

生成 AI を用いた広告クリエイティブにおける ジェンダー関連炎上の可能性検出の検討

A Study of Detecting the Potential for Gender-Related Flaming in Ad Creative Using Generative AI

平野徳麻¹ * 笹嶋宗彦¹
Tokuma Hirano¹ Munehiko Sasajima¹

¹ 兵庫県立大学
¹ University of Hyogo

Abstract: 近年、企業や自治体の広告が炎上し、その企業や自治体の評判が低下することが問題となっている。そこで本研究では、ジェンダーの観点に着目し、広告の種類をオンライン広告のバナー広告に限定した広告の炎上防止を目的とする。その目的の達成のためにジェンダー関連の広告炎上原因を構造化したオントロジーを構築し、生成 AI (ChatGPT-4o) を用いてバナー広告やオントロジーの解析を行い、最終的にバナー広告の炎上可能性の検出を行う。

1 はじめに

バナー広告とはインターネット上の広告枠に画像や動画などで表示される形式の広告で、私たちがウェブサイトなどを閲覧する際によく見ることができる。バナー広告は企業や自治体にとって重要なマーケティング手段となっており、松田ら (2007) は、バナー広告の単純接触効果が商品評価と購買意図に与える影響を実験的に検討し、バナー広告の反復呈示により好意度および購買意図が向上することを示した [1]。また松田ら (2008) は、バナー広告の反復呈示により商品記憶および好感度や再認の評価が長期的に保持されることを示した [2]。しかし、広告表現の内容が不適切であることが原因で炎上し、メディアに取り上げられることが増加している。広告では、電車内に貼られていた栄養ドリンクの広告で男女それぞれ 2 種類の広告がある中、女性側の広告にだけ「仕事、育児、家事。3 人自分が欲しくないですか？」とあり、その内容が女性への家事や育児の押し付けと見られ、時代錯誤で性差別だと SNS を中心に炎上した例がある [4]。企業や自治体の広告が炎上することにより、その企業や自治体の商品やサービス、または企業や自治体そのものの評判が低下することが問題であると考えた。

そこで本研究では、炎上の原因をジェンダーの観点に着目し、対象をバナー広告に限定した広告の炎上防止を目的とし、生成 AI を用いた炎上可能性の検出を行う。具体的には、オントロジーを用いて体系化した

ジェンダー関連での広告の炎上原因をマイ GPT に知識としてアップロードし、バナー広告の画像の炎上可能性をカスタマイズしたマイ GPT を用いて検出する。その結果、オントロジーで体系化した炎上原因に沿った、広告の炎上可能性の検出をすることができた。

2 関連研究

2.1 バナー広告の効果

前節で示した通り、バナー広告は企業や自治体にとって重要なマーケティング手段である。ユーザーが様々なウェブサイトを閲覧する際に、バナー広告を何回も反復して目にすることで単純接触効果により商品やサービスの好感度や認知度が上がり購買意欲に結びつく。さらに中岡ら (2008) は人がバナー広告にどのようなイメージを抱いているかを分析するために、SD 法による感性評価実験を行い、バナー広告を評価する基準として、「目にとまる」、「心をひかれる」の 2 つが重要視されることを示した。バナー広告の背景色や文字色、背景色の彩度、アニメーションやイラストが感性に影響を与え、そのデザインを工夫することで、ユーザーにクリックされやすくなる [3]。

2.2 炎上のメカニズム

ウェブで発生する炎上について平井 (2012) は、「炎上とは、ブログ、ミクシィ (mixi)、ツイッター (Twitter) などに投稿されたメッセージ内容、ならびに投稿者に

*連絡先：兵庫県立大学社会情報科学部社会情報科学科
〒 651-2197 兵庫県神戸市西区学園西町 8 丁目 2-1
E-mail:sasajima@sis.u-hyogo.ac.jp

対して批判や非難が巻き起こる現象」と定義し、炎上が起こる背景として電子掲示板の2ちゃんねるを中心に形成された文化と、若年層による携帯電話のやりとりを通じて形成された文化の2つの文化がウェブサービスの発展に伴い、双方が拡張し、接続してきたことであると考察している[5]。また、山口(2016)は、炎上を「ある人や企業の行為・発言・書き込みに対して、インターネット上で多数の批判や誹謗中傷が行われること」と定義した。さらに、炎上加担行動の動機について定量的な分析を行い、「頻発している炎上について、ごく一部の人が、正義感から加担しており、かつ、そのような人が複数の炎上事例に加担し、何度も書き込んでいる」ことを示した[6]。ソーシャルメディア上の個人攻撃を類型化して背後にある構造を理解することを目的とした瀬川ら(2021)の研究では、Sentence-BERTを利用して日本語大規模ツイートデータから網羅的に攻撃事例を抽出し、攻撃パターンが趣味趣向が近いユーザー間で起こる場合と、そうでないユーザー間で起こる場合の2つがあることを明らかにした。趣味趣向が近いユーザー間の攻撃は、一部のコミュニティで起こっている場合が多く、そうでないユーザー間の攻撃は、政治的な思想が強いユーザーや、普段の投稿内容がネガティブなユーザーによって起こされやすく、また、集団攻撃となっている傾向が強く、普段の会話においてインタラクションを行っていることが多いことを示した。[7]。

2.3 広告の炎上

広告の炎上について水村(2023)は、人種・性差別な表現で炎上した事例を基に、広告を見てSNS上で批判・攻撃モードを発動させる人の心の中はどうなっているのかについて様々な角度から検討した。また、不適切な広告表現については、「広告主・広告代理店・媒体社各社が予防的な措置を講じるべきである。」と述べた上で、「実際問題として、広告を見て何を連想するかは三人三様で、どのような表現が人の心の機微に触れるのかを正確に予測することは難しい。」と広告の炎上を防ぐことは困難であると言及している。[8]

2.4 ジェンダー関連の炎上

ジェンダー関連で炎上したCMについて瀬地山(2020)は、過去炎上したCMのパターンを分類し、4象限に図式化した。縦軸に「女性」「男性」といった「商品や描写内容の訴求対象」をとり、横軸には「性役割」「容姿・性的メッセージ」といった「炎上ポイント」をとった。以上の4象限の図式を図1に示す。

まず、右上の第1象限は、「女性」を応援する意図の広告であったはずが「性役割」の固定化・強化と受け取られ炎上したパターンである。次に、左上の第2象限は、「女性」を応援したつもりなのに「外見・容姿」の面で性差別と受け取られ炎上したパターンである。左下の第3象限は、一般的に受け取られるように作られたが、「男性」へ性的なメッセージが強く炎上したパターンである。最後に右下の第4象限は、「男性」への共感がかえって「性役割」の固定化・強化につながってしまい炎上したパターンである。[9]

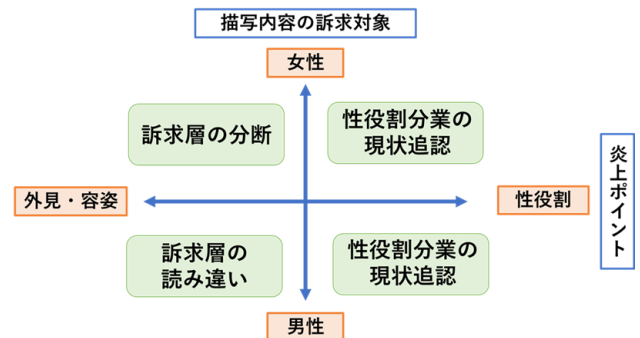


図 1: ジェンダー関連で炎上した CM の 4 象限

これまでに発表された研究では、炎上の背景や原因といったメカニズムに関する研究は行われているが、炎上を予見したり、防いだりするためのシステムの研究はされておらず、炎上を防ぐためには広告主や広告代理店、媒体社各社などの現場で対策しなくてはならない。

3 提案手法

本研究では、バナー広告がジェンダーの観点で炎上する可能性の有無を、ジェンダー関連での広告の炎上原因を体系化したオントロジーの知識を持たせた生成AIを用いて検出するシステムを提案する。バナー広告の制作者やインターネット広告の運用者が広告を作成して出稿する際に、事前に提案システムを利用することで、炎上の可能性の有無が分かるようになり、予見された炎上に備えることができるようになると考えている。本研究では生成AIとしてChatGPT-4oを用いる。2024年11月時点のモデルである。また、生成AIを用いた一連の処理についてはChatGPTに搭載されているGTPsという機能で、特定の処理に特化したチャットボットを作成することのできるマイGPTを作成した。マイGPT内に構築したオントロジーをXMLファイル形式でアップロードし、プロンプトもプロンプトの原則26ヶ条[10]を参考に事前に記述した。それにより、提案手法は作成したマイGPT上にてバナー広告

の画像を1枚ずつアップロードし入力され、検出の結果が出力される。

本章では、提案手法について、オントロジーの構築、バナー広告の解析、バナー広告の炎上可能性の検出、検出の結果の4節に分けて示す。提案手法で用いたプロンプトは付録に示す。

3.1 オントロジーの構築

ジェンダー関連での広告の炎上原因を体系化したものとして、ジェンダー広告炎上原因オントロジーを構築した。オントロジーとは、共通語彙(概念)を提供する体系化された辞書のようなもの、つまり、対象の知識の意味を共通する語彙で表し体系化したものである。オントロジーが知識のバックグラウンドにある暗黙的な情報を明示することで、知識が表している内容が明確になり、相互理解の助けになる[11]。ジェンダー広告炎上原因オントロジーは、瀬地山(2020)が図式化した4象限[9]を基に構築した。ジェンダー広告炎上原因オントロジーを構築することで、個人的な感想に偏らず、ジェンダーに着目した広告の炎上原因の知識を明確化できた。以上のジェンダー広告炎上原因オントロジーを図2に示す。

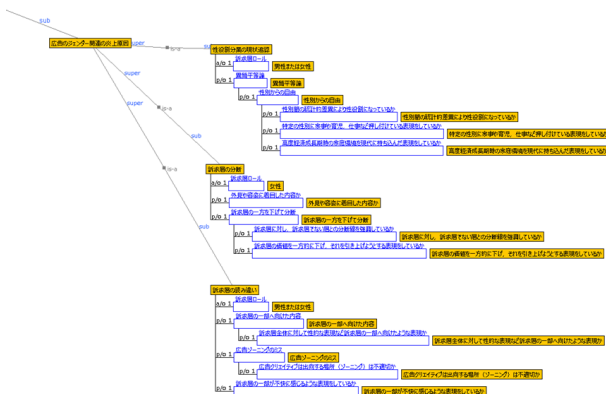


図 2: ジェンダー広告炎上原因オントロジー

3.2 バナー広告の解析

バナー広告の炎上可能性を検出するためには、バナー広告の内容を適切に解析し、テキスト情報や視覚的要素を抽出する必要がある。バナー広告の画像データを対象に解析を行い、炎上可能性を評価するための特徴量を抽出した。本節では、バナー広告解析の具体的な手法と手順について示す。

バナー広告の画像は主に、商品説明やコピーなどのテキストと商品自体や背景のイラストで構成されてい

る。そのテキストやイラストの情報を解析するために、付録??で示したプロンプトでは、まず、バナー広告内のテキストを抽出させ、文字集合ごとにリスト化させる。次にバナー広告のイラストについてレビューさせ、その詳細を形状、色、構図などに着目して説明させる。また、炎上可能性の検出には、それが誰に向けた広告であるかを表す、訴求対象の情報が必要なので、得られたテキストとイラストの情報を基にバナー広告の訴求対象を判断させた。

3.3 バナー広告の炎上可能性の検出

バナー広告に含まれる要素をもとに、炎上可能性の有無を判断する。本節では、炎上可能性の検出の手順とプロンプト設計について示す。

前節で示した、バナー広告のテキストとイラストの情報。そして、その2つの情報を基に取得した訴求対象の情報の、合わせて3つの情報と3.1で構築したジェンダー広告炎上原因オントロジーを用いてバナー広告のジェンダーの観点での炎上可能性の検出を行う。プロンプトでは、オントロジーの解析の際に構築したオントロジーの「広告のジェンダー関連の炎上原因」の概念の下位概念を用いることを指示した。また、プロンプトの原則26ヶ条[10]に従い、指示と情報の分割やFew-shotプロンプトの使用、「ステップバイステップで考えてください」プロンプトの使用といった工夫をプロンプトに反映させた。

3.4 検出の結果

提案手法を用いて、バナー広告の炎上可能性を検出した。本節では、検出の詳細と結果について述べ、手法の有効性および課題を考察する。

検出に用いたバナー広告の画像は

1. 炎上原因に該当する自作バナー広告(図3)
2. 炎上原因には該当しない自作バナー広告(図4)
3. クローリングして収集した炎上原因に該当する既存のバナー広告(図5)
4. クローリングして収集した炎上原因には該当しない既存のバナー広告(図6)

の4種類を用いた。炎上原因に該当する、または該当しないの判断は筆者の予測である。

検出の結果をバナー広告と共に図3から順に示す。図3, 図4, 図5の炎上可能性の検出に関しては、筆者の予測通りの結果になった。しかし、図6の炎上可能性の検出に関しては、筆者の推測とは違う結果になった。

図6の化粧品の広告は図1では「訴求層の分断」という炎上原因に該当することが多いが、図6のバナー広告は訴求層の一方を落とすような表現は見られず、炎上原因に該当しないと考えた。しかし、容姿に着目した内容かつ女性しか広告内に登場していないため、女性が美しくあるべきという性役割の固定・強化をしないと ChatGPT が判断したと考えられる。



図 3: 炎上原因に該当する自作バナー広告

- 炎上可能性: あり
- 炎上可能性の原因:
 - 「25歳すぎると顔にでる」という表現が「若さ」を過剰に強調し、年齢による美の価値観を押し付けていると捉えられる可能性があります。
 - 年齢を基準に美しさを評価することが、訴求層の一部に不快感を与える可能性があります。また、外見や容姿への過剰なフォーカスは、ジェンダー関連の炎上要因である「外見や容姿に着目した内容」に該当します。



図 4: 炎上原因には該当しない自作バナー広告

- 炎上可能性: なし、または著しく低い
- 炎上可能性の原因: このバナー広告にはジェンダー関連で炎上する原因となる要素は含まれていません。しかし、困惑した表情がターゲット層にとって誤解を招く可能性はあるため、文脈を十分に伝える補足テキストを加えることを検討すると良いでしょう。



図 5: クローリングして収集した炎上原因に該当する既存のバナー広告

- 炎上可能性: あり
- 炎上可能性の原因:
 - 「女性が育児・教育を行う」という構図が性別役割分業を暗に支持していると受け取られる可能性。
 - 男性やその他のジェンダーが全く描かれていない点で、ジェンダー的な偏りを感じる視聴者がいるかもしれない。



図 6: クローリングして収集した炎上原因には該当しない既存のバナー広告

- 炎上可能性: あり
- 炎上可能性の原因:
 - － バナー広告は「外見や容姿に着目した内容」となっており、ジェンダー広告炎上オントロジーの「性役割分業の現状追認」や「外見に依存した表現」の項目に該当するリスクがあります。
 - － 女性のみをターゲットとし、美しさや肌の透明感を前提とするメッセージが一部の層にステレオタイプ的であると認識される可能性があります。
 - － さらに、スキンケアが女性専用であるかのような暗黙のメッセージが含まれており、「性別による役割の固定化」と受け取られるリスクが存在します。

4 おわりに

本研究では、ジェンダー関連の広告の炎上原因をオントロジーを用いて体系化し、生成 AI によるバナー広告の炎上可能性の検出することを提案した。本研究は予測が困難と言われる炎上がある程度検出可能であることを示すことができた点で、広告の炎上防止に貢献していると言える。

しかし、検出の結果が筆者の予測と異なる場合もあるため、不適切な広告表現の検出は、確実であるとは言えない。また、炎上の原因をジェンダーの観点に、広告の種類をバナー広告に限定して考えたため、それ以外の原因で起きる炎上には対応出来ない。今後は、オントロジーやプロンプトを改善することで、人種差別など他の不適切な表現、映像広告といった広告形態など、幅広い範囲で広告の炎上を防げるようなシステム構築を目指す。

参考文献

- [1] 松田憲, 平岡斉士, 杉森絵里子, 楠見孝 (2007). バナー広告への単純接触が商品評価と購買意図に及ぼす効果. *Cognitive Studies*, 14(1), 133-154.
- [2] 松田憲, 小林剛史 (2008). バナー広告の反復呈示が短期・長期記憶と商品評価におよぼす効果. *日本心理学会第 72 回大会発表論文集*. 2AM097.
- [3] 中岡伊織, 氏平淳子, 串田淳一, 亀井且有 (2008). 「ついクリックしたくなる」バナー広告の感性評価およびそのモデル構築. 第 24 回ファジィシンポジウム. TD2-4.
- [4] やまもとゆか. 「鬼すぎない？」 大正製薬の広告が“性差別”と物議…… 男女の“非対称性”に「昭和かな?」「時代にあってない」. ねとらぼ. 2024-07-04. <https://nlab.itmedia.co.jp/nl/articles/2407/03/news185.html>. (参照 2024-11-24).
- [5] 平井智尚 (2012). なぜウェブで炎上が発生するのか—日本のウェブ文化を手がかりとして. *情報通信学会誌*, 29 巻, 4 号, 61-71.
- [6] 山口真一 (2016). 炎上加担動機の実証分析. 2016 年社会情報学会 (SSI) 学会大会. https://www.sgu.ac.jp/soc/ssi/papers/32.pdf?utm_source=chatgpt.com
- [7] 瀬川友香, 浅谷公威, 坂田一郎 (2021). ユーザーに着目した SNS 上の攻撃とそのメカニズムに関する分析. 2021 年度人工知能学会全国大会 (第 35 回). 1I4-GS-4c-04.
- [8] 水村典弘 (2023). 不適切な広告表現の研究-人種・性差別的な表現で炎上した事例の分析-. *日本経営倫理学会誌*, 第 30 号, 61-74.
- [9] 瀬地山角 (2020). 炎上 CM でよみとくジェンダー論, 光文社新書.
- [10] Sondos Mahmoud Bsharat, Aidar Myrzakhan, Zhiqiang Shen (2023). "Principled Instructions Are All You Need for Questioning LLaMA-1/2, GPT-3.5/4". <https://doi.org/10.48550/arXiv.2312.16171>
- [11] 溝口理一郎, 古崎晃司, 来村徳信, 笹島宗彦 (2006). *オントロジー構築入門*, オーム社.

付録

【提案手法で用いたプロンプト】

バナー広告画像から指定された情報を抽出し、その情報をもとにバナー広告のジェンダー関連での炎上可能性を検出し、指摘してください。

- バナー広告内の文字を抽出し、文字集合ごとにリスト化してください。
- バナー広告内のイラストや描写の詳細な説明を提供してください。
- 抽出した文字とイラスト描写を基にバナー広告の訴求対象を考えてください。
- 入力されている『ジェンダー広告炎上オントロジー v1.xml』という XML ファイルからオントロジーの構造を解析してください。
- バナー広告の文字集合、イラスト説明、訴求対象と XML ファイルのオントロジーからバナー広告の炎上可能性を検出し、指摘してください。

Steps

1. バナー広告画像を解析し、文字を抽出します。
2. 検出された文字を一つの文字集合として整理し、一つのリストとして記載します。
3. バナー広告のイラストをレビューし、その詳細を説明します。形状、色、構図などに着目してください。
4. step 2 と step 3 で得られた情報を基にバナー広告の訴求対象を考えてください。
5. step 2, step 3, step 4 で取得した情報をまとめ、バナー広告の情報を取得してください。
6. 入力されている『ジェンダー広告炎上オントロジー v1.xml』という XML ファイルを解析し、オントロジーの構造を取得してください。
7. step 5 で取得したオントロジーの構造のうち、「広告のジェンダー関連の炎上原因」の概念の下位概念のオントロジーを用いて、step 5 でまとめたバナー広告がジェンダー関連で炎上する可能性があるか考えてください。
8. 炎上の可能性がある場合は、その炎上原因についてオントロジーを基に考えてください。炎上の可能性が無い場合はその旨を記述してください。

Output Format

- 炎上可能性: テキスト形式
- 炎上可能性の原因: 詳細なテキスト記述形式

Examples 1

****入力:****

- バナー広告画像

****出力:****

- 炎上可能性: あり
- 炎上可能性の原因: 「このバナー広告では訴求対象が女性で、イラストに女性だけが家事を行っている描写があり、昭和のステレオタイプである『家事は女性がするもの』ということを追認しています。すなわち女性に対して家事をすることを押し付けることを肯定してしまっており、『性役割の現状追認』の点で炎上する可能性があります。」

Examples 2

****入力:****

- バナー広告画像

****出力:****

- 炎上可能性: なし、または著しく低い
- 炎上可能性の原因: このバナー広告にはジェンダーに関して炎上する原因となる描写は見つかりませんでした。しかし、ジェンダー以外の観点で炎上してしまう可能背はゼロではないので、文字・文章表現やイラストの描写などの確認をしましょう。

Notes

- テキスト解析の精度は画像の質に依存することがあります。
- イラスト説明には、主観が入らないように具体的な要素にフォーカスすることが重要です。
- 両方の情報がバランスよく提供されるように努めてください。
- バナー広告の広告ゾーニングは「不特定多数の人の閲覧の可能性がある WEB」です。
- オントロジーは「広告のジェンダー関連の炎上原因」の概念の下位概念を用いて、炎上の検出はそれぞれの下位概念の最下層の葉の概念から準備検討してください。
- 入力されたオントロジー知識の情報で炎上の可能性を判断してください。
- 出力は日本語で出力してください。
- ステップバイステップで考えてください。
- 出力は Examples の出力例に沿った出力をしてください。