

# AI and Inclusive Marketing: Detecting and Reducing Stereotypes in Algorithmically Generated Creative Content

Dira Efriani<sup>1\*</sup>

<sup>1</sup> Universitas Ahmad Dahlan, Yogyakarta, Indonesia

## Abstract

### Article history:

Received: 24 October 2025

Revised: 13 November 2025

Accepted: 04 December 2025

Published: 30 December 2025

### Keywords:

AI-Generated Content, Algorithmic Bias, Inclusive Marketing, Representation Bias, Stereotypes.

### Identifier:

Nawala

Page: 122-136

<https://nawala.io/index.php/iraim>

This article examines how AI generated creative content interacts with inclusive marketing goals, asking whether and how generative systems encode, detect, and reduce stereotypes in brand communications. Positioned within debates on diversity in advertising and algorithmic bias, the study conducts a systematic literature review of peer reviewed work published between 2019 and 2024 on text to image and other creative AI tools used in marketing contexts. The synthesis shows that common generative models tend to overrepresent majority identities in high status roles, under represent marginalized groups, and reproduce traditional gendered and racialized associations, with these patterns flowing into campaigns that rely on AI assets. Audience research further indicates that AI based diversity initiatives can be perceived as inauthentic or tokenistic, dampening trust and brand attitudes. By integrating technical audits, consumer studies and inclusive branding research, the article highlights key methodological gaps and outlines research and governance priorities for designing creative AI that supports more equitable representation.

\*Corresponding author:  
(Dira Efriani)

©2025 The Author(s).

This is an open-access article under CC-BY-SA license (<https://creativecommons.org/licence/by-sa/4.0/>)



## 1. Introduction

Artificial intelligence has rapidly moved from a back-end analytics tool to a visible creator of marketing content, with generative models now producing images, copy, and video assets at scale. In advertising and digital campaigns, these systems promise efficiency, personalization, and creative experimentation, allowing brands to tailor messages to micro-segments and adapt creatives in near real time (Davenport et al., 2020; Huang & Rust, 2021). However, as AI systems increasingly assume the role of “creative partners,” concerns have grown that they may reproduce and even intensify harmful stereotypes related to gender, race, age, or disability, thereby undermining inclusive marketing efforts and eroding consumer trust (Akter et al., 2022; van Giffen et al., 2022).

Inclusive marketing has simultaneously emerged as a strategic imperative. It calls for brands to reflect diverse identities, amplify under-represented groups, and design communications that are accessible and respectful across demographic and cultural lines (Eisend et al., 2023; Grieco, 2024). Research shows that inclusive and diverse portrayals in advertising can enhance perceived fairness, brand authenticity, and long-term equity, while also contributing to broader social goals (Viglia et al., 2023; Masiello et al., 2024). Yet much of this literature has focused on human-crafted campaigns and legacy media, with limited attention to how algorithmically generated creative content may reshape the landscape of diversity and representation in practice.

Emerging empirical work on AI-generated images indicates that text-to-image systems do not simply mirror existing social patterns but can amplify occupational

and gender stereotypes. Studies of generative AI tools such as DALL·E and Bing Image Creator show that depictions of professions frequently overrepresent men and reproduce traditional role expectations, reinforcing biased associations in visual outputs (Sandoval-Martin & Martínez-Sanzo, 2024). At the same time, mapping studies on AI, advertising, and generative models point to a rapidly expanding ecosystem of tools that marketers can deploy to design personas, compose visuals, and localize campaigns, often with limited transparency regarding training data, bias audits, or debiasing procedures (Lim et al., 2024). This combination of powerful creative automation and opaque bias dynamics makes the intersection of AI and inclusive marketing both promising and fraught.

Recent work begins to address these tensions by examining how AI-generated “diverse” models are perceived by consumers. Evidence suggests that when inclusion efforts appear instrumental or technologically driven, they can be seen as inauthentic and may even backfire, harming trust and brand attitudes (Sands et al., 2024). These findings imply that mitigating bias in AI outputs is not only a technical challenge of adjusting data or algorithms but also a strategic and ethical challenge of aligning AI-assisted creativity with credible inclusive marketing intentions (Akter et al., 2022; Lim et al., 2024). Nevertheless, existing research remains fragmented across computer science, ethics, and marketing, with few integrative accounts of how stereotypes are detected, evaluated, and reduced in algorithmically generated creative content that is actually deployed in campaigns.

This article addresses that gap through a systematic literature review of peer-reviewed studies published between 2019 and 2024 on AI, inclusive marketing, and

stereotypes in algorithmically generated creative content. The review synthesizes insights on three core issues: how AI-generated advertising and brand content encode or challenge stereotypes; which conceptual and methodological approaches are used to detect and measure these patterns; and what technical, organizational, and communicative strategies have been proposed to reduce biased portrayals while maintaining creative effectiveness. By integrating these streams, the study aims to clarify the role of AI in inclusive marketing, identify blind spots in current practice, and delineate an agenda for future research on building creative AI systems that support, rather than undermine, more equitable representation in digital marketing.

## **2. Literature Review**

Research on inclusive marketing has long documented how advertising reproduces and legitimizes social hierarchies by normalizing certain bodies, roles and lifestyles as more desirable or suitable than others. Classic work on stereotypical portrayals shows that women, ethnic minorities and non-conforming bodies are often depicted in constrained or subordinate roles, while majority and high-status groups are associated with competence, authority and leadership (Eisend et al., 2023; Viglia et al., 2023). Recent reviews argue that these representational patterns do not simply mirror society but also shape what audiences come to see as normal, aspirational and trustworthy, making stereotype dynamics central to brand meaning and consumer response (Grieco, 2024; Masiello et al., 2024). At the same time, inclusive branding and diversity-oriented campaigns have been shown to generate both positive outcomes, such as increased identification and loyalty among

underrepresented groups, and accusations of tokenism when inclusivity is perceived as superficial or inconsistent with corporate behavior (Sands et al., 2024; Sandoval-Martin & Martínez-Sanz, 2024).

Parallel streams in marketing and information systems examine how AI is transforming creative production. Studies on AI in marketing strategy describe a progression from decision support tools to autonomous systems that generate and optimize messages, visuals and offers in near real time (Davenport et al., 2020; Huang & Rust, 2021; Akter et al., 2022; Lim et al., 2024). Consumer-facing work on AI-generated advertising suggests that audiences react differently to AI versus human authorship depending on appeal type and perceived social role of the AI agent, with consequences for attitudes and perceived authenticity (Chen et al., 2024). Yet much of this literature focuses on effectiveness and efficiency rather than on whether AI-generated content perpetuates or mitigates harmful stereotypes in its visual output.

A growing body of research outside mainstream marketing directly investigates bias in generative models that are now used to create advertising-like imagery. Large-scale audits of text-to-image systems show systematic distortions in how occupations, personality traits and everyday situations are represented across gender, race, age and geography, with models overrepresenting majority identities in high-status roles and aligning neutral prompts more closely with Western contexts (Bianchi et al., 2023; Naik & Nushi, 2023). Medical and professional imagery studies similarly find that AI-generated depictions of clinicians and specialists fail to reflect workforce diversity and instead emphasize narrow notions of attractiveness,

masculinity and whiteness, raising concerns about their suitability for educational and promotional communication (Gisselbaek et al., 2024). These findings imply that when the same models are used in commercial creative workflows, brand visuals may inadvertently encode and amplify representational inequalities even when marketers intend to be inclusive.

Despite these advances, several gaps remain. First, empirical studies rarely connect technical audits of generative bias with marketing outcomes such as trust, perceived inclusivity or willingness to engage with AI-created campaigns. Second, most work treats ethical AI, fairness and diversity as abstract design goals rather than examining how marketers actually integrate bias detection or mitigation into their day-to-day creative processes. Third, there is limited synthesis across disciplines on concrete mechanisms, such as prompt engineering, model selection, post-editing or governance structures, that can help brands detect and reduce stereotypes in AI-generated visuals without sacrificing relevance or performance. Addressing these gaps requires integrating insights from inclusive marketing, algorithmic fairness and human-AI collaboration to understand how AI-mediated creativity can be steered toward more equitable representations.

### **3. Methods**

The study uses a systematic literature review design to synthesize current knowledge on how AI generated creative content encodes, detects and mitigates stereotypes in the context of inclusive marketing. A structured search was conducted in major academic databases such as Scopus, Web of Science, ScienceDirect and

Google Scholar for peer reviewed publications between 2019 and 2024. Search strings combined terms related to artificial intelligence and generative models (for example “AI generated”, “generative AI”, “text to image”, “creative AI”) with terms related to marketing and advertising (“advertising”, “brand communication”, “marketing content”) and inclusion or bias (“stereotypes”, “bias”, “diversity”, “inclusive marketing”). The search was limited to English language articles. After removing duplicates, titles and abstracts were screened to identify studies that addressed algorithmically generated or AI assisted creative content used in marketing, advertising or brand communication and that explicitly examined stereotypes, bias or representation.

Full texts of the remaining articles were assessed using predefined inclusion and exclusion criteria. Studies were included if they (1) involved AI generated or AI mediated creative assets such as images, videos or ad copy, (2) examined stereotype related phenomena such as representation, bias, diversity or inclusion in those outputs and (3) reported empirical findings, audits or analytical frameworks relevant to marketing or communication practice. Purely technical papers on model training without any link to representation, ethical or marketing implications, and work on non-creative AI applications (such as pricing or recommendation systems) were excluded. For each included study, a coding template captured information on research context, type of AI system, creative modality, operationalization of stereotypes and inclusion, methods for detection or assessment, mitigation strategies and reported outcomes. This structured extraction enabled cross study comparison

and synthesis around how AI and inclusive marketing intersect in the design and governance of algorithmically generated creative content.

## **4. Results and Discussion**

The studies included in this review converge on the finding that generative AI systems used for creative production systematically encode existing social hierarchies in their outputs. Audits of text-to-image models show that prompts related to high-status or technical occupations disproportionately yield images of white, male and Western appearing figures, while prompts for caring or service roles are more likely to produce women or feminised presentations (Bianchi et al., 2023; Naik & Nushi, 2023). Medical and professional imagery displays similar patterns, with AI-generated representations of clinicians and specialists under-representing women and ethnically diverse professionals relative to real workforce data (Gisselbaek et al., 2024). When such models are inserted into marketing workflows to generate or adapt campaign visuals, these biases are effectively imported into brand communications, even if marketers do not explicitly request stereotypical content. This first cluster of findings suggests that, left unconstrained, generative systems tend to amplify rather than attenuate the representational imbalances documented in earlier advertising research (Eisend et al., 2023; Viglia et al., 2023; Zayer et al., 2023).

A second cluster of studies examines how audiences respond when AI-generated creative assets are used to signal inclusion or diversity. Evidence from consumer experiments indicates that AI-created or AI-enhanced models can be perceived as less authentic than traditional photography or human-led casting,

particularly when the use of AI is disclosed or becomes salient (Chen et al., 2024; Sands et al., 2024). In some cases, attempts to showcase diversity through synthetic models are interpreted as cost-saving or opportunistic, leading to accusations of tokenism and lowering attitudes toward the brand, despite the presence of more diverse faces on screen (Grieco, 2024; Sandoval-Martin & Martínez-Sanz, 2024). This pattern resonates with broader work on inclusive campaigns that use real human diversity, where perceived sincerity and alignment with organisational behavior are critical to positive outcomes (Qayyum et al., 2023; Masiello et al., 2024). The contribution of AI is therefore ambivalent: it can lower the technical barriers to visual diversification, but it can also undermine perceived authenticity if audiences suspect that inclusion is generated rather than lived.

On the technical and organizational side, the review shows that work on bias detection is far more developed than work on mitigation in real marketing contexts. Computer science and AI ethics studies propose a range of diagnostic tools, from stereotype benchmarks and prompt-based audits to embedding-space analyses of demographic associations (Akter et al., 2022; Bianchi et al., 2023; Naik & Nushi, 2023). In parallel, research on algorithmic fairness in marketing has begun to formalize fairness criteria and evaluation procedures tailored to marketing scenarios, highlighting trade-offs between equity, efficiency and targeting precision (Yang et al., 2024). However, few marketing-focused papers document how such tools are operationalized in day-to-day creative processes, for example in agency workflows, brand-side content studios or platform creative tools. Generative AI is often introduced primarily for its efficiency and scalability benefits, consistent with

broader narratives around AI-driven automation in marketing (Davenport et al., 2020; Huang & Rust, 2021; Lim et al., 2024), with inclusion and fairness considerations added after the fact, if at all. The result is a gap between the availability of bias diagnostics and their systematic use in governing creative outputs.

The evidence on mitigation strategies is still limited but points to several promising directions and important constraints. Some studies suggest that prompt engineering and stronger specification of desired diversity can partially counteract default biases, but they also show that such fixes require sustained expertise and may be fragile across model updates and deployment settings (Bianchi et al., 2023; Naik & Nushi, 2023). Others emphasize human-in-the-loop arrangements, where creatives and brand managers review and curate AI outputs to ensure that they align with inclusive guidelines and brand values before deployment (Zayer et al., 2023; Grieco, 2024; Masiello et al., 2024). Yet few articles provide detailed evidence on the effectiveness or cost of these approaches at scale. In addition, there is almost no empirical work comparing different governance models, such as centralized AI review boards versus distributed team-level controls, despite growing recognition that algorithmic bias in marketing is a socio-technical issue involving tools, people and organizational incentives (Akter et al., 2022; Yang et al., 2024).

Taken together, the results suggest that AI and inclusive marketing currently intersect in a contradictory way. On one hand, generative systems have the capacity to expand the representational palette available to brands and to support more flexible, localized creative experimentation. On the other hand, without deliberate design and governance, these systems tend to reproduce and amplify stereotypes,

while their visible use in diversity initiatives risks being read as inauthentic. The overall evidence base is also constrained by methodological limitations: most studies use short-term experiments with narrow outcome measures, and there is little longitudinal or field-based work linking AI-generated content to real marketplace behavior. These gaps underscore the need for future research that evaluates mitigation strategies in operational settings, integrates bias diagnostics and fairness evaluation frameworks into creative toolchains, and investigates how consumers across different cultures and identities interpret AI-mediated representations in the longer term.

## **5. Conclusion**

This review shows that AI generated creative content currently sits in tension with the goals of inclusive marketing. Across audits and experiments, generative systems tend to reproduce and often intensify existing patterns of underrepresentation and stereotyping, particularly in how gender, race and professional status are visually portrayed. When these systems are used to create or adapt advertising assets, their biases flow directly into brand communications, even where marketers intend to be more inclusive. At the same time, consumer studies suggest that using obviously AI generated or AI enhanced models to signal diversity can be perceived as inauthentic or cost driven, weakening trust and brand attitudes despite more visibly diverse imagery. Together, these findings indicate that AI has the capacity to expand the creative palette and reach, but without careful design and governance it risks undermining both representational equity and perceived sincerity.

The conclusions drawn from this literature are constrained by several important limitations in the underlying studies. Most of the evidence comes from short term laboratory or online experiments, often with convenience samples, narrow product categories and limited outcome measures. Many audits rely on specific models, prompts and time points, which may not generalize across tools, markets or cultural contexts. There is little longitudinal or field-based work linking AI generated content to actual marketplace behavior, nor systematic comparison of different mitigation strategies or governance models. These gaps reduce the strength and external validity of current findings and invite readers to consider how far results would hold under different model architectures, organizational practices or regulatory environments.

Future research should therefore treat AI and inclusive marketing as a socio technical problem rather than a purely technical or purely creative one. Conceptually, there is a need to integrate theories of stereotyping, authenticity and brand meaning with emerging work on algorithmic fairness and human AI collaboration in creative teams. Empirically, longitudinal and field studies could test how different combinations of prompts, human oversight and governance structures affect both representational patterns and consumer responses over time. Methodologically, cross cultural and intersectional analyses are needed to capture how diverse audiences interpret AI mediated inclusion efforts, and participatory approaches could involve marginalized groups in evaluating and redesigning creative pipelines. By addressing these shortcomings, future work can move from documenting the biases of generative systems to developing concrete principles and tools that help

brands detect and reduce stereotypes in AI generated content while preserving creative effectiveness and trust.

## **References**

- Akter, S., Dwivedi, Y. K., Sajib, S., Biswas, K., Bandara, R. J., & Michael, K. (2022). Algorithmic bias in machine learning-based marketing models. *Journal of Business Research*, *144*, 201-216.
- Bianchi, F., Kalluri, P., Durmus, E., Ladhak, F., Cheng, M., Nozza, D., Hashimoto, T., Jurafsky, D., Zou, J., & Caliskan, A. (2023). Easily accessible text-to-image generation amplifies demographic stereotypes at large scale. *In Proceedings of the 2023 ACM conference on fairness, accountability, and transparency* (pp. 1493-1504).
- Chen, Y., Wang, H., Rao Hill, S., & Li, B. (2024). Consumer attitudes toward AI-generated ads: Appeal types, self-efficacy and AI's social role. *Journal of Business Research*, *185*, 114867.
- Davenport, T. H., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*, *48*(1), 24-42.
- Eisend, M., Muldrow, A. F., & Rosengren, S. (2023). Diversity and inclusion in advertising research. *International Journal of Advertising*, *42*(1), 52-59.
- Gisselbaek, M., Minsart, L., Köseleli, E., Suppan, M., Meco, B. C., Seidel, L., Albert, A., Barreto Chang, O. L., Saxena, S., & Berger-Estilita, J. (2024). Beyond the stereotypes: Artificial Intelligence image generation and diversity in anesthesiology. *Frontiers in Artificial Intelligence*, *7*, 1462819.

- Grieco, C. (2024). Conceptualizing inclusive marketing: A synthesis of theory and practice. *European Management Journal*, 43(6), 1049-1064.
- Huang, M. H., & Rust, R. T. (2021). A strategic framework for artificial intelligence in marketing. *Journal of the Academy of Marketing Science*, 49(1), 30-50.
- Lim, C. V., Zhu, Y. P., Omar, M., & Park, H. W. (2024). Decoding the relationship of artificial intelligence, advertising, and generative models. *Digital*, 4(1), 244-270.
- Masiello, B., Garofano, A., Izzo, F., & Bonetti, E. (2024). Inclusive marketing and disability: value creation strategies for organisations and society in the toy industry. *Journal of Marketing Management*, 40(9-10), 851-876.
- Naik, R., & Nushi, B. (2023). Social biases through the text-to-image generation lens. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 786-808).
- Qayyum, A., Jamil, R. A., Shah, A. M., & Lee, K. (2023). Inclusive advertising featuring disability on Instagram: Effect on consumer well-being, brand engagement, and purchase intention. *Journal of Retailing and Consumer Services*, 75, 103515.
- Sandoval-Martin, T., & Martínez-Sanzo, E. (2024). Perpetuation of gender bias in visual representation of professions in the generative ai tools dall· e and bing image creator. *Social Sciences*, 13(5), 250.
- Sands, S., Demsar, V., Ferraro, C., Campbell, C., & Cohen, J. (2024). Inauthentic inclusion: Exploring how intention to use AI-generated diverse models can backfire. *Psychology & Marketing*, 41(6), 1396-1413.

- van Giffen, B., Herhausen, D., & Fahse, T. (2022). Overcoming the pitfalls and perils of algorithms: A classification of machine learning biases and mitigation methods. *Journal of Business Research*, 144, 93-106.
- Viglia, G., Tsai, W. H. S., Das, G., & Pentina, I. (2023). Inclusive advertising for a better world. *Journal of Advertising*, 52(5), 643-646.
- Yang, M., Guo, J., Zhu, L., Zhu, H., Song, X., Zhang, H., & Xu, T. (2024). Fairness evaluation of marketing algorithms: A framework for equity distribution. *Journal of Electronic Business & Digital Economics*, 3(3), 251-274.
- Zayer, L. T., Coleman, C. A., & Gurrieri, L. (2023). Driving impact through inclusive advertising: An examination of award-winning gender-inclusive advertising. *Journal of Advertising*, 52(5), 647-665.