

AI 検索サービスにおける企業優遇バイアス：

市場競争への潜在的リスク

学籍番号：2301330039

氏名：安田 直也

1. 研究テーマ

AI 検索サービスにおける企業優遇バイアス：市場競争への潜在的リスク

2. 概要

本研究では、AI 検索サービスにおける企業優遇バイアスの有無とその影響を明らかにすることを目的とする。従来の検索エンジンと異なり、AI を用いた検索サービスは情報の生成・推薦を行うため、検索結果が特定の企業に偏るリスクが存在する。研究では、感情評価によるバイアス検出、オススメ順ソート、Google 検索との比較といった手法を用い、一部システム化することでバイアスの定量化を試みる。さらに、市場集中度指数（HHI）を導入し、検索バイアスが競争市場に与える影響を評価する。

本研究では、以下の AI 検索サービスを調査対象とする。

- ChatGPT
- Perplexity
- Genspark
- Felo
- Google AI Overviews
- Microsoft Copilot

3. 研究の背景

近年、ChatGPT や Perplexity などの AI 検索サービスが急速に普及している。しかし、それらが生成する検索結果の透明性や公平性は十分に検証されておらず、規制や倫理的ガイドラインがその進化のスピードに追いついていないのが現状である。

AI 検索サービスは利便性を提供する一方で、検索結果の決定プロセスが不透明な「ブラックボックス化」が進んでいる。ニューラルネットワークの重みを分析しても仕組みを完全に理解することは難しく、さらに学習データは通常開示されないため、その影響を直接検証することが困難である。

この結果、特定の企業が優遇されていても利用者が気づきにくく、公正な市場競争が損なわれるリスクがある。また、検索結果の偏りは消費者の意思決定に影響を与え、多様な選択

肢を狭める要因となるだけでなく、新規参入企業や中小企業が市場に参入しづらくなり、技術革新や多様なサービスの提供が阻害される可能性もある。

こうした AI のバイアス問題への対応として、各国では対策が進みつつある。例えば、シンガポールは「A. I. Verify」という AI ガバナンスフレームワークを導入し、AI システムのリスク評価を推進している。また、米国政府は AI のリスク管理に関する政策を進めており、レッドチーミングを通じて AI システムの脆弱性を評価することが推奨されている。日本においても、経済産業省が「AI 事業者ガイドライン」を策定し、AI の公平性や透明性を確保するための指針を示している。こうした取り組みは、AI のバイアスを是正するための重要な一歩ではあるが、企業優遇バイアスに関する具体的な対応策は依然として不足しているのが現状である。

こうした問題に対応するためには、AI 検索の透明性向上や公平性の確保に向けた規制の整備、技術的な対策の強化が求められる。公正な市場競争を維持し、技術革新を促進するとともに、多様な選択肢を確保するためにも、適切なルール作りと監視体制の確立が急務である。そのためにはまず、AI 検索サービスにおける企業優遇バイアスの実態を定量的に評価し、その影響を明らかにすることが不可欠である。

4. 先行研究

先行研究として、以下の知見が得られている。

- LLM (大規模言語モデル) は、ポジティブ・ネガティブといった感情評価を持つこと。
- LLM が特定のブランドや経済圏に対して偏った情報を提示すること。
- AI 推薦システムのバイアスが音楽、映画、書籍の推薦にも影響を与えていること。

また、最近では、Google AI Overviews が白人の遺伝的優位性を示唆するような IQ に関する偏った回答を生成したとの報告があり、Copilot AI や Perplexity も同様に、偏った統計情報を参照しているとの指摘がある。すでに AI 検索のバイアス問題は表面化しつつあり、企業優遇バイアスについても今後同様に問題視される可能性がある。

人種、地域、身体的特徴、年齢、性別などのバイアスについては多くの研究が行われているが、企業優遇バイアスに関する研究はほとんど存在しない。さらに、これまでの研究では AI 検索サービスと企業優遇バイアスを直接結びつけたものは少なく、日本語環境での分析も不足している。本研究は、企業優遇バイアスの定量化と日本語環境での分析を補うものとなる。

5. 研究の目的

本研究では、AI 検索サービスが特定の企業や商品を優遇するバイアスを持つかどうかを明確にし、その影響を定量的に評価する。バイアスの有無を検証するための簡易システムを開発し、GitHub で公開することで、後続研究者による検証を可能にする。さらに、市場集中度 (HHI) を指標として導入し、AI 検索サービスが市場競争に与える影響を評価する。また、

AI 検索サービスが与える競争市場への潜在的リスクについて考察し、政策提言や倫理的ガイドラインの形成に寄与することを目的とする。

6. 研究の方法

本研究では、以下の3つの手法を用いる。

1. 感情評価によるバイアス検出

- 商品・サービスに関する例文を作成し、企業名のマスク有無で感情評価を比較。
- 評価の変動を企業優遇バイアスの指標とする。

2. オススメ順ソートによるバイアス検出

- あるサービスカテゴリ（例えばクラウドサービスやストリーミングサービス）に関して、オススメ順でサービス名を列挙してもらい（AWS、Azure、Google Cloud…）プロンプトを実行する。
- 第二案として AI 検索サービスに「オススメのクラウドサービスは？」などのサービス名は提示せずに列挙してもらい、結果を比較。
- オススメ順ソート結果の表示前後でのアンケートを実施し、五段階評価の変化を分析することでバイアスの影響を評価する。
- 他の AI 検索サービスや市場シェアと比較し、偏りを評価。

3. Google 検索との比較によるバイアス検出

- AI 検索サービスと Google 検索の結果を比較。
- 公式情報の割合やネガティブ情報の比率の変動を分析。
- Google の検索品質評価ガイドラインでは、権威性（Authority）、専門性（Expertise）、信頼性（Trustworthiness）（E-E-A-T）を重視しているとされています。公式情報の評価は高く、その割合の変動がバイアスと捉えることができると考えます。

これらの手法を組み合わせ、バイアスの指標を定量化し、HHI を用いて市場競争への影響を評価する。

7. 現在までの進行状況

現在までに以下の成果を得ている。

● データ収集

- 感情評価用のプロンプトテンプレートを作成。（完了）
- AI 検索サービスのオススメ順位データを収集。（10%）
- 感情評価に基づく企業バイアスデータセットを作成。（20%）

● システム試作

- Perplexity の API を活用し、バイアス検出システムを Google Colab 上で試

作。

- 感情評価用のプロンプトテンプレートを適用し、カテゴリごとにサービス名をマスクした場合とそうでない場合の感情評価値を取得。
- 収集したデータを辞書型変数に格納し、JSON 形式で S3 バケットに保存。データは時系列で管理し、変化を追跡可能なフォーマットとする。これはアルゴリズム変更やパーソナライズの影響があった場合に検知するためである。
- 簡易システムによる検証では、Perplexity において楽天市場の評価のみ「2」と著しく低く表示していることが確認されており、特定の企業に対するバイアスの可能性が示唆されている。
- 今後、処理の自動化や他の AI 検索サービスとの比較機能を追加予定。

● 評価方法の設計

- **感情評価の比較**: 特定企業の感情評価スコアのマスクあり・なしの差を、全体の感情評価スコアの平均で正規化。
- **オススメ順ソート**: 検索結果の提示前後で消費者の印象を 5 段階で評価するアンケートを実施しその変化を評価。
- **検索結果の比較**: AI 検索サービスと Google 検索での公式情報リンクの表示割合の差。
- 各 AI 検索サービスの上記のバイアス傾向を直感的に把握しやすくするために、バイアス指標をヒートマップで可視化する。
- 市場集中度指数 (HHI) を用いた市場競争への影響分析の試行。

● 次のステップ

- システムによる残りのデータ収集と機能追加、評価指標の妥当性を確認していく。
- 言語処理学会などの研究会への参加・発表も視野に入れる。

※市場集中度指数 (HHI) について

市場競争への影響を分析するために、市場集中度指数 (HHI) を試験的に適用する予定である。HHI は金融市場の研究などで用いられる指標であり、仮想的な市場シェアの偏りを定量的に評価することが可能である。特に、検索結果の上位に表示される企業の分布を「仮想的な市場シェア」として扱うことで、AI 検索サービスが特定の企業を優遇しているかどうかを定量的に評価できる可能性がある。

また、米国司法省の基準では、HHI が 1,500 未満を競争的市場、1,500~2,500 を適度に集中した市場、2,500 以上を高集中市場と定義されており、この基準を AI 検索サービスの検索結果に適用することで、企業優遇の度合いを客観的に測定できると考えられる。さらに、Google の検索市場の独占度を分析する際にも HHI が用いられた事例があり、AI 検索サービスの市場支配力を測る指標としての妥当性が示唆されている。

一方で、検索結果の市場シェアをどのように定義するか、また、アルゴリズムの変更やパーソナライズによる影響をどのように考慮するかといった課題が残る。検索クエリの選定や結果の変動幅を踏まえた補正が必要となる可能性があり、AI 検索サービスにおける企業優遇バイアスの評価に適用する妥当性については、追加調査が必要である。

8. 参考文献

- 中川 慧、平野 正徳、藤本 悠吾：「大規模言語モデルを活用した金融センチメント分析における企業固有バイアスの評価」、第 21 回テキストアナリティクス・シンポジウム、2024。
- Kamruzzaman, M., Nguyen, H. M., Kim, G. L.: *"Global is Good, Local is Bad?" : Understanding Brand Bias in LLMs*, arXiv preprint, 2024
- Sakib, S. K., Das, A. B.: *"Challenging Fairness: A Comprehensive Exploration of Bias in LLM-Based Recommendations"* arXiv preprint, 2024
- Gilbert, D. "Google, Microsoft, and Perplexity Are Promoting Scientific Racism in Search Results." WIRED, 2024-10-28. アクセス日 : 2025-02-06. <https://www.wired.com/story/google-microsoft-perplexity-scientific-racism-search-results-ai/>
- Fujitsu Research. LLM Bias Diagnosis: 大規模言語モデルのバイアス診断と緩和に関する最新動向. FLTech Blog; 2024-03-22. アクセス日 : 2025-02-06. <https://blog.fltech.dev/entry/2024/03/22/LLM-Bias-Diagnosis-ja>.
- AI Verify Foundation. What is AI Verify. 公開日不明. アクセス日 : 2025-02-06 <https://aiverifyfoundation.sg/what-is-ai-verify/>.
- JETRO/IPA New York. ニューヨークだより 2024 年 5 月号 —NIST の取組を中心に—。独立行政法人情報処理推進機構; 2024. アクセス日 : 2025-02-06. <https://www.ipa.go.jp/digital/chousa/trend/m42obm000000ic9a-att/ny-dayori202405.pdf>
- 経済産業省, 総務省. 「AI 事業者ガイドライン (第 1.0 版)」を取りまとめました. 経済産業省 2024. アクセス日 : 2025-02-06. <https://www.meti.go.jp/press/2024/04/20240419004/20240419004.html>
- Zivanta Analytics. Herfindahl-Hirschman Index: Using the Regulator's Tool to Marketer's Advantage in Title Insurance Industry. Zivanta Analytics; アクセス日 : 2025-02-06. <https://zivanta-analytics.com/herfindahl-hirschman-index-using-the-regulators-tool-to-marketers-advantage-in-title-insurance-industry/>
- 野木村忠邦. 米国の 2010 年司法省・連邦取引委員会「水平的合併ガイドライン」の概要. 政経研究. 2013 アクセス日 : 2025-02-06; 49(3):1044-1040.

https://www.publication.law.nihon-u.ac.jp/pdf/political/political_49_3/each/33.pdf