

# インタラクティブマップ型 SNS 検索システムの LLM 設定人格を用いたユーザ評価実験

神場知成<sup>†1</sup>

**概要:** フィルターバブルの軽減を目的として、多様な視点を提示するインタラクティブマップ型 SNS 検索システムの開発を進めている。本研究ではその効果を検証するための予備的評価として、大規模言語モデル (LLM) によって設定した複数のペルソナを被験者とし、システムを継続的に利用した際に生じる閲覧投稿の選択傾向の変化を観察した。具体的には、指定した検索ワードに対する SNS 検索結果 100 件から LLM が自動生成したインタラクティブマップの軸名や各投稿の座標値を、プロンプトでキャラクター設定したペルソナに提示し、ペルソナが興味のある投稿 10 件を順次閲覧する形式とした。閲覧後には、理解・共感・反感などの反応があれば、自身の設定に反映するよう指示した。また、一部のペルソナには実験前に「マップ全体を広く見るように」という追加指示を与えた。7 日間の実験の結果、この追加指示は初期段階では一定の効果を示したものの後半では薄れていき、ペルソナが元来のキャラクター設定に基づく嗜好へと回帰する傾向が確認された。LLM 設定ペルソナによる評価は、実際の人間の行動と異なる可能性を含む一方で、多様な属性・人格を容易に設定できるという利点がある。今後は、人間被験者による評価結果との比較を通じてこの方法の妥当性と限界をさらに検討する。

## 1. はじめに

近年、ソーシャル・ネットワーキング・サービス (SNS) は人々がニュースや意見を収集・発信する主要な媒体となっているが、利用者が自分の関心や立場に近い情報ばかりに触れることで、異なる視点が遮断され思考や態度が極端化する「フィルターバブル」や「エコーチェンバー」が懸念されている[1,2]。特にレコメンデーション機能やタイムライン最適化は既存嗜好を強化し、多様な情報接触を阻害する要因とされる[3,4,5]。

筆者等はこの課題に対し、SNS 検索結果から大規模言語モデル (Large Language Model: LLM) を用いてリアルタイムで複数視点を抽出し、2 軸マップ上に可視化することでバランスの取れた閲覧行動をとれるようするインタラクティブマップ型 SNS 検索システムを開発した[6]。

本研究では、このシステムの有効性を効率的かつ疑似的にユーザ評価するために、LLM に仮想的な人格を設定した複数の「被験者」を設定し、その被験者が本システムを複数回利用していく中での閲覧傾向の変化を観察するとともに、実際の人間による利用との相違についても考察する。

## 2. 関連研究

### 2.1 フィルターバブルやその軽減策について

フィルターバブルやエコーチェンバーにより、イデオロギー的な論争や偽情報の拡散が強化されるリスクがあり、これらは世界の分断や衝突を招く要因ともなる。これらはもちろん政治的に解決すべき問題もあるが、技術的にもさまざまな対策の可能性がある。

例として、Opinion Space [7]、DaCENA [8]のような可視化とインタラクションを組み合わせた手法や、セレンディピ

ティと呼ばれる発見的な情報との出会いを促す方法などが検討されている[9-12]。筆者等のシステムは、SNS 検索結果に対し、LLM による多様な視点をインタラクティブマップ形式で提供することで利用者が自然に視野の拡大ができることを目指している。

### 2.2 LLM 設定のペルソナによる UI 評価について

一般に、システムのユーザインタフェース評価において、実ユーザ調査・実験はコストと時間がかかるため、LLM にペルソナを設定して疑似的にユーザの応答や行動を生成して評価する手法も検討されている。特にこの手法は、条件統制や繰り返し実験を容易にし、初期段階の評価や仮説検証に役立つ可能性が示されている[13-15]。

本研究は、インタラクティブマップを用いた SNS 検索システムが人間の閲覧行動に与える影響を調べるための初期実験として、LLM で設定したペルソナを被験者とする。

## 3. SNS 検索システム概要

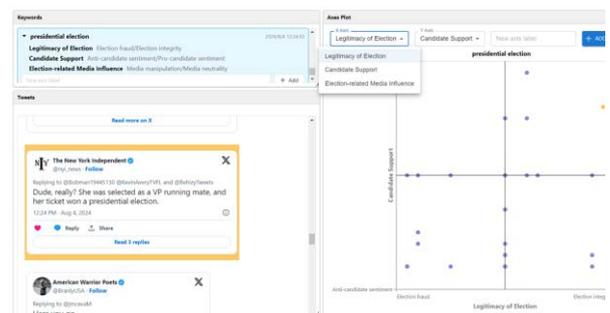


図 1. インタラクティブマップ型 SNS 検索システム

<sup>†1</sup> 東洋大学 情報連携学部 (INIAD)

本研究でユーザ評価に用いるインタラクティブマップ型 SNS 検索システムは、著者らが先行研究として開発・報告したものである[6]。ユーザインタフェース画面例を図 1 に示す。現時点で検索対象は X (twitter) である。

ユーザが検索キーワード (例：総裁選) を入力すると、システムは X の検索結果から LLM が自動的に複数の軸 (視点) と、その軸の両極の特徴 (例：経済対策, 積極的/消極的) を抽出し、各投稿をその視点から見たときの位置づけをスコア化し、任意の 2 軸から成るマップ上にノードとして配置する。ユーザは軸の切り替えや自分で考えた言葉による軸の追加も可能である。マップ上のノードに対するクリックと、自動スクロールして画面左側に表示される投稿は連動する。図 1 では、presidential election というキーワードの検索結果から抽出された Legitimacy of election および Candidate support という 2 軸のマップ上に X の投稿が配置されている。こうした可視化とインタラクションにより、ユーザが情報の多様性やバランスを意識した閲覧行動を促すことを狙っている。

## 4. LLM 設定人格による評価実験

### 4.1 実験方法

システムを利用した際の被験者の行動変化を確認するため、人間被験者の代わりに LLM (ChatGPT-4o) に設定したペルソナを用いて予備的な評価を行った。検索ワードを入れると、約 10 秒で SNS の検索結果 100 投稿から LLM によってインタラクティブマップの軸名、各投稿の座標値などが計算される。その結果を、プロンプトでキャラクター設定されたペルソナに渡すと、ペルソナは数秒で興味のある 10 件を選んでマップ上に表示し、内容に理解・共感・反感などがあれば自身のキャラクター設定に反映する。これは、人間が投稿を読めば内容によって自分の考え方にも影響があることをシミュレートしている。

実験において一部のペルソナには実験前に「マップ全体を広く見るように」という指示を与え、与えないペルソナとの行動の相違を観察した。

検索ワードには一般に議論が賛否に分かれやすい「vaccination (ワクチン接種)」という単語を選び、下記の 2 種類の対照的なペルソナを設定した。いずれも米国人の設定である。

- **P1**: 米国北東部に住む 45 歳男性会社員。科学的根拠を重視する中道の立場で、ワクチン接種済。義務化には慎重。

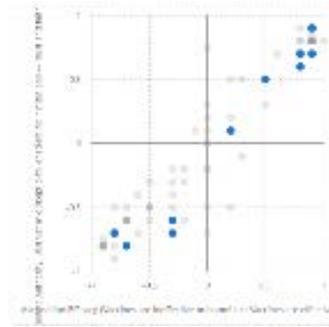
- **P2**: 米国南部在住の 35 歳女性フリーランスデザイナー。個人の自由を重視する立場で、未接種・政府メッセージに懐疑的。

さらにそれぞれのペルソナに、次の 2 条件を設定した。

- **Control 条件**: 特別な指示を与えない
- **Nudge 条件**: 初回の実験前に「マップ全体を広く見るように」と促す

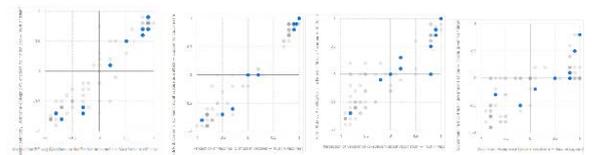
これらの 2 ペルソナ×2 条件=計 4 パターンについて、2025 年 9 月 25 日から 10 月 1 日まで 7 日間連続で毎日 1 回、その時点での X の検索結果を渡した。

### 4.2 実験結果

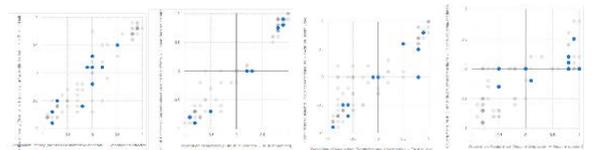


選択理由: 本ペルソナはワクチン有効性にやや肯定的、科学的権威に中立的で、既存の信念を確かめやすい傾向。選んだ投稿はこれら価値観とおおむね一致し、信念を強化する内容に加えて、接種体験などの適度な新規性も含む。

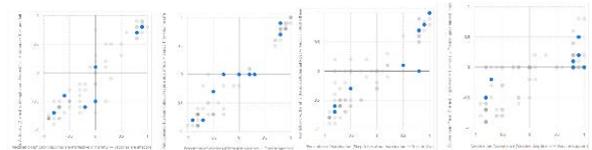
図 2. ペルソナによる投稿選択結果の表示例 (選択された投稿が青色)



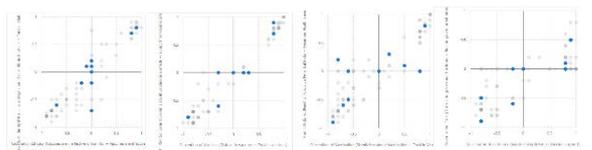
(a) P1-Control



(b) P1-Nudge



(c) P2-Control



(d) P2-Nudge

図 3. ペルソナによるマップ上の選択投稿変化 (7 日間のうちの最初 2 日と最後 2 日)

図2に、ペルソナの投稿選択結果の表示例を示す。システムは投稿選択の理由も表示する。図3に、実験期間中の選択記事の位置変化を示す（7日のうちの最初の2日と最後の2日）。

この変化を評価する指標として次の2つを考える。

・**カバー率 (Coverage)**：選択された投稿がマップ全4象限にどの程度分布しているか。最低で1象限(25%)、最高で4象限(100%)となる。カバー率が高ければさまざまな象限の投稿(さまざまな視点や、肯定否定など異なる意見を持つと思われる投稿)を見ていることを示す。

・**中庸率 (Balance)**：選択された投稿がマップ中心から半径0.6以内にある比率を示す。中庸率が高ければ、極端な立場の投稿ばかりではなく中庸に近い投稿を多く見ていることを示す。

図4に、7日間のカバー率および中庸率の推移を示す。各日の数値は離散的なので、回帰直線を示している。濃い黒の線がP1(中道的ペルソナ)でグレーの線がP2(自由重視ペルソナ)を示し、実線が「Control条件(何も指示なし)」で点線が「Nudge条件(「マップ全体を広く見るように」という指示あり)」である。この結果から、いくつかの特徴が見える。

#### (1) ペルソナ差

P1(中道的ペルソナ、濃い黒)では、カバー率・中庸率ともに、日数経過につれ不変か高めに移行する傾向が見られた。つまり、繰り返し利用する中で徐々にバランスの取れた投稿選択へと収束していく。

P2(自由重視ペルソナ、グレー)では、実験開始初期は比較的多様で中庸な投稿も含んでいたが、時間の経過とともにカバー率・中庸率が下がり、選択が偏る傾向が明確に見られた。これは、初期の探索的行動が時間とともに失われ、特定の傾向の投稿を見るようになることを示す。

#### (2) Control と Nudge の差

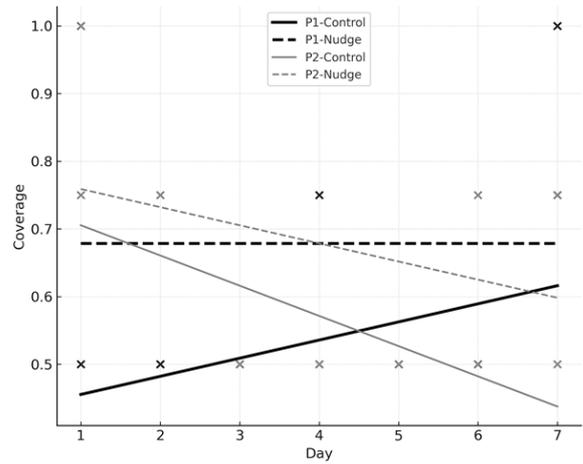
Nudge条件(指示あり、点線)では、Control条件に比べて全体的にカバー率・中庸率がやや高く、特にP2においてその効果が顕著に見られた。つまり、「マップ全体を広く見るように」という単純な介入でも、ある程度は多様な情報接触を促進することが確認できる。

一方で、その効果は強い持続性を持つわけではなく、初期には大きな差が見られるが、数日経つと縮小する傾向も確認された。

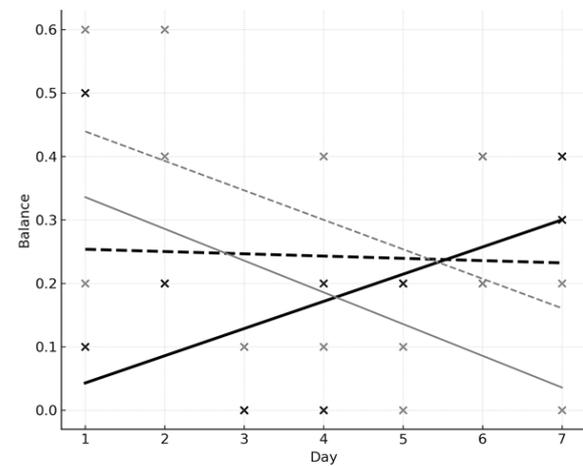
## 5. 考察

### 5.1 閲覧傾向の変化について

本研究ではカバー率、中庸率という2つの指標で確認した。いずれも、ここでは比較的高い方が好ましいという想定で設定している。つまり「なるべくさまざまな視点の投稿を読む方が望ましい」「極端な見解のものばかりでなく、なるべく中庸な投稿を読む方が望ましい」という想定であ



(a) カバー率



(b) 中庸率

図4. カバー率と中庸率の変化(7日間)

る。ただし、極端な投稿であっても一つの視点の両極にあるものを読むことにより中立的な立場が保たれることもあるし、常にさまざまな立場の投稿を読まなくても特定期間は特定の傾向の投稿を読んでも十分というような場合もあり、両方の比率が高い方が良いと一概に言うことはできないだろう。

また、ここではワクチン接種という特定の用語に対する一週間の変化を調べたものだが、さらに政治的、技術的、エンタテインメントなどさまざまな分野の用語について、長期に渡る変化を調べることを計画している。

### 5.2 LLM設定のペルソナを被験者にすることについて

今回、LLM設定のペルソナを被験者とする実験は、実際の人間による評価の前段階と位置づけたが、この評価が人間を対象とするものと類似、あるいは強い関係性があることがわかった場合、事実上、本格的な被験者評価の代替とすることも考えられる

LLMであれば、ペルソナ設定を行うにあたり、たとえば国籍、性別、年齢などだけではなく職業、趣味、嗜好、考え方の傾向（頑固あるいは柔軟など）も設定可能で、そのようなペルソナを任意の人数設定することもできる。これは実際の人間を被験者とする際には事実上不可能である。人間による評価を行った場合でも、よほど大規模かつ慎重な調査でない限り、被験者選択の過程などで特定の傾向に偏ることは十分あり得るので、そのような被験者とLLM設定ペルソナによる評価が同等の有効性を持つことは十分あり得るだろう。

ただし、LLM設定のペルソナに発生する技術的な限界については考慮しておく必要がある[16, 17]。

まず、LLMには記憶保持の構造的制約がある。現状のLLMは与えられたコンテキスト内でのみ推論する仕組みであり、長期的な利用経験を内的な状態として蓄積することができない。UIを繰り返し使用することでユーザが徐々に慣れたり、誤操作が減ったりするような「長期的行動変化」は、十分には再現されないと考えられる。

第二に、LLMは人間と異なり、視認性の悪さや物理的操作の煩雑さといった感覚的負荷を直接経験しない。UIの細部が引き起こす微妙なストレスや、非合理的な操作・思考の揺らぎを自然に再現することは難しいであろう。

さらに、LLMが生成する判断や行動は、学習データに起因する統計的・文化的バイアスがある可能性も知られており、設定したペルソナ属性が適切に反映されていない可能性もある。

これらを考慮すると、LLMによる評価と人間被験者による評価を組み合わせるハイブリッドな方法が、現段階では最も現実的と考えられる。

## 6. おわりに

フィルターバブルの緩和を目的とするインタラクティブマップ型SNS検索システムは、LLM設定のペルソナによる被験者評価としては、「マップ全体を広く見るように」という補足的な指示を行えば多様な投稿の閲覧に役立つ可能性があるという結果となった。今後、さらにペルソナ設定や検索ワードの幅、実験期間の拡大などを行うとともに、実際の人間被験者を用いた検証との比較などを進める予定である。本研究はLLMを用いたペルソナ評価手法が、インタラクティブ検索システムの初期的な有効性を評価するための有望なアプローチとなり得ることも示したと考えている。

## 謝辞

本研究は、JSPS 科研費 23K11205 の助成を受けたものです。同助成に感謝いたします。

## 参考文献

- [1] E. Pariser (2011): The Filter Bubble: What the Internet is hiding from you. Penguin UK.
- [2] C. D. Ruiz, T. Nilsson (2023): Disinformation and Echo Chambers: How Disinformation Circulates on Social Media Through Identity-Driven Controversies, *Journal of Public Policy & Marketing*, 42(1), 18-35, <https://doi.org/10.1177/07439156221103852>
- [3] N. Thurman (2018): Personalization of News. the *International Encyclopedia of Journalism Studies*. Massachusetts, USA: Wiley-Blackwell.
- [4] E. Bozdog (2013): Bias in Algorithmic Filtering and Personalization. *Ethics Inf Technol* 15, pp. 209-227 (2013). Available from: <https://doi.org/10.1007/s10676-013-9321-6>
- [5] M. D. Conover, J. Ratkiewicz, M. Francisco, B. Goncalves, A. Flammini, F. Menczer (2011): Political Polarization on Twitter, *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media* 5(1), 89-96
- [6] Tomonari Kamba (2025): SNS Search System that Uses a Large Language Model to Display Posts on an Interactive Map from Various Perspectives, *HCI International 2025*, S303 (to be printed)
- [7] S. Faridani, E. Bitton, K. Ryokai, K. Goldberg (2010): Opinion Space: A Scalable Tool for Browsing Online Comments, *Proceedings of CHI 2010*, 1175-1184, 2010, <https://dl.acm.org/doi/10.1145/1753326.1753502>
- [8] M. Palmonari, G. Ubaldi, M. Cremaschi, D. Ciminieri, F. Bianchi (2015). DaCENA: Serendipitous News Reading with Data Contexts, *ESWC (Satellite Events) 2015*, pp. 133-137. [https://doi.org/10.1007/978-3-319-25639-9\\_26](https://doi.org/10.1007/978-3-319-25639-9_26)
- [9] A. Thudt, U. Hinrichs, and S. Carpendale (2012): The bohemian bookshelf: supporting serendipitous book discoveries through information visualization. *CHI '12 Proceedings*, pp. 1461-1470. <https://doi.org/10.1145/2207676.2208607>
- [10] L. Björneborn (2008): Serendipity dimensions and users' information behavior in the physical library interface, *Information Research*, 13(1), paper 370.
- [11] L. McCay-Peet & E. Toms (2011): Measuring the dimensions of serendipity in digital environments, *Information Research*, 16(3) paper 483. <https://informationr.net/ir/16-3/paper483.html>
- [12] M. Jasim, C. Collins, A. Sarvghad, N. Mahyar (2022): Supporting Serendipitous Discovery and Balanced Analysis of Online Product Reviews with Interaction-Driven Metrics and Bias-Mitigating Suggestions, *CHI'22 Proceedings*, Article No. 9, pp. 1-24, <https://doi.org/10.1145/3491102.3517649>
- [13] Krisztian Balog, ChengXiang Zhai (2025): User Simulation in the Era of Generative AI: User Modeling, Synthetic Data Generation, and System Evaluation, *arXiv:2501.04410*, <https://doi.org/10.48550/arXiv.2501.04410>
- [14] Krisztian Balog, ChengXiang Zhai (2023): User Simulation for Evaluating Information Access Systems, *SIGIR-AP '23*, pp. 302-305, <https://doi.org/10.1145/3624918.3629549>
- [15] Hang Jiang, Xiajie Zhang, Xubo Cao, Cynthia Breazeal, Deb Roy, and Jad Kabbara. 2024. PersonaLLM: Investigating the Ability of Large Language Models to Express Personality Traits. In *Findings of the Association for Computational Linguistics: NAACL 2024*, pages 3605–3627, Mexico City, Mexico. Association for Computational Linguistics.
- [16] Mirjana Prpa, Giovanni Maria Troiano, Matthew Wood and Coady, Yvonne (2024): Challenges and Opportunities of LLM-Based Synthetic Personae and Data in HCI, *CHI EA'24*, No. 461, pp.1-5, <https://doi.org/10.1145/3613905.3636293>
- [17] C. K. Lazik et al. (2025): The Impostor is Among Us: Can Large Language Models Capture the Complexity of Human Personas?, *MuC '25*, pp.434-451, <https://doi.org/10.1145/3743049.3743057>