

高校生を対象とした生成 AI 型キャリア教育支援システムの有効性評価

Effectiveness Evaluation of a Generative AI-Based Career Education Support System for High School Students

松浦瑛太 *

Eita Matsuura

雲居玄道 *

Gendo Kumoi

Abstract— This study aims to evaluate the human behavior imitation capabilities of generative AI through a career education support system. We constructed and conducted a proof-of-concept experiment with a system where generative AI proposes multiple career paths based on student profile information. Quantitative analysis confirmed high alignment with student intentions, while system usage logs enabled the categorization of decision-making processes. This demonstrated the effectiveness of generative AI's human behavior imitation capabilities in career education.

Keywords— Generative AI, Career Education, Human Behavior Mimicry Capability

1 研究背景と目的

近年、大規模言語モデル（LLM）を用いた生成 AI 技術の発展により、人間の意思決定や行動を模倣したパターンを分析する「Simulation with Generative Agents (SGA)」が注目されている。例えば、Bougie らは大規模シミュレーションにより、都市における人間の行動をモデル化できることを示した [1]。また、Park らは、約 2 時間にも及ぶインタビュー結果を入力することで、大規模言語モデルを用いて個人の態度・行動パターンを学習し、一般的な社会調査の回答を約 85% の精度で再現することに成功している [2]。このように、生成 AI が人間の行動選択を高い精度で模倣できる可能性が示されたことにより、企業における A/B テストや災害時の多様な行動分析など、あらゆる分野での応用が期待される。本研究では、SGA をキャリア教育へ活用することを目的とする。

現在、日本の高校教育におけるキャリア教育では、多様化するキャリアパスについて、生徒一人ひとりに対するきめ細かな進路支援が課題となっている。文部科学省はキャリア教育を「一人一人の社会的・職業的自立に向け、必要な基盤となる能力や態度を育て

ることを通して、キャリア発達を促す教育」と定義し [3]、その推進のために「キャリア・パスポート」を含む各種教材や手引きを提供している [4]。しかし、現場では配布された冊子への記入で終わる場合や、教師が生徒全員に個別対応する時間が不足するなどの運用上の課題が指摘されている。具体的には、教員の時間的制約による個別指導の限界、形式的な資料配布のみで終わる実態、学年間での情報引継ぎ不足、取組効果を定量評価する仕組みの欠如等が挙げられる。これらの課題の結果、生徒の興味・適性に合致した進路提示や主体的な進路選択の促進が十分に行われていないのが現状である。

本研究では、これらの背景を踏まえ、SGA を用いた高校生向け個別キャリア支援システムを開発し、実際の教育現場においてその有効性を評価する。Park らにより示された SGA の高精度な行動模倣能力 [2] を手掛かりに、高校生のプロフィールアンケート回答から 18 歳から 30 歳までの詳細なキャリアパスを自動生成し、生徒一人ひとりにパーソナライズされた進路オプションを提示する。従来の適性検査や進路指導は画一的かつ短期的な助言にとどまりがちであったが、本研究のアプローチは生徒一人ひとりのプロフィール情報に基づく、長期的視点に基づくキャリアプランの提案を特徴としている。

さらに、本システムを実環境での大規模実証として新潟県立高校 245 名を対象に実施するため、主体性を支える UI 設計を実装し、生成 AI の教育分野への応用可能性を実証する。

2 提案手法

本研究では、高校生の将来キャリア選択を支援するキャリアパス提案手法を提案し、ユーザが利用可能なシステムを開発した。このキャリアパスの提案を SGA としての人間の選択行動模倣能力として捉える。この際、Park ら [2] の着想を援用し、インタビューの

* 〒 940-2137 新潟県、長岡市上富岡町 1603-1 長岡技術科学大学,
Nagaoka University of Technology, 1603-1 Kamitomioka-cho,
Nagaoka City, Niigata 940-2137, Japan E-mail: s243366@stn.
nagaokaut.ac.jp, kumoi@vos.nagaokaut.ac.jp

代替として自由記述アンケートを実施し、ユーザのパーソナリティを獲得する。得られたパーソナリティをもとに、ユーザが将来選択する可能性があるキャリアパスを提案するというプロセスである。具体的に、本システムは以下の 3 つの主要コンポーネントから構成される。

(1) アンケートフォーム：文科省の「キャリア・パスポート」例示資料 [4] を基盤として、生徒の過去の体験・現在の興味・未来の希望を体系的に自由記述で回答させる設問を用意。質問内容はユーザ自身の振り返りを促し、自己理解を深めるよう設計されている。詳細のアンケート項目は付録付録 A に示す。このアンケート結果はパーソナリティとして、後述の生成 AI への入力プロンプトとなる。

(2) キャリアパス生成 (SGA)：生成 AI を用いたテキスト生成エンジンである。前段のアンケート回答をプロンプトとして構造化し、18 歳から 30 歳までの詳細なキャリアパス案を生成する。各ユーザに対して 1 回の処理で複数パターン（本研究では 3 パターン）のキャリアプランを出力し、提案内容が固定化しないよう試行ごとにプロンプトの一部内容を変化させる。さらに、SGA は学習データに対して平均的な回答を行うため実際の人間の選択行動よりも分散が少なくなることが知られている [5]。そこで、同一ユーザが再生成を試みることで多様性のあるキャリアパスが選択肢として提示されるような工夫を加えた。

(3) Web アプリケーション：ユーザが生成されたキャリアパス提案を閲覧、比較、選択できるインターフェースである。UI は提供対象となる高校生が直感的に操作できるようシンプルなデザインとし、各提案プランの詳細（例：職業名、必要な学歴や資格、将来展望など）を見やすく提示した。生徒は提示された複数のキャリアパスの中から興味に沿うものを選択または保存でき、不要な提案をスキップすることで自分に合った候補を絞り込める。選択・閲覧履歴および操作ログは自動記録され、後述の行動分析に活用した。また、教員用画面も実装し、生徒の選好傾向や進捗を把握できるようにしている。本システムはフロントエンドに Next.js/React、バックエンドに Firebase（認証・データベース）を用いた Web 技術で構築され、クラウド上でスケーラブルに動作する。生徒のアンケート回答送信から AI 応答生成、結果表示・選択、ログ記録までの一連の処理はすべてリアル

タイムで行われ、学校のタブレットで利用可能な形で提供した。

3 実験方法と評価指標

本研究では、新潟県立長岡大手高等学校の 1 年生 245 名を対象にシステムを導入し、その有効性を検証した。実験は 2025 年 5 月 21 日（水）の探究学習の授業内で実施し、生徒は学校配布タブレットからシステムにアクセスした。

3.1 実験手順

1. アンケート入力：自己紹介、興味・関心、将来の希望などをフォームに記入。
2. キャリアパス生成：生成 AI が 3 通りのキャリアパスを即時生成。
3. 提案の取捨選択：生徒は「保存」「スキップ」「再生成」を自由に操作。
4. 最終確定：最も納得できる 10 件以内のキャリアパスを最終的に選定。
5. アンケート回答：7 項目の主観評価と自由記述を入力。
6. キャリアパスレポートの出力：キャリア教育のため、入力したアンケート内容、選択したキャリアパスに加えて、全体のサマリーレポート、選択したキャリアパスに基づくバックキャスティングレポートを生成 AI が作成し出力。

なお、実験の公平な比較と多様性評価を確保するため、事前に生徒へ「少なくとも 6 回以上の再生成（計 6 試行以上）を行う」旨を明示的に指示した。これは、6 回目以降で前述の多様性ある出力が有効化されるシステム設計に整合する運用である。

また、すべての操作はログから記録され、欠損のないログとアンケートがそろった 216 名を分析対象とした。

3.2 評価指標および特徴量

提案の受容度と探索行動を各特徴量に基づき定量的に評価する。

本システムは 1 試行ごとに 3 種類の異なるキャリアパスを提示するシステムである。また、システム設計において、最大試行回数 $\alpha = 50$ と設定している。

ユーザ u_i の試行回数を T_i とするとき、 T_i はユーザがキャリアパスの出力について深く考えたり、好むものが選択されなかったりと、試行錯誤した結果

を定量的に表すものである。

本システムではキャリアパス提案の受容判断が 2 段階に分かれている。初期段階は、試行ごとに提示されるキャリアパスについて「Accept」, 「Reject」の 2 種のボタンで受容をユーザが判断するものである。全試行におけるユーザ u_i が Accept と受容の判断をした数を N_i とするとき、Accept 率 A_i を以下のように定義する。

$$A_i = \frac{N_i}{T_i \times 3} \quad (1)$$

Accept 率が高いほど、ユーザはシステムから提示された候補の中でより多くの提案を受容したこと意味する。

2 段階目の選択は、ユーザが第 1 段階で受容したキャリアパス N_i 件の中から、最終的に $L_i \leq 10$ 件に絞り込む作業である。これはキャリア教育の観点から選択したキャリアパスに基づいたバックキャスティングによる分析を各自が実施するため、受容したキャリアパスについて詳細に検討し選択させた。この選択結果に基づき、選択シーケンス特徴ベクトル x_i をユーザ特徴量とする。 x_i は最大試行回数 α 次元のベクトルであり、 $x_{ij} \in x_i$ は、 j 回目の試行で出力されたキャリアパスのうち、最終的な選択に残ったキャリアパスの数を意味する。このため、 $L_i = \sum_{j=1}^{N_{\max}} x_{ij}$ となる。

3.3 行動クラスタリングと主観評価

x_i は、ユーザがどの試行でのキャリアパスを最終選択に残したかを意味する特徴量である。この x_i について、Ward 法による階層クラスタリングを用いてユーザの行動タイプを分析する。また、7 項目アンケートから提案品質とユーザ満足度を分析した。

4 結果と考察

4.1 試行回数

各ユーザが最終的なキャリアパスを決定するまでに要した試行回数の分布を図 1、統計量を表 1 に示す。

以上の結果から、ユーザは平均して約 13 回の試行を経て最終決定に至っていることがわかる。中央値(11 回)と最頻値(10 回)が近い値となっており、多くのユーザは約 10~11 回程度の試行で納得のいく結果に達していることが示唆される。また、分布は右に裾を引いており、一部のユーザは 40 回以上にわたる

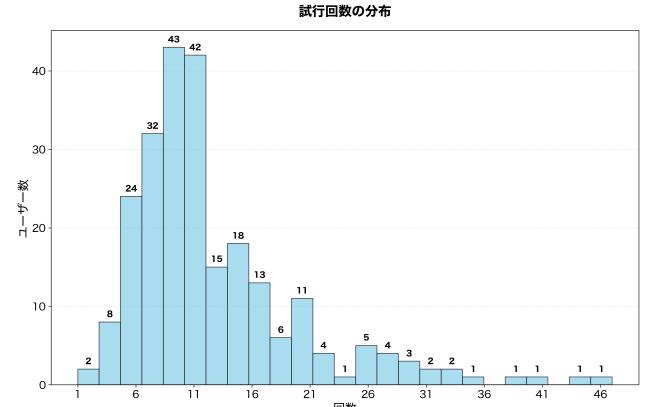


図1 全ユーザの試行回数の分布

表1 試行回数の統計量

統計量	値
平均	12.84 回
中央値	11 回
最頻値	10 回
最小	1 回
最大	47 回

表2 Accept 率の比較

	1~5 回目	6 回目以降	全期間
平均値	35.29%	29.91%	33.10%
中央値	33.33%	26.67%	30.56%

非常に多くの試行を重ねて最終決定している傾向が確認された。これはユーザの意思決定プロセスに個人差があり、一部のユーザが慎重な意思決定プロセスを採用していたことが明らかとなった。

4.2 Accept 率

SGA によるキャリアパス提案とユーザの選択行動との関係性を明らかにするため、ユーザ毎の Accept 率分布を図 2 に示す。

また、本システムでは 6 回目以降でキャリアパス提案から新しい発見を促すため多様性を增幅する多様性ブーストを行っている。そこで、前期(5 回目まで)と後期(それ以降)で層別したデータに基づいたユーザ毎の Accept 率分布を図 3, 4 に示す。さらに、Accept 率の統計量を表 2 に示す。

これらの結果より、全期間を通じて Accept 率は約 30% 程度の水準にあり、システムから提示された候補の約 3 分の 1 を最終的に受容していることがわかる。また、前期(35.29%)から後期(29.91%)にかけて平均 Accept 率が 5.38 ポイント低下している。これは、6 回目以降に導入された多様性ブーストによ

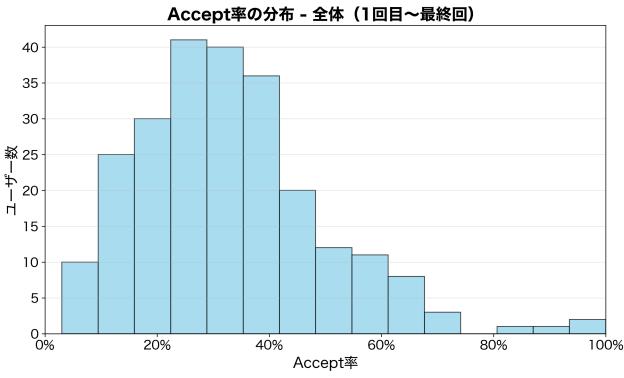


図2 全期間（1回目～最終回）におけるユーザAccept率の分布

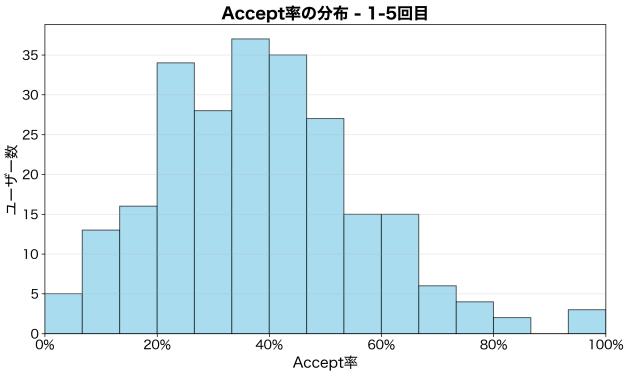


図3 前期（1～5回目）におけるAccept率の分布

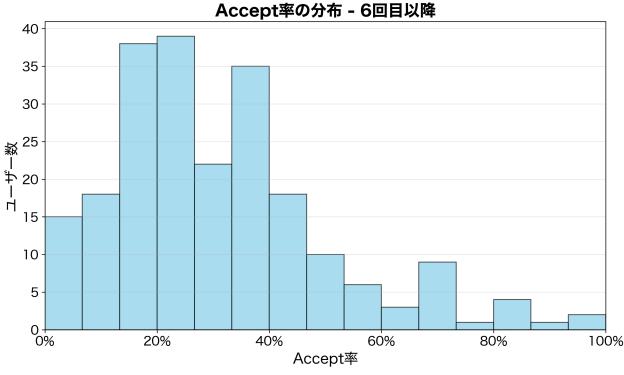


図4 後期（6回目以降）におけるAccept率の分布

り、より幅広い選択肢が提示されるようになったため、ユーザは多くの候補から選択する機会を得た一方で、受容できない提案が増えたことが原因であると考えられる。一方で、多様性が増幅されたとしてもAccept率の分布も大きな偏りは生じず、平均的に約30%のAccept率であったことから、ユーザにより多様かつ受容可能な選択肢を提供できたと考えられる。

ただし、これらの結果は、被験者へ6回以上の試行を案内したため、後期の探索幅が意図的に拡張されている。さらに、最終選択は10件以内としたため、ユーザの受容件数は10件程度を選択することが1つの基準となっている。そのため、Accept率や前期・後

表3 ユーザ行動パターンのクラスタ分類 ($N=216$)

クラスタ	人数(割合)
超早期決定型	50人(23%)
中期探索型	64人(30%)
広域探索型	102人(47%)

期の単純比較には介入効果（設計上の操作）の影響が含まれる点に留意する必要がある。

以上より、本システムの提案受容行動に関して、ユーザは提示された候補を一定の水準で選択的に受容していた一方で、一部ユーザは多くの提案を試しながら取捨選択を行っていたことがわかる。

4.3 ユーザ行動パターンのクラスタリング分析

各ユーザの選択行動パターンを解析するため、 x_i について、ユーフリッド距離に基づく階層的クラスタリング結果を図5に示す。この結果、ユーザの行動パターンは表3に示すように大きく3種類にクラスタリングされた。各クラスタの概要は以下の通りである。

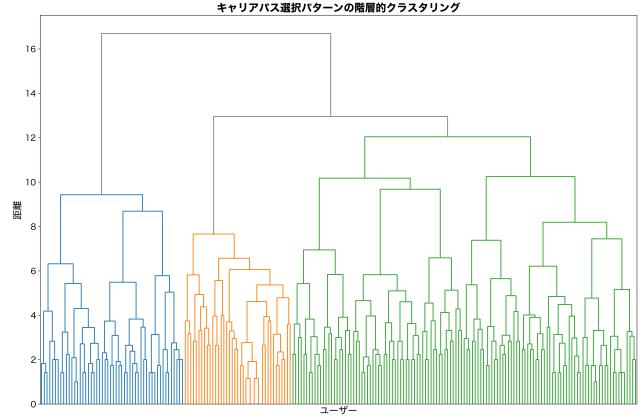


図5 Ward法によるユーザ行動パターンの階層的クラスタリング結果

図5に示す青色のクラスタ1「超早期決定型」は全ユーザの約23%を占め、初期の1～3回程度の試行で必要な選択をほぼ終えるグループである。オレンジ色のクラスタ2「中期探索型」は約30%のユーザが属し、1～8回目あたりの中期にかけて計画的に探索を進めて最終決定するタイプである。緑色の最大のクラスタであるクラスタ3「広域探索型」（約47%）は、全期間にわたって提案の探索と見直しを繰り返し、最後まで慎重に検討を続ける傾向が見られる。

以上の結果から、ユーザのキャリアパス選択行動には画一的ではなく少なくとも3通りの異なるプロセスが存在することが明らかになったと言える。本シ

表4 ユーザアンケート主要項目（5段階）

評価項目	平均値	中央値	標準偏差
システム再利用意向	3.1	3	1.0
提案の多様性満足度	3.3	3	0.9
提案の実用性満足度	3.0	3	0.8

表5 主観評価アンケートの分布

評価項目	< 3	= 3	3 <
システム再利用意向	16.10%	56.59%	27.32%
提案の多様性満足度	13.17%	56.59%	30.24%
提案の実用性満足度	13.17%	62.44%	24.39%

システムは多様な意思決定パターンに対応する必要があり、各クラスタの特徴に応じた支援策の検討が今後の課題となる。

5 主観評価アンケート結果

実験後に実施した主観的なユーザ評価アンケートの結果について述べる。アンケートでは、本システムに関するいくつかの設問（5段階評価）に回答を得た。主な評価項目として「システムを今後も継続利用したいか（再利用意向）」「提案されたキャリアパスの多様性」「提案されたキャリアパスの実用性」などが含まれる。表4に各項目の平均値・中央値・標準偏差を示す。

全体的に、ユーザの主観評価は5段階中3前後とおおむね中程度の値に留まった。全項目で中央値は3となっており、過半数のユーザが評価を「どちらともいえない（3）」もしくはそれ以上と回答することを示す。一方で最高評価の「5」を付けたユーザは限定的で、平均値も3をやや上回る程度にとどまっている。これは、授業の最後に実施され短時間の中でのアンケートであり、デフォルト設定として「3」を設定したため、多くのユーザが評価を変更せずに回答したことが原因であると考えられる。

そこで、デフォルト設定の「3」から意図的に変更した分布を表5に示す。この結果から、8割以上はシステムへ負の評価を持たず、高評価をつけた生徒が3割程度いることがわかる。特に多様性の満足度は高く、キャリアパスの提案パターンとしては十分であったことが伺える。一方で実用性の満足度や再利用の意向は多様性よりも低下した。このため、多様性はあるが、キャリアパス提案の方向性が生徒にとっての満足度としてそこまで高いものではなかった。

しかし、本システムの目的は高校1年生という視野がまだ狭い段階において、多様なキャリアの可能性を提示し、視野を広げるサポートをすることであった。この観点においては、十分にその成果を示すことができたと考えられる。

この点は、生徒へ実施した別のアンケートにおいて、「探究に取り組むことを通して、自己を理解し、将来の在り方生きを考えることができる」という項目の回答が、入学前は2.84であったことに対し、本システム利用後には3.40と大きく増加したことからも伺える。

6 まとめと今後の課題

本研究では、生成AIを用いたキャリアパス提案システムを開発し、高校生を対象としたフィールド実験を通じてその利用実態と有効性を分析した。ユーザは平均13回程度の提案試行を経て10件のキャリアパス候補を選定しており、システムから提示された全候補に対する受容割合（Accept率）の平均値は33.1%であった。これは、ユーザがシステムから提示された多様な選択肢の中から慎重に選別を行い、自身の興味や適性に合致するものを選択的に採用していたことを示している。また、ユーザの選択行動パターンは少なくとも3種類にクラスタ分類でき、一部のユーザは初期段階で迅速に決定する一方、他のユーザは最後まで慎重に探索を続けるなど、意思決定プロセスに個人差が見られた。

一方で、主観評価アンケートの結果からは、システム利用後のユーザ満足度は概ね中程度であり、客観的な利用状況との間にギャップが見られた。ユーザはシステムから事実上大きな恩恵を得ていたものの、その価値を十分に実感できていなかった可能性がある。このギャップを解消するためには、システムの利便性向上や提案内容の質のさらなる向上、ユーザへのフィードバック機能の強化などが考えられる。

以上より、本システムはキャリア選択支援のツールとして一定の有用性を持つことが示された。しかし、ユーザ体験を向上させる余地があることも明らかとなった。今後の課題として、ユーザが提案の価値を直観的に理解できるインターフェースデザインの工夫や、個々のユーザの探索スタイルに応じたパーソナライズされた支援（例：広域探索型のユーザには提案履歴の整理を支援する機能を提供する等）が挙げられ

る。また、長期的なキャリア形成への影響を評価するため、提案したキャリアパスに沿った進路決定の実現性や、ユーザの自己効力感の変化について追跡調査を行うことも有益であろう。本研究の知見を踏まえ、今後さらなるシステム改良と実証を重ねることで、AI技術を活用した効果的なキャリア支援の実現を目指す。

謝辞

本研究の一部は科研費 JP24K20833 の助成を受けたものである。

付録 A アンケート項目

• 過去の経験について

– 小学校時代で一番心に残っていること

- * 出来事
- * 理由
- * それを通して成長したこと

– 中学校時代で一番心に残っていること

- * 出来事
- * 理由
- * それを通して成長したこと

• 自己分析について

– 自分の長所

- * 特性・性格・個性
- * 自分の強み
- * 今後より成長させたい強みとその理由

– 自分の興味関心

- * 特技・趣味・興味があること

• 高校生活について

– 高校生活で大切にしたいこと

- * 大切にしたいこととその理由
- * 身につけたい特技や能力とその理由

• 将来の展望について

– 進路・将来

- * 高校卒業後のイメージ
- * 希望する職業やライフスタイル

– 1年後の自分

- * できるようになりたいこととその理由
- * 1年生のうちに成長させたい部分とその理由
- * この1年で特に努力したいこととその理由

– 30歳の自分

- * できるようになりたいこととその理由
- * 高校生のうちに成長させたい部分とその理由
- * 高校生活で特に努力したいこととその理由

• 進路選択の価値観について

– 進学・就職先の決断で重視したいこと

- * 重視する項目
- * 具体的に重視すること
- * 重視する理由

参考文献

- [1] N. Bougie and N. Watanabe, “Citysim: Modeling urban behaviors and city dynamics with large-scale llm-driven agent simulation,” 2025. <https://arxiv.org/abs/2506.21805>
- [2] J.S. Park, C.Q. Zou, A. Shaw, B.M. Hill, C. Cai, M.R. Morris, R. Willer, P. Liang, and M.S. Bernstein, “Generative agent simulations of 1,000 people,” 2024. <https://arxiv.org/abs/2411.10109>
- [3] 中央教育審議会, “今後の学校におけるキャリア教育・職業教育の在り方について（答申）,” 2011. 答申日：平成23年1月31日.
- [4] 文部科学省, “「キャリア・パスポート」例示資料等について：文部科学省,” https://www.mext.go.jp/a_menu/shotou/career/detail/1419917.htm.
- [5] J. Bisbee, J.D. Clinton, C. Dorff, B. Kenkel, and J.M. Larson, “Synthetic replacements for human survey data? the perils of large language models,” Political Analysis, 2024.