

# AI Biases in Marketing Practice: Perception and Action Strategies

Caterina Fox<sup>1</sup> and Gabriele Schuster<sup>2</sup>

<sup>1</sup>IU Internationale Hochschule, Mannheim, Germany

<sup>2</sup>IU Internationale Hochschule, Hamburg, Germany

[caterina.fox@iu.org](mailto:caterina.fox@iu.org)

[gabriele.schuster@iu.org](mailto:gabriele.schuster@iu.org)

**Abstract:** While the integration of artificial intelligence (AI) into marketing practices can improve efficiency and personalization, there are concerns about inherent biases that might produce unfair outcomes. Inherent biases refer to systematic, unintended preferences or exclusions embedded within AI algorithms, often originating from training data or design choices. Through qualitative interviews, this study examines how marketing professionals manage AI bias, offering actionable recommendations and best practices. Initial findings indicate a spectrum of attitudes towards AI bias in marketing, ranging from proactive engagement with the issue to avoidance.

**Keywords:** Artificial intelligence (AI), Bias, Marketing, Expert Interviews

---

## 1. Introduction

The rapid development of artificial intelligence (AI) has made significant progress in recent years, touching many sectors of our society. In marketing especially, AI technologies offer numerous applications that enhance marketing efficiency and effectiveness (Haleem et al., 2022), from automation to personalized content generation. However, AI-driven strategies often operate within a "black box," limiting marketing professionals' oversight and potentially disrupting strategic control (Kozinets and Gretzel, 2021). Given that marketing decisions impact people, it is crucial to uphold principles of fairness, transparency, and equality (Jones and Middleton, 2007). Delegating these decisions to AI inherently transfers ethical considerations to algorithms as well. Unfortunately, algorithmic decision-making sometimes falls short of meeting our fairness expectations. Studies show that AI can inadvertently replicate and amplify human biases, resulting in unintended negative consequences in marketing (Rivas and Zhao, 2023; Wach et al., 2023). Bias refers to a situation where systematic differences emerge despite lacking an objective basis for disparate treatment (Turner Lee et al., 2019).

This research project addresses this unexplored issue by posing the following questions:

- What efforts are marketing professionals currently making to recognize, manage and mitigate algorithmic biases in AI-driven marketing, including biases related to demographic representation, cultural sensitivity, and stereotype propagation?
- What strategies can be developed to address these challenges and design responsible, diversity-oriented marketing concepts?

## 2. AI and Marketing Ethics: Examples of Marketing Decisions Distorted by AI Bias

The primary applications of AI in marketing currently lie in execution (SRH Berlin University of Applied Sciences, 2023). Content creation is likely to constitute a significant portion of this (Chui et al., 2023) as generative AI is already widely used for writing texts and creating images. Generative AI refers to algorithms that can create new content, such as text, images, or music, often based on patterns learned from existing data (e.g., GPT models). Given the profound influence that marketing content can have on individual perceptions, AI wields substantial social impact in this field (Zhou et al., 2023).

Biases in AI are well-documented in numerous studies within the field. For example, AI-driven facial recognition systems have consistently shown weaknesses related to gender and skin color, performing better for men and individuals with lighter skin tones (Buolamwini and Gebru, 2018). This limitation likely extends to generative AI as well. For instance, when prompted with "happy family," image suggestions often predominantly feature white individuals. Such outputs perpetuate stereotypes and reflect underlying cultural biases (Bianchi et al., 2023).

Bias in AI manifests not only in content generation, but also extends to operational processes fundamental to marketing. A particularly significant example is the automated display of advertisements, where similar biases can be replicated and amplified (Varsha, 2023). Turner Lee et al. (2019), for instance, found that algorithms are more likely to show exploitative credit card offers to African-American consumers than to white Americans. Furthermore, based on gender stereotypes, Facebook displays different job advertisements to men and women (Hao, 2021).

Biases have also surfaced in product and distribution policy. Amazon, for example, excluded certain geographical areas from same-day Prime delivery for reasons tied to profitability. The decision employed parameters such as the number of Prime members in an area, proximity to a distribution center, and the number of service providers available for delivery. Ultimately, these exclusions affected predominantly lower socioeconomic neighborhoods with largely Black populations (Ingold and Soper, 2016).

This discussion naturally extends to marketing ethics, which provides a lens for evaluating the fairness and inclusivity of AI-driven marketing practices. Marketing ethics refers to principles and standards guiding fair, equitable, and transparent marketing decisions. A marketing measure's ethicality can be assessed through two dimensions: (1) the product's potential harm and (2) the target group's vulnerability (Jones and Middleton, 2007). The aforementioned examples refer to target group vulnerability and thus highlight the ethical quandary posed by AI biases in marketing. These biases negatively impact certain groups, even if there was no programmer intent to discriminate. Thus, it is imperative and timely to investigate the unintentional (re)production of biases by AI in marketing (Turner Lee et al., 2019). Flowers (2019) suggests that biased algorithms should not be seen solely as failures but as starting points for examining societal biases.

### **3. AI Bias and its Causes**

Fui-Hoon Nah et al. (2023) outline potential biases impacting marketing decisions in general, such as favoritism towards certain groups/individuals, the underrepresentation of groups, cultural misunderstandings, language preferences, and stereotype propagation. These biases are often reproduced and amplified by AI through various processes:

1. Stereotypes and prejudices are deeply entrenched in society (Kuck, 2023). Existing human biases and implicit prejudices are replicated and amplified by machines through feedback loops (Turner Lee, 2018).
2. AI-inherent biases also arise due to inadequate training data (Jain et al., 2023). An over- or under-representation of certain groups in the data can have significant consequences. Additionally, programming teams often lack diversity, which can result in unnoticed problems. The opacity in data origins further complicates efforts to identify potential biases (Zhou et al., 2023).
3. Even when characteristics like gender or racial identity are deliberately omitted from AI systems to ensure equal treatment, algorithms may still use proxy variables that indirectly reveal group membership. For instance, height and weight might be indicators of gender (Angwin and Parris, 2016).

### **4. Recognizing and Avoiding AI Bias in Marketing**

Beyond ethical concerns, marketers should be interested in avoiding biased decisions for legal and liability reasons (Turner Lee et al., 2019). Ensuring AI-supported decisions are inclusive and fair is crucial (Rivas and Zhao, 2023). While marketers have very limited influence over the AI itself (Kozinets and Gretzel, 2021), they should be equipped to recognize and rectify biases.

Typically, algorithms can be evaluated by testing their outputs for inconsistencies, such as by comparing results across different groups (Turner Lee et al., 2019). With generative AI, this becomes more challenging due to the absence of objectively correct outputs (Zhou et al., 2023). Instead, outputs need to be scrutinized for bias-containing patterns. However, this is complicated further as outputs can reinforce existing stereotypes and cement biases. Turner Lee et al. (2019) propose increasing team diversity, evaluating training data quality, and generally reflecting cultural sensitivity in decision-making as potential recommendations.

### **5. Methodology**

The study utilized a qualitative methodology, employing semi-structured interviews with a sample of 12 marketing professionals across various roles, including content creators and communication agency representatives. Each interview lasted between 45 and 60 minutes and followed a thematic guide exploring topics such as perceptions of bias in general, AI bias in particular, mitigation strategies, and ethical considerations in AI-driven marketing. Data was analyzed using thematic analysis, with codes and themes iteratively refined to ensure reliability and validity.

### **6. Preliminary Results**

Initial findings from the ongoing interview process reveal a diverse range of perspectives and strategies. Opinions vary from those who resolutely believe in taking a stand against bias and communicating clear values

to those who prefer to avoid the topic altogether. Conscious strategies to avoid bias in AI usage are scarce. Those practitioners who prioritize bias avoidance often cite personal values and life experiences as their primary safeguards against bias. They do not feel that the use of AI necessitates a change in approach. In contrast, some interviewees even express dissatisfaction when AI adopts a perceived politically correct stance.

The preliminary results reflect broader challenges identified in the literature, such as the unintentional perpetuation of stereotypes and the opacity of algorithmic decision-making (Buolamwini and Gebru, 2018; Turner Lee et al., 2019). This underscores the urgent need for marketing-specific frameworks to detect and mitigate bias, integrating both technical and ethical considerations.

## References

- Angwin, J. and Parris, T. Jr. (2016) "Facebook lets advertisers exclude users by race", [online], ProPublica, <https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race>.
- Bianchi, F., Kalluri, P., Durmus, E., Ladhak, F., Cedheng, M., Nozza, D., Hashimoto, T., Jurafsky, D., Zou, J., and Caliskan, A. (2023) "Easily accessible text-to-image generation amplifies demographic stereotypes at large scale", Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency, pp 1493–1504, <https://doi.org/10.1145/3593013.3594095>.
- Buolamwini, J. and Gebru, T. (2018) "Gender Shades: Intersectional accuracy Disparities in commercial gender classification", Proceedings of the 1st Conference on Fairness, Accountability and Transparency, Proceedings of Machine Learning Research, <https://proceedings.mlr.press/v81/buolamwini18a.html>.
- Chui, M., Hazan, E., Roberts, R., Singla, A., Smaje, K., Sukharevsky, A., Yee, L. and Zempel, R. (2023) "The economic potential of generative AI: The next productivity frontier", [online], McKinsey and Company, <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier>.
- Ferraro, C., Hemsley, A. and Sands, S. (2023) "Embracing diversity, equity, and inclusion (DEI): Considerations and opportunities for brand managers", *Business Horizons*, Vol 66, No. 4, pp 463–479, <https://doi.org/10.1016/j.bushor.2022.09.005>.
- Flowers, J. C. (2019) "Rethinking algorithmic bias through phenomenology and pragmatism", Computer ethics-philosophical enquiry (CEPE) proceedings, <https://doi.org/10.25884/MH5Z-FB89>.
- Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K. and Chen, L. (2023) "Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration", *Journal of Information Technology Case and Application Research*, Vol 25, No. 3, pp 277–304, <https://doi.org/10.1080/15228053.2023.2233814>.
- Haleem, A., Javaid, M., Qadri, M. A., Singh, R. P. and Suman, R. (2022) "Artificial intelligence (AI) applications for marketing: A literature-based study", *International Journal of Intelligent Networks*, Vol 3, pp 119-132, <https://doi.org/10.1016/j.ijin.2022.08.005>.
- Hao, K. (2021) "Facebook's ad algorithms are still excluding women from seeing jobs", *MIT Technology Review*, <https://www.technologyreview.com/2021/04/09/1022217/facebook-ad-algorithm-sex-discrimination>.
- Ingold, D. and Soper, S. (2016) "Amazon doesn't consider the race of its customers, Should it?", [online], Bloomberg, <https://www.bloomberg.com/graphics/2016-amazon-same-day>.
- Jain, V., Rai, H., Parvathy, P. and Mogaji, E. (2023) "The prospects and challenges of ChatGPT on marketing research and practices", *SSRN Electronic Journal*, <https://doi.org/10.2139/ssrn.4398033>.
- Jones, J. L. and Middleton, K. L. (2007) "Ethical decision-making by consumers: The roles of product harm and consumer vulnerability", *Journal of Business Ethics*, Vol 70, No. 3, pp 247–264, <https://doi.org/10.1007/s10551-006-9109-2>.
- Kozinets, R. V. and Gretzel, U. (2021) "Commentary: Artificial intelligence: The marketer's dilemma", *Journal of Marketing*, Vol 85, No. 1, pp 156–159, <https://doi.org/10.1177/0022242920972933>.
- Kuck, K. (2023) "Generative artificial intelligence: A double-edged sword", 2023 World Engineering Education Forum - Global Engineering Deans Council (WEEF-GEDC), pp 1–10, <https://doi.org/10.1109/WEEF-GEDC59520.2023.10343638>.
- Paul, J., Ueno, A. and Dennis, C. (2023) "ChatGPT and consumers: Benefits, pitfalls and future research agenda", *International Journal of Consumer Studies*, Vol 47, No. 4, pp 1213–1225, <https://doi.org/10.1111/ijcs.12928>.
- Rivas, P. and Zhao, L. (2023) "Marketing with ChatGPT: Navigating the ethical terrain of GPT-based chatbot technology", *AI*, Vol 4, No. 2, pp 375–384, <https://doi.org/10.3390/ai4020019>.
- SRH Berlin University of Applied Sciences (2023) "Wie sehr treffen die folgenden Aussagen auf Ihre Marketingabteilung zu? Wir nutzen KI heute für ... [Graph]", [online], Statista, <https://de-statista-com.pxz.iubh.de:8443/statistik/daten/studie/1446333/umfrage/einsatzgebiete-von-kuenstlicher-intelligenz-im-marketing-in-der-dach-region>.
- Turner Lee, N. (2018) "Detecting racial bias in algorithms and machine learning", *Journal of Information, Communication and Ethics in Society*, Vol 16, No. 3, pp 252–260, <https://doi.org/10.1108/JICES-06-2018-0056>.
- Turner Lee, N., Resnick, P. and Barton, G. (2019) "Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms", [online], Brookings, <https://www.brookings.edu/articles/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms>.

- Varsha, P.S. (2023) "How can we manage biases in artificial intelligence systems – A systematic literature review", *International Journal of Information Management Data Insights*, Vol 3, No. 1, <https://doi.org/10.1016/j.jjime.2023.100165>.
- Wach, K., Duong, C. D., Ejdy, J., Kazlauskaitė, R., Korzynski, P., Mazurek, G., Paliszkiwicz, J. and Ziemia, E. (2023) "The dark side of generative artificial intelligence: A critical analysis of controversies and risks of ChatGPT", *Entrepreneurial Business and Economics Review*, Vol 11, No. 2, pp 7–30, <https://doi.org/10.15678/EBER.2023.110201>.
- Zhou, M., Abhishek, V. and Srinivasan, K. (2023) "Bias in Generative AI" (Work in Progress), [online], [https://www.andrew.cmu.edu/user/ales/cib/bias\\_in\\_gen\\_ai.pdf](https://www.andrew.cmu.edu/user/ales/cib/bias_in_gen_ai.pdf).