

生成 AI を活用した卒業論文執筆支援システムの開発

林 勇 樹・上 石 圭 一

Development of a Generative AI-Based Writing Support System for Graduation Theses

Yuki HAYASHI, Keichi AGEISHI

要 約

本研究は、教員一人あたりの学生数（ST 比）が高い大学環境における卒業論文指導の効率化を目的として、生成 AI（GPT-4）を組み込んだ執筆支援システムを開発・導入し、その有用性と課題を検証したものである。従来、卒業論文の指導では文法修正や引用形式の確認など、形式面での添削に時間が取られ、研究の本質的な議論に割ける時間が十分に確保されにくいという問題があった。本システムは、学生が段階的に章立て・推敲を進める過程でリアルタイムに AI からコメントを得られるよう設計し、誤字脱字や文体統一、論理の飛躍などを指摘する。一方で、研究テーマの選定や考察の深掘りといった学術的思考の核心部分は学生と教員の対話に委ねることで、「教員の添削負担軽減」と「学生の学習効果向上」を両立する設計を追求した。

学生が本システムを利用して卒業論文を作成した際に AI が提供したコメントログを解析した結果、誤字脱字などの形式的な問題は早期に解決される一方、論理・考察面やデータの引用整理など追加作業が必要な項目は未解決や部分的な修正に留まるケースが多いことが確認された。また、教員による研究の核心部分の指導が十分に行われている学生ほど、AI からの提案を積極的に活用して論文全体を洗練させている傾向が示唆された。

本研究は、ST 比の高い大学・学部における卒業論文必修化を継続しながら、教育の質保証を図るうえで、AI を活用した執筆支援環境の整備とその有効な活用法を提示するものである。

キーワード：システム開発, 論文執筆支援, 生成 AI, 高 ST 比, 教育工学

Key words: System Development, AI-Assisted Writing, Generative AI, High Student-Teacher Ratio, Educational Engineering

1. 研究背景

大学における卒業論文は、学生にとって大学生活における学習の集大成であり、通常の講義で課される課題やレポートと異なり、独自の研究テーマに基づいて、論理的に構成した文章で表現することが求められる。卒業論文の執筆は多くの学生にとって初めての体験であり、論文構成、学術的な文章表現、引用方法など、論文特有のルールを理解することも重要な学習の要点である。従来、卒業論文の指導では、研究の内容に対するアドバイスに加えて、学生から提出された文章をチェックし、文法の誤りや、主語述語の対応、誤字の指摘などを行う必要がある。この際、学生の指導に使用する時間のうち、相対的に文章の修正に多くの時間が割かれてしまうと、指導の本質である研究の意義や方法についての議論まで十分に行えなくなるという問題がある。例として、市販されているレポートや卒業論文向けの添削スタンプの文字には、「段落冒頭 1 字下げ」、「体言止め」、「論理の飛躍」、「根拠・出典の明示」、「一文が長い」といった指摘項目が用意されている。これらの指摘項目は、教員が添削において形式面で繰り返し指導している内容を反映している^[1]と考えられる。

文部科学省の第 3 回教育未来創造会議における資料によれば、ST 比（教員一人あたりの学生数）が高い学部は、卒業論文・卒業研究必修化の実施率が比較的低いと報告されている^[2]。報告では ST 比が比較的高い社会科学系学部の中でも ST 比が特に高い学部で卒業論文・卒業研究必修化の実施率が低いことが報告されており、必修化を学部全体で実施している割合は、理工農学系で 92%、人文科学系で 78% であるのに対し、社会科学系学部で 44% にとどまっている。このことから、ST 比が高い学部における卒業論文執筆指導においては、教員の指導における負荷の軽減が求められていると考えることができる。

西野は、ゼミナール教育・卒業論文等から考える『出口の質保証』の中で、全国の大学教員に対して卒業研究・論文の質の実態を調査し、文字数や体裁などの規定が守られるレベルであるものが 52.4%、学会報告は難しいまでも研究論文として適切な中身であると言えるレベルであるものは 40.6% であるという結果を報告している^[3]。

近年生成 AI に代表される人工知能技術により、自然言語処理の分野では大規模言語モデル (Large Language Model: LLM) が実用化され、一般ユーザでも容易に利用できるレベルまで到達している。このような技術を活用することで教員が担ってきた形式面の添削を AI に代替させることが期待される。しかしその一方で、学生が生成 AI に依存し、研究論文の内容や文章を自動生成させてしまうことがあっては、学生自身が思考し文章を練り上げるという、卒業論文執筆本来の目的を損なう懸念がある。

さらに、AI の自動生成した文章が、意図せず他者の著作権を侵害するリスクも存在する。現在、国内各大学において、AI 利用に関するガイドライン整備が進められている^[4]。しかし、多くのガイドラインでは不正行為の防止に主眼が置かれており、学生が AI をどのように活用する

ことが効率的な学習につながるかという指針が示されているものは少数である。例えば、大阪公立大学高等教育開発センターが提供する「生成 AI ツールと教育についての教員向けガイド」では、学生による AI ツールの利用について、大学教育の目的や評価の公平性を損なわないように配慮することが強調されている一方で、AI ツールのもつ教育的可能性についても言及されている^[5]。従来の検索エンジンよりも踏み込んだサポートを提供できる AI ツールは、学生の自己学習を促す「家庭教師的」役割を担い得るとされ、アイデアを深めるプロセスへ利用できれば、教育の質を高める可能性があると述べられている一方、学生のデジタルリテラシーにも格差があることにも注意が必要であり、学生間の格差が成績評価に影響しすぎないように留意する必要があるとされている。このようなことから、AI からの支援を受けつつ研究論文を執筆するための統一的な環境を提供することは重要な意義がある。

以上を踏まえれば、AI を活用した執筆支援システムを開発することは、教員の負担を軽減するとともに、学生の学習効率を高める可能性があるが、論文執筆の目的である「自分で考えること」を損なわないように設計・運用する必要がある。

次章では、この課題意識を踏まえて、本研究が目指す具体的なシステム開発の方向性と目的を明確化する。

2. 研究目的

本研究の目的は、学生の卒業論文執筆プロセスにおいて、AI が学生を適切に支援しながらも、研究内容や論理構成自体は学生の自律的な思考によって構築される執筆環境を実現するシステムを開発することである。具体的には、垣田が指摘するような、学生が論文を執筆する際に生じる形式面の様々な課題（論文の構成や論理の飛躍、文法・表現の誤りなど）を AI が学生の執筆中にリアルタイムに指摘し、必要に応じて修正案を提示できるようにする^[6]。一方、研究テーマの選定や論点の深掘り、仮説の検証など学術論文における本質的な思考プロセスはあくまで学生自身が主体となって進められる設計を行う。本システムを通じて、教員が文章表現の初歩的な添削作業に煩わされる時間を削減し、研究の意義や方法論に関する本質的な指導に集中できるよう支援することで、大学現場における論文指導の質保証に寄与することが期待される。

さらに、本研究では、実際に開発した執筆支援システムを用いて学生が約3ヶ月間執筆した際に AI がどのような指摘や修正案を提示したか、学生がどの程度それを取り入れたかを解析する。これにより、システムの提示するフィードバックが卒業論文の形式面の修正に寄与しているかを検証する。特に、ST 比の高い学部において生じると考えられる教員負担の軽減を目指し、新しい論文指導のあり方を提示することを最終的な目的とする。こうした成果を踏まえ、卒業論文必修化の可否が問われる大学現場に対して、AI 活用による論文指導方法を提案することが本研究の狙いである。

次章では、AI を活用した執筆支援ツールの現状や、大学での AI 利用方針を中心に既存研究を概観し、本研究の独自性を位置づける。

3. 先行研究

AI システムを利用した支援執筆システムの実装は、英語圏向けの論文執筆支援ツールがトレンドの主流となっている。Langsmith Editor や Paperpal といった国外企業が提供するサービスは、英語の文法や語彙、スタイルを矯正することに特化しており、論文執筆そのものを支援してほしいと思うユーザが主たる対象ではない^{[7][8]}。また、大規模言語モデル (LLM) の開発自体が OpenAI や Google など巨大 IT 企業によって主導されている現状を踏まえると、彼らが提供する API は国際的に広く利用されている一方、その学習データの大半が英語中心である可能性が高く、適切な指示をせずに、提供されている LLM サービスをそのまま学生が使用することは好ましくない。

広島大学による生成 AI の活用状況に関する調査結果では、学生の多くが生成 AI を認知しているにもかかわらず、教員からの具体的な指導を受けずに独学で試行錯誤している実態が示唆されている^[9]。また、武田 (2023) の横断調査によれば、多くの大学はガイドラインを整備しつつあるものの、その多くは不正行為の防止や倫理的配慮に焦点が当てられており、教育現場における具体的な活用方法や学生の学習効率向上につながる運用指針が明確には示されていない^{[10][11]}。さらに、すでに存在する生成 AI 活用の研究として、AI に自律的に研究させ論文を執筆させるものや英語非話者に生成 AI が論文執筆プロセスに与える影響を検証したもの、が存在するが、大学生向けに卒業論文執筆環境を包括的に整備し、教員の負担軽減と学生の自律的な思考を両立させる執筆支援システム開発を行った事例は存在しない^{[12][13]}。

以上を踏まえると、先行研究の課題としては (1) 学生が論文執筆に LLM を使用することへの不安感、(2) 既存の国内ガイドラインや調査研究が、不正防止や倫理面の対応に比重を置いている一方、具体的な学習支援方法が未成熟であること、(3) 学生が生成 AI を手探り状態で利用している現状に対して、効果的な学習環境が十分に提供されていないことが挙げられる。

次章では、このような課題を解決するために、本研究がどのような方法論でシステム開発を進めるかを詳細に述べる。

4. 研究方法

本研究は、学生が卒業論文執筆において自律的に思考しながらも、形式面の課題や論理的な飛躍を AI から指摘を受けながら執筆できる環境を整備することを目的としたシステムを構築するものである。その際、論文の章立てや文章表現のスタイル、段落間の因果関係などを総合的に検

出・評価する仕組みを設計し、指摘と推敲を繰り返すプロセスを提供する。以下では、本研究において実装した機能の全体像と、学生が本質的な研究活動に専念しつつも AI の支援を適切に活用できるように配慮するための方法論について述べる。

まず、卒業論文執筆に不可欠とされる典型的な章節構成をあらかじめ提示し、それぞれのセクションで期待される内容を学生へ示す設計とした。本研究では、「背景」「先行研究」「方法」「結果」「考察」「結論」といった一般的なフローを基本形としながら、学生が独自に節や項目を追加・変更できる柔軟性を残すことで、学術論文としての体裁を保ちつつも個別の研究領域に合わせた構成調整が可能となるよう工夫している。例えば、研究の背景に相当するセクションに本来含むべき内容が別のセクションに紛れ込んでいれば、指摘と提案が生じるようにし、適切な位置に収まっているかをシステムが確認する。加えて、段落間の関係性を LLM が得意とするテキスト解析によって把握し、文意に飛躍がある場合に「前文と後続文の接続に論理的な断裂がある」などの具体的なコメントを提示する。こうした機能は、学生にとって見落としがちな論理の破綻を早期に指摘し、論文全体の整合性を保つための基盤となる。

また、文法や表現のチェックを行い、論文スタイルとして望ましい修正案や適切な言い換えの候補を自動で提示するようにする。論文作成では敬体と常体の混在、口語的表現、一文が冗長になる傾向などが頻出すると想定されるため、本研究では学生が指摘内容を容易に理解し、適宜修正を施せるよう配慮する。ただし、システム上での内容の完全な自動修正ではなく、学生が AI からの修正提案内容を参照したうえで、最終的に修正を採用するか否かを選択できるプロセスを重視することで、学習者の独自性や文章表現を損なわないよう配慮している。

これらの指摘をリアルタイムフィードバックとして実現するために、OpenAI の提供する API (GPT-4 モデル) を使用する。具体的には、学生が最後に内容を入力・編集し終えてから 10 秒後に API を介してモデルとの通信が随時行われ、AI が執筆内容を解析したコメントを即時に返すプロセスを繰り返す。学生は提示された校正・修正案を参考に推敲し、執筆を継続することで、新たな提案を得る。学生は執筆ファイルを明示的に保存することもできるが、ローカルストレージへの自動保存機能を組み合わせる設計とした。保存される内容は学生の執筆内容だけでなく、各セクションで AI からフィードバックされたコメントを含み、どのような指摘に基づいて執筆を進めてきたかを確認できるようにする。また、執筆プロセスを可視化する機能として、セクションごとの字数および論文全体の合計字数をリアルタイムに表示する。

最後に、学生のオリジナリティを守るための仕組みとして、AI に依存しすぎない執筆プロセスを確立するよう注意を払う。システムは学生が書いた文章を対象とした校正・推敲の提案を主目的とし、研究内容そのものを自動生成する機能は取敢えず設けない。具体的には、学生がセクションについてほとんど執筆していない状態では、そのセクションについては執筆を続ける旨のコメントのみが表示され、AI からの提案が行われないよう制限をかけ、あくまでユーザの入力に基づいたコメントや修正例を提示するに留める実装としている。こうした制限により、学生が思

考を放棄してしまうリスクを抑え、自らの研究内容に対する考察や論理展開を深める工程を確保して、過剰な AI 依存や不適切な利用を防ぐとともに、システムの目的である「形式面を効率的に補完し、研究内容への注力を可能にする」ことを最大限に活かす運用を目指す。

次章では、これらを踏まえた具体的なシステムの機能実装や画面構成、技術的要件を詳細に述べる。

5. システム設計と開発

5-1. システムの機能と画面構成、技術的要件

本節では、「AI を組み込んだ卒業論文執筆システム」の開発プロセスと、そのシステムに採用した設計思想、実装の概要、画面構成などについて詳述する。冒頭で示したように、高い ST 比のもとで教員が担う論文添削の負荷を軽減し、学生が段階的かつ主体的に論文を執筆できる環境を提供することが、本システムの中心的な目標である。設計にあたっては、文字数や段落構成といった形式面を AI が支援し、教員が論文内容の意義や方法論など本質的な指導に注力できる構造を念頭に置いた。さらに、論文執筆の初心者である学生が自分の研究内容を論理的にまとめあげるために、章や節を順序立てて書き進めるフローを意識した画面設計を採用している。

The screenshot displays the user interface of the AI-assisted thesis writing system. The main content area is titled "親ガチャ現象が若者の価値観に与える影響" (Impact of Parent Lottery Phenomenon on Young People's Values). Under the heading "1. 研究背景" (Research Background), there is a text input field containing the following text: "近年、日本社会では「親ガチャ」という言葉が若者の間で広く使われるようになり、注目を集めている。親ガチャとは、自分の人生や幸福が親からの遺伝的要素や生育環境などに大きく左右されるという考え方を表現したスラングである。この言葉は、若者たちが感じる社会的不平等や不確実性、人生に対する不安や不満を象徴しているとも考えられる。親ガチャという言葉の広がりや、若者の価値観や社会観にどのような影響を与えているのだろうか。この研究では、親ガチャ現象が若者の価値観や社会に対する意識に与える影響を社会学的な視点から明らかにすることを目的とする。" To the right of this text is a grey box with instructions: "研究背景の説明として、現象の概要を説明し、その後、研究問題の明確な設定が重要です。現象の例やデータを追加して説得力を強化してください。例えば、具体的な調査結果や統計データを用いて、親ガチャの影響がどの程度広範囲に感じられているかを示すと良いです。" On the far right, a sidebar contains a checklist for writing the background section: "1. 研究目的の明確化 - 現在の研究背景の最後の部分に、研究の目的を述べていますが、もっと具体的な問いを設定することをお勧めします。「親ガチャ現象が若者の価値観や社会に対する意識にどのような影響するのか」をより詳細に定義しましょう。 2. 先行研究のレビュー - 「親ガチャ」に関連する過去の学術論文や報告書をレビューし、研究の位置づけを明確にします。このセクションを追加することで、理論的な基礎を強固にすることができます。 3. 研究方法 - アンケート調査やインタビューなど、具体的な調査手法を詳述しましょう。調査対象、サンプルサイズ、質問項目、分析方法などを明確に記載することが重要です。データの収集と分析の方法についても具体的に述べることが求められます。このセクションがなければ、研究の再現性が担保されません。 4. 仮説設定 - 具体的な仮説を立て、それを検証する枠組みを設定するとよいでしょう。例えば、「親ガチャ」という認識が高い若者ほど、将来の不確実性に対する不安が強い、など具体的な仮説です。 5. 結果と考察 - 結果の提示方法やそれに対する考察を詳細に説明し、研究の意義をさらに強調しましょう。 ## 依然として存在する課題 - 曖昧な仮説設定 - 具体的な仮説を設定することが求められます。現状では問題提起に止まっており、研究の焦点がややぼけています。"

図 1：実装した論文執筆支援システムの画面構成

図 1 は実装したシステムの画面概要である。システムは、フロントエンドでの執筆画面と、サーバ側で実行される AI モジュールによるフィードバック機能から成る。学生は学習管理システム (LMS) の卒業論文演習ページから本システムを起動し、最初に論文タイトルを入力する。

システムは論文タイトルにそって、大まかな執筆方針を学生に提案する。その後、画面上の「セクションの追加」を選択すると、研究背景や研究目的といった執筆すべき項目（章）が順次提案される。提案された各セクションを選択すると、学生が記入を進めるためのテキスト入力欄が表示される。この入力欄では、Web ブラウザの localStorage 機能を用いた自動保存機能により、定期的にブラウザ内部に原稿が蓄積され、通信が不安定でも文章が失われにくいよう配慮している。執筆中は、テキスト入力欄のすぐ横で、執筆内容の変更点を踏まえた AI からのコメントが随時提供される。学生が別のセクションに移ったり、執筆をやめたと判断されると、論文全体を通してチェックする別の AI からのコメントが右側の講評エリアに提示され、内容面の評価、形式や論理構成の不備などに関する提案を受ける。また、システムの右上に「保存」および「開く」ボタンが配置されており、保存ボタンを選択すると、その時点までの文章や AI コメント情報が JSON 形式でダウンロードされる。これにより、学生は別のパソコンにデータを移し、普段と異なる環境であっても同じ論文編集環境で作業を再開できる。ファイル名は「卒論データ―日時.json」のように一意化されるので、バージョン管理が容易になり、教員への提出やバックアップがスムーズに行える。最終的な卒業論文提出においては JSON データを Tex フォーマットに変換できるプログラムを持ち、学生自身も見出しの生成や論文のスタイル修正などに煩わされることがない。論文執筆におけるすべての作業がこのシステム内で完結させられるため、教員がこれまで行っていたフォーマット整形に費やす時間が大幅に削減されることが期待できる。

本システムにおいて、教員にとって学生が書いた文面の初歩的な誤りの修正に時間を取られず、研究内容への深いコメントに時間を振り分けられるという利点が多い。しかしその一方で、AI に過度に依存してしまう学生が出ないように、利用マニュアルや画面中の説明文において、AI コメントを自分なりの考えで取り入れるよう指導することが大切である点に注意が必要である。

本章で示した設計思想と実装を踏まえ、「学生がどのように本システムを利用し、どのような成果や課題が得られたか」を検討する必要がある。次節では、学生の執筆に対する AI からのコメントデータを分類し、具体的な利用状況を分析する。

5.2. AI モデル (GPT-4) に対するプロンプト

本研究で開発したシステムでは、学生が入力したテキストや論文全体の構成を、OpenAI の API を通じて解析することで、リアルタイムのフィードバック（コメントや修正指摘）を生成している。ここで重要となるのが「プロンプト (prompt)」の設計である。大規模言語モデルはプロンプトとして与えられた文章をもとに出力を決定するため、どのように指示を与えるかが、生成されるフィードバックの品質や内容に大きく影響する。

本システムでは、GPT-4 を呼び出す際に 3 つのエージェント（フロー）を想定した設計を行い、それぞれ異なる視点・役割の「system」メッセージ（および関連パラメータ）を GPT-4 に

与えている。以下では、それぞれのエージェントとプロンプトの役割について詳述する。

5-2-1. GPT-4 におけるプロンプトの役割

大規模言語モデル (LLM) である GPT-4 は、事前学習された膨大なテキスト知識を活用し、与えられたプロンプトに応じて続きのテキストや推論を生成する仕組みをもつ。プロンプトとは、モデルに対する“文脈”や“指示”を含んだ入力メッセージのことであり、OpenAI API では主に以下のような形式で送信される。

system : システムや開発者側からモデル全体に与える基本的なルールや役割指定
assistant : モデル自身 (GPT-4) がこれまでに出力した発話
user : ユーザ側からの発話 (学生の執筆内容)

リスト1 : OpenAI API におけるプロンプト指示の役割

本システムでは、特に **system** の役割を重視し、モデルに対して「何をするエージェントとして振る舞うべきか」「どのような形式で回答すべきか」などの指示を詳細に与えている。さらに、学生の入力テキストを含む“**user**”メッセージとの組み合わせにより、校閲や論文全体の整合性チェックなど、様々な視点から執筆文章に対するフィードバックを得ている。

5-2-2. 3種類のエージェント (フロー) の概要

本研究の執筆支援システムでは、リスト2に示す3通りのエージェントを GPT-4 上でコントロールしている。それぞれに対応した **system** プロンプトを用意し、API 呼び出しごとに適切な指示を与えることで、モデルが異なる視点からフィードバックを生成できるように設計している。

1) 章立て管理エージェント (セクション構成を提案する AI)

役割 :

「新たに章やセクションを追加する際に、学生が次にどのような内容を書くべきか」をガイドする。学生が卒業論文全体を大きく見渡せるように、標準的な論文構成 (研究背景, 研究目的, 文献レビュー, 研究方法, 結果, 考察, 結論, 謝辞, 参考文献など) と照らし合わせた助言を行う。

期待する動作 :

章立て (セクション) の定型例を参照しつつ、今後執筆すべきパートについて簡潔な説明やヒントを与える。あくまで「章の構成や内容の概要」を促す立場であり、文章の具体的な執筆

や内容そのものを自動生成しないようにする。

2) ミクロ視点エージェント（セクション内の変更を監視する AI）

役割：

学生が「直前で執筆・修正した文章」の差分を解析し、文法・表現・構成などを細かく指摘する校閲・校正的な役割を担う。「直前の文章」(previous)と「今回の文章」(current)の差分を踏まえ、必要に応じて修正案を提示し、さらなる改善点を示唆する。

期待する動作：

1度に大量の指摘を羅列するのではなく、1つの指摘ポイントに絞り、学生の負担を小さくしながら段階的に推敲を促す。修正の必要がない場合は「続けて執筆してください。」と出力し、過度な自動生成を行わないように制限する。

3) マクロ視点エージェント（論文全体の変更を監視する AI）

役割：

論文全体の構成を総合的にチェックし、セクション間のつながりや論理の飛躍の有無、書き足りない章の提案などを行う。前回チェック時からの変更点によって、全体的な流れがどのように改善・改悪されたかを評価する。

期待する動作：

仮定の「教員」として、論文を通読した上での総評や、まだ執筆されていないセクションの提案を行う。タイトルや研究目的から逆算して、どの章が不足しているかを助言し、研究方法やデータの不整合があれば指摘する。部分的な文法チェックは行わず、あくまで研究全体の説得力や構成を見直す視点に特化する。

リスト2：3種類のエージェントに対する役割と期待する動作

6. データ収集

6-1. 評価のための分類

以下では、本研究で開発した AI を活用した卒業論文執筆支援システムが学生の文章に指摘した内容を分析し、その具体的傾向と分類方法を検討する。AI による卒業論文の添削や指導コメントを改めて、入力された文章の文脈や論理構造を解析したところ、様々な視点からのコメント

が行われていたことが確認された。これらのコメントを体系的に把握するための分類枠組みとして、AIからの指摘を大きく(1)構成・論旨、(2)内容・情報量、(3)論理・考察、(4)文法・表現、(5)データ・引用の5つのカテゴリーに分類した。たとえば「研究背景」「主張の説得力」「主観表現の修正」「具体性の不足」といった多様なキーワードを、この五分類にもとづいて振り分けることで、学生の論文のどの部分にAIが指摘を行ったかを分析できる。また、執筆後の文章に指摘内容が反映されたかどうかをコメント毎にA. 修正済み (Resolved) B. 一部のみ修正 (Partially Resolved) C. 未修正 (Not Resolved) の3つの基準で評価し、AIからの指摘内容に対して、学生が最終的にどのように対応したかを評価する。

実際に収集したコメントの大半は文法・表現(表記上のミス)や構成・論旨(論理の流れ)の指摘に集中しており、次いで具体例や背景説明の不足(内容・情報量)、考察の浅さ(論理・考察)が多く、引用スタイルやデータ出典の不備(データ・引用)がこれに続く構造がみられた。以下ではより具体的に学生の執筆内容と執筆に際してAIシステムから与えられたコメントログを分析し、学生にの執筆行動に対してAIシステムがどのように機能したか、またAIシステムの提案内容を学生がどの程度取り入れていたのかについて分析する。

6.2. 学生毎の執筆内容とログの解析

6-2-1. 学生 A

文章構成や論理展開への指摘に着目すると、AIは各セクションの内容が研究背景や研究目的といった定型的な流れに合致しているかを的確に検証していた。たとえば、「研究背景」で扱うべきデータが「研究目的」の部分に配置されているなど、学生の初稿に見られる全体構成の不整合に対しては「段落の移動が必要」といった指摘が行われ、学生はそれに応じて論理のつながりや段落の配置を修正していた。また、論理の飛躍を抱える箇所には「データの引用元を先に示すべき」といった助言が投げかけられており、修正後のテキストではデータの導入タイミングが適切な順序で示されるようになった。結果として、研究背景から研究目的へと至る流れが自然になり、論文全体の主張がいっそう明確になっている。

文体や表現の問題に関しては、ログデータのコメントによると、AIは口語調や曖昧表現を適切に学術的な表現に置き換えるよう提案していた。具体的には、「～だと思う」「～ではないか」といった主観的な文末に対し、「～と考えられる」「～が示唆される」といった客観的な書き方への修正を促していた。加えて、同一段落内で敬体と常体が混在しているケースにも着目し、文末表現を統一する提案が示されていた。また、データや情報量の観点から、もともと学生はデータの説明を概念的に述べることにとどめていたが、AIは具体的な例示や数値の提示が十分でない点を指摘したため、学生は具体的な数値や出典を追記していた。教員によるフィードバックを待たずして、その場で文章の問題点に気づき、修正できる点は非常に有益であると感じられた。学生がAI提案を鵜呑みにするだけでなく、「この指摘はどこが問題なのか」を自分なりに考えた

うえで再編集したケースも多く見受けられ、論文執筆を通じて文章力と論理構成力を身に付ける効果もあると推察される。

6-2-2. 学生 B

本学生の執筆ログでは、「1. 研究背景」に関する指摘が集中的に行われていた。とくに、学生が研究背景を記述する段階で、「実際の事例や統計をもう少し詳しく示すべきである」「具体的な困難を数値や事例とともに説明することが望ましい」「時系列や事例を交えて示すと説得力が増す」といった指摘がなされた。これを受け、学生は具体的な事情を追加で記述しデータとともに示すようになった。また、文中では研究ターゲットとした企業・団体名についての表記ゆれがあり、読み手にとって混乱を招く可能性があるとの指摘があった。「SNS サービス名を記述するには正式名称か略称かを統一すること」などのコメントが示されており、サービス名を「Instagram, X (旧 Twitter), YouTube, TikTok」などと明確に記載するよう修正を行った。

以上のように、学生が本システムから受け取った指摘も有効に機能したと言える。とくに、誤字脱字や文面の繰り返しといった基本的なミスがリアルタイムで指摘され、即時に修正できたことは、文章の質を高めるうえで大きな助けとなったと考えられる。さらに、研究目的や結論をどのように繋げるかに関しても、システムのコメントが章構成の整理や結論部でのまとめに寄与したといえる。

6-2-3. 学生 C

学生は既存研究を参照しながら文献の内容をまとめ、研究背景を示そうと試みていた。しかし、学生自身の研究目的との関連性の点で、どの文献がどのように自分の研究に寄与するのかが明確に示されていなかった。これに対し、AI 支援システムは「文献同士の関連性や差異を示す必要性」、「追加すべき関連文献の提案」、「研究背景と研究目的を明示的に結びつける文章構成の重要性」の指摘を行っていた。学生はシステムからのコメントを受け、一部の文献内容を追記したり新たな文献を引用したりするなど、指摘を部分的に取り入れていた。しかし、システムが強調していた「文献同士を比較・対比させる」「自らの研究課題との関連性を明確に示す」といった要点には十分に対応できていないようであった。システムは段階的に「この文献と別の文献との関係を明示する」「自分の研究がどのようなギャップを埋めるのかを示す」などの具体的コメントを提示していたにもかかわらず、学生はこれを部分的にしか活用しきれていなかった。結果として、論理の流れを整理し論文全体の説得力を高めるという本質的な改善には至らなかった。

システムからのフィードバックが十分に反映されなかった要因としては、学生の執筆プロセスにおける主体性の不足が指摘できる。コメントを受容する姿勢自体は見られるものの、「なぜその提案が必要なのか」「自らの研究のどの部分を補強できるのか」という視点を深める段階に踏み込めていなかったのではないかと考えられる。

本システムの課題として、AI システムによるコメントの履歴や内容を意識的に振り返り、それらをどのような優先度で取り入れるかを検討する手順を示していない点が挙げられる。AI シ

システムから提示された提案のうち、どれが自身の研究目的に最も直結するのか、あるいはどのコメントを優先的に反映させるべきかを整理することができず、継ぎはぎの追加情報になってしまい、全体としての論理の流れが曖昧になってしまった可能性がある。

6-2-4. 学生 D

学生の使用履歴において、自動生成されたコメントをそのまま文面に反映し、結果として冗長な表現になる場面が散見された。例えば、提案された記述を複数箇所に挿入し、段落冒頭と末尾で似た言い回しが重複しているケースが指摘された。システムのコメントは方向づけや段落構成の指針としての役割を果たすが、そのまま引用するのではなく、執筆者の文脈に合わせて調整・要約する必要がある。ログ上からは、学生が何度かシステムの指摘を修正前後で比較し、自身の文章との整合性を確認している様子が伺え、最終段階に近づくにつれて重複が減る傾向がみられた。この経緯を総合的に評価すると、本システムは書き手に必要な情報を提示できている一方で、書き手が自立的に再構成する余地を十分に理解していない初期段階では、文章構造がかえって複雑になるリスクもあることが示唆される。

本システムは論文執筆における複合的な視点の提供や段落構成には有用であるものの、利用者のリテラシーによっては、必ずしも最適な結果が得られるわけではないことが明らかになった。学生の執筆ログを検討した結果として、システムが提供する助言を鵜呑みにしないよう指導する必要性が示された。

6-2-5. 学生 E

学生は目的変数に影響を与える複数の要因を整理し、各要因の観点から説明変数の影響を分析する方法を検討していた。AI システムからは、研究方法の章でデータ収集と分析に関する情報が記載されていたものの、どのようにホームページや外部データを参照し、その正確性をどのように担保しているのかが不十分だという指摘があった。また、適用した分析手法について、結果セクションにおいて単にスコアや順位を提示するだけで、どのような意味を持ち、どのような改善策が考えられるのかが曖昧であり、実施した解析の成果が十分に活かされていないという点の指摘があった。主成分分析やクラスター分析に関する結果の解釈は、AI から「結論づけが曖昧」と指摘されていたことに対し、学生はクラスターごとの特徴の記述を追加することで、説明変数と分析結果との関係を説明しようとした。全体としては段階的に改善が試みられ、研究の妥当性を高める方向に修正が行われている点は評価できる。しかしながら、一連の指摘の中でも特に強調されていた「統計解析の説明」「参考文献の統一」「考察の具体化」といった内容への改善は部分的にとどまった。本システムを用いて研究を進める際、こうした繰り返しの校閲機能やログの活用が、研究論文の完成度を高めるうえで有効であると考えられる。

6-3. 学生の執筆内容とログ解析のまとめ

本章では、開発した AI 執筆支援システムを学生が利用して卒業論文を執筆した際に AI が提

供したコメントを、学生の最終提出稿とともに検証した。具体的には、各学生がどのような視点で AI からの指摘を受け、その内容をどの程度修正・反映できたかを分析し、文章構成や論理展開・文法表現・引用など多面的な課題解決に AI がどのように寄与しているかを考察した。

いずれの事例でも、誤字脱字や文体の統一、冗長表現の削除といった校閲的な面については、AI が提示する修正提案を比較的容易に取り入れるケースが多く見られた。文法や表現上の誤りに関しては修正が明確であり、学生も「提案をそのまま導入するかどうか」を選択しやすいため、未解決が少なく解決率が高い傾向がある。これにより、教員が従来の添削で割いていた基礎的な文章チェックの時間を削減し、研究内容に関するより本質的な指導へ時間を充てやすくなる利点を確認された。

一方、論文の構成・論旨や論理・考察といった、より高度な思考プロセスを要する面では、学生によって AI コメントの取り入れ方や修正の度合いが異なり、改善にも差が見られた。たとえば「研究背景」や「研究目的」の段階的な組み直し、先行研究との関連づけ、考察の深掘りなどは、システムが複数回コメントを提示しても、学生が「どのように優先順位をつけて修正するか」を主体的に判断しなければ十分に解決しきれない傾向がある。一部の学生は、AI 指摘を部分的にしか反映できずに終わり、論文全体の流れがやや曖昧なまま提出に至ったケースも確認された。これは、AI コメントがあくまで「どう直せばよいかの指針」を示すにとどまり、大幅な論理構成の再構築には学生自身の理解と取捨選択が求められることを示している。

また、データの根拠づけや引用スタイルの不備は、AI からの指摘があっても修正が後回しになりやすいという特徴が確認された。こうした課題では、AI が必要な修正案を提示していても、追加の調査や引用整合性のチェックなど学生側の実務的作業が増えるため、解決を先送りしてしまう状況がうかがえる。これらは教員との進捗確認や優先度付けを組み合わせることで改善が期待されるが、完成に向けたチェックリストを提示するなど、学生のタスク管理を支援するための体系的な仕組みづくりも今後の課題といえる。

学生のケースの中には、AI コメントをそのまま大量に取り入れた結果、同じような表現が段落冒頭と末尾で重複するなど、逆に冗長化する例が確認された。リアルタイムに多くの提案が与えられる環境は裏を返せば学生の安易な判断に起因する「指摘の鵜呑み」というリスクもある。逆に指摘を踏まえて論文の構成や論理を自分なりに再編集することで、段階的に文章が洗練され、具体的なデータ追記や事例比較を行うなど、より説得力のある論述につながった事例もあった。したがって、AI のコメントがすべて正解というわけではなく、学生が必要な助言を選択しながら執筆を進める必要があるというリテラシー的な指導も並行して行うことが重要であると言える。

次章の定量分析でも扱うように、本質的な研究内容を教員と議論する機会の多い学生ほど、AI コメントの解決率が高くなる傾向が示唆された。これは、研究の着眼点や考察の方向性について学生が理解を深めていけば、AI からの形式面・構成面の助言を的確に取り込めるためと考

えられる。逆に、研究内容への理解や思考が浅いまま本システムを使っても、どのコメントを優先すべきか判断が難しく、最終的には部分的・表層的な修正に終わる可能性がある。このことは、AI システムだけに頼らず、教員やゼミの仲間との対話など、「研究の内容を掘り下げる」ことに意義があることを示している。

今回のログ解析からは、AI が段落構造・文法表現・引用スタイルなど形式面を中心とした幅広い指摘を即時に行うことで、学生が初歩的なミスや構成上の欠陥に自発的に気づき、修正を繰り返す学習プロセスが活性化する様子が確認された。一方で、論理展開の大幅な再構築や研究内容の深掘りといった、論文執筆の本質的かつ高難度の部分には、依然として学生自身の主体的取り組みと教員の個別指導が不可欠であることも示された。ログ解析から、その目的が一定程度達成されていることが示唆された。

7. データ解析

7-1. AI コメントと学生の対応について

本研究で開発した AI 執筆支援システムは、研究背景や文献レビュー、研究方法、考察など各章の執筆過程で、学生が入力した文章に対しリアルタイムで指摘コメントを提示し、修正案を参照できるようにするものである。本節では、システム上で生成された「AI コメント」に対して、学生がどのように対応したかを定量的に把握するため、学生の最終執筆論文と AI から与えられた段階的なコメントの対応状況を OpenAI の GPT-4 モデルに入力し、課題が未解決 (Not Resolved) / 部分的解決 (Partially Resolved) / 解決済み (Resolved) のいずれかであるかを判定した。それら 3 種類ステータスと、その指摘が属する項目を分析した。具体的には、24 名の学生が約 3 ヶ月間にわたって本システムを用いて行った卒業論文執筆データに含まれる AI コメント履歴から、(1) どの「章」でどのようなステータスになりやすいか (図 2)、(2) どの「指摘分類」でどのようなステータスになりやすいか (図 3) をクロス集計し、さらにそれらの解決状況にグループ毎で偏りがあるかを、カイ二乗検定と標準化残差を用いて検証した。

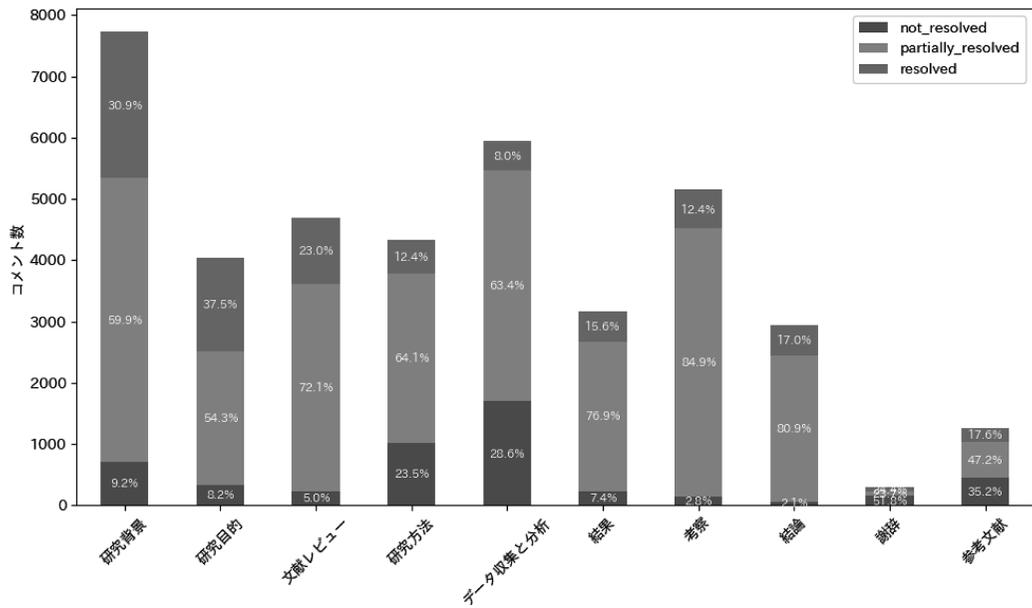


図 2：章ごとの AI からのコメント数と学生の対応状況

特に、AI からのコメント内容は多岐にわたるため、評価を容易にするため、「構成・論旨」「内容・情報量」「論理・考察」「文法・表現」「データ・引用」の 5 つの項目に再分類を行った。

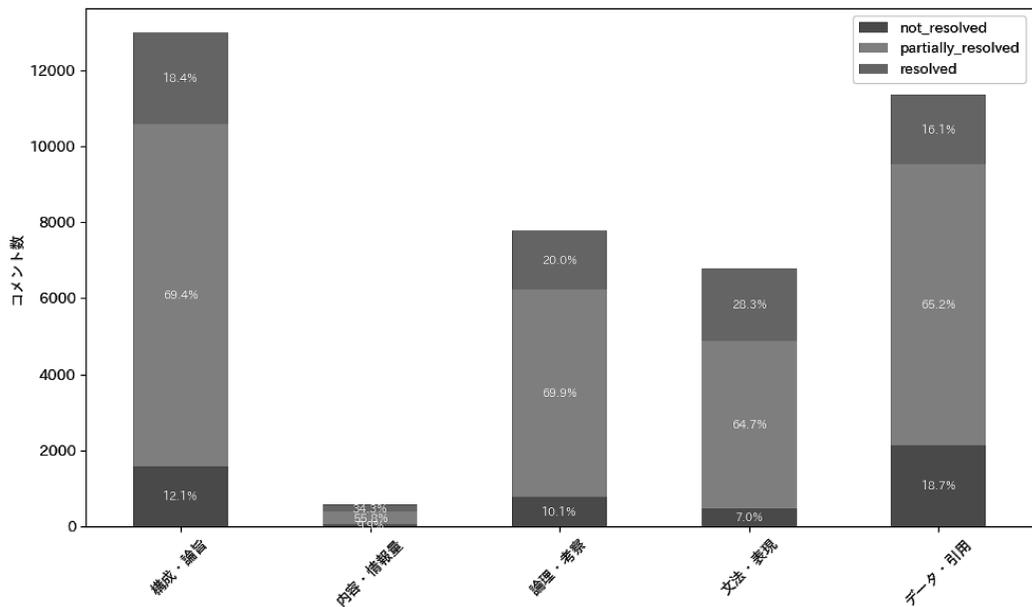


図 3：AI からの指摘内容に対する学生の対応状況

表1：各章×ステータス別クロス集計の標準化残差

章	Not Resolved	Partially Resolved	Resolved
研究背景	-8.69	-7.86	21.32
研究目的	-8.01	-10.04	24.79
文献レビュー	-14.87	4.01	4.5
研究方法	19.87	-2.54	-11.18
データ収集と分析	34.28	-3.62	-20.67
結果	-8.32	6.64	-5.54
考察	-20.03	15.41	-12.28
結論	-16.19	9.01	-3.62
謝辞	18.98	-9.17	1.69
参考文献	22.42	-8.7	-1.91

表2：指摘分類×ステータス別クロス集計の標準化残差

章	Not Resolved	Partially Resolved	Resolved
構成・論旨	-1.8	2.97	-4.01
内容・情報量	-1.88	-3.35	7.64
論理・考察	-6.44	2.81	-0.01
文法・表現	-13.27	-2.59	15.33
データ・引用	17.96	-2.74	-9.29

7.1.1 章ごとのステータス傾向

はじめに、「研究背景」「研究目的」「文献レビュー」「研究方法」「データ収集と分析」「結果」「考察」「結論」「謝辞」「参考文献」の10章について、ステータス分布（Not Resolved, Partially Resolved, Resolved）をクロス集計しカイ二乗検定を実施した。その結果、 $\chi^2(18)=6089.392$, $p < .001$ となり、章によって解決状況の構成比に大きな差があることが示された。また、Cramer's V（クラメル連関係数）は0.278と算出され、中程度の関連が認められると考えられた。表1の標準化残差を参照すると、たとえば「研究背景」「研究目的」はResolvedの残差が大きくプラス（21.32, 24.79）で、Not ResolvedとPartially Resolvedがマイナスの値をとる。これは、学生が論文執筆初期に背景・目的をある程度しっかり書き込み、AIコメントを取り入れながら修正を加えた結果、早い段階で「Resolved」ステータスに至ったと考えられる。

一方、「研究方法」「データ収集と分析」「参考文献」などはNot Resolvedの残差が非常に高く（研究方法：+19.87, データ収集と分析：+34.28, 参考文献：+22.42）、未解決のまま残るケースが多いことが示唆された。特にデータ収集と分析は、具体的な実験計画や分析作業が必要

となり、学生が AI コメントを参照しても実行の難しさなどから修正できなかった場合が多いと考えられる。また、「謝辞」や「参考文献」のように本質的内容（研究の意義や考察）と離れた部分でも、AI コメントからの修正が取り込まれにくい傾向がうかがわれる。

さらに、「考察」に関しては Partially Resolved の残差が高く（+15.41）Resolved の残差は低い（-12.28）。考察においてあいまいな言い回しや具体例不足を指摘されても、一部のみの修正にとどまり、最終的に解決に至らなかった可能性がある。

7.1.2 指摘分類ごとのステータス傾向

次に、AI コメントを「構成・論旨」「内容・情報量」「論理・考察」「文法・表現」「データ・引用」といった指摘分類別に集計したところ、カイ二乗検定の結果は $\chi^2(8) = 984.740, p < .001$ となり、指摘分類ごとにもステータスに有意な差異があることが示された。Cramer's V は 0.112 となり、弱い関連があると認められた。表 2 の標準化残差を参照すると、たとえば「文法・表現」は Resolved が多く（+15.33）、Not Resolved が少ない（-13.27）。これは、表記の統一や主語述語の修正など、学生が AI からのコメント指摘を取り入れやすい分類で解決へ至るケースが多かったことを意味する。対照的に、「データ・引用」は Not Resolved（+17.96）と非常に高く、Resolved が少ない（-9.29）。この要因としては、データ不足や引用形式の誤りを修正するには、研究データそのものを取得し直したり、文献調査をしたりと追加の手間を要するため、修正を後回しにしたまま放置された可能性がある。

7.1.3 解釈と今後のシステム開発への知見

以上の結果から、AI コメントと学生の対応状況には「どの章を執筆しているか」「指摘内容がどの分類に属するか」で大きな偏りがあることがわかった。Cramer's V の値（章ごと：0.278、指摘分類：0.112）から、特に章間のステータス差はある程度大きく、指摘分類別の差異はやや小さいながらも有意な関連が認められる。具体的には、「研究背景・目的」は解決されやすい一方、「研究方法」「データ収集と分析」、そして「参考文献」などは未解決が多く残る。学生へのヒアリングでは「修正を行ってもまた新しい指摘が来るのでどこまで反映していいかわからない」という声が得られており、AI からの指摘を取り入れようとする学生ほど負担が大きくなっている可能性があり、一つの修正が終わるまで別の項目を提示しないなどの制限を行う必要性が示唆された。また、指摘分類の観点では、文法や表現上の問題は修正難易度が低く解決率が高い反面、データ不足や引用不備のように追加調査を要するケースでは未解決で終わりがちである。また、「構成・論旨」や「論理・考察」のコメントは、ある程度対応がなされるものの、最終的に完全解決には至らず Partially Resolved 止まりとなる傾向がみられる。論理展開を大きく再構成する場合、学生の思考負荷が高まるため、AI コメントの存在だけでは十分に対処しきれない面があると考えられる。

これらの知見を踏まえると、システム面では「研究方法」「データ・引用」など修正に手間がかかる領域について、段階的に対処を促すよう UI や学習タスクを組み込むことが有効であると

考えられる。また、論理・考察の深掘りが不十分なまま提出されるケースを減らすには、AIが提示するコメントを優先度順に提示する機能が必要だと考えられる。一方で、学生が執筆内容や研究そのものについて抜本的な再構成を行うには、教員との対話やアドバイスが必要であり、AIコメントを適切に活かすための教員指導も重要であるといえる。

7.2. 学生の使用状況と教員の指導状況および評価

本節では、コメント毎に学生の最終執筆論文とAIコメントを照らし合わせて「学生がAIコメントをどの程度解決 (Resolved) または部分的に解決 (Partially Resolved) したか、解決できなかったか (Not Resolved)」の3段階でGPT-4モデルが評価したデータを用いて、学生がAIコメントにどれだけ対応できたかを示す「解決率 (Resolved Rate)」(コメント総数のうち、AIが解決したと判定した割合)を、説明変数とした分析結果を報告する。また本研究の目的である「教員の学生への指導の中心が研究における本質的なものであったか」が最終的な論文評価にどのような影響を及ぼしているかを考察するため、「本質的な指導」という変数を作成した。これは教員からの指導の中心が研究活動の本質に関わるものであったと言える場合を1、形式的な点への指導が中心とならざるを得なかった場合を0として、教員の主観で評価した指標である。以上の2つの説明変数を用いて、論文の出来映えを示す教員による5段階評価の「最終的な判定」を目的変数とした統計的解析を実施した。

7.2.1 相関分析と検定結果

本研究で得られたデータに対して、まずスピアマンの順位相関係数を算出した結果を表3に示す。「最終的な判定」と「Resolved Rate」に対して中程度の正相関 ($r \approx 0.36$) が得られたが、いずれも5%有意水準を満たさず ($p \approx 0.08 \sim 0.11$)、「有意な相関」と言い切るには至らなかった。一方、「本質的な指導ができた (0/1)」と「Resolved Rate」間では、スピアマン相関で+0.44程度とやや高い相関があり、さらにマン・ホイットニーのU検定(ウィルコクソンの順位和検定)では $p \approx 0.037 < 0.05$ という結果が得られている。つまり、本質的な指導を受けた学生ほどAIコメントに対してより多くの修正を行い、結果的に未解決が少ない(resolved率が高い)傾向が確認された。また、「最終的な判定」と「Resolved Rate」の間には弱め～中程度 ($r \approx 0.38$) の正相関があるものの、マン・ホイットニーのU検定では $p \approx 0.069 > 0.05$ であり有意とは言えない。これは、今回の分析対象者が24名と少ないこと、あるいは「AIコメントを解決しても最終評価を左右するのは研究内容そのものの質」といった他要因が存在する可能性が考えられる。

表 3：スピアマンの順位相関係数

	本質的な指導	最終的な評価	Resolved Rate
本質的な指導	1	0.384455	0.441386
最終的な判定	0.384455	1	0.335204
Resolved Rate	0.441386	0.335204	1

7.2.2 解釈と指導上の示唆

本質的な指導 (=1) が行われた学生ほど Resolved Rate が有意に高い (マン・ホイットニーの U 検定: $p < 0.05$) ことから, AI コメントを活かした修正・解決が進んでいる学生は, 形式的な指摘にとどまらず, 論理構成や内容面でも適宜 AI の助言を受け取ることで, 自身の考えを深めるプロセスも時間を費やしていたと考えられる. 一方で, Resolved Rate が高いほど最終的な判定も若干上がる傾向はあるものの, 有意な差とまでは言えない結果だった. Resolved Rate を上げるには, AI コメントの指摘を逐一修正していく必要があるが, 研究の意義やオリジナリティは学生の思考に加え, 教員からの指導で深める部分であるため, AI からのコメントに対する Resolved Rate だけではカバーしきれない要素が最終評価を左右しているとも考えられる.

今回のログ分析と統計検証によって, 本システムは教員の補助ツールとして有効に機能し, 学生の執筆プロセスを円滑化する効果があることが示唆された. 教員が「指導の中心が研究の本質的なものであった」と認識した学生は解決率 (Resolved Rate) が有意に高かった点を踏まえると, ST 比の高い学部においては, 教員がすべての学生に研究の核心部分での指導を行う時間を確保するためにも, 本研究で開発した執筆支援システムのような環境を整備した上で, AI コメントによる指摘を学生自身が優先順位をつけて活用できる指導を通常の指導に加えて行うという AI システムの利用を前提とした運用が, 本システムの利点を最大化し, 学生の学習効率を最大化する方法と言える.

7-3. 本システムの運用コストとスケール

本研究で開発したシステムでは, OpenAI が提供する GPT-4 (gpt-4o-2024-08-06) を継続的に利用した. API へのリクエストは, トークン数に応じた料金が発生する. 学生が本システムを使用した期間の費用を集計したところ, \$641.58 となった. (日本円で約 10 万円)

教育現場や研究室単位で本システムを運用する場合, トークン単価×使用回数で運用コストが積み上がるため, 本研究においては 24 名の学生が 3 ヶ月間使用したが, より大きな規模のゼミや大人数の演習でシステムの運用を行うと, さらにコストかかる点に注意が必要である.

8. 考察

本研究では、生成 AI を活用した卒業論文執筆支援システムを開発・運用し、AI からのコメントや学生の執筆状況を分析することで、システムの有効性と課題を検証した。従来、教員が多くの時間を割かざるを得なかった文法表現の指摘や段落構成の確認といった形式面の指導を AI が担うことで、教員が「研究の本質」への指導に注力しやすくなる点は、ST 比（教員一人あたりの学生数）が高い学部における卒業論文指導の効率化に大きく寄与すると考えられる。その一方で、システムから得られるデータや AI コメントの活用方法によっては、学生の学習効果や論文評価が左右されるため、教員ごとの指導方針や学生のデジタルリテラシーを踏まえた運用に注意が必要である。

8.1. システムにおける AI コメントの活用について

7章の分析結果から、AI コメントに対する学生の「解決率」(Resolved Rate) は「文法・表現」カテゴリで比較的高く、一方「研究方法」「データ収集と分析」「データ・引用」といった追加作業や思考が必要な領域では「未解決 (Not Resolved)」が多く残る傾向が明らかになった。これは、文章の表記揺れや誤字といったすぐに修正ができる問題ほど解決率が高く、統計処理や文献調査といった手間のかかる指摘は後回しになりやすいことを示唆する。また、「構成・論旨」や「論理・考察」のコメントは部分的な対応のみで終わってしまうケースが多かった。論理展開を大幅に作り直す作業は学生にとって負担が大きいため、AI の提案を十分に取り入れられないまま最終提出を迎えてしまう場合があると推察される。また、「研究背景」「研究目的」など初期段階の章は解決率が高めである一方、後半の「考察」や「結論」でコメントが未解決に終わる割合が上昇する傾向も確認された。これは論文後半で研究全体のまとめや、思考の言語化など、単なる執筆にとどまらない複合的な作業を要するため、学生が自力で修正しきれなかった可能性があると考えられる。また「謝辞」「参考文献」のように「本質的な内容からは離れた」章も未修正のまま放置されることがあり、細部の調整に対して学生のモチベーションが下がる点がかかわれる。

これらの結果は、リアルタイムに提示される AI コメントを学生だけが確認するのではなく、教員との個別指導の中で学生自身の目的意識として取り入れて、一緒に解決を目指すことで学生の自主性を残しつつ、論文の完成度が高くなることを示唆する。

8.2. 「本質的な指導」の重要性について

7.2節で示した統計解析によると、AI の指摘（コメント）をどの程度解決できたかを示す「Resolved Rate」と学生の最終的な論文評価（5段階の教員評価）とのあいだに、中程度の正相関は確認されたものの、有意な相関とは言えない結果となった。一方、教員が「研究の核心的部

分を指導できた」と判断した学生ほど解決率が高くなる傾向が、マン・ホイットニーの U 検定を通じて有意に示された ($p < 0.05$)。これは、学生が AI コメントを効果的に取り入れて論理構成や先行研究との関連づけを深めるために、研究内容について教員と繰り返し議論する機会、すなわち、本研究の目的である本質的な指導のための時間の確保が必要であることを示唆している。AI からの部分的な提案だけでは「なぜその修正が大事なのか」まで理解できない学生も少なくないため、教員が論文全体の流れを踏まえて学生を導くプロセスを省いては、AI の支援効果が限定的になる。AI 支援により教員の添削負担が軽減された分、議論の要点や思考の深掘りに時間を割きやすくなることが期待される。AI による論文執筆上の形式面での指導と、教員による研究における本質的な指導との有機的な組み合わせが重要であると考えられる。

8.3. デジタルリテラシーの格差について

本研究では、AI コメントへの対応度合いを分析することで、システムの効果検証を試みた。しかし、このようなシステム生成の数値指標を学生評価に直接反映すると、リテラシーによって評価が左右される可能性がある。たとえば、AI コメントを素早く取り入れる能力や修正プロセスのコツを掴む力は、学生ごとのデジタルリテラシーの差によって変動する可能性があるため、「高い解決率 = 高評価」として評価基準に組み込むことは望ましくない。教員側は AI からのコメントや解決率をあくまで学生の執筆プロセスを把握する指標として位置づけ、本来の研究の質・考察内容・探究プロセスといった本質的観点を軸に評価を行う必要がある。

8.4. システム活用と持続的な教育モデルの構築

本研究の結果からは、AI 支援型の執筆システムが論文の形式面における質向上と教員の負担軽減に大きく貢献する一方、学生の主体的思考力を十分に育成する指導体制や、AI コメントを単なる自動修正ではなく「研究目的と結びつける」プロセスへと導く教員の働きかけが不可欠であることが確認された。AI の持つ校正能力や論文構成支援の利点を活かしつつ、学生同士や教員との会話による、研究の本質的議論を疎かにしない仕組みを整えることが必要である。

8.5. まとめと今後の展望

本研究で開発した執筆支援システムが示す AI コメントは、文法・表現の誤りや文章構成上の問題点を早期に可視化する点で高い有用性を持ち、高い ST 比に起因する教員の添削負担を軽減しながら、学生にとって学習効率の向上をもたらすことが確認された。一方、考察の深掘りや研究の独自性を伸ばすには、AI の提示を自動修正にとどめず、指摘内容を研究全体と関連づける主体的な学習が不可欠であり、教員の本質的指導と補完し合う形で初めて最大の効果が得られる。特に、学生のデジタルリテラシーの格差にも注意が必要であり、AI コメントに頼りすぎず、研究テーマや論理性を深めるための指導とともに、システムを使用した演習機会を提供し、

AIによる指摘内容を取り入れつつ執筆する経験を卒業論文執筆よりも前の段階で経験させるなど、学生のICTスキル差が不公平につながらないように配慮する必要がある。

本研究が提示したシステムを卒業論文指導の現場で安定的に運用し、継続的な検証を重ねることで、高いST比の下でも卒業論文必修化を維持しつつ、論文指導の質と効率を同時に高めるモデルが確立できると期待される。また、研究成果を超えて学生自身の主体的な思考を鍛える場として卒業論文の意義を再確認するうえでも、AIツールとの共存は今後ますます重要性を増すであろう。今後は、大学全体の生成AIに対するガイドラインを注視しつつ、学生・教員双方にとって利点の大きい指導環境の構築を行っていくことが重要である。

9. 結論

本研究では、教員一人あたりの学生数（ST比）が高い学部における卒業論文指導の負荷軽減と学生の学習効率向上を目的として、生成AI（GPT-4）を組み込んだ執筆支援システムを開発・導入し、ログ分析と教員評価を通じてシステムの有効性を検証した。その結果、以下の知見と今後の課題が得られた。

1) 形式面の添削の有効性

システムがAIコメントをリアルタイムに提示することで、学生が平易な修正を実施しやすい環境が構築された。これにより、教員の添削作業を削減し、研究内容そのものの意義や方法論の指導に注力できる可能性が示唆され、本研究の目的の一部が達成された。

2) 論理・考察面へのAI支援の限界

考察部分の論理展開など、高度な思考プロセスを要する領域に関して、AIが示すコメントの取り入れ方や優先度が学生ごとにばらついた。特に考察の深さやデータ解析に関する指摘は未解決のまま残ることが多く、本システムの支援の限界を示唆するものとなった。

3) AIコメントと論文評価の関連性

教員が研究の核心部分を十分に指導できた学生ほどAIコメントの解決率が高くなる傾向が有意に示されており、AI支援と本質的な指導が機能することで、論文の完成度が向上する可能性が示唆された。

4) 持続的なシステム運用とコスト面の課題

GPT-4の利用にはトークン量に応じた費用が発生し、運用規模が拡大するほどコストも増加する。ログ分析では約3か月・24名の学生利用で10万円程度のAPI使用料が発生しており、

学部全体への導入を検討する際には予算配分や利用頻度のコントロールを考慮する必要がある。

謝辞

本研究の論文執筆にあたり、OpenAI 社の ChatGPT (o1-pro/ o1/ GPT-4) を用いてデータ解析／文章表現推敲を行った。ただし、卒業論文執筆支援システムの設計、システム運用におけるデータの取得は、すべて著者自身が行ったものであり、最終的な論文の内容に関する責任は著者にあることを明示する。

参考文献

- [1] レポート・卒業論文向け添削スタンプ「文章ゼミ」,
<https://store.shopping.yahoo.co.jp/hankos/bunshozemi-01.html> (2025 年 1 月 12 日閲覧)
- [2] 「出口における質保証」に関する参考資料, 文部科学省,
https://www.mext.go.jp/content/221220-mxt_koutou01-000026647_10.pdf (2025 年 1 月 13 日閲覧)
- [3] ゼミナール教育・卒業論文等から考える『出口の質保証』, 西野毅朗,
https://www.mext.go.jp/content/20221027-mxt_kikakuka01-000025675-4.pdf (2025 年 1 月 13 日閲覧)
- [4] 生成 AI の利用について, 文部科学省,
https://www.mext.go.jp/a_menu/other/mext_02412.html (2025 年 1 月 19 日閲覧)
- [5] 生成 AI ツールと教育についての教員向けガイド, 大阪公立大学 高等教育開発センター,
https://www.omu.ac.jp/las/highedu/publication/generative_ai/index.html (2025 年 1 月 13 日閲覧)
- [6] 垣田裕介, 2014, 「論文作成の形式集」『福祉社会科学』, 第 4 号, 39-56 頁
- [7] Langsmith Editor, <https://ja.langsmith.co.jp/> (2025 年 1 月 13 日閲覧)
- [8] Paperpal (2025). 英語論文執筆 AI – 研究者のための学術論文執筆ツール, <https://paperpal.com/ja>
- [9] 卒論・修論作成における生成 AI 活用状況に関する調査結果報告, 広島大学, 2023,
<https://www2.media.hiroshima-u.ac.jp/public/survey/gaiusage2023/gai-usage-2023.html> (2025 年 1 月 13 日閲覧)
- [10] 武田俊之, 2023, 大学は生成系 AI の影響をいかに認識しているか?. 日本教育工学会研究報告集, JSET2023-2-A14, 87-94.
- [11] 武田俊之, 日本の大学の AI 利用方針, ガイドライン, <https://osf.io/esju3> (2025 年 1 月 13 日閲覧)
- [12] The AI Scientist: Towards Fully Automated Open-Ended Scientific Discovery, Lu, Chris and Lu, Cong and Lange, Robert Tjarko and Foerster, Jakob and Clune, Jeff and Ha, David, arXiv preprint arXiv: 2408.06292, 2024
- [13] Wang, Chaoran. (2024). Exploring Students' Generative AI-Assisted Writing Processes: Perceptions and Experiences from Native and Nonnative English Speakers. *Technology, Knowledge and Learning*. 1-22. 10.1007/s10758-024-09744-3.