

# 生成AIの強みを活かす対話型歴史学習アプリの開発と授業デザイン —— 概念形成・批判的思考を促す歴史人物シミュレーターによる 個別最適化授業の可能性 ——

Development of an Interactive History Learning App Leveraging Generative AI Strengths and Instructional Design

—— The Potential for Individualised Optimisation in Teaching Through Historical Figure Simulators Fostering Concept Formation and Critical Thinking ——

濱田 英毅\*・佐藤 雄太\*\*・鈴木 博幸\*\*\*

玉川大学教育学部\*・株式会社みんがく\*\*・福島県立光南高等学校\*\*\*

生成AIは歴史学・歴史学習との親和性が高いと指摘されてきたが、英語教育などに比べその研究・実践の蓄積は十分とはいえない。本稿では、筆者らが生成AIの利点と限界を踏まえて設計した「歴史人物シミュレーター」を取り上げ、概念形成と批判的思考の促進、ならびに学習者に応じた個別最適な学びの実現可能性を検討する。あわせて、授業デザインの重心を教授 (Teaching) から伴走・支援 (Coaching/Facilitation) へ移すスタイル転換を提起する。

キーワード：生成 AI, シミュレーション教材, ソクラテス問答, リフレクション

## 1. はじめに

生成AIと歴史学・歴史学習の潜在的親和性が高いとされる一方で、生成AIを歴史学習で活用するには若干の問題点が残る。本稿は、歴史学習で活用する場合の課題や、生成AIの技術的特長・限界を明らかにし、教育的に配慮した設計を加えた「歴史人物シミュレーター」を活用することで、概念形成・批判的思考を促す個別最適化された歴史学習を提唱し、それを活用するための授業デザインを提案する。

## 2. 生成 AI と歴史学・歴史学習

### 2.1. 従来の AI 観と近年の生成 AI の進化

2011年、教育とAIの本格的なかかわりを模索する「ロボットは東大に入れるか」プロジェクトが始まった。当初、2021年度に東京大学の合格ライン突破を目標としたが、2016年に一定の水準に達したとしてプロジェクトは一段落。以後、AIの弱点であると同時に人間の強みであると考えられた読解力と意味理解を伸ばすためのリーディングスキルテストの実施へと軸足が移った。

しかし、2022年11月に生成AI (Generative AI) のChatGPTが一般公開され、2024年9月に論理的推論モデル (Reasoning Model) が登場すると、

2025年4月には論理的推論モデルのChatGPT o1とDeepSeek R1が東京大学理科Ⅲ類の合格ラインを突破。「知識に比べ幼稚な知性」(新井 (2016)) という従来のAI観は覆されるに至った。

なお、2025年8月末時点の生成AIのIQレベルはGPT-5 Proが123、Gemini 2.5 Pro /Claude-4 Opusが118で、博士号取得者に相当するレベルに達した。もはや生成AIの知識レベルは、高度な専門性を有する教師に比肩、もしくはそれを超えるレベルにある。ゆえに、今後の教育は、生成AIが教師や生徒より高い知識レベルにあることを前提に、生成AIといかに向き合い、生成AIの知識に飲み込まれることなく、どのように主体性を保ちながら活用できるかが問われているといえる。

### 2.2. 生成 AI と歴史学・歴史学習の潜在的親和性

あらゆる職業の中で、歴史に関連した職業が強い影響を受けるという予測はあまり認識されていない。Felten (2023) は、774の職業が生成AIの進歩で影響を受ける度合いをAI職業エクスポージャー (AIOE: AI Occupational Exposure) として指標化。第4位を「History Teachers, Postsecondary」とした。また、Tomlinson (2025) は、Microsoft copilotの会話データ分析から、生成AIがどの職業 (作業活動) でどの程度使われたかAI適用性スコア (AI applicability

score) に算出。Top40の第2位を「Historian」とした。

いずれの研究も、歴史に関連した職業の特性、ならびに作業活動が、生成AIの機能的特長であるデータベースの探索、様々な情報源から情報を得ること、情報整理（統合や要約）、文章表現力等と重なるために、生成AIと歴史学・歴史学習の潜在的親和性が非常に高いことを示している。

以上の予測とは別に、米国歴史学会（American Historical Association）は生成AIが歴史学の教育と学習に与える影響を分析。2025年8月5日に「指導原則」を発表した。生成AIの生成物を鵜呑みにせず、歴史家が従来行ってきた出典の批判的吟味、バイアスの特定、文脈の理解といった批判的なレンズを通して内容を評価する必要性を指摘した。言い換えれば、AI時代の歴史学習は読解力や批判的思考力の育成、AIリテラシーの育成が鍵で、生成AIを批判的思考の代替ではなく、あくまで補助ツールとして活用する方法を身につける必要がある、との指摘である。

一方、日本の歴史学で、生成AIと歴史学との関係分析は遅れている。大知（2025）の2024年8月時点の研究では、日本の歴史学はハルシネーションや解釈の精度に冷笑したり懸念したりするばかりで、生成AIと向き合い、その機能や特質を論じようとしていない。デジタルヒューマニティーズ分野では歴史学を含む人文学における利用価値について大いに注目されているが、いずれにせよ歴史家の生成AIの機能や特質に関する理解度が低いことは問題である。

### 2.3. 生成 AI と授業科目としての歴史の関係

生成AIと歴史学・歴史学習の潜在的親和性の高さが示唆される一方で、実際に授業科目としての歴史の学習で生成AIを活用するには、若干の課題が残る。前出・生成AIが東京大学理科Ⅲ類の合格レベルを突破した際の解答の分析によると、他教科と比べ日本史の正解率が低い。共通テストはo1が文系合計91%、理系合計が90%に対し、日本史探究と化学が最も低く79%。R1が文系合計82%、理系合計が80%に対し、数学1Aの59%、物理の67%に次いで日本史探究が70%であった。ただし、世界史探究が両モデルとも97%であり、LLM (Large Language Model) の日本語データの少なさが問題とも考えられる。

また、二次試験の正解率はo1が文系科目で64%、理系科目で63%に対し、日本史と理系数学が最も低い32%。R1が文系科目で60%、理系科目で64%に対し、文系数学23%、理系数学41%の次に低い42%であっ

た。ただし、共通テストと同様に、世界史がo1で57%、R1で67%、地理がo1で68%、R1で62%と、地理歴史という教科全体で考えれば、他の教科と比べて決して低い水準ではない。問題点は次の通りである。

- ▶ 「それっぽい文章」を作るが、要点を欠き、抽象・具体のバランスが取れない
- ▶ 問題に忠実すぎる答案を作成するため、日本語的に不自然な箇所ある
- ▶ 文章の表層をまとめ、翻訳することはできても、出題者の意図を汲み取って解答する問題が弱い
- ▶ ただし、教科書通りの知識を持った状態なら、模範解答の作成に役立つと考えられる
- ▶ ただ特徴を羅列するのではなく、それを噛み砕き論理的に整合性が取れた文章を作成する必要
- ▶ 細かい絵の読み取りは、大幅に人間より劣る
- ▶ 独自の視点を持たないAIは世界中の不合理的な問題を考えられない。人間が補助として部分的に人間以上の知能を持つAIを使用すればよい

以上の問題を克服した時、授業科目としての歴史の学習に最適な生成AIツールとなるであろう。（濱田）

## 3. 生成 AI の特長と問題点の克服方法

### 3.1. 生成 AI の技術的特長と構造的限界

生成AIの技術的特性が学習支援に適する理由は、次の点に整理できる。

- ▶ LLMは膨大な文献・史料を瞬時に参照し、学習者の多様な質問に個別対応できる
- ▶ 自然言語による双方向対話を通じて、学習者の理解度に応じた柔軟な説明が可能である
- ▶ 24時間利用可能な個別指導の機能を有し、時間・場所の制約なく個別最適な学習支援を提供できる
- ▶ 多様な表現形式で情報提示し、学習者の認知特性や学習段階に応じた最適な情報提供ができる

ただし、Bender et al. (2021) らが指摘するように、教育における生成AIの活用を考える場合には、生成AIの内部構造に起因する構造的限界についても理解しておく必要がある。

- ▶ 事実に基づかないもっともらしい情報を生成するリスク（ハルシネーション）が一定程度存在する
- ▶ 特に指定しない限り、特定の情報がどの史料・研究に基づいているのか明示できず、情報源が不透明で、史料批判や出典確認の習慣形成を阻害する
- ▶ 文脈理解が浅く、表層的な情報統合は可能でも深層的な因果関係や構造的連関の把握に限界がある

### 3.2. RAGの導入による信頼性の向上

上記の限界を克服する最も有効な手法が、RAG（Retrieval-Augmented Generation・検索拡張生成）の導入である。RAGは外部の信頼できるデータベースから関連情報を検索し、それに基づき回答する。

検定教科書、史料集、学術論文データベース等をRAGの知識源として組み込めば、生成AIの回答に明確な根拠を付与し、ハルシネーションリスクを大幅に軽減できる。また、出典表示機能により、学習者が情報の出典を常に確認する習慣を獲得できる。

### 3.3. 生成AIの教育的活用のための設計原則

Lee et al. (2025) は、生成AIへの信頼度が高いと批判的思考力が少なくなり、長期的で過度な依存が問題解決能力の低下につながる可能性を指摘する。また、エストニア政府が高校生に提供する生成AI活用スキルを学ぶAI Leapプログラムも、批判的思考力の育成とAI活用リテラシーの育成を主眼とする。

ゆえに、生成AIの教育的活用を進めるに際しては、上記の技術的特長と構造的限界をふまえつつ、批判的思考力の育成につなげる必要がある。そこで、濱田・鈴木が開発した「歴史人物シミュレーター」ではRAGを導入し、ソクラテス問答のモードを採用することで、学習者が自ら思考し、批判的に検証しながら学ぶ仕組みを取り入れた。（佐藤）

## 4. 学校現場の課題と授業デザイン

### 4.1. 次期「学習指導要領」を見据えた課題

現行「学習指導要領」では、ICTを最大限に活用し、これまで以上に「個別最適な学び」と「協働的な学び」を一体的に充実し、「主体的・対話的で深い学び」の実現にむけた授業改善が求められている。しかし、学校現場には二つの構造的課題がある。

- ▶ 地理歴史科は教員の専門領域や Pedagogical Content Knowledge (PCK：教科内容と教授法の統合知) の濃淡で、授業の質にばらつきが生じる
  - ▶ 一斉授業の中で、適切な足場かけと評価的フィードバックを全員に同時提供するには限界がある
- 加えて、次期「学習指導要領」の議論（有識者検討会の論点整理）では、「概念的知識」と「学習方略」の獲得が重視されている。教師が研修等で自身の教育的専門性を高めることは前提でありつつも、個別最適な学習支援を可能とする生成AIを補助ツールとして活用しない限り、上記の課題の解決は難しい。

### 4.2. 「歴史人物シミュレーター」の提案

授業に関連した上記の諸課題の解決につながる生成AIの補助ツールとして、佐藤の設計原則をふまえつつ濱田・鈴木で「歴史人物シミュレーター」を開発した。2024年度にVer.1を発表し、2025年度はソクラテス問答の仕組みを入れたVer.2の実装に取り組んでいる。

歴史上の重要人物（Ver.1で86名）等に関する知識は、旺文社から提供を受けたデータのRAG化でハルシネーションを防止しつつ、適度なレベルの知識で回答されるような配慮をした。また、適切な足場かけや評価的フィードバックで、学習者の習熟度に応じた個別最適な学びを実現しつつ、学習者が時代像や人物像という「概念的知識」を身につけることができる仕組みとした。また、「概念的知識」の具体的なレベル設定のため「思考コード」を参考とした。

本ツールの根底にある教育観は、Wineburgが目指す、史料に基づく探究で「歴史的思考」を培う学びの実現である。歴史学習は、どうしても学習者の習熟度（既有知識の確かさ）が前提となるが、生成AIとの問答で「概念的知識」を身につけるハードルが下がれば、より高次元の学びへと進む足掛かりになり得る。また、学習者や教師はどうしても現代の価値観や視点から判断・解釈してしまう（Presentism）傾向がある。これは歴史理解を妨げる最大の障害であると、開発者としても認識を共有している。

そこで、本ツールは、過去をその時代の文脈・その歴史的人物の文脈で回答するようにプロンプトを設定し、あくまでも当時の人々の環境や立場を理解した上で（Historical Empathy：歴史的共感）、学習者自身の解釈・判断を問う仕組みとした。感情移入とは異なる、より知性的な判断を促す学びの構想である。

### 4.3. 生成AIを活用した授業デザイン

以上の設計思想に基づき開発した「歴史人物シミュレーター」を活用する授業デザインを提案する。

#### ① 予習：反転学習【ソクラテス問答モードの使用】

特に家庭学習では、フィードバック可能な教師がいないため、単に解答を出さない、学習者の習熟度に応じて問いの抽象度・ヒントの粒度を個別最適に自動調整したソクラテス問答モードが必須である。適切な追問を途切れさせず提示する対話相手として機能し、生徒は自分の言葉で因果と文脈の骨格を組み立てて、授業の前段で概念の土台を暫定形成する。

また、生成AIとの対話のみによるのではなく、引

用・出典の明示と簡易ファクトチェック（複数史料・教科書記述との照合）を必須タスク化する。

## ② 教室【シミュレーション・モードの使用】

教室では、同時代の複数AIペルソナ（例：豊臣秀吉／徳川家康）を切り替えることで、一つの歴史的事象に対する多面的・多角的な視点の形成が可能である。具体的には、因果（何が・どのように・どこへ）や利害（誰に何がもたらされ、何が損なわれたか）の観点等を生徒に提示して、何を問えばよいのかという「学習方略」の学びにつなげることが重要である。

この際に注意すべきは、Presentismへの配慮である。あくまでも当時の人々の価値観や、当時の仕組みを、まず理解する。その上で、当時の人々の課題解決である行動の意味や意義を問うと、「自分ならばこうする」というシミュレーションの意義が生じる。

また、シミュレーションで得られた学びの成果が前提となり、史料に基づく探究で「歴史的思考」を培うような、高次元の学びに接続することができる。

## ③ 教室【振り返りモードの使用】

授業の終盤、ないし単元学習の最後に、メタ認知ドリルとしての振り返りモードを活用する。生成AIが学びに関する短問をテンポよく提示し、生徒は自己説明→反証提示→再主張の短サイクルを数セット繰り返し、概念の定着と転移可能性の準備を整える。

上記①～③の学びが示唆するのは、授業レベルの標準化である。教員の専門性の濃淡による授業格差、生徒の学力格差の是正といった、学校現場の構造的課題は解決されるであろう。その上で、教員の力量次第では、「歴史的思考」を培うような、より高次元の学びへと進むことが容易になる。また、学習者の習熟度に応じた適切な足場かけと評価的フィードバックは、生徒が歴史的概念（歴史像・人物像）を自ら構築し、当事者意識をもって歴史的事象を判断・評価できるようになる助けとなる。（鈴木・濱田）

## 5. おわりに

本稿は、生成AIの歴史学習における課題を指摘し、その解決策としてRAGとソクラテス問答を導入した「歴史人物シミュレーター」の活用を提案した。本ツールはPresentismを避け歴史的共感を育みながら、主体的に歴史的概念を形成する支援をする。予習から授業、振り返りまで一貫した活用で、教員の専門性や生徒の習熟度によらない個別最適な学びを実現し、より高次元な歴史的思考への扉を開くものである。

## 参考文献

- サム・ワインバーグ (2017). 歴史的思考 その不自然な行為, 春風社.
- 大知聖子 (2025) 【展望：デジタル人文学の可能性】歴史学と生成AI, 文明動態学, Vol.4 (March 2025), pp.110-122.
- Wineburg, S. (2001). Historical Thinking and Other Unnatural Acts: Charting the Future of Teaching the Past. Temple University Press.
- Bender, E. M. et al. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? , Conference on Fairness, Accountability, and Transparency (FAccT '21), pp. 610-623. <https://doi.org/10.1145/3442188.3445922>
- Felten, E. et al. (2023). How will Language Modelers like ChatGPT Affect Occupations and Industries?, arXiv:2303.01157 [econ.GN], [v2] 18 Mar 2023.
- Lee, H-P. et al. (2025). The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers, CHI Conference on Human Factors in Computing Systems (CHI '25), <https://doi.org/10.1145/3706598.3-713778>
- Tomlinson, K. et al. (2025) Working with AI: Measuring the Occupational Implications of Generative AI, arXiv:2507.07935 [cs.AI], [v3] 22 Jul 2025.
- American Historical Association. (2025). Guiding Principles for Artificial Intelligence in History Education.
- 文部科学省 (2024). 今後の教育課程、学習指導及び学習評価等の在り方に関する有識者検討会 論点整理 (2024年9月18日), 文部科学省, 東京.
- 新井紀子 (2016). 知識に比べ幼稚な知性 アンバランスなAI, 日経産業新聞 2016年11月10日付, <https://www.nikkei.com/article/DGXXKZO09343840Z01C16A1X12000/?msockid=1bae8aaf785d6a061bd29f8e79b76be3>, 2016年11月15日公開, 2025年8月31日最終閲覧
- 株式会社ライフプロンプト (2025). 【東大理3合格】ChatGPT o1とDeepSeek R1に2025年度東大受験を解かせた結果と答案分析【河合塾採点協力】, <https://note.com/lifeprompt/n/n0078de2ef36b>, 2025年8月31日最終閲覧