて（周知）」という文書で注意喚起をしているが、本学もそれに先立って 2023 年 4 月 20 日付で学生向けに「生成系 AI (ChatGPT 等) の活用について」という文書で生成系 AI を適切かつ慎重に利用するようにという学生向けメッセージを公開した。

生成系 AI についてはさまざまな問題が指摘されてはいるが、大学教育の現場で、つまり授業や成績評価に関係するところで問題となるのは、以下の三点である。1. 使用の可否について、2. (使用可とするのであれば) 実際の運用方法について、3. 成績評価の妥当性・公平性について。以下ではこの三点を念頭におきつつ、授業での実践例を紹介する。

「日本の文学 A」(半期 30 回 4 単位)では日本近代文学を精読する授業をし、そのうちの一つないしいくつかの作品を取り上げて論じさせるレポート課題を出している。そのレポートで、小説の内容を全くでっち上げたようなレポートが数例あったので、manaba course を通じて「内容が支離滅裂なので、失格とします。原因は生成系 AI の不適切な使用にあると考えられます。今後二度とおなじ過ちを繰り返さないようにしてください」というコメントを送付した。これは、報告者が油断しており、学生が生成系 AI をこのように不適切な形で利用してくると思っておらず、完全に失敗した事例である。

「日本の文学 C」(半期 15 回 2 単位)では対策を講じた。この講義では初回授業で期末レポートの課題を公表しており、そこで生成系 AI で作成したレポート例を学生に示して、不適切であることを伝えた。生成系 AI では課題に対してまったく一般的で浅いことしか出力できず、授業で扱うことをまるで反映できていないので、そのままレポートを提出してもほぼ 0 点となること、特に課題では授業内容をまとめることも求めているので、生成系 AI では対応できないことを説明した。それが功を奏したのか、この授業では失格者は出なかった。

「文章表現法 A」(半期 15 回 2 単位)は上の二つの講義系科目とは異なり、アクティブ・ラーニング系科目で、創作(クリエイティブ・ライティング)を学生に実践させる授業である。この授業の最後の課題では、従来からの「短編小説を書く」「自由なテーマで自由に書く」という選択肢に加えて、「生成系 AI を使って、小説を創作させる。できあがった作品に対して、自分のコメントを付ける」という選択肢を新たに入れた。すると約 2 割の学生(21 人)がこの選択肢を選び、さまざまな反応が見られた。なかにはもちろん、生成系 AI に簡単な指示を与えただけで出てきたものをそのまま提出し、「あまり面白くない」といったコメントを付けるにとどまった学生もいたが、相当な努力(格闘)をして生成系 AI を長時間使って課題を提出した学生も数名いた。後者の学生からは、結局のところ AI に指示する能力を問われるので人間の側が AI よりも 2 歩先くらいにいないと納得のいく作品は作れないというコメントや、普通に書くことの 100 倍難しいといった意見が寄せられた。

以上の実践から暫定的にまとめると、講義系科目については、生成系 AI の利用は現段階では授業の到達目標に資するものとは言い難い(もともと、報告者の分野(文学)が生成系 AI の苦手な分野であることは留意すべき)。不適切に利用すれば大きな過誤につながることを周知徹底する必要がある。ただ、適切に扱えば、一般的な知識や基本的なリテラシーを身につけるのに役立つ可能性があるし、言語論・認知科学・テクノロジー・倫理など人文学にとって本質的な問題を考える契機となりうる。アクティブ・ラーニング系科目については、生成系 AI が現段階でも授業の到達目標に資する可能性がある。もちろん、

依存や丸投げは逆効果にしかならないが、試行錯誤を繰り返すことで主体的・批評的に生成系 AI と向き合う姿勢を学生が獲得できれば、学生自身の表現を広げる可能性を見出せるかもしれない。

【議論概要】

本報告に関する議論の概要を以下に提示する。

1. 文章表現法 A のレポート(テーマ：短編小説を書く)について、ChatGPT の使用を認めたものがあり (ChatGPT による作品と、それに対する自分のコメントを示す)、その中の一つに、ChatGPT と喧嘩をしてその実況自体を報告している学生がいて、一種のインタラクティブなやり取りとなっていて興味深かった。ただ他方で、その報告が ChatGPT 同士を議論させたものである可能性も否定できないのでは、というコメントがフロアからあった。
2. ChatGPT の裏をかくような条件を付けても、その条件をふまえたアウトプットを ChatGPT がするかもしれない→イタチごっこのような状況を呈するのでは、という疑問。
3. 文章表現法 A の成績分布について、ChatGPT を使った学生とそうでない学生との間で、成績の違いは見られたか→報告者からは、特に把握はしていない、といった回答。報告者としては、自分で書いた方が面白いとは思いますが、将来は分からない。
4. 人間の学習と生成 AI の学習とはどう違うのか、という疑問→哲学的にいうと、全く違うとは思えないというフロアの声もあった。
5. 生成 AI のヘンテコなテキストを面白いと評価する声が出るかもしれない。アバンギャルドな文学を AI が書く時代の到来?→報告者としては、読者である「私」の身体性を通すという実存的で、ある意味古典的な評価基準をしばらくは堅持したい。
6. 生成 AI の活用にかんして、龍谷大学の指針は雑であると思うがどうか、というフロアからの声もあった。

以上

2. 第 2 回研究会

日 時：2024 年 2 月 21 日 (水) 15:00~17:30

場 所：深草学舎和顔館 B108 教室 (Zoom 併用)

報告者：藤原崇人 (文学部准教授)

「人文系授業における生成系 AI の使用事例—情報精度の確認—」

妻木進吾 (経営学部准教授)

「成績評価と生成系 AI 不正利用レポート／合理的選択としての生成系 AI の不正利用」

(1) 藤原報告

【報告概要】

本報告は報告者 (藤原崇人) が担当する「人文科学セミナー」(通年, 水 3) の授業課題を事例として、生成系 AI の提供情報の精度を検証したものである。

本授業においては時事・歴史的な教養を身に着けるとともに、それぞれの個別的課題に

関する調査・考察の結果を文章化（レポート作成）することを課している。言うまでもなく学術的文章は関連する各種の情報（論説・データなど）に依拠して作成するものである。現在、WEB から情報を入手することが一般化しているが、当該の媒体から入手した情報は内容の確実性（信憑性）と持続性が必ずしも担保されておらず、WEB 上の情報に依拠する生成系 AI も当然ながら同様の問題をはらむことになる。

かかる問題点を念頭に置いて、本授業では以下のような課題を提示した。

「生成系 AI を使用して歴史上の人物や事件・事象について調べ（対象地域・時代は不問）、その内容を各種の図書や信頼できる WEB 情報を用いて精査してください」

- ・ 2,000 字程度で回答するように AI に指示すること。
- ・ 発表資料の基本構成は以下の通り。
 - ア 「題目」
 - イ 「生成系 AI の回答の全文」
 - ※修正を要する箇所には下線を引くか文字色を変える。
 - ウ 「修正した回答の全文」
 - ※検証に用いた資料情報を提示。

本報告ではこの課題のレポートを提出した受講生 12 名のうち 2 名（ともに経済学部 2 年生）のものをサンプルとして提示した。ひとは「合従軍の戦いの全容について」と題して、中国の戦国時代における所謂「函谷関の戦い」（前 241 年）について ChatGPT-3.5 に説明を指示している。当該学生は生成系 AI の回答に 6 ケ所の間違いを見つけ、資料に基づいてそれぞれ訂正を施した。生成系 AI の回答は、史実としてこの戦いに参加していない人物（司馬遷など）を参加者として挙げるなど致命的なミスが目立つ内容であった。

いまひとは「田中角栄について」と題して、故田中角栄の履歴とキャリアについて同じく ChatGPT-3.5 に説明を指示した。当該学生は生成系 AI の回答に 5 ケ所の間違いを見つけ、資料に基づいてそれぞれ訂正を施している。生成系 AI の回答は、田中の生没年・生誕地・出身校を誤るほか、有名な「日本列島改造論」についても異なる施策名を提示するなど、こちらも信ぴょう性を著しく欠く内容となっていた。

上記の課題を実施したのち受講生に対して以下のアンケートを実施した。回答者 12 名、〔 〕内は回答人数である。

- Q1. 生成系 AI の回答文の修正箇所
 ①0-2 ケ所〔6 人〕 ②3-5 ケ所〔5 人〕 ③6-8 ケ所〔1 人〕 ④9 ケ所以上〔0 人〕
- Q2. 回答文の修正に用いた資料の種類
 ①図書〔0 人〕 ②WEB〔9 人〕 ③図書と WEB〔3 人〕
- Q3. 回答文の修正に用いた WEB 資料（複数回答可）
 ①事典サイト〔7 人〕 ②学術機関サイト〔2 人〕 ③政府・自治体サイト〔5 人〕
 ④新聞データベース〔4 人〕 ⑤その他〔3 人〕
- Q4. 生成系 AI は大学での学びに役立つか。
 ①大いに役立つ〔2 人〕 ②役立つ〔7 人〕 ③あまり役立たない〔3 人〕
 ④全く役立たない〔0 人〕

- Q5. 上記 Q4 で「大いに役立つ」または「役立つ」と答えた方は、どのように役立てることができるのか具体的に教えてください。
- ・おおまかな情報をすぐに知ることができる。
 - ・全体的な知識をすばやく取り込むことができる。
 - ・文章の組み立てや構成の参考になる。
 - ・自分で間違いを見つけることで新たな知識が身につく。
 - ・文章の始まりや結びにどのようなことを書けばよいのか参考になる。
 - ・効率的に資料を探してくれる。
 - ・自分で文章を書く際に真似をすることができる。 etc.
- Q6. 上記 Q4 で「あまり役立たない」または「全く役立たない」と答えた方は、なぜ役立たないのかその理由を教えてください。
- ・自分の知りたい正しい情報をピンポイントでくれるわけではないから。結局は自分で調べる方が早い。
 - ・今回の授業でみんなの発表を聞くと、思っているよりも間違いがあり、抽象的な内容が多いから。
 - ・でたらめな情報を取り入れてしまう可能性があるから。
- Q7. 生成系 AI を大学の授業（講義・ゼミ・実習など）で活用する場合、どのような方法があるのか、学生目線での考えを聞かせてください。
- ・発表の際のテンプレートとして利用する。ただし具体的な内容は自ら補足する。
 - ・自分の知らない概念を先生が提示したときに AI に説明してもらう。
 - ・情報の真偽を見分ける能力の向上に利用できると思う。
 - ・授業の最初に大まかな情報を提供するために使用する。
 - ・生成系 AI から得た情報を自分で深掘りしていく。 etc.

本授業の課題を通して生成系 AI の提供情報を精査することの重要性は受講生も理解してくれたようである。とはいえ現段階においては概して生成系 AI に「使われている」感が否めない。とりわけ情報を精査するための知識や方法を十分に修得しているとは言い難く、丁寧な指導が必要になると感じた。ただしこの指導に関しても「人文科学セミナー」のような少人数制の演習系科目であれば比較的スムーズに行うことができるが、大人数講義科目では困難である。この種類の科目における指導方法の検討と確立が今後の課題と言える。

【議論概要】

本報告に関する主要な議論を以下に掲示する（敬称略）。

妻木：生成系 AI の提供情報の検証の内容はとても面白いと思い、ゼミか何かでやってみたいと思った。いくつか例を見せてもらったが、かなりめちゃくちゃな結果を AI が出してきていて、それをみなさんが「これはおかしい」と調べている。そのプロセスを経てもなお生成系 AI が役立つという割合がかなり高い。その内容も「おおまかな情報をすぐに知ることができる」などと言うように齟齬があるように思える。どうしてこうなるのか、不思議に思った。

藤原：「役立たない」と回答した学生もいるが、大体は間違っているにもかかわらず「役立つ」と答えている。その辺の認識のずれをどう捉えていいのか難しい。

妻木：課題はとても面白いと思ったが、この回答を評価するのはとても難しいと思った。汎用性がないのではないかな。

藤原：例えば田中角栄の件などは専門ではないので、確かにチェックするのは難しいと思う。お題によっては教員側の見落としが確かにあると思う。

田口：作家の生没年などが間違いだらけであることは気づいていた。歴史の授業で生成系 AI を使うことの危険性があると思う。ただ、学生は常識の範囲では生成系 AI が使えると思って「役立つ」と答えているのではないかな。例えば哲学的な概念だとざっと教えてくれるが、AI にも得意分野と不得意分野があるように見受けられる。

竹内：ウィキペディアの方がまだ信頼できるように思った。ウィキペディアを引用するなどと教えられてきたし、教えてきたが、生成系 AI と比べたらウィキペディアの方が圧倒的に良いように思った。

藤原：分野によるが、歴史の場合はウィキペディアも多少ましという程度のものも多い。

竹内：アンケートの中で「文章の組み立てや構成の参考になる」や「文章の始まりや結びにどのようなことを書けばよいのか参考になる」、「自分で文章を書く際に真似をすることができる」とあったが、内容をつかむというよりは体裁を参考にするというのにはあり得るかもしれない。英作文が正しいか不安な時に Google や DeepL を使うのと似ているのではないかな。

藤原：その程度の使い方に留めておいた方がいいのかもしれない。

東山：学生へのアンケートのところで、AI の回答文の修正箇所数のグラフがあったと思うが、0~2 カ所と回答した学生が 6 名いてその中で 0 だったと回答した人は 2 名という話だったと思う。この中には自分では修正箇所を見つけることができなかったという意味の 0 も含まれるのかな。

藤原：私の方ですべて検証できているわけではないが、0 と回答した学生のうちの 1 人はウォルト・ディズニーについて調べていた。ディズニーは色々な所に情報があって割と正確な情報が引き出せたと考えられる。学生も色々調べたが間違いは見つからなかったと言っていた。つまり情報の多い対象であれば割と正確なものも無いわけではない。

田口：得意分野だったということかもしれない。日本語は苦手で英語の情報がたくさんあるウォルト・ディズニーは得意分野なのだと思う。プロンプトの出し方も重要で一回だけではうまくいかない場合が多く、出てきた回答にさらに指示を加えなければならない。しかし、それを学習してくれないのが問題。

藤原：実際、学生も指示の出し方が難しいと回答している。

以上

(2) 妻木報告

【報告概要】

本報告では、報告者（妻木進吾）が担当する複数の授業で得られた情報を基に、生成系 AI 不正利用をめぐる現状が紹介され、また今後の課題が提起された。

まず生成系 AI で目下もっとも知られている OpenAI のサービス「ChatGPT」の使用状況が説明された。それによれば、2022 年 11 月の日本国内サービス開始から間がない 2023

年度前期においては、履修生がレポートにおいて生成 AI を使用する例はさほど多くなかった。また使用しても、問いへの回答に明らかな誤情報が含まれるなど、文章生成の精度が低く、実用に適したレベルになかった。しかし 2023 年度後期になると、学生による不正使用の割合が急増した。ただし内容的には、論述が全体に抽象的であり、授業で求められた設問内容に答えていないケースが目立った。なお、学年による不正使用割合の偏りは見られなかった。

つぎに生成系 AI を不正利用したと思われるレポートの成績評価をめぐる現状と課題が説明された。これまで外部情報の不正利用として多く見られたウェブからのコピペは、その証明が可能であり、不正発覚を理由として容易に 0 点評価を下すことが可能だった。それに対し生成系 AI の不正使用については、その事実を確定する証拠の入手が困難であり、疑いの域を出ない。履修生からの成績疑義に回答できる十分な根拠がないため、生成系 AI 不正利用の疑いがあっても、内容のみで評価するほかない。ただし、そうしたレポートを提出する履修生の多くは平常点が低く、結果として該当者の多くは不可となった。ただし、平常点が普通レベル以上であった一部の履修生は、レポートにおいて生成系 AI 不正利用の疑いがあっても単位は認定された。

ここで報告者から、現在提供されている複数の生成系 AI 使用確率判定ツール・サービスの試用経験が紹介された。それによれば、現場教員の実感に照らして、各ツールの間で精確性に違いが看取される。試用したもののなかでは「Smodin」というツールが、最も教員の実感と近い結果を検出したという。

さらに生成系 AI 不正使用への対応策につき二つの方法が紹介された。一つ目は「使用しても作成しにくい課題設定にする」ということである。具体例として、授業の具体的な学習内容と結びついた課題を出すことが挙げられた。二つ目は「過去学期における優秀レポートのサンプルを事前に示す」ということである。具体的には、内容が興味深く、表現面でも手本となるようなレポートを集めた資料を共有し、それにキャッチアップする意欲を喚起する。結果として生成系 AI の不正利用に頼ることを抑止することが挙げられる。

最後に報告者による今後の見通しが述べられた。ここでは生成系 AI の不正利用が現在の学生にとって「合理的選択」となっていることが指摘された。そのことを示す一例として、レポート向けの文章生成ツール「レポートお助けくん」サイトにあるサービス提供の趣旨文が紹介された。それによれば、現在の大学生は学業以外の多様な経験をもつことが重要であり、それが将来のためになるという。こうした考え方はおそらく多くの大学生に共有されている。卒業単位を出来るだけ低コストで集めたい、と考える学生にとって、生成 AI の不正利用はレポート作成のコストを下げるとともに、不正行為として認定されるリスクも低く、当然に選択される行動となる。こうした学生側のニーズの高まり、および生成系 AI のさらなる質的向上により、今後は生成系 AI の不正使用が、より拡大していくという予見が提示された。

【議論概要】

成績評価の際に生成系 AI の利用をいかに考慮すべきかについて活発な議論がなされた。その概要を以下に提示する。

1. 画像生成を用い、コンセプトを説得的に表現するなど、積極的に生成系 AI を活用し表現の幅を広げる学生もいれば、効率的に単位を取るために不適切な仕方で生成系 AI を用いる学生もいる。それゆえ、生成系 AI の利用を一律に禁じるわけにはいかず、コピペとは異なる複雑な問題があるのではないかという指摘がなされた。

2. 報告者からは、生成系 AI の不正利用に関して、AI コンテンツを検出するソフトも

あくまで推定でしかなく、それで不正認定をすることはできないという考えが述べられた。

3. これに対してフロアから、評価基準の設定の段階で、生成系 AI が作成したものに類似する文章について、実際に生成系 AI が作成したか否かを問わず、低い評価しか与えないことはできるのではないかという意見がだされた。しかし、それは難しいだろうという意見もあった。他方で、生成系 AI の文章には、具体性がない、抽象的な概念の羅列、中身がなく一般的な見解に留まるなどの特徴があり、そのような文章はいずれにせよ、高い評価は与えられないという議論となった。

4. また生成系 AI だけに頼ることのできない課題の工夫などが議論された。授業で触れた内容を取り上げさせたり、講義内容の要約をさせたりなどの、各教員の試みが紹介された。

以上

おわりに

近年の生成系 AI の急速な普及は大学教育にも大きな影響を与えており、大学教員は好むと好まざるとこれに対応せざるを得ない状況になっている。本 FD プロジェクトはかかる状況を背景に企画したものである。プロジェクト研究会の各報告によっていくつかの新たな知見を得ることができたと考えるが、現在も生成系 AI は進化を続けており、これらの知見が今後もそのまま通用するとは限らない。本 FD プロジェクトのコンセプトに基づく研究課題に継続的に取り組むことが必要と考える。

(文責：藤原 崇人)