

新聞記事中に意図的に付与された 認知バイアスの抽出方法の提案と評価

日大生産工(院) ○清家 光生
日大生産工 古市 昌一

1. 研究背景

近年、インターネットやSNSの発展により、個人が日常的に多様なニュース記事に接する機会が増えている。一方で、記事内容には発信者の意図や立場が反映された表現が含まれる場合があり、読者が無意識のうちに特定の見方へ誘導される危険性がある。これらは「認知バイアス」と呼ばれる心理的傾向と深く関係しており、事実と意見、報道と印象の境界を曖昧にする要因となっている。近年我が国でも活動されるようになったファクトチェック機関は誤情報や虚偽情報の訂正に有効である一方で、記事中の言語表現に潜む意図的な偏りの検出は不十分である。

1.1 同一事象における表現差の例

たとえば、現立憲民主党の野田佳彦代表と現石破茂首相のやり取りを報じた各紙の記事の引用を表1に示す。

同一の事実関係を扱いながらも、産経新聞の記事中下線部では、筆者の推測が断定的に提示されている。この記述は「フレーミング効果」に該当する可能性があり、他の4紙に比べて明確な意図的バイアスを含む構成となっている。一方、他社の記事は各社記事中の下線部のように事実の経過説明にとどまっている。このように、同一の事象を報じていても各社の焦点や記述の方向性には違いが見られ、具体的な語彙選択や文構造の違いが、読者に与える印象の差異につながっている可能性がある。

1.2 本研究の目的

本研究では、このような意図的に付与された認知バイアスを自動的に抽出するための分析手法を構築し、その精度を検証を目的とした。具体的には、生成AIとしてChatGPT(以降AIと表記する)を用い、AIに認知バイアスの分類辞書を提示し、新聞記事本文から該当箇所・バイアス名・根拠を出力させるシステムを設計し、人手による基準データとの比較によりその実現可能性を評価した。

表1 同一内容に対する主要5紙の表現

毎日新聞 (2025-08-07) 給付付き税額控除 自民と立憲が協議 立憲の野田佳彦代表から『与党として真剣に検討する可能性はあるか』と問われた石破茂首相(自民総裁)が『一つの解だが、問題点について議論をさらに深めすることが重要だ』と 協議に応じる姿勢を示し 、この日の会談となった。
日本経済新聞(2025-08-06) 立民が給付・税制で自民と接近 石破首相続投に助け舟、維国と距離 野田氏は政治改革でも前進のために譲歩する考えを示した。
読売新聞(2025-08-06) 自民と立民、「給付付き税額控除」週内にも協議開始へ…医療・年金「負担のあり方」も議題に 石破茂首相(自民総裁)は4日の衆院予算委員会で立民の野田代表から導入に向けた協議を呼びかけられ、『その通りにしたい』と応じた。『社会保障も合わせて議論する機会をみいだしたい』とも語っていた。
朝日新聞(2025-08-06) 自民、「給付付き税額控除」めぐり立憲と意見交換 政調会長が会談 物価高対策をめぐっては、4日の衆院予算委員会で立憲の野田佳彦代表が給付付き税額控除にまつわる協議体の設置を要求し、石破茂首相も応じる姿勢を示していた。
産経新聞(2025-08-06) 自民と立民、物価高対策の協議に着手 首相は連携模索も「政権継続か先行き不透明」と懐疑 自民内で退陣論が強まる首相としては、参院選で最大の争点となった物価高対策で歩み寄る姿勢をみせることで立民を取り込み、政権の延命を図りたいと考えが透ける。

2. 提案手法

2.1 認知バイアスの定義と分類

認知バイアスとは、人間の判断や意思決定が無意識のうちに特定の方向へ偏る心理的傾向のことである。代表的なものとして、自己の信念に沿った情報のみを重視する「確証バイアス」¹⁾や、情報の提示の仕方によって印象が変化する「フレーミング効果」²⁾などが挙げられる。

本研究では、心理学および認知科学の知見に基づき、115種類の認知バイアスを体系的に整理し、それぞれに定義を付与した分類辞書を作成した。これにより、AIが文章を分析する際に、特定の語句や文構造をどのバイアスに該当するものとして判断するかを一貫して参照できるようにした。

なお、本研究で用いるAIには、出力の多様性を制御するパラメータとして温度が設定され

Proposal and evaluation of a method for extracting intentionally added cognitive biases in newspaper articles

Koki SEIKE

ている。温度は生成の確率分布を調整するもので、値が高いほど出力のランダム性が高く、低いほど確定的な応答となる。

2.2 AIを用いた抽出方法

本研究では、新聞記事に含まれる認知バイアスをAIがどの程度正確に抽出できるかを検証した。分析対象は、日本の主要五紙（朝日新聞、読売新聞、毎日新聞、日本経済新聞、産経新聞）から、政治・経済・社会・国際・文化などの分野を抽出した100記事である。

記事本文をAIにつづつ入力し、AIに対して「該当箇所」「バイアス名」「理由」の三要素を出力させた。出力のランダム性をなくすため、温度設定は0.0とした。これは一般利用者がAIを用いる際のデフォルト設定に等しく、その条件でどの程度正確に認知バイアスを抽出できるかを検証することを目的とした。また、人手による基準データ（正解データ）は筆者自身が作成し、温度0.8で得られた幅広い出力を参考にしながら、内容を吟味して最終的に確定した。

本提案手法の特徴は、AIモデルを再学習させるのではなく、あらかじめ構築した心理学的分類辞書をそのまま入力情報として利用し、AIがそれに基づいてどの程度一貫した判断を下せるかを評価できる点にある。これにより、AIが記事内容をどの程度意味的に理解し、心理的・論理的な偏りを識別できるかを定量的に把握できる。すなわち、本研究は「AIが人間と同様の基準でバイアスを検出できるか」という観点から、その分析能力の実用的限界と精度を明らかにすることを目的としている。

3. 実験方法

提案手法の有効性を検証するため、本研究ではAIによる認知バイアス抽出結果と、人手による基準データとの一致度を比較し、AIの抽出精度を定量的に評価した。分析対象は、日本の主要五紙（朝日新聞、読売新聞、毎日新聞、日本経済新聞、産経新聞）から選定した計100記事であり、政治・経済・社会・国際・文化など複数の分野を均等に含めた。記事は特定の論調や政治的立場に偏らないよう配慮し、内容の多様性を確保した。

3.1 実験1：模範解答作成のための比較

基準データを人手で作成する際に、AIの出力を補助的に活用しても問題がないかを確認するために実験1を実施した。本実験の目的は、①AIの出力が人間の判断に近い内容を含み、手動データ作成時の参考資料として利用できることを確認すること、②AI出力を参照するこ

とで筆者自身の判断が偏ったり、精度が低下したりしないかを検証することである。

AIの温度設定は出力の多様性と実用的な再現性の両立を目的に0.8とし、3記事を対象に3回の繰り返し出力から抽出結果を集計した。温度0.0や0.2では出力が画一的で抽出範囲が狭く、逆に1.0では文脈に沿わない過剰な候補が増えたため、0.8が最も適切なバランスを示した。筆者が同一記事について完全手動で抽出作業を行い、両者のデータを比較して出力傾向や整合性を確認した。この結果をもとに、手動データ作成時の補助資料としてAI出力を利用する方針を定めた。

3.2 実験2：100記事を対象とした評価

実験2では、AIの抽出精度を本格的に評価することを目的とした。主要五紙から抽出した100記事を対象に温度設定は0として分析を行った。これは一般利用者がAIを利用する際のデフォルト条件を想定し、その標準的な設定下でどの程度の精度が得られるかを確認するためである。

AIの出力はすべて記録し、筆者が作成した手動の基準データと比較を行った。比較では、記事ごとに抽出されたバイアスの件数や種類の一一致・不一致を確認し、AIと人手の判断傾向の差を整理した。本実験では辞書内の115種類のみを対象とし、AIに独自の判断を行わせる自由回答形式は採用していない。

4. まとめ

本研究では、AIを用いて新聞記事中の認知バイアスを抽出する分析手法を提案した。主要5紙の記事を対象に分析を行い、AI出力と人手データの整合性を確認した結果、媒体や記事構成により抽出傾向が異なる可能性が示された。今後は、より広範なデータを用いて精度を検証し、AIによるバイアス分析の有効性を高めることが課題である。

参考文献

- 1) 情報文化研究所・高橋 昌一郎（監修），
情報を正しく選択するための認知バイアス事典，フォレスト出版，(2021) pp164-167
- 2) 情報文化研究所・高橋 昌一郎（監修），
情報を正しく選択するための認知バイアス事典 行動経済学・統計学・情報学 編，
フォレスト出版，(2023) pp54-57