

研究課題	大規模言語モデルのAPIを活用した中等生向け個別最適化英語ライティング指導の革新
副題	～1人1台端末で実現するリアルタイムフィードバックによる学習成果向上のための拡張可能なシステム設計～
キーワード	生成AI、LLM API、英語ライティング、形成的フィードバック、個別最適化
学校/団体名	公立群馬県立中央中等教育学校
所在地	〒370-0003 群馬県高崎市新保田中町184
ホームページ	https://chuo-ss.gsn.ed.jp

1. 研究の背景

近年、生成AIは教育現場において急速に存在感を高めており、英語ライティング指導においても、文法、語彙、構成、論理展開などに対する即時フィードバックの可能性が注目されている。一方で、中等教育段階の通常授業において、こうした技術を安全かつ継続的に運用し、学習成果の向上と授業改善の双方に結び付けた実践研究は十分に蓄積されているとは言い難い。

本校では、生徒の英語パラグラフ・ライティングに対して、授業内で十分な添削と書き直しの機会を確保することが難しく、教師による個別フィードバックの頻度や即時性に限界があった。そのため、生徒が「指摘を受ける→修正する→再提出する」という反復的な改善サイクルを十分に回しにくい状況が続いていた。

こうした背景のもと、本研究では、本プロジェクトの研究者自身が OpenAI の API を活用して開発した教育支援プラットフォーム「Arizona AI」を授業内に導入し、1人1台端末環境のもとで、生徒が自らの英文に対して即時かつ反復可能なフィードバックを受け、赤ペン修正と清書へとつなげる学習循環を構築した。Arizona AI は、単一の英語ライティング添削ツールではなく、将来的に他校へ展開可能であるだけでなく、ライティング支援以外のAIツールも組み込める拡張的な基盤として設計された。実際、本年度の研究では、ライティング支援機能に加えて、チャットボット型支援や定期考査振り返り支援ツールなど、複数の並行ツールの試行も進められた。

以上のことから、本研究は、生成AIを英語ライティング指導に導入すること自体を目的とするのではなく、中等教育において、生成AIを活用した個別最適化フィードバックがどのような条件のもとで学習成果の向上と授業改善に資するのかを明らかにする点に意義がある。

2. 研究の目的

本研究では、大規模言語モデルの API を活用した英語ライティング支援システムを中等教育の授業内に実装し、その教育的有効性と実装上の条件を明らかにすることを目的とした。特に、1人1台端末環境において、生徒が自らの英文に対して即時かつ反復的なフィードバックを受け、修正・再考・清書へとつなげる学習循環が、英語ライティングの質的向上にどのように寄与するかを検証することを目指した。

具体的には、次の三点を明らかにすることを目的とした。

1. 生成AIを活用した授業内フィードバック循環が、生徒の英語ライティング得点の向上に寄与するか。
2. 生徒はAIフィードバックをどのように受け止め、どのような点に価値や課題を見いだしたのか。
3. 生成AIを活用した英語ライティング支援を学校現場で継続的に運用するためには、どのような設計上・運用上の条件が重要か。

本研究の独自性は、生成AIを単発的な教材としてではなく、授業内に組み込まれた反復的・即時的・個別最適化フィードバックの仕組みとして検証した点にある。また、学習成果だけでなく、生

徒の認識変化、自由記述、実装上の失敗と改善過程まで含めて検討することで、他校が参照可能な実践モデルを提示することを目指した。

3. 研究の経過

本研究では、授業実践、システム開発、質問紙調査、ライティング評価、校内外への発信を並行して進め、量的・質的データの双方を蓄積した。研究の途中では、AI採点機能の実装も試みたが、安定した構成を確立するには至らず、本研究期間内では範囲を超える課題であると判断した。加えて、当初の計画では高校1年生および2年生を対象としてデータ収集を行う予定であったが、同一年度内に両学年を十分な精度で追跡することはやや過大であることが明らかになった。そこで、本研究では対象を高校2年生に絞り、年度後半以降は、複数モデルを用いた二段階・カテゴリ別プロンプトによる質的フィードバック改善と、その教育的効果の検証に重点を移した。

なお、統制群については、同一出版社の同一教科書を用い、対応するライティング課題に取り組んだうえで、教師による書き込み型の訂正フィードバックを中心とする通常の指導を実施した。ただし、多忙な校務の中で、全生徒の全課題に対して即時かつ詳細なフィードバックを継続的に行うことは現実的に困難であり、指導は可能な範囲にとどまる場面も少なくない。本研究では、このような時間的制約とフィードバック頻度の限界を含む授業内指導を対照条件として位置づけ、AIによる常時的・即時的フィードバックとの対比を行った。

時期	取り組み内容	評価のための記録
R7.5.1～5.16	Arizona AI運用開始。Prompt v1導入；ベースライン・ライティング評価実施	授業観察記録、生徒提出物、システム運用記録；ベースライン得点（2名の盲検評価者による採点）
R7.5.23～5.27	ライティング実践①；S01質問紙調査	生徒ワークシート、AIフィードバックログ、授業観察記録、質問紙回答（n=101）
R7.6.16～6.17	ライティング実践②；S02質問紙調査	生徒ワークシート、AIフィードバックログ、授業観察記録、質問紙回答（n=109）
R7.7.14～7.15	ライティング実践③；S03質問紙調査	生徒ワークシート、AIフィードバックログ、授業観察記録、質問紙回答（n=96）
R7.7.31まで	React.js/Firebase構成への移行完了	システム更新記録、授業実施記録
R7.8～R7.9	システム調整、授業内運用の安定化、校内外での実践共有	研修記録、参加者反応、授業観察記録
R7.9.1～R7.9.9	実験群ミッドライン評価	ミッドライン得点（実験群）
R7.9.19～R7.9.24	ライティング実践④；S04質問紙調査	生徒ワークシート、AIフィードバックログ、授業観察記録、質問紙回答（n=93）
R7.10.20～R7.10.24	統制群ミッドライン評価	ミッドライン得点（統制群）
R7.10.31～R7.11.19	ライティング実践⑤；S05質問紙調査。Prompt v2.1導入。不具合によりフィードバックが曖昧・一般化	生徒ワークシート、AIフィードバックログ、授業観察記録、質問紙回答（n=94）、自由記述
R7.11.25	Prompt v2.1の不具合修正	システム修正記録、授業記録
R7.12.19	Prompt v2.2を安定運用版として本格導入。ライティング実践⑥期間開始；S06期間開始	システム更新記録、生徒ワークシート、AIフィードバックログ、授業観察記録、質問紙回答（n=95）、自由記述

R8. 1. 21	OCR機能を授業内で導入	授業観察記録、生徒反応、自由記述
R8. 2. 17～ R8. 3. 1	ライティング実践⑦; S07質問紙調査	生徒ワークシート、AIフィードバックログ、授業観察記録、質問紙回答 (n=109)
R8. 2. 16～ R8. 2. 20	エンドライン評価	エンドライン得点、BME比較、ICC
R8. 3	データ分析および報告書作成	統計分析結果、自由記述分析、図表、報告書草案

4. 代表的な実践

本研究では、研究期間全体（2025年5月～2026年2月）を通して安定的に実施した「3サイクル型ライティング指導」を代表的実践として取り上げる。対象は、群馬県立中央中等教育学校の5年生（高校2年生）に相当する実験群生徒であり、教科書の単元内容やパラグラフ・ライティング課題に対応した通常授業内で実施した。

本実践のねらいは、AIを用いた即時フィードバックを授業内に位置付けることで、生徒が自らの英文の弱点を把握し、それを具体的な修正行動へとつなげる力を育成することである。

(1) **授業の時系列構成**：本実践は、3つの授業時限にまたがるライティング循環として実施した。第1時限では、ウォームアップとして 1st Try を行い、通常15分程度で初稿を書かせた。第2時限では、2nd Try として再考・再作成を行い、ここでも通常15分程度を用いた。第3時限では、Arizona AI によるフィードバックの確認、赤ペン修正、Final Draft の作成、ならびに質問紙回答までを含む一連の活動を、55分授業1時間を使って実施した。

(2) 学習の流れ

① 1st Try (初稿作成)

生徒は与えられたトピックについて、まず自力で英文パラグラフを書いた。

② 2nd Try (再考・再作成)

生徒は自分の英文を読み返し、内容や表現を見直した。

③ Arizona AI によるフィードバック

生徒は2nd Try の文章を Arizona AI に入力し、文法、語彙・表現、つなぎ表現、複文・重文、構成・論理など複数カテゴリに分けてフィードバックを受けた。OCR導入後は、手書き英文を撮影して入力できるようになり、入力負担が軽減された。

④ 赤ペン修正

AIから得た助言をもとに、生徒は自分の原稿に赤ペンで修正を加えた。

⑤ Final Draft (清書)

最後に、修正を反映した最終稿を黒のボールペンで清書した。



1st Try
If schools promote volunteer work among students, there are many advantages and disadvantages. When advantages, students can have responsibility to work as volunteer. In schools, we have little responsibility for our behavior. It's not good to don't have disadvantages, which has more things to do. In addition, students should study and they don't have time to do volunteer work.



2nd Try
If schools promote volunteer work among students, there are more advantages and disadvantages. So we have to consider in this. About advantages, students can have responsibility to work as volunteer. In schools, we have little responsibility for our behavior. It's not good to do them. Furthermore, they can meet new people. Now we have friends or classmates, people. If we have volunteer, we can meet many people who have different background. Many disadvantages, which have more things to do. In addition, students should study and they don't have time to do volunteer work. When they are students, they study at the most and they can do volunteer work after they graduate school.



2nd Try
If schools promote volunteer work among students, there are more advantages and disadvantages. So we have to consider in this. About advantages, students can have responsibility to work as volunteer. In schools, we have little responsibility for our behavior. It's not good to do them. Furthermore, they can meet new people. Now we have friends or classmates, people. If we have volunteer, we can meet many people who have different background. Many disadvantages, which have more things to do. In addition, students should study and they don't have time to do volunteer work. When they are students, they study at the most and they can do volunteer work after they graduate school.

If schools promote volunteer work among students, there are many advantages and disadvantages. So we have to consider in this.
One of advantages is the students can get responsibility by working as volunteer, but, it's not good to do them. Furthermore, they can meet many people who have different background. Many disadvantages, which have more things to do. In addition, students should study and they don't have time to do volunteer work. When they are students, they study at the most and they can do volunteer work after they graduate from school.
Because of the disadvantages, we should consider about promoting volunteer work.



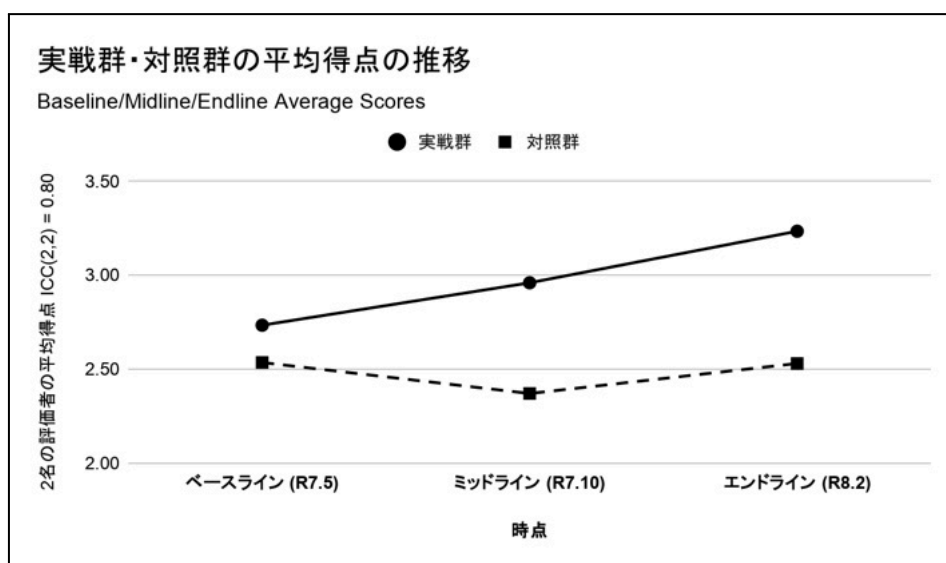
(3) **本実践にみられた特徴**：本実践の特徴は、第一に、AIが単なる誤り検出器ではなく、授業内の書き直し循環を支える装置として機能した点にある。第二に、生徒の価値づけが、初期の文法・自然表現中心から、後期には接続表現、段落構成、論理、内容展開など、より高次の支援へと移行した点である。第三に、S05の一時的失敗とその後の回復を通して、AIフィードバックの有効性が、どの程度具体的かつ修正可能な形で提示されるかに強く依存することが明らかになった。

5. 研究の成果

本研究では、ライティング得点、質問紙調査、自由記述という量的・質的データを組み合わせて成果を検証した。その結果、生成AIを活用した授業内フィードバック循環は、単に生徒の満足度を高めるだけでなく、英語ライティングの得点向上、自信の向上、および書き直しに対する理解の深化に結びつく可能性を示した。以下、主要な結果を示す。

5.1 **ライティング得点の変化**：ライティングの成果については、TOEFL iBT holistic rubric に基づき、2名の評価者が匿名化された答案を用いて盲検採点を行った。分析は、Baseline・Midline・Endline の3時点すべてにデータがそろった完全ケース分析を主分析とし、欠測を含む全使用可能データによる分析を補足的に行った。

完全ケース分析では、統制群は97名、実験群は113名であった。平均得点は、統制群が Baseline 2.54、Midline 2.37、Endline 2.53、実験群が Baseline 2.73、Midline 2.96、Endline 3.23 であった。群間比較では、Baseline 時点では有意差は認められなかったが、Midline および Endline では、いずれも実験群が統制群を有意に上回った。事前テスト (Baseline) の結果、実験群 (2.73) と統制群 (2.54) の間に統計的な有意差は認められなかったものの、数値上は実験群がやや高い状態で開始された。そこで本研究では、時点ごとの絶対値の比較のみならず、各群におけるスコアの推移および変化量にも着目して考察を行った。



統制群は Baseline から Endline にかけて実質的な変化を示さず、年度全体としてほぼ横ばいであった。一方、実験群は Baseline から Endline にかけて 0.50 ポイント、18.3% の有意な上昇を示した。また、Baseline から Endline までの変化量を群間で比較すると、実験群は統制群より 0.50 ポイント大きく改善しており、この差は有意であった。

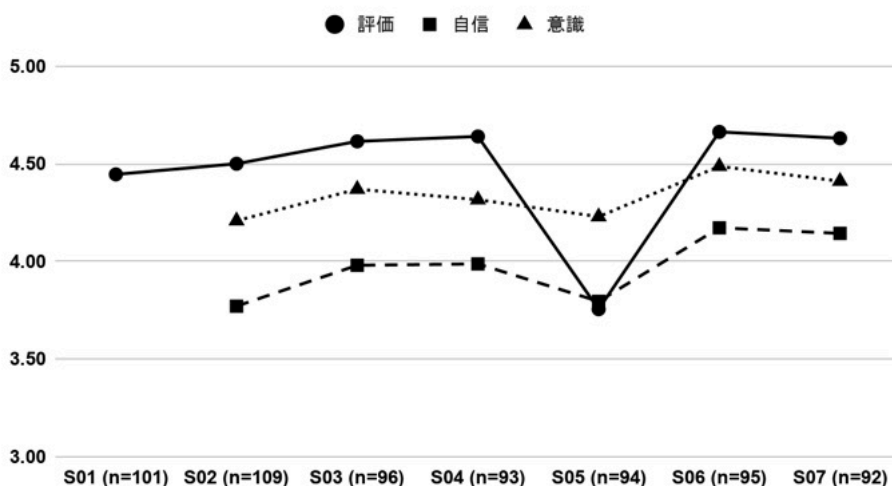
全使用可能データを用いた分析でも同様の傾向が確認された。実験群は Baseline 2.71 から Endline 3.22 へと 0.51 ポイント、19.0% の有意な上昇を示したのに対し、統制群は有意な変化を示さなかった。混合効果モデルでも、Midline および Endline における Group × Time 効果が有意であり、結果の頑健性が確認された。

5.2 評価の信頼性：2名の評価者平均による得点の信頼性を ICC(2, k) により確認したところ、完全ケース全体では 0.799、全使用可能データ全体では 0.801 であり、概ね許容可能な水準であった。特に Endline では、統制群 0.863、実験群 0.855 と高い値を示し、最終時点の平均得点の信頼性は十分に確保されていた。一方、単独評価者ベースの ICC(2, 1) は全体として .80 を下回っていたため、本研究では2名の盲検評価者による平均得点を分析単位とした。

5.3 質問紙調査にみられた学習者認識の変化：実験群を対象とした縦断的質問紙調査では、S01～S07の7波にわたり、フィードバック評価、意識、自信の推移を追跡した。各波の回答者数は92～109名であった。なお、S02～S07の対応分析は、同一回答者の対応データに基づいて実施した (n = 52)。

複合得点の推移を見ると、フィードバック評価は全体として高い水準を維持し、意識も当初から高水準で安定していた。これに対し、自信は年度を通して最も明確な上昇を示した。S02からS07の対応のあるデータに基づく分析でも、フィードバック評価は高水準で安定し、意識は緩やかに上昇した。一方、自信は3.84から4.13へと明確に上昇し、対応のある t 検定の結果、有意な上昇が確認された (p = .00067)。特に伸びが大きかったのは、文法 (+0.33)、複文・重文 (+0.33)、つなぎ表現 (+0.31)、語彙・表現 (+0.28)、段落構成 (+0.20) に関する自信項目であった。

AIフィードバックに対する評価と各観点における自信・意識の平均値



5.4 S05 における一時的低下とその回復：S05では「評価・自信・意識」の各指標に一時的な低下が見られた。これは、2025年11月4日に導入した Prompt v2.1 に不具合があり、生成されるフィードバックが曖昧で一般的すぎるものとなっていた時期と一致している。その後 Prompt を修正した結果、S06 で「評価」が大きく回復し、S07 でも高い水準が維持された。したがって、S05 は単なる偶然的変動ではなく、プロンプト設計上の実装失敗を反映したものであり、同時に、生徒が実際のフィードバック品質に敏感に反応していたことを示す証拠でもある。

5.5 自由記述にみられた質的变化：自由記述の分析からは、生徒による Arizona AI の評価が、年度を通してより精緻化していったことがうかがえた。初期段階では、文法訂正、不自然な表現の修正、即時性といった文レベルの利点に注目する記述が多く見られた。しかし、後期になるにつれて、生徒は接続表現、段落構成、論理展開、内容の補強といった、より高次の文章構成支援に価値を見いだすようになった。

自由記述では、初期には例文、翻訳、語彙提案、文法説明などの追加支援を求める声が多かったが、後期にはチャット機能、追質問、修正版提示、進捗表示など、より対話的で個別化された支援を求める声が増加した。これは、生徒がシステムを単なる誤り指摘ツールではなく、対話的なライティング・コーチとして捉え始めたことを示している。

修正版S06の分析では、S05の不具合が継続していたのではなく、S06では明確な回復が見られた。「特になし」とする回答が増え、具体的な訂正や例示の不足に対する不満は大きく減少した。

その一方で、複文・重文に関する助言は、文章の流れをよくし、表現の幅を広げる点で有用と感じる生徒がいる一方で、無理につなげることで不自然になったり、かえって読みにくくなったりするという指摘も見られた。

5.6 AI と教師の役割分担に関する認識：AI vs. Human の自由記述では、生徒は AI と人間教師を単純な優劣関係で捉えていなかった。AI は「速い」「細かい」「何度も使える」「気軽に見直せる」といった点で評価されており、一方で人間教師は「ニュアンスが分かる」「柔軟である」「背景を踏まえた具体例を示せる」「曖昧な場合の最終判断が分かりやすい」といった点で高く評価されていた。特に S05 では AI に対する信頼が一時的に低下し、人間教師の方が正確で頼れるという記述が増したが、S06 以降は再び AI の即時性・詳細性が評価されるようになった。

5.7 成果の総括：以上より、本研究の成果は次の三点に整理できる。第一に、実験群のライティング得点は年度を通して有意に向上し、統制群との差も Midline・Endline で明確に確認された。第二に、質問紙調査では、フィードバック評価は高く維持され、「意識」も高水準で安定していたが、特に「自信」の上昇が顕著であった。第三に、自由記述からは、生徒が AI フィードバックを文法訂正中心の支援から、構成・論理・内容展開・修正方略まで含むより高次の支援として価値づけるようになったことが示された。

6. 今後の課題・展望

本研究を通して、生成AIを活用した英語ライティング支援が、中等教育段階において一定の教育的有効性をもち得ることが示された。一方で、その有効性は、プロンプト設計、授業内での足場掛け、フィードバック量、学習者の特性など、複数の条件に強く依存することも明らかになった。

第一の課題は、フィードバックの質と足場掛けの程度の最適化である。S05における一時的な低下は、フィードバックが曖昧で一般的すぎる場合、生徒がすぐにその弱さを見抜くことを示した。一方で、修正版導入後のS06では、具体例、ヒント、整理されたカテゴリ別助言など、より足場掛けの明確な支援が高く評価された。今後は、「考えさせる」とことと「修正のための具体的手掛かりを示す」とことのバランスをより慎重に調整する必要がある。

第二の課題は、学習者層に応じた適用可能性の検証である。本研究の実践は、比較的学力水準の高い学校環境の中で行われたため、現行のフィードバック量や情報密度が、他の学校やより支援を必要とする学習者層に対しても同様に機能するとは限らない。今後は、フィードバックの長さ、複雑さ、カテゴリ数を調整し、より多様な学習者に対応可能なモード設計を検討する必要がある。

第三の課題は、評価設計のさらなる精緻化である。本研究では、BMEライティング評価と縦断的質問紙調査によって一定の効果を確認したが、個票レベルの利用ログが十分に整備されていなかったため、個人ごとの利用頻度や修正量と成果との対応を詳細に検証するには限界があった。今後は、生徒ごとの利用履歴や修正過程をより体系的に記録し、どのような利用パターンがどのような伸びにつながるのかを分析できる仕組みを整える必要がある。

今後の展望としては、Arizona AI をより安定的かつ対話的なライティング支援システムへと発展させることが挙げられる。具体的には、追質問機能、ヒント提示、正答・改善例の段階的提示、進捗表示など、生徒が求めていた対話型支援を取り入れながら、単なる誤り指摘ツールではなく、反復的な修正を支える学習パートナーとしての機能を強化したい。さらに、学校単位での導入を見据え、組織単位で生徒をオンボードできる構成や、学校管理者レベルでの運用管理機能を整備することで、より広い実装可能性を探りたい。

7. おわりに

本研究では、大規模言語モデルのAPIを活用した英語ライティング支援システムを、中等教育段階の通常授業に実装し、その教育的有効性と実装上の条件を検討した。結果として、実験群のライティング得点は年度を通して有意に向上し、生徒の「自信」も明確に高まった。また、自由記述の分析からは、生徒が AI フィードバックを単なる便利な機能としてではなく、文法訂正から構成・論理・修正方略に至るまで、多面的な支援として捉えるようになっていったことが確認された。

一方で、本研究は「AIを入れればうまくいく」という単純な結論を支持するものではない。むしろ、S05における一時的な実装失敗が示したように、AI支援の有効性はプロンプト設計や足場掛けのあり方に強く依存している。生徒は、フィードバックが具体的に修正可能であるときには高く評

値し、逆に曖昧で一般的な助言には敏感に否定的反応を示した。この点は、AI活用の成否が技術の新しさそのものではなく、学習者が実際に書き直しに生かせる形で支援を受け取れるかどうかにかかっていることを示している。

また、生徒の記述からは、AIと教師が置き換え可能な存在ではなく、異なる役割を担う補完的な存在として理解されていたことも重要である。AIは速さ、詳細さ、反復性という強みを持ち、人間教師は文脈理解、柔軟性、身近な具体例、ニュアンスのある支援において優れていた。このことから、本研究の実践的意義は、生成AIを教師の代替として位置づけるのではなく、教師の指導を補完し、授業内における個別最適化フィードバックの量と質を高める手段として位置づけた点にある。

本研究の遂行にあたっては、パナソニック教育財団による助成をはじめ、多くの支援を受けた。校内外の研究会・研修会で助言を寄せてくださった先生方、公開研究授業や各種発表の機会を通して本実践を支えてくださった関係者の皆様、そして何より、試行錯誤を重ねながら真摯に書き直しに取り組んだ生徒たちに深く感謝申し上げたい。

また、本研究で用いた Arizona AI の各バージョン、プロジェクトコード、課題プリント、BME課題、授業案等の関連資料は参考文献・関連資料欄に示す予定であり、読者が実践の内容を具体的に参照できるようにしている。本研究の意義は、単一の授業実践の成果にとどまらず、生成AIを活用した個別最適な学び、形成的評価、および1人1台端末環境を前提とした授業改善の具体的なモデルを提示した点にある。これは、今日の学校教育が求められている教育DX、情報活用能力の育成、個別最適な学びと協働的な学びの一体的充実という方向性にも接続するものである。今後も、生成AIを用いた英語ライティング指導の可能性を、授業実践に根ざした形で継続的に探究するとともに、他校でも参照・応用可能な形で成果の共有を進めていきたい。

8. 参考文献

- Nhan, L. K., Hoa, N. T. M., & Quang, L. V. N. (2025). “Leveraging AI for Writing Instruction in EFL Classrooms: Opportunities and Challenges.” *Educational Process: International Journal*, 15, e2025158.
<https://doi.org/10.22521/edupij.2025.15.158>
- Qais Hweidi, ai-assisted-development-guide, GitHub repository,
<https://github.com/Qais-Hweidi/ai-assisted-development-guide.git> (2026年3月13日参照)
- 森 紗良 (2025年7月2日) 「生成AI『自分で使う』、日本の高校生は低調 学校の紙重視が影響？」『日本経済新聞 電子版』
<https://www.nikkei.com/article/DGXZQ0UD290ZT0Z20C25A6000000/>
- 中島 大介 (著)・西 宏章 (協力) 『この一冊で全部わかる ChatGPT & Copilotの教科書』SBクリエイティブ、2024年。
- クジラ飛行機 『大規模言語モデルを使いこなすためのプロンプトエンジニアリングの教科書』マイナビ出版、2024年。

関連資料欄

- [Arizona AI v.1](#) GAS版ウェブアプリ
- [Arizona AI v.1](#) GAS版プロジェクトコード (プロンプト等を含む)
- [課題プリント・BME課題](#)
- [授業案](#)
- [Arizona AI v.2 React.js版](#)

※パナソニック教育財団関係者向けのデモ用ログイン情報は、実施報告書を参照。