

【田中頼人特別研究】

第8回レポート

2301330039：安田直也



研究テーマのタイトル：

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

先行研究との違い：

- ①AI検索サービスに焦点を当てていること
- ②企業優遇バイアスに焦点を当て競争市場への影響を調査すること

AI検索サービスの対象は画像からGeminiを除いたサービス + Google AI Overview + Microsoft Copilot
検索代わりに使うユースケースを想定し、無料プランのデフォルトモデルで検証を予定

	Gemin	ChatGPT (SearchGPT)	Perplexity	Felo	Genspark
開発元	Google	Open AI	Perplexity AI	Sparticle	Genspark AI
回答内での 参照元の表示	なし	あり	あり	あり	あり
一覧形式での 参照元の表示	なし	あり	あり	あり	なし
ファクトチェック	あり (Googlで再確認)	なし	なし	なし	あり
出力結果を まとめる	なし	なし	ページ (テキスト+画像)	スライド マインドマップ	Sparkpage (テキスト+画像)
有料プラン	2900 円 / 月	20ドル / 月	20ドル / 月	2099 円 / 月	なし

AI検索サービスについての追加調査TODO：

- ①各AI検索サービスがそれぞれどのような情報を使って検索しているか
特にGoogleやMicrosoftは個人情報を使ってパーソナライズしているかもしれません
- ②検索ロジックは何を使っているのか
Google検索を基本にしているのか、独自ロジックなのか
- ③各サービスの検索範囲
- ④検索結果がどのように順位付けされているか
- ⑤デフォルトLLMモデルとその学習データ

前提：企業優遇が一般ユーザーに与える影響度を研究したいため
なるべくロール情報やコンテキストを与えず、端的なやりとりだけでバイアスが確認できる方法が望ましい

- **深津式プロンプト**
 - 背景・コンテキスト
 - **ロールプレイング**
 - 多段階プロンプト
- **Zero-Shotプロンプティング**：通常の記事ではなく、「命令」と「入力・出力」をラベリングして分けたものをLLMに投げかけるだけで解答の精度が向上
- **Few-Shotプロンプティング**：いくつかの例を提示してからタスクを指示する手法
- **Chain-of-Thought (CoT) プロンプティング**：「まず、問題を理解し、次に解決策を考え、最後に結論を出す。」といった形で、AIに思考のステップを明示
- **Zero-Shot CoTプロンプティング**：「ステップバイステップで」と追記するだけで、CoTを再現する手法
- **Tree of Thoughts (ToT) プロンプティング**：各ラウンドで一番正解に近いアイデアから、次のラウンドのアイデアが「樹状」に派生 方向性刺激プロンプティング
- **自動プロンプトエンジニアリング (APE)**：生成AIに対して効果的なプロンプトの作成をお願いする

以下のプロンプトリストを作成する必要があるそう。

- ①企業名をマスクして答えてもらうプロンプト
- ②特定サービスのランキングづけ、能力を比較するようなプロンプト
- ③単純にある領域での主なサービス名を列挙してもらうプロンプト※企業名を挙げる順番、出てくる回数を確認
- ④上記がうまく機能しない場合、ロールやコンテキスト、例を記載して回答を引き出すようなプロンプト

データはLLMの学習に使用するのではなく、企業優遇バイアスを検証する裏付けとして活用します。

具体的には、企業の財務状況や比較対象のサービスシェアなど、実際のビジネスデータが必要です。
これらのデータは、LLM専用のデータセットではなく、企業のコーポレートサイトや各種調査結果から取得します。

政府機関や公的データベース

- 総務省統計局 (<https://www.stat.go.jp/>)
- 経済産業省の調査レポートや白書
- 世界銀行 (World Bank) やOECDのデータポータル

市場調査会社のレポート

- Statista (<https://www.statista.com/>) GartnerやIDC (技術系の市場調査で有名)

Relative Exposure Index（相対的露出指数）

定義：各企業の露出頻度を対象企業の平均と比較。

Market Consistency Ratio（市場一貫性比率）

定義：検索結果の表示順位が市場シェアと一致するかを測定

Fair Share Index（公平シェア指数）

定義：検索結果がすべての企業に均等に分配されているかを測定。

Individual Fairness（個別公平性）

定義：似たような意味を持つ異なるクエリ（例: 「最高のスマホ」と「一番人気のスマホ」）に対して、検索結果や推薦される企業が著しく異ならないことを測定

Weights & Biases Japan：日本語LLMのバイアス評価に関するリーダーボード「Nejumi」を開発し、バイアスの定量的評価を行っている。今の所バイアスを専門で扱っている組織の中で一番参考になりそう。

https://note.com/wandb_jp

東京カレッジ：AI安全性レッドチーミング活動

https://www.tc.u-tokyo.ac.jp/ai1ec_event/13327/

slackコミュニティ

→JDLA、松尾研LLMコミュニティに参加しているがバイアスに関する投稿は少ない