

PAPER DETAILS

TITLE: Generative Artificial Intelligence in Marketing: The Invisible Danger of AI Hallucinations

AUTHORS: Burak Yaprak

PAGES: 133-158

ORIGINAL PDF URL: <https://dergipark.org.tr/tr/download/article-file/4382734>

Generative Artificial Intelligence in Marketing: The Invisible Danger of AI Hallucinations

Burak YAPRAK

Orcid.org/ 0000-0001-9831-0813

Dr. Öğr. Üyesi., İstanbul Ticaret Üniversitesi

Asst. Prof, Istanbul Commerce University

byaprak@ticaret.edu.tr

Abstract

This study explores the transformative impact of Generative Artificial Intelligence (GAI) on the marketing highlighting both its significant opportunities and inherent challenges. GAI enhances marketing strategies through automated content creation, personalized customer experiences, and advanced data analytics, thereby increasing efficiency and engagement. However, the phenomenon of AI hallucinations—where AI models produce realistic yet incorrect or misleading information—poses substantial risks, including damage to brand reputation, erosion of consumer trust, and potential legal ramifications. To mitigate the risks associated with AI hallucinations, the study proposes comprehensive risk management strategies that include technical solutions to detect and correct erroneous outputs, human oversight to ensure accuracy, and adherence to ethical and regulatory frameworks. By balancing the advantages of GAI with robust measures to address AI-generated inaccuracies, organizations can harness its full potential while safeguarding their brand integrity and maintaining trust with customers.

Keywords: Generative Artificial Intelligence, GAI, Marketing, Marketing Strategies, AI Hallucinations, ChatGPT

Pazarlamada Üretken Yapay Zekâ: Yapay Zekâ Halüsinasyonlarının Gizli Tehlikesi

Öz

Bu çalışma, pazarlama alanında Üretken Yapay Zekânın (ÜYZ) dönüştürücü etkisini hem önemli fırsatlarını hem de barındırdığı zorlukları vurgulayarak incelemektedir. ÜYZ, otomatik içerik üretimi, kişiselleştirilmiş müşteri deneyimleri ve ileri düzey veri analitiği yoluyla pazarlama stratejilerini geliştirerek verimliliği ve etkileşimi artırmaktadır. Ancak, Yapay Zekâ modellerinin gerçekçi ancak yanlış veya yanıltıcı bilgi üretmesi olarak tanımlanan Yapay Zekâ halüsinasyonları, marka itibarına zarar verme, tüketici güveninin sarsılması ve olası yasal sonuçlar gibi önemli riskler taşımaktadır. Çalışma, literatür incelemesi ve vaka analizleri kullanarak Üretken Yapay Zekânın pazarlamadaki çift yönlü etkilerini, yani potansiyel avantajları ve ilişkili riskleri ele almaktadır. Çalışma, Yapay Zekâ halüsinasyonları ile ilişkili riskleri azaltmak için, yanlış çıktıları tespit etmek ve düzeltmek için teknik çözümler, doğruluğu sağlamak için insan denetimi ve etik ile yasal düzenlemelere uyum gibi kapsamlı risk yönetimi stratejileri önermektedir. Bu sayede, ÜYZ'nin avantajlarını dengeli bir şekilde kullanarak marka bütünlüğünü korumak ve müşteri güvenini sürdürmek hedeflenmektedir.

Anahtar Kelimeler: Üretken Yapay Zekâ, ÜYZ, Pazarlama, Pazarlama Stratejileri, Yapay Zekâ Halüsinasyonları, ChatGPT

Introduction

In recent years, the rapid development of artificial intelligence (AI) has significantly transformed various industries, particularly reshaping marketing strategies by enabling more personalized and effective engagement with target audiences (Chaffey & Ellis-Chadwick, 2019). Generative Artificial Intelligence (GAI), a subset of AI, has emerged as a transformative technology, offering innovative solutions such as automated content creation, personalized customer experiences, and advanced data analytics. By leveraging tools such as OpenAI's ChatGPT, Google's Bard, Anthropic's Claude, and Meta's Llama 2, businesses can generate human-like text, visuals, and videos, thereby enhancing the efficiency and creativity of marketing strategies. Each of these models presents unique strengths: ChatGPT is renowned for its general linguistic capabilities and widespread application support, Bard excels in real-time data integration and knowledge retrieval, Claude focuses on ethical and safety-driven responses, and Llama 2 stands out for its open-source accessibility, making it ideal for research and customization.

GAI applications enable businesses to craft tailored marketing campaigns, ranging from personalized email content to dynamic social media strategies, that resonate with specific audience preferences. For instance, Sherly Steffi et al. (2024) highlight how AI technologies can analyze consumer behaviors and provide personalized content, positively influencing customer satisfaction and brand loyalty. Additionally, Bhattacharya (2024) emphasizes strategies to address hallucinations in large language models, an issue central to the reliability of AI-generated content. Furthermore, GAI's predictive data

analytics capabilities allow businesses to better forecast market trends and make data-driven decisions (Bernard, 2023). These advancements underscore the transformative role of GAI in fostering more impactful and engaging marketing strategies.

However, alongside these advantages, there exists an inescapable risk: AI hallucinations. These occur when AI models generate plausible but incorrect or misleading outputs, posing significant challenges in marketing where accuracy and reliability are critical. For example, Wen and Laporte (2024) note that while AI can evoke positive emotional responses in audiences, its lack of diversity and emotional depth can undermine the authenticity of marketing communications. Misleading or inaccurate content, such as false product information or inappropriate campaign messaging, can damage brand reputation, erode consumer trust, and result in legal challenges (Bender et al., 2021). Moreover, as Christensen et al. (2024) explain, consumers' awareness of AI hallucinations can influence their decision-making processes, particularly in trust-sensitive industries like tourism.

The purpose of this study is to provide a comprehensive examination of the phenomenon of AI hallucinations encountered in GAI applications. By investigating the causes and implications of these inaccuracies, the study seeks to balance the opportunities offered by GAI with the associated risks, offering practical strategies for marketers to utilize this technology responsibly. The structure of the study begins with an exploration of GAI's specific applications in marketing, including content generation, personalization, and data analytics, followed by an in-depth analysis of AI hallucinations, their

causes, and their manifestations in marketing. It then discusses strategies for risk mitigation, including technical, ethical, and human-centered approaches. By addressing these critical issues, this research aims to guide marketing professionals in leveraging GAI effectively while safeguarding brand integrity and customer trust.

1. Generative Artificial Intelligence and Marketing

GAI is a subfield of artificial intelligence aimed at creating new data samples resembling existing data. Unlike discriminative models, GAI models focus on learning the underlying distribution of inputs to generate similar new outputs rather than classifying or predicting data. One of the fundamental architectures of GAI is Generative Adversarial Networks (GANs), which consist of two neural networks: a generator that produces data and a discriminator that evaluates whether the data is real or fake. These two networks compete against each other, enabling the generator to enhance its ability to produce increasingly realistic data with each iteration (Goodfellow et al., 2014). Another significant advancement in GAI is the development of the Transformer architecture, which has accelerated progress in the field of Natural Language Processing (NLP). Models such as GPT-3, leveraging deep learning and large datasets, can generate human-like texts and demonstrate advanced proficiency in understanding the context and nuances of language (Brown et al., 2020). These developments have enabled the establishment of more natural and personalized interactions with target audiences in the marketing field.

GAI models are trained on large datasets to learn the probability distributions of inputs. This training process involves

adjusting the model's parameters to minimize the difference between the generated data and real data. For instance, in GANs, while the generator creates data, the discriminator determines whether the data is real or artificial. Through iterative training phases, the generator enhances its capacity to produce outputs that can deceive the discriminator, leading to the generation of increasingly realistic data (Goodfellow et al., 2014). This process facilitates the creation of more realistic and impactful content in the marketing domain. GAI has various applications that transform how businesses engage with consumers and design strategies.

1.1. Content Generation

GAI models can produce high-quality content such as articles, social media posts, visuals, and videos. For example, AI-powered tools can prepare personalized e-mail campaigns tailored to specific target audiences or generate product descriptions that resonate with user groups (Kietzmann et al., 2018). In the visual content domain, GAI can be used to design logos, advertising materials, or pre-production product simulations (Radford et al., 2019). These capabilities enhance marketing practitioners' ability to produce content quickly and diversely (Elgammal et al., 2017).

GAI has enabled innovative marketing campaigns across various industries. For instance, Heinz used AI to create unique, visually appealing images of ketchup bottles for their advertisements, which highlighted the brand's innovative approach (Speedy Brand, 2024). Similarly, Mango adopted AI-generated imagery to streamline its content production for marketing campaigns, reducing costs while maintaining creativity (FT, 2024). Coca-Cola also showcased the

creative potential of GAI by integrating AI-generated music and art into its advertising campaigns. In 2023, Coca-Cola launched the "Create Real Magic" platform in collaboration with OpenAI and Bain & Company, allowing digital artists to generate original artwork using the brand's iconic assets (Coca-Cola Company, 2023). Additionally, in 2024, the company released an AI-generated Christmas commercial that reimaged its classic "Holidays Are Coming" ad, further blending tradition with modern AI technologies (NYP, 2024). These examples demonstrate how GAI can transform content creation processes while reinforcing brand identity and engaging consumers.

1.2. Personalized Customer Experiences

Personalization has become a critical element in modern marketing, and GAI enables scaling these experiences to be highly individualized. By analyzing consumer data, AI can provide personalized recommendations, adapt website content to individual users, and create marketing messages tailored to each consumer. This level of personalization has been shown to enhance customer engagement and loyalty (Lambrecht & Tucker, 2019). Personalization through GAI fosters deeper connections between consumers and the brand, significantly boosting the success of marketing activities (Butler, 2023). For example, Nike's "By You" campaign utilized AI-driven tools to allow customers to design their own shoes, creating a highly engaging and personalized customer experience (Digital Agency Network, 2024). This demonstrates how GAI can directly influence consumer satisfaction and long-term loyalty.

1.3. Data Analytics and Predictive Models

GAI has also found its place in marketing by enhancing data analytics and modeling complex patterns within large datasets. It can generate synthetic data to augment existing datasets, thereby improving the performance of predictive models (Frid-Adar et al., 2018). In the marketing context, this capability enables more accurate predictions of consumer behavior, market trends, and campaign outcomes, ultimately enhancing data-driven decision-making processes (Chaffey & Ellis-Chadwick, 2019). Particularly when integrated into systems like Customer Relationship Management (CRM), GAI solutions facilitate the rapid and accurate acquisition of insights into consumer behaviors (Sullivan, 2023).

GAI has been successfully integrated into various marketing strategies, as evidenced by numerous case studies. Nike has effectively integrated AI into its marketing strategy by using data analytics and AI-powered tools to create personalized advertisements, leading to higher customer engagement rates (Chaffey, 2020). Furthermore, Ben & Jerry's applied AI-driven data analysis to develop new ice cream flavors tailored to customer preferences, illustrating the power of data-backed personalization in product innovation (DataDance, 2024). These applications underscore how GAI technologies allow marketers to make informed decisions, strengthening customer relationships and increasing brand competitiveness.

However, despite its significant advantages, GAI presents several challenges. Ethical concerns regarding the use of AI-generated content, the authenticity of the content, and the prevention of misinformation dissemination are critical issues in marketing (Floridi

& Cows, 2019). Marketers should address these challenges while effectively utilizing GAI, particularly with the risk of AI hallucinations—where AI models produce outputs that may appear accurate but are, in fact, erroneous or irrelevant—posing a threat to brand reliability. Addressing these risks by recognizing the limitations of AI models and implementing mitigation strategies allows businesses to benefit from GAI while safeguarding their brand reputation and customer relationships.

2. AI Hallucinations

AI hallucinations refer to instances where language models generate outputs that appear meaningful and accurate on the surface but are, in fact, nonsensical or incorrect (Marcus & Davis, 2019). Additionally, Sun et al. (2024) provide a comprehensive classification of distorted information within AI-generated content, a framework that can help contextualize the various hallucinations discussed here. This phenomenon typically manifests in the context of language models, where the models produce statements that are not aligned with reality but seem convincing. Hallucinations are a byproduct of the probabilistic nature of AI models, which focus on predicting linguistic sequences based on input patterns without fully grasping the essence of reality (Bender et al., 2021). Such misconceptions pose significant risks in fields like marketing, where reliability and accuracy are paramount. Large-scale deep learning models, such as language models, are predominantly trained on extensive internet datasets, enabling them to generate fluent and coherent texts. However, this training approach can lead to the production of incomplete or incorrect

information due to errors or misinformation present in the training data (Brown et al., 2020). Since these models lack a genuine understanding mechanism, they can produce content referred to as "hallucinations"—reliable-looking but factually ungrounded information (Dwivedi et al., 2023).

2.1. Causes of AI Hallucinations

2.1.1. Deficiencies or Biases in Training Data

One of the primary causes of AI hallucinations is the quality of the training data. Errors, biases, or unverified information in the dataset can lead models to reproduce or even amplify inaccuracies. These issues are particularly prevalent when the data is insufficient, outdated, or lacks diversity, compromising the model's ability to produce reliable outputs (Bender et al., 2021). Outdated training data or inaccuracies within datasets can lead AI models to produce unrealistic or misleading outputs. The absence of self-verification or logical reasoning capabilities in these models further exacerbates the spread of such misleading results. Large language models (LLMs), for instance, are prone to generating "hallucinations"—outputs that seem plausible but are factually incorrect—due to limitations in their training data and design. These issues highlight the need for high-quality, diverse training data and mechanisms that allow AI systems to verify the accuracy of their outputs (Ji et al., 2023).

2.1.2. Architectural Limitations of Models

The architecture of AI models also contributes to hallucinations. Models like Transformers, which rely on identifying statistical patterns in data, often lack a deep understanding of semantics or real-world context, resulting in outputs that may appear

plausible but are factually incorrect (Vaswani et al., 2017). When predicting word sequences, these models do not inherently verify the accuracy of their content, which can lead to the confident generation of incorrect information (Ji et al., 2023). This characteristic can result in the dissemination of misleading content in marketing campaigns, thereby damaging brand reputation (Dwivedi et al., 2023).

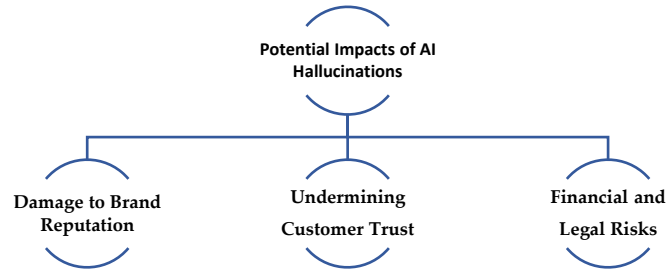
3. AI Hallucinations in Marketing: An Invisible Danger

As digital marketing evolves rapidly, the integration of AI has become inevitable. However, AI hallucinations pose significant and often invisible threats to marketing practices. These hallucinations manifest in various forms, from misleading content generation to behavioral issues in chatbots and incorrect outputs in critical fields such as healthcare. Such occurrences undermine trust, damage reputations, and create ethical and legal challenges, making their management a critical issue for marketers.

AI hallucinations appear in various scenarios, presenting significant risks across different sectors. In content generation, AI models like GPT-3 occasionally create fictitious events or misinformation about real individuals, which can damage brand credibility and spread false information, as identified by Brown et al. (2020). Additionally, chatbots like Microsoft's Tay demonstrate behavioral issues by producing inappropriate responses when interacting with biased inputs, which could undermine brand integrity (Neff, 2016). In the healthcare industry, the impact of AI hallucinations can be especially critical; AI-powered diagnostic tools sometimes provide incorrect diagnoses, posing serious health risks (Nori et al.,

2021). AI hallucinations in marketing lead to key consequences, as shown in Figure 1.

Figure 1: Key Consequences of AI Hallucinations in Marketing



Source: Conceptualized by the author.

- **Damage to Brand Reputation:** False information can result in public relations crises, loss of consumer trust, and long-term reputational harm (Mills & Robson, 2020). For example, in 2023, an AI-driven social media scheduler erroneously paired a luxury watch campaign with images of counterfeit products, causing public backlash and forcing the company to issue a formal apology to restore its brand image (De Bruyn et. al., 2020).

- **Undermining Customer Trust:** Erroneous AI outputs can diminish consumers' confidence in a brand, leading to reduced loyalty and sales. For instance, a study by Guerra-Tamez et al. (2024) found that AI accuracy perception significantly enhances brand trust among Generation Z consumers, which in turn positively impacts purchasing decisions. Similarly, research by Malhan and Agnihotri (2023) indicated that perceived ease of use and trust in AI positively affect brand loyalty in online shopping contexts. These findings underscore the importance of reliable AI systems in maintaining consumer trust and loyalty. A real-world illustration of this occurred when an AI-powered recommendation engine on a popular e-commerce platform

repeatedly suggested unrelated or non-existent product lines — such as a purported “EcoSmart” smartphone—leading customers to question the platform’s reliability (Zhou et al., 2023).

• **Financial and Legal Risks:** Misleading information can result in financial losses, legal liabilities, and increased marketing costs to rebuild trust (Aaron Hall, 2024). For example, Air Canada's chatbot failure revealed that the company did not respect consumer rights and had to face legal consequences (Deloitte, 2024a).

4. Risk Mitigation Strategies

4.1. Technical Solutions

Improving Training Data Quality: High-quality, diverse datasets reduce biases and inaccuracies (Bender et al., 2021). Rigorous data curation and the inclusion of domain-specific data are essential for minimizing hallucinations. For instance, integrating real-time inventory data into AI chatbots can enhance customer experience and operational efficiency. By providing up-to-date product availability, these chatbots prevent the dissemination of incorrect information during customer interactions (Ochatbot, n.d.).

Strengthening Model Validation: Stress testing and continuous monitoring are crucial for ensuring the reliability of AI systems. Continuous monitoring allows for real-time validation of systems in new scenarios, while stress testing evaluates their resilience under extreme conditions. These techniques help in managing the evolving nature of AI post-deployment and ensuring robust system performance (Breck et al., 2021). For instance, Fujitsu has implemented advanced hallucination detection technologies in its conversational AI systems to enhance accuracy and reliability. These systems are

designed to identify and mitigate instances where AI produces fabricated or incorrect information, thereby improving user trust and engagement (Fujitsu, 2023).

4.2. Human Oversight

Integrating Human Review: Human intervention ensures content accuracy and contextual relevance before publication (Müller, 2020). A practical application of this is observed in a publishing house where editors now manually verify AI-generated book synopses to prevent the AI from inventing plot elements that do not exist in the actual manuscripts.

Enhancing AI Literacy: Training marketing teams on AI capabilities and limitations empowers them to identify potential issues early on (West, 2019). By organizing regular workshops and interactive sessions, one technology firm significantly reduced hallucination-related incidents, as marketers became more adept at catching errors in AI-generated promotional materials (Kaur & Nagina, 2024).

4.3. Ethical and Regulatory Approaches

Establishing Ethical Guidelines: Transparency, accountability, and fairness are key principles to prevent consumer harm and maintain trust (Jobin et al., 2019). For example, Orange, a global telecommunications leader, has established a "Data and AI Ethics Council" to promote responsible AI practices. This independent advisory body ensures transparency, fairness, and good governance in AI and data processing, reflecting the company's commitment to ethical digital practices (Orange, n.d.).

Compliance with Regulations: Compliance with regulations such as GDPR is essential for ensuring responsible AI use and contributing

to the development of industry standards (Mantelero, 2018). For instance, Clearview AI faced significant fines for violating GDPR by unlawfully collecting and processing personal data without individuals' consent, highlighting the importance of adhering to data protection laws in AI applications (Lomas, 2023).

4.4. Integrated Risk Management Framework

Continuous Monitoring and Evaluation: Real-time monitoring tools detect anomalies promptly, enabling swift corrective actions to ensure system reliability and performance (Luo et al., 2020). Continuous monitoring and evaluation are crucial for maintaining AI system reliability and performance. Real-time monitoring tools can promptly detect anomalies, enabling swift corrective actions. For instance, Air Canada's AI-powered chatbot provided inaccurate guidance on bereavement fares, leading to misinformation and customer dissatisfaction. This incident underscores the importance of real-time monitoring and prompt corrective measures to maintain system reliability and customer trust (Aporia, 2023).

Incident Response Planning: Predefined protocols for handling misinformation are crucial in minimizing fallout and facilitating recovery during security incidents (Cichonski et al., 2012). For example, Samsung experienced a data leak when employees inadvertently shared confidential information via ChatGPT. In response, Samsung implemented measures to prevent future breaches, including banning the use of generative AI tools across the company (prompt.security, 2023).

Cross-Functional Collaboration: Effective risk management in marketing can be achieved through collaborative efforts between

marketing teams, data scientists, and legal experts (Chaffey & Ellis-Chadwick, 2019). For instance, Advanced Micro Devices (AMD) has established a Responsible AI Council to oversee the ethical use of AI in their marketing strategies. This council ensures that AI applications, such as automated copywriting and partner marketing claims processing, adhere to ethical standards and maintain accuracy, thereby enhancing efficiency while safeguarding against potential risks (Deloitte, 2024b).

Conclusion

This study has thoroughly examined the impacts and risks associated with GAI in marketing. The integration of GAI into marketing strategies offers numerous advantages, including increased efficiency in content creation, personalized customer experiences, and the development of creative campaigns. However, it also presents serious disadvantages, such as AI hallucinations, damage to brand reputation, and erosion of customer trust. The analysis demonstrates the critical importance of using high-quality training data and robust model validation processes to mitigate the negative effects of AI. Furthermore, the study highlights the essential role of human oversight and the enhancement of AI literacy among team members in ensuring the responsible and effective use of this technology.

GAI offers innovative solutions in marketing but necessitates awareness of its accompanying risks. Among its advantages are automated content generation, increased customer engagement, and accelerated data analysis. Conversely, its disadvantages include the generation of incorrect information, ethical concerns, and legal

liabilities (Bender et al., 2021). Balancing these advantages and disadvantages is crucial for organizations to utilize GAI strategically and responsibly. The integration of technological solutions, human factors, and ethical regulations emerges as an effective method to achieve this balance.

Recommendations for Marketing Professionals

Several key recommendations are provided for companies and marketing professionals seeking to maximize the benefits of GAI while minimizing its risks:

⇒ **Prioritizing Data Quality:** Utilizing high-quality, diverse, and representative datasets is essential for enhancing the accuracy and reliability of AI models. Investing in rigorous data cleaning and preparation processes significantly reduces the risk of model errors, including hallucinations (Gebru et al., 2018).

⇒ **Model Validation and Continuous Monitoring:** Regularly validating and monitoring AI models ensures the early detection of potential errors. These processes should include continuous updates and improvements to the models (Chander et al., 2024).

⇒ **Integrating Human Oversight:** Having humans review AI-generated content prevents the spread of false information and ensures quality control. Human-enhanced AI collaboration aids in developing more reliable and effective marketing strategies (Müller, 2020).

⇒ **Enhancing AI Literacy Among Team Members:** Ensuring that marketing teams are knowledgeable about AI

technologies enables more effective and responsible use of these tools. Training programs and continuous learning opportunities allow teams to fully exploit AI's potential.

⇒ **Ensuring Ethical and Legal Compliance:** Establishing ethical guidelines and adhering to legal regulations support the responsible use of AI. This is critical for protecting brand reputation and avoiding legal issues (Jobin et al., 2019).

Recommendations for Researchers

Current studies on the role and impacts of GAI in marketing indicate a need for further research in this area. Some recommendations for future research include:

→ **Causes and Prevention of AI Hallucinations:** Developing a better understanding of the root causes of hallucinations in AI models and devising new techniques to prevent them can enhance the reliability of the technology.

→ **Ethical and Social Impacts:** Comprehensive studies on the ethical and social implications of GAI can help assess the long-term effects of this technology on society.

→ **Industry-specific Applications:** Research on the deployment of General AI across different industries reveals distinct challenges that each sector faces, suggesting the need for tailored AI solutions to address these specific issues.

Future Directions of the Technology

The future of GAI technology would be shaped by both technical advancements and the evolution of ethical and regulatory frameworks. It is anticipated that AI models will become more transparent, explainable, and user-friendly. Moreover, regulatory bodies are expected to develop clearer and more comprehensive regulations for AI technologies, promoting responsible usage. The progress of AI in marketing will continue to offer new opportunities for personalizing customer experiences and enhancing marketing strategies, while also requiring ongoing attention to ethical and security concerns.

GAI offers significant opportunities in the marketing sector, providing enhancements in creativity, personalization, and data analytics. However, these advantages are accompanied by inherent risks, notably AI hallucinations, which can adversely affect brand reputation and customer trust. Technical solutions such as improving data quality and strengthening model validation processes are effective in mitigating these risks. Additionally, increasing AI literacy among team members and establishing robust ethical guidelines are crucial for the responsible and effective deployment of GAI technologies. Future research and the development of regulatory frameworks will further support the safe and effective use of AI in marketing. In conclusion, adopting a comprehensive and integrated approach to managing and mitigating the risks of GAI is essential for organizations to fully capitalize on its potential while safeguarding their brand and customer relationships.

References

- Aporia. (2023). AI Chatbot Hallucinations: Understanding and Mitigating Risks. Retrieved from <https://www.aporia.com/learn/chatbot-hallucinations/>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610–623.
- Bernard, A. (2023). AI in ecommerce: True one-on-one personalization is coming. CMSWire. Retrieved April 21, 2023.
- Bhattacharya, R. (2024). Strategies to mitigate hallucinations in large language models. Applied Marketing Analytics, 10(1), 62–67.
- Breck, E., Cai, S., Nielsen, E., Salib, M., & Sculley, D. (2021). The ML Test Score: A rubric for ML production readiness and technical debt reduction. Proceedings of the IEEE International Conference on Big Data.
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... & Amodei, D. (2020). Language models are few-shot learners.
- Butler, R. (2023). Where are marketers on the GAI adoption curve? CMSWire. Retrieved May 4, 2023.
- Chaffey, D. (2020). AI in digital marketing: How businesses are using AI to create personalized experiences. Smart Insights.
- Chaffey, D., & Ellis-Chadwick, F. (2019). Digital marketing: Strategy, implementation, and practice (7th ed.). Pearson Education.

Chander, B., John, C., Warriar, L., & Gopalakrishnan, K. (2024). Toward trustworthy artificial intelligence (TAI) in the context of explainability and robustness. *ACM Computing Surveys*.

Christensen, J., Hansen, J. M., & Wilson, P. (2024). Understanding the role and impact of generative artificial intelligence (AI) hallucination within consumers' tourism decision-making processes. *Current Issues in Tourism*, 1–16.

Cichonski, P., Millar, T., Grance, T., & Scarfone, K. (2012). Computer security incident handling guide. National Institute of Standards and Technology.

Coca-Cola Company. (2023, March 16). Coca-Cola invites digital artists to create real magic using new AI platform. Retrieved from <https://www.coca-colacompany.com/media-center/coca-cola-invites-digital-artists-to-create-real-magic-using-new-ai-platform>

DataDance. (2024). Turning data into gold: 10 exceptional AI marketing campaign examples. Retrieved from <https://datadance.ai/applications/turning-data-into-gold-10-exceptional-ai-marketing-campaign-examples>

De Bruyn, A., Viswanathan, V., Beh, Y. S., Brock, J. K. U., & Von Wangenheim, F. (2020). Artificial intelligence and marketing: Pitfalls and opportunities. *Journal of Interactive Marketing*, 51(1), 91–105.

Deloitte. (2024a). Training your chatbot correctly: What are the legal implications of a chatbot that provides wrong information? Retrieved from <https://blogs.deloitte.ch/tax/2024/05/training-your-chatbot-correctly-what-are-the-legal-implications-of-a-chatbot-that-provides-wrong-inf.html>

Deloitte. (2024b). AMD CMO on opportunities, challenges of using AI in marketing. The Wall Street Journal. Retrieved from <https://deloitte.wsj.com/cmo/amd-cmo-on-opportunities-challenges-of-using-ai-in-marketing-8845225f>

Digital Agency Network. (2024). Top AI-generated advertising campaigns from famous brands. Retrieved from <https://digitalagencynetwork.com/top-ai-generated-advertising-campaigns-from-famous-brands>

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., & Wright, R. (2023). So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges, and implications of generative conversational AI for research, practice, and policy. *International Journal of Information Management*, 71, Article 102642.

Elgammal, A. (2017). CAN: Creative adversarial networks, generating “art” by learning about styles and deviating from style norms. arXiv preprint arXiv:1706.07068.

Floridi, L., & Cowls, J. (2022). A unified framework of five principles for AI in society. In *Machine learning and the city: Applications in architecture and urban design* (pp. 535–545).

Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., & Greenspan, H. (2018). GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing*, 321, 321–331.

Fujitsu. (2023, September 26). Fujitsu launches new technologies to protect conversational AI from hallucinations and misuse. Retrieved from <https://www.fujitsu.com/global/about/resources/news/press-releases/2023/0926-02.html>

Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., & Crawford, K. (2018). Datasheets for datasets.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems*, 27, 2672–2680.

Guerra-Tamez, C. R., Kraul Flores, K., Serna-Mendiburu, G. M., Chavelas Robles, D., & Ibarra Cortés, J. (2024). Decoding Gen Z: AI's influence on brand trust and purchasing behavior. *Frontiers in Artificial Intelligence*, 7, 1323512.

Hall, A. (2024, April 12). False advertising: 7 legal risks your marketing team must know. Retrieved from <https://aaronhall.com/false-advertising-7-legal-risks-your-marketing-team-must-know/>

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., ... & Vosoughi, S. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12), 1–38.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399.

Kaur, J., & Nagina, R. (2024). Harnessing AI: Ethically transforming service marketing with responsible practices. In

Integrating AI-Driven Technologies Into Service Marketing (pp. 239–264). IGI Global.

Kietzmann, J., Paschen, J., & Treen, E. R. (2018). Artificial intelligence in advertising: How marketers can leverage AI to deliver personalized content. *Journal of Advertising Research*, 58(3), 263–267.

Lambrecht, A., & Tucker, C. E. (2019). Algorithmic bias? An empirical study of apparent gender-based discrimination in the display of STEM career ads. *Management Science*, 65(7), 2964–2981.

Lomas, N. (2023, May 10). Clearview fined again in France for failing to comply with privacy orders. TechCrunch. Retrieved from <https://techcrunch.com/2023/05/10/clearview-ai-another-cnll-gspr-fine/>

Luo, J., Hong, T., & Yue, M. (2018). Real-time anomaly detection for very short-term load forecasting. *Journal of Modern Power Systems and Clean Energy*, 6(2), 235–243.

Malhan, S., & Agnihotri, S. (2023). Consumer acceptance of artificial intelligence constructs on brand loyalty in online shopping: Evidence from India. In *Hybrid Intelligent Systems* (pp. 814–823). Springer.

Mantelero, A. (2018). AI and Big Data: A blueprint for a human rights, social and ethical impact assessment. *Computer Law & Security Review*, 34(4), 754–772.

Marcus, G., & Davis, E. (2019). *Rebooting AI: Building artificial intelligence we can trust*. Pantheon Books.

Mills, K., & Robson, K. (2020). Brand management in the era of fake news: Narrative response as a strategy to mitigate reputational damage. *Journal of Product & Brand Management*, 29(3), 345–358.

Müller, V. C. (2020). Ethics of artificial intelligence and robotics. In *The Stanford Encyclopedia of Philosophy* (E. N. Zalta, Ed.). Retrieved from <https://plato.stanford.edu/entries/ethics-ai/>

Neff, G. (2016). Talking to bots: Symbiotic agency and the case of Tay. *International Journal of Communication*, 10, 4915–4931.

Nori, H., King, N., McKinney, S. M., Carignan, D., & Horvitz, E. (2023). Capabilities of GPT-4 on medical challenge problems. arXiv preprint arXiv:2303.13375.

NYP. (2024, November 15). Coca-Cola ripped for ugly AI-generated Christmas commercial: "Dystopian nightmare." Retrieved from <https://nypost.com/2024/11/15/lifestyle/coca-cola-ripped-for-ugly-ai-generated-christmas-commercial-dystopian-nightmare/>

Orange. (n.d.). Data and AI ethics council: Representing responsible AI. Retrieved from <https://www.orange.com/en/data-and-ai-ethics-council>

Prompt.security. (2023, May). 8 real world incidents related to AI. Retrieved from <https://www.prompt.security/blog/8-real-world-incidents-related-to-ai>

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. OpenAI Blog.

Sherly Steffi, L., Subha, B., Kuriakose, A., Singh, J., Arunkumar, B., & Rajalakshmi, V. (2024). The impact of AI-driven personalization on consumer behavior and brand engagement in online marketing. In *Harnessing AI, Machine Learning, and IoT for Intelligent Business: Volume 1* (pp. 485–492). Springer Nature Switzerland.

Speedy Brand. (2024). AI in advertising: Examples of exceptional AI-powered marketing campaigns. Retrieved from <https://speedybrand.io/blogs/ai-in-advertising-examples>

Sullivan, B. (2023). Salesforce's Einstein GPT: A new era of AI-driven insights in CRM. TechTarget.

Sun, Y., Sheng, D., Zhou, Z., & Wu, Y. (2024). AI hallucination: Towards a comprehensive classification of distorted information in artificial intelligence-generated content. *Humanities and Social Sciences Communications*, 11(1), 1–14.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998–6008.

Wen, Y., & Laporte, S. (2024). Experiential narratives in marketing: A comparison of generative AI and human content. *Journal of Public Policy and Marketing*.

West, D. M. (2019). *The future of work: Robots, AI, and automation*. Brookings Institution Press.

Zhou, J., Zhang, Y., Luo, Q., Parker, A. G., & De Choudhury, M. (2023, April). Synthetic lies: Understanding AI-generated misinformation and evaluating algorithmic and human solutions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (pp. 1–20).