

令和7年度 生成AIに関する教員向け研修動画シリーズ 第1回

生成AIの基礎と 教育における活用可能性



吉田 墾

東京大学 大学院工学系研究科 准教授

2025年6月25日

自己紹介 吉田塁

・東京大学 大学院工学系研究科 准教授

- 専門は教育工学（アクティブラーニング、生成AI、オンライン学習、ファカルティ・ディベロップメント）
- コロナ禍における教育のオンライン化支援の貢献が認められ、オンライン授業等におけるグッドプラクティス総長表彰
- 研究室で開発した意見交換ツール「LearnWiz One」が世界最大の EdTech コンペ GESAAwards 2021 R&D 部門で世界大会優勝
- 教育における生成AI利活用に関する活動が評価され、日本工学教育協会 第29回 工学教育賞（業績部門）受賞（2024）

・生成 AI との関わり

- コロナ禍のように教員が困るだろうと感じ、できるだけサポートしたいと思い活動
- 2023年5月13日実施の4時間の教員向け講座は参加者が650名以上、最高評価が8割以上、公開している動画は再生数が9万回以上
- 文部科学省学校DX戦略アドバイザー。文部科学省による「初等中等教育段階における生成AIの利用に関する暫定的なガイドラインver 2.0」の作成に委員として関与



@luiyoshida

自己紹介 吉田塁 生成AI関連の主な活動

公開講座

- 大規模公開講座 ([2023年5月23日](#)、[2025年3月2日](#))
- 毎月末 [オンラインイベント](#)



講演

- 文部科学省、国会図書館、教育委員会、学校関係団体、学校、企業など[計60件以上](#)

官公庁

- 文部科学省「初等中等教育段階における生成AIの利活用に関する検討会議」委員
- 文部科学省「G7教育イニシアティブ・シンポジウム」講演者・コーディネーター
- 経済産業省 令和6年度学びと社会の在り方改革推進事業 メンター

コミュニティ

- 「[教育現場の生成AIなんでも相談所](#)」運営（テキスト情報共有・月1ミーティング）

ポータルサイト

- 教育×生成AI ポータルサイト [Manabi AI \(まなびあい\)](#) 運営

研究

- 生成AIのエッセイ評価能力検証（教育AI国際トップ会議 AIED 2024, 2025 に論文採択）
- 生成AIの多言語能力評価

目次

- **生成AIについて**
 - 生成AIに関する基本的な知識
- **教育と生成AI**
 - 活用事例、研究、国内の動向、教育への影響

生成 AI について

生成 AI とは

- ・学習データをもとにテキストや画像などのコンテンツを生成できる人工知能
- ・関連技術やサービスの進展が早い（最近はマルチモーダル化が進んでいる）

対話型生成 AI (テキストがメインの生成AI)

- ChatGPT
- Gemini
- Claude
- Copilot
- Llama
- ...

画像生成AI

- Midjourney
- Stable Diffusion
- DALL-E3
- Image Creator
- Adobe Firefly
- ...

動画生成AI

- Veo3
- Sora
- Wan 2.1
- ...

生成 AI との対話にあたって

- **用語説明**
 - プロンプト: 生成AIに入力する指示出しのこと
- **対話をする上で重要なこと**
 - プロンプトに行ってほしいことを目的なども含めて具体的に入れる
(必要に応じて出力形式やデータも指定・アップロードする)
 - 思い通りの出力にならなくても、対話を繰り返し、自分がほしい出力のイメージを具体的に伝えてみる
 - プロンプトは改善するもの、生成AIは対話するものと認識する
 - **出力を鵜呑みにしない** (自分で判断する。AIは責任を持たない)

生成 AI の性能（一部ピックアップ）

- 2023年2月 米医師試験 合格ライン（ChatGPT GPT-3.5）
- 2023年3月 米司法試験 上位10%スコア^{*1}（GPT-4, GPT-3.5は下位10%）
- 2023年4月 日本医師試験 合格ライン^{*2}（ChatGPT GPT-4）
- 2024年3月 多様な分野の多肢選択問題、数学問題、プログラミング課題などで GPT-4よりも高スコア^{*3}（Claude 3 Opus）
- 2024年5月 扱えるトークン数が200万（約100～200万文字）に達する^{*4}
(Gemini 1.5 Pro、GPT-4 は12.8万)
- 2024年12月 人間のトッププログラマーレベル（世界約6万人中175位相当）、博士レベルの科学問題 正答率 87.7%^{*5} (o3)

*1: GPT-4 Technical Report <https://arxiv.org/abs/2303.08774>

*2: Performance of Generative Pretrained Transformer on the National Medical Licensing Examination in Japan | medRxiv <https://www.medrxiv.org/content/10.1101/2023.04.17.23288603v1>

*3: Introducing the next generation of Claude ¥ Anthropic <https://www.anthropic.com/news/clause-3-family>

*4: Gemini 1.5 Pro and 1.5 Flash GA, 1.5 Flash tuning support, higher rate limits, and more API updates - Google Developers Blog <https://developers.googleblog.com/ja/gemini-15-pro-and-15-flash-now-available/>

*5: OpenAI o3 and o3-mini—12 Days of OpenAI: Day 12 - YouTube <https://www.youtube.com/watch?v=SKBG1sqdylU>

ただし、人間が正答できる簡単な問題で間違うこともあることから、全知全能ではない

生成 AI でできること

- ・説明可能な課題に対して何かしら回答を出力してくれる
(出力の質はモデルや分野などによって異なる)
- ・例
 - ・ 読書感想文作成
 - ・ レポート作成
 - ・ 選択問題への回答
 - ・ AI 家庭教師
 - ・ 授業案作成
 - ・ 選択問題作成
 - ・ 文章の英訳
 - ・ Excel の数式作成
 - ・ Excel のマクロ作成

生成 AI でできること～読書感想文作成～

- ・プロンプト
 - ・夏目漱石の「坊っちゃん」に関する読書感想文を、具体的な例も交えながら作成して下さい。
- ・ChatGPT 4o 出力
 - ・夏目漱石の『坊っちゃん』を読んで、改めて人間の正直さや正義感の大切さを考えさせられました。主人公の坊っちゃんは、江戸っ子気質で短気だが真っ直ぐな性格を持つ青年です。彼が松山の中学校に教師として赴任し、さまざまな人間模様に巻き込まれながらも、自分の信念を貫く姿に共感を覚えました。
 - ・特に印象に残ったのは、坊っちゃんが教頭の「赤シャツ」と対立する場面です。赤シャツは表向きは知的で上品な人物に見えますが、実際には陰で姑息な手を使って周囲を操る狡猾な人物でした。坊っちゃんはこうした偽善的な態度を許せず、最終的に赤シャツとその取り巻きを懲らします。この場面は、正義感の強い坊っちゃんの性格が最もよく表れており、読んでいて痛快でした。
 - ・…（略）

絶妙なデタラメが入っている。表面的な文章である印象
(プロンプトを工夫すれば異なる出力になる)

生成 AI でできること～選択問題への回答～

- 問題例: Hendrycks et al. (2020) の問題を日本語訳

- 静止状態からボールを落とすと、 9.8m/s^2 の加速度で下向きに加速される。代わりに、空気抵抗がないと仮定して投げる場合、手から離れた直後の加速度は
 - (A) 9.8m/s^2 である。
 - (B) 9.8m/s^2 より大きい。
 - (C) 9.8m/s^2 より小さい。
 - (D) 投げるときの速度が与えられていなければ言えない。

- ChatGPT 4o 回答 正解

- (前略) 空気抵抗がない場合、手から離れた後は重力のみが作用するため、加速度は常に 9.8 m/s^2 (下向き) となります。よって、正解は(A) 9.8m/s^2 である。

- ChatGPT 4o mini 回答 正解

- (前略) 手から離れた直後の加速度は、重力加速度によるものであり、加速は 9.8 m/s^2 です。これは、投げる速度が与えられていなくても変わりません。したがって、正解は (A) 9.8m/s^2 であるです。

スクリーンショットも活用可能…

生成 AI でできること ~AI家庭教師~



出典: Math problems with GPT-4o - YouTube https://youtu.be/_nSmkyDNulk 性能への過信には注意

生成 AI でできること ~授業案作成~

- プロンプト
 - 中学2年生向けの数学の授業（確率）について、効果的なアクティブラーニングを交えた授業案（50分）を作成してください
- ChatGPT 4o mini 回答（一部）
- 【授業案：確率の基本と体験学習】（50分）
 - 目標：
 - 確率の基本的な概念（事象、確率）を理解する。
 - 確率の計算方法を学び、実際の事例でその計算を体験する。
 - 日常生活の中で確率をどのように使うかを考える。
 - 準備物：
 - サイコロ（人数分）
 - コイン（人数分）
 - ホワイトボード、マーカー
 - プリント（問題シート）

生成 AI でできること ~座席表作成: Excel のマクロ作成~

・プロンプト（一部）

- 学生シートのA列に学生の氏名が入っているエクセルファイルがあります。座席シートに対して、ランダムに学生の氏名を重複せずに配置するマクロを作成してください。
座席は、1つの机に3名横並びに座れて、その机が、横に3つ、縦に7つあります。
表示する際、机の横は1セル空けてください。
また、座席に対して学生の数が少ないときは、前（上）に詰めるようにしてください。

・Copilot 回答（一部）

- Excelで「Alt + F11」を押してVBAエディターを開きます。
- 「挿入」→「標準モジュール」を選択して新しいモジュールを作成します。
- 以下のコードを新しいモジュールに貼り付けます。

Vba

□ コピー

```
Sub AssignSeats()
    Dim students As Range
    Dim studentList() As String
    Dim seatSheet As Worksheet
    Dim i As Long, j As Long
    Dim maxRow As Long, maxCol As Long
```

生成 AI が苦手なこと・使う上で注意すること

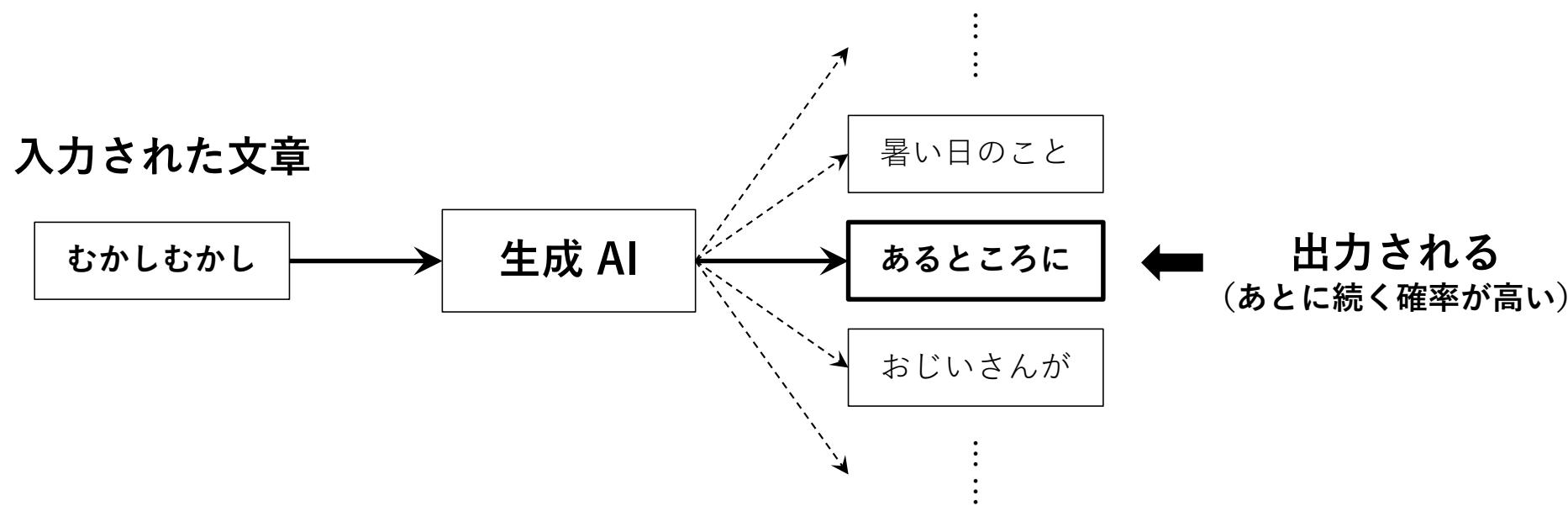
- 必ずしも正確な情報を述べるわけではない
 - デタラメを言うことがある（幻覚: ハルシネーション）
 - GPT-4.5 で改善されているようだが完全ではない ([OpenAI 2025](#))
 - 専門知識を持つ、1次情報にあたる、情報の信頼性を検討することが重要
 - 出力の利用にあたっては人間が責任を持つ（AIは責任を持たない）
- 分野やモデルによってできることが異なる
 - それぞれの分野でどの程度活用できそうかを検討する必要あり
- 自動生成された文章を検出することは難しい
 - ChatGPT によって生成された文章かを ChatGPT に聞いてしまい、その結果を信じて学生の半分以上を不合格にする事案も発生（[参考記事](#)）
 - 実用上、自動生成された文章の検出は難しい（Sadasivan et al. 2023）

生成 AI の仕組み ~テキスト生成（簡易）~

*推論モデルでは異なる（後述）

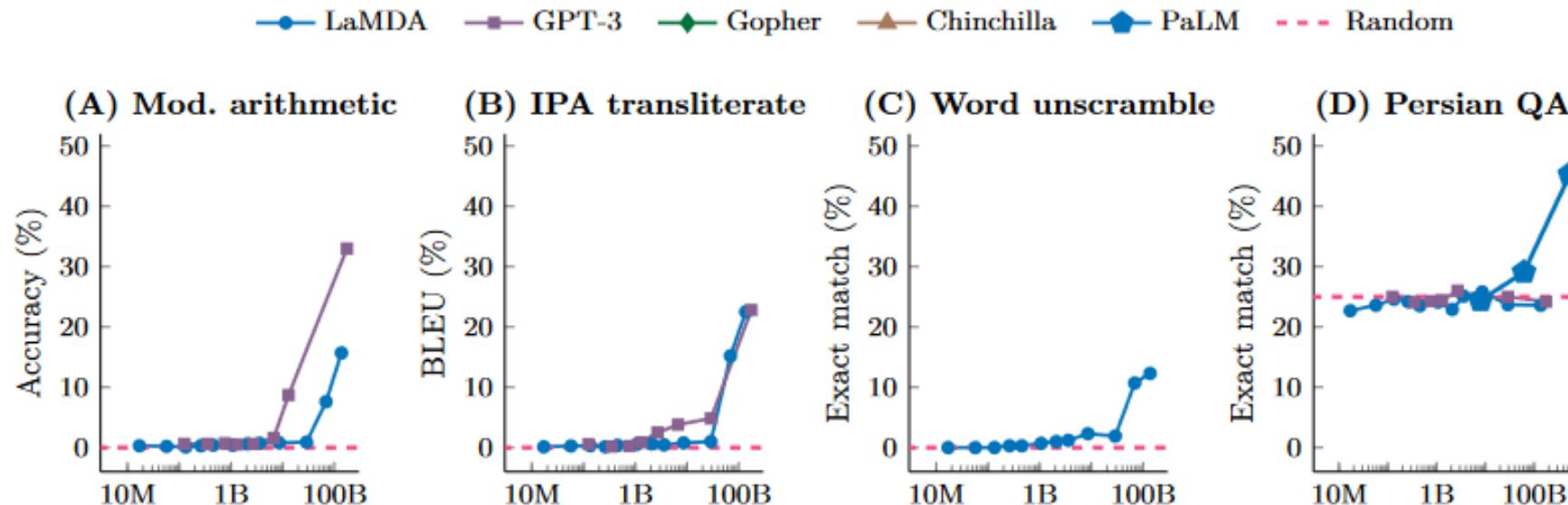
- ・過去のテキストを参考に次に来る文字や単語を予測して出力する*
- ・単純な機構と思いきや、モデルのサイズ（パラメータ数など）を大きくすることで、能力が高くなり、多様なタスクをこなすことが可能（Brown et al. 2020, Wei et al. 2022a, OpenAI 2023）

あとに続く言葉の候補



生成 AI の仕組み ~テキスト生成（簡易）~

- ・スケーリング則: パラメータ数、学習計算量、データサイズを上げるほど性能が上がる (Kaplan et al. 2020)
- ・創発的能力: あるパラメータ数（学習計算量）を超えると急に性能が上がる (Wei et al. 2022a)



出典: Wei et al. 2022a

最近では小さなモデルでも性能を高める方法（蒸留など）が開発されている

生成 AI の仕組み ~プロンプトの重要性 (プロンプトエンジニアリング)~

- 例を載せると性能が向上 (Few-shot prompting) (Brown et al. 2020)
- 思考過程を出力させると性能が向上 (思考の連鎖: Chain of Thought) (Wei et al. 2022b)
- 例を載せなくても「Let's think step by step」とプロンプトに入れるだけで性能が向上 (Zero-shot prompting) (Kojima et al. 2022)

No.	Category	Template	Accuracy
1	instructive	Let's think step by step.	78.7
2		First, (*1)	77.3
3		Let's think about this logically.	74.5
		:	
14	irrelevant	By the way, I found a good restaurant nearby.	17.5
15		Abrakadabra!	15.5
16		It's a beautiful day.	13.1
-		(Zero-shot)	17.7

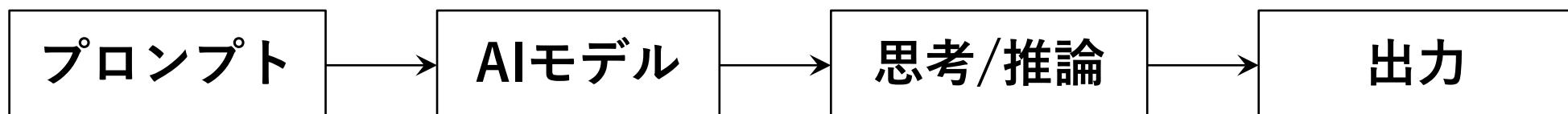
生成 AI の仕組み ~推論モデル~

- ・ モデル自身が思考の連鎖を行い、出力の質を高められるように学習させたもの
(雑に言えば自身でプロンプトエンジニアリングできるようになったもの)

従来のテキスト生成 (例: OpenAI GPT-4o, DeepSeek-V3)



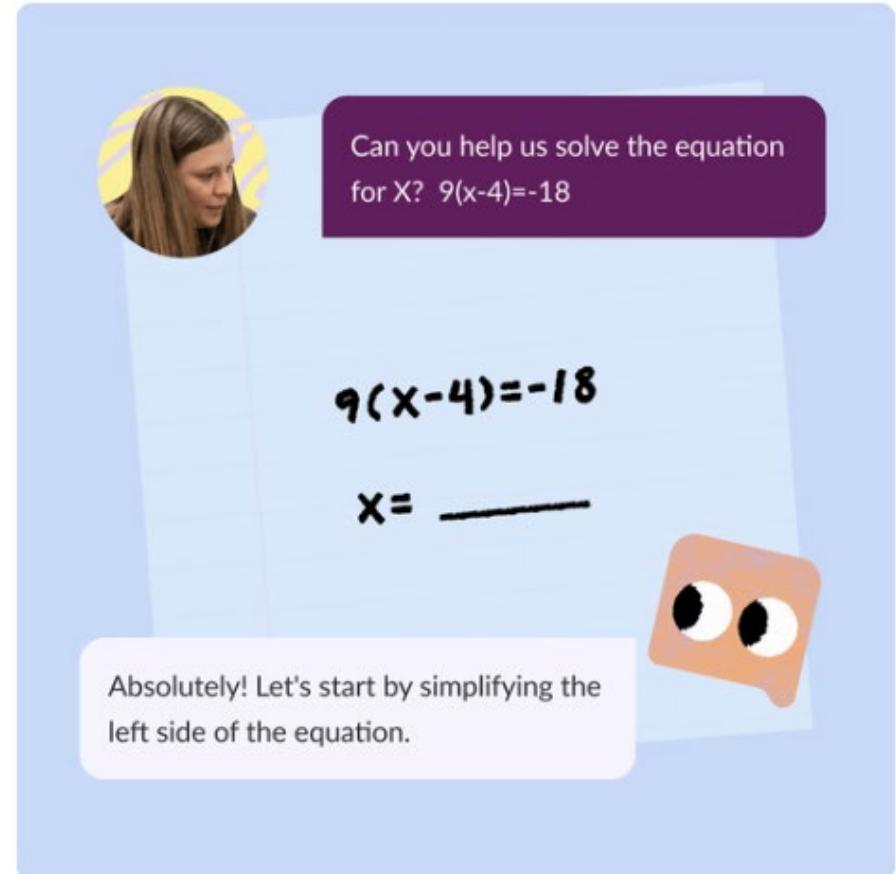
推論モデルのテキスト生成 (例: OpenAI o3, DeepSeek-R1)



この部分を
追加でトレーニング

生成 AI の可能性

- 学習者による活用例
 - **個別学習支援:** 個別指導、個別教材・問題作成、メンター・相談相手としてのコメント・フィードバック作成、ドラフト作成
 - **グループ学習支援:** チームメンバー、ファシリテーター、メンターとしてのコメント・フィードバック作成
 - **課外活動支援:** 部活動、学園祭運営支援
- 教職員による活用例
 - **授業支援:** 授業案作成、教材作成、問題・課題作成、評価補助
 - **校務・事務支援:** 資料・報告書作成補助、データ処理補助
 - **研究支援:** 英訳、フィードバック生成



出典: Meet Khanmigo: Khan Academy's AI-powered teaching assistant & tutor <https://www.khanmigo.ai/>

生成 AI のリスク ~例~

- ・出力が不正確な場合がある（幻覚: ハルシネーション）
- ・評価の妥当性が低くなり得る
- ・バイアス・毒性が存在する (Zhuo et al. 2024, Deshpande et al. 2023)
- ・言語格差が存在する
- ・意図せずデータが学習される可能性がある
- ・個人情報の捏造、流出の可能性がある
- ・機密情報の流出の可能性がある
- ・著作権を侵害する可能性がある
- ・公平性が保てない可能性がある（有料版と無料版の性能差）
- ・過度な依存を促す可能性がある
- ・消費電力が大きい

教育と生成 AI

活用事例～文部科学省 生成AIパイロット校 事例集～

1人1台端末とクラウド環境を
活用した教育実践

1人1台端末とクラウド環境を
活用した実践動画

生成AIの活用方法に関する教育実践

生成AIパイロット校の最終報告に掲載されている「生成AIの活用方法に関する教育実践」を、絞り込み検索することができます。

利用種別

教育利用

校務利用

具体的なプロンプトの記載

あり

学年

小学校1年

小学校3年

小学校4年

小学校5年

小学校6年

中学校1年

中学校2年

中学校3年

高等学校

教科

国語・書写

社会・地理歴史・公民

算数・数学

理科

生活

音楽

図画工作・美術

体育・保健体育

技術・家庭

外国語・外国語活動

情報・専門教科情報

道徳

総合的な学習の時間・総合的な探究の時間

※学年タグは教育利用の事例のみ付与されています。

検索

活用事例～札幌市立中央小学校 国語での活用～

- ・**学年・科目:** 小学6年生・国語 俳句
- ・**内容:** 生成AIが「秋」をテーマに出力した俳句を批判的に検討した後、小学生がそれぞれ俳句を作成し、最後に画像生成AIを用いて俳句の挿絵を出力する
- ・**実際:** 生成AIの詩人としての能力の限界への気づき（2つの季語、直接的な表現、リズムの悪さ）、自分なりの俳句の工夫
- ・生成AIが出力した俳句
「紅葉舞う 秋風さらさら 心落ちつ」
- ・**感想:** 「AIで画像生成するのはとても便利であるが、間違った使い方をすると自分の学びにとってよくないかもしれないから気を付けて使いたい。」「今後も授業で生成AIを使ってさらに理解を深めていけるのではないかと思った。」



活用事例～アリゾナ州立大学（AIネイティブ大学へ）～

- ChatGPT Edu を教育、研究、大学運営に導入
- 導入プロセス
 - 教職員や学生からChatGPTの革新的な活用法に関する提案を募集
 - 80%以上の学部から400件以上の提案、200以上のプロジェクトが進行中
- 活用事例
 - 教育：学生の論文作成支援するボット、医療系学生が患者との対話スキルを練習できるボット「Sam」など
 - 研究・運営：博士課程の学生が、研究参加者を倫理的かつ分かりやすく募集するための文章作成補助

研究 ~プログラミング学習支援 (Brender et al. 2024)~

- ・スイスの大学院におけるロボット工学コースにおいて実験
 - ・演習を行う際にChatGPT利用群（40名）と非利用群（24名）に分けて、演習の成績と事前事後のテスト結果を比較
- ・全体的にはChatGPTの使用は演習の成績を向上させるが学習は向上させない
- ・ただしChatGPTの使い方によって演習とテスト成績への影響が異なる

ChatGPT の使い方	結果	
概念探求者: 概念理解を使う	演習成績: やや低い	テスト成績: 高い
実践的開発者: 課題解決を使う	演習成績: 平均的	テスト成績: 平均的
デバッガー: エラー修正を使う	演習成績: 高い	テスト成績: 低い

研究 ~エッセイ評価 (Yoshida 2024)~

- **目的: GPTモデルのエッセイ評価能力を検証**
 - GPT-3.5, 4 の各3バージョン（2023年6月～2024年1月）、計6種類の GPT を用いてエッセイ評価能力、プロンプト内の正答例の影響を比較
- **発見1: マイナーバージョンを含むモデルによって評価能力が異なる**
 - → マイナーバージョンも含めてモデル間で影響が異なるため、それぞれのモデルで自動エッセイ採点性能を評価する必要がある
- **発見2: 最新モデルではない GPT-4 (2023年6月) が最も安定性・性能が高い**
 - → 最新のモデルが必ずしも全ての分野で最も性能が良いわけではないため、それぞれ評価する必要ある
- **発見3: GPT による評価と専門家による評価の一致率は高いとは言えない**
 - → エッセイ評価能力はそこまで高くないため、GPT による文章評価の結果を過信してはいけない

研究～AI利用が批判的思考に与える影響 (Gerlich 2025)～

- ・目的: AIへの依存が「批判的思考力」に与える影響を検証
 - ・666人を対象にアンケート調査と50人を対象に半構造化インタビューを実施し、「認知的オフロード」(思考の外部委託)」が鍵となるかを検証
- ・発見1: AI利用と批判的思考力に「負の相関」
 - ・AIツールを頻繁に利用する人ほど、情報を分析・評価する能力が低い傾向
- ・発見2: 背景に「認知的オフロード」の存在
 - ・AIに頼ることで自ら深く考える機会が減少し、思考力の低下につながる可能性あり
- ・発見3: 若者と高学歴層で異なる傾向
 - ・若年層はAIへの依存度が高く影響を受けやすいが、高学歴層はAIを使いつつも高い思考力を維持
- ・提言: AIとの「賢い付き合い方」が重要
 - ・AIの利便性を享受しつつも、教育などを通じて自律的な思考力を養うバランスが不可欠
 - 思考の「丸投げ」には要注意。AIリテラシーの教育が重要

研究 ~ChatGPTが学習に与える影響 (Wang & Fan 2025)~

- **目的: ChatGPTは学生の学習に本当に効果的か?**
 - 51件の研究データ（6,621件の論文からスクリーニング）をメタ分析し、学習パフォーマンス・認識・高次思考への影響を科学的に検証
- **発見1: 「学習パフォーマンス」が大幅に向上**
 - テストの点数などに代表される学習成果に、大きなプラス効果（効果量 $g=0.867$ ）が確認
- **発見2: 「使い方」が効果を左右する**
 - 「問題解決型学習」で「4～8週間」使うと効果が最大化。また、ChatGPTを「指導役（チューター）」として活用すると、学生の思考力育成に特に有効
- **発見3: 「学習意欲」や「思考力」にもプラス効果**
 - 学習へのポジティブな態度や、創造的・批判的思考といった高次の能力にも中程度の効果
- **結論: ChatGPTは強力な教育ツールだが、戦略的な活用が不可欠**
 - その効果を最大限に引き出すには、学習の目的や期間、方法に合わせて使い方を工夫する必要あり

AI の教育への組み込み方も重要 (教員の力の見せ所)

国内の動向～文部科学省の取り組み～

- ・2023年7月 「初等中等教育段階における生成AIの利用に関する暫定的なガイドライン」公表
- ・2023年10月 初年度（令和5年度）生成AIパイロット校事業開始
- ・2024年7月 初等中等教育段階における生成AIの利活用に関する検討会議
- ・2024年12月 「初等中等教育段階における生成AIの利活用に関するガイドライン（Ver.2.0）」公表
- ・以降、様々な生成AIに関する実証事業を実施（以下は例）
 - ・生成AIの活用を通じた教育課題の解決・教育DXの加速「学びの充実など教育課題の解決に向けた教育分野特化の生成AIの実証研究事業」（令和6年度補正予算事業）
 - ・令和6年度 次世代の学校・教育現場を見据えた先端技術・教育データの利活用推進（実証事例を踏まえた先端技術の活用方法・諸外国の先端技術の動向に関する調査研究）

国内の動向～生成AI活用に関するガイドライン～

- ・「初等中等教育段階における生成AIの利活用に関するガイドライン（Ver.2.0）」が公開
 - ・本体、概要1枚、概要資料の3点セットが2024年12月26日に公開されている
- ・内容
 - ・ガイドラインの策定に当たって
 - ・ガイドライン策定の背景、目的、構成、位置付けが記載
 - ・1. 生成AIについて
 - ・生成AIの定義、技術的な仕組み、急速な進化と普及の現状、教育分野における活用の可能性とリスク
 - ・2. 基本的な考え方
 - ・学校現場での生成AI利活用に関する基本的な考え方（「人間中心の原則」など）
 - ・3. 学校現場において押さえておくべきポイント
 - ・5つのポイント（安全性、情報セキュリティ、個人情報等の保護、公平性、透明性）が記載
 - ・3-1. 教職員が校務で利活用する場面
 - ・3-2. 児童生徒が学習活動で利活用する場面
 - ・3-3. 教育委員会等が押さえておくべきポイント
 - ・参考資料編
 - ・教職員・児童生徒向けのチェックリスト、生成AIパイロット校の先行事例、学校現場で活用可能な研修教材、国内外の関連資料など

生成 AI が与える教育への影響 ~全般~

※ 現時点の技術レベルでの見解

- ・生成AIが教育へ与える影響は複層的である（分野、モデル、マイナー含むバージョン、使い方、プロンプトなどで性能が異なる）
 - ・要約、文章作成、アイデア出し、プログラミング、翻訳など全般的に得意な分野もあるため、その強みを活かす活用法は汎用的になりやすい
- ・生成AIモデルやサービスの進展スピードを考慮すると今後も臨機応変な対応、課題の収集・知見の蓄積が求められる（臨機応変に対応できる体制づくりが必要）

実際に利用して試行錯誤することが大事

技術を用いた思考/協働を促す授業設計/環境整備が重要

手段が目的化しないように注意！

生成 AI が与える教育への影響 ~授業・学習~

- **評価の妥当性を下げ得る**

- 筆記に関連する学習成果物を自動生成させられると評価が揺らぐ
(誤りが含まれることもあり、必ずしもそのまま使えるわけではない)
- 評価方法、波及的に授業目的や内容、カリキュラムを再検討する可能性も
- 対策として、対面評価、生成AIが成果物を自動生成しにくい課題設定など

- **学習プロセスを支援し得る**

- 例) 家庭教師、相談相手、チームメイトなどの役割を与えて活用する

- **教員の授業作りや校務・事務を支援し得る**

- 例) 授業案、問題、教材、資料などのドラフトを作成してもらう

本質的には、生成 AI で楽したいと思われるような授業・課題ではなく、
もっと自分で学びたい、自分で取り組みたいと思ってもらえるような
授業・課題作りが重要だと吉田は思っています

おわりに

おわりに

- ・ご清聴いただきありがとうございました！
- ・まずはご自身で使ってみることをおすすめします
- ・あくまでも AI は「副操縦士」であり、「操縦士」であるあなた・生徒の知識、思考、意思決定が重要です

参考文献 (ファイル内のURLは全てアクセス日 2025年6月23日)

- Brender, J., El-Hamamsy, L., Mondada, F., & Bumbacher, E. (2024, July). Who's Helping Who? When Students Use ChatGPT to Engage in Practice Lab Sessions. In *International Conference on Artificial Intelligence in Education* (pp. 235-249). Cham: Springer Nature Switzerland.
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, 33, 1877-1901.
- Chen, L., Zaharia, M., & Zou, J. (2024). How is ChatGPT's behavior changing over time?. *Harvard Data Science Review*, 6(2).
- DeepSeek-AI. (2025). Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948*.
- Deshpande, A., Murahari, V., Rajpurohit, T., Kalyan, A., & Narasimhan, K. (2023). Toxicity in chatgpt: Analyzing persona-assigned language models. *arXiv preprint arXiv:2304.05335*.
- Gerlich, M. (2025). AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking. *Societies*, 15(1), 6.
- Hendrycks, D., Burns, C., Basart, S., Zou, A., Mazeika, M., Song, D., & Steinhardt, J. (2020). Measuring massive multitask language understanding. *arXiv preprint arXiv:2009.03300*.
- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., ... & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM computing surveys*, 55(12), 1-38.
- Kaplan, J., McCandlish, S., Henighan, T., Brown, T. B., Chess, B., Child, R., ... & Amodei, D. (2020). Scaling laws for neural language models. *arXiv preprint arXiv:2001.08361*.
- Kojima, T., Gu, S. S., Reid, M., Matsuo, Y., & Iwasawa, Y. (2022). Large language models are zero-shot reasoners. *Advances in neural information processing systems*, 35, 22199-22213.
- Liu, R., Zenke, C., Liu, C., Holmes, A., Thornton, P., & Malan, D. J. (2024, March). Teaching CS50 with AI: leveraging generative artificial intelligence in computer science education. In *Proceedings of the 55th ACM technical symposium on computer science education V. 1* (pp. 750-756).
- Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education sciences*, 13(4), 410.
- OpenAI (2023). Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*.
- OpenAI (2025). OpenAI GPT-4.5 System Card, <https://openai.com/index/gpt-4-5-system-card/>
- Sadasivan, V. S., Kumar, A., Balasubramanian, S., Wang, W., & Feizi, S. (2023). Can AI-generated text be reliably detected?. *arXiv preprint arXiv:2303.11156*.
- Yoshida, L. (2024, July). The Impact of Example Selection in Few-Shot Prompting on Automated Essay Scoring Using GPT Models. In *International Conference on Artificial Intelligence in Education* (pp. 61-73). Cham: Springer Nature Switzerland.
- Wang, J., & Fan, W. (2025). The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking: insights from a meta-analysis. *Humanities and Social Sciences Communications*, 12(1), 1-21.
- Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., ... & Fedus, W. (2022a). Emergent abilities of large language models. *arXiv preprint arXiv:2206.07682*.
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., ... & Zhou, D. (2022b). Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35, 24824-24837.
- Zhou, M., Abhishek, V., Derdenger, T., Kim, J., & Srinivasan, K. (2024). Bias in generative AI. *arXiv preprint arXiv:2403.02726*.