Journal of Media, Journalism & Communication Studies

Vol. 1 No. 1 2025

Article Open Access

The Role of Generative AI in the Evolution of Digital Advertising Products

Xiao Liu [1,*]

[1] Meta Monetization, Meta, Bellevue, Washington, 98004, USA

* Correspondence: Xiao Liu, Meta Monetization, Meta, Bellevue, Washington, 98004, USA

Received: 31 May 2025
Revised: 07 June 2025
Accepted: 23 June 2025
Published: 25 June 2025

Abstract: With the widespread adoption of Generative Artificial Intelligence (GenAI) across various industries, digital advertising has emerged as a key sector demonstrating the urgency and potential of GenAI integration. Recent reports indicate that over 90% of advertisers have begun incorporating AI tools into their campaign targeting strategies. Currently, a significant portion of digital advertising content is generated using GenAI technologies. In the context of increasingly diverse and abundant data channels — including video, image, and text — GenAI is playing a pivotal role in aggregating and analyzing data to enable more precise audience targeting. Some studies suggest that GenAI-driven campaigns can achieve average cost savings of up to 20%. This paper provides a comprehensive overview of GenAI's innovative applications in digital advertising, with a particular focus on structured prompt design, multimodal content generation, behavior-driven feedback optimization, and intelligent content auditing. It explores the implementation of these technologies in automated ad content creation, personalized content delivery, and cross-lingual adaptation. Through this review, it becomes evident that GenAI is reshaping the digital advertising landscape — transforming content production workflows, optimizing targeting strategies, enhancing user engagement mechanisms, and significantly boosting industry-wide efficiency and economic value.

Keywords: generative AI; digital advertising; multimodal generation; prompt project; personalized advertising

1. Introduction

Generative AI has emerged as a fundamental driving force in the development of digital advertising technologies. In response to the limitations of traditional advertising methods — characterized by low efficiency, content homogeneity, and limited precision in audience targeting — generative AI offers transformative solutions. Leveraging advancements in natural language processing and multimodal generation, this technology enables the automated production of diverse content formats, including text, images, and videos, thereby addressing a wide range of advertising needs with greater speed and scalability.

The paradigm of advertising content creation has shifted from a "manual creation + static dissemination" model to a "model-driven generation + intelligent distribution" framework. This evolution highlights the significant advantages of generative AI in enhancing content creativity, improving targeting accuracy, and optimizing customer interaction processes. This paper conducts an in-depth analysis of the technical underpinnings and practical implications of generative AI in advertising, and further explores its strategic role and future potential within intelligent marketing ecosystems.

2. Overview of Generative AI Technology

As previously discussed, one of the most significant applications of Generative AI (GenAI) in the digital advertising industry is automated content creation. Powered by deep learning algorithms, GenAI models can produce a wide range of media types — including text, images, audio, and video — in response to user-defined prompts. These models are trained on large-scale datasets consisting of millions of samples, allowing them to replicate semantic, stylistic, and contextual patterns to generate high-quality, context-specific content. By leveraging this training data, GenAI can be fine-tuned to meet specific advertiser requirements, enabling personalized and brand-aligned content generation. Moreover, GenAI-generated content can adapt to interactive scenarios, such as real-time responses based on user behavior and engagement metrics, thus enhancing audience interaction and campaign effectiveness. This marks a departure from traditional advertising workflows, where static content was manually created and often lacked real-time adaptability. GenAI significantly reduces production time and costs, transforming processes that previously required days of effort into ones achievable within minutes. This automation not only increases operational efficiency but also fosters richer and more diverse advertising content, thereby accelerating innovation and contributing to the overall growth of the digital advertising industry [1].

3. Key Technologies of Generative AI in the Evolution of Digital Advertising Products

3.1. Structured Prompt Control Technology

In the era of digital advertising shaped by generative AI, prompts serve not merely as command inputs to the model but as essential control mechanisms for guiding content output. Structured prompt engineering has emerged as a critical technique, enabling advertisers to define template-based and parameterized input specifications that direct the content generation process with greater precision. This approach addresses key challenges in traditional prompt usage, such as ambiguity, inconsistent output, and deviations in style or tone [2]. At its core, a structured prompt involves the insertion of variable control elements within a fixed semantic framework. This design allows advertisers and marketers to generate content that adheres to predefined stylistic, contextual, and functional parameters, thereby enhancing consistency and alignment with brand identity. Structured prompts facilitate scalability, reproducibility, and customization in AI-generated advertising content, ultimately improving the efficiency and effectiveness of creative workflows.

$$P(x) = t_0 + \sum_{i=1}^{n} t_i\,(v_i) \quad (1)$$

In this framework, t_0 denotes the task-level instruction (e.g., "generate advertising copy"), t_i represents the i-th attribute prompt template, and v_i corresponds to the actual content variable assigned to each attribute. A representative example of a structured prompt in advertising can be expressed as:

P(x) = "Generate an advertisement about "+ (product name) +", the target audience is "+ (user group) +", and the style is "+(tone).

To enhance both the credibility and diversity of the generated content, weighting mechanisms can be introduced for each control element within the prompt structure. By assigning differential weights to specific attributes — such as product relevance, audience targeting, or stylistic tone — advertisers can fine-tune the output to prioritize key campaign goals or brand messaging constraints. This structured and weighted prompting approach significantly improves the alignment of generative outputs with marketing objectives.

$$P_w(x) = t_0 + \sum_{i=1}^{n} w_i \cdot t_i(v_i) \quad (2)$$

Among them, $w \in [0,1]$ represents the proportion of the impact of each parameter on the final output, from which the generated content focus can be adjusted according to the advertising strategy.

In practical implementation, reinforcement learning mechanisms can be employed to dynamically adjust the weights of prompt parameters, optimizing content generation outcomes based on utility functions — such as click-through rate (CTR) or conversion rate (CVR). This optimization objective can be formally expressed as:

$$\max_{w} \quad \mathbb{E}[R(G(P_w(x)))] \tag{3}$$

where R denotes the advertising effectiveness evaluation function, *G* represents the output function of the generative content model, and *Pw(x)* is the weighted prompt function with parameter vector w.

The structured prompt regulation mechanism ensures consistency and accuracy in advertising copy style, while supporting multivariate testing to generate diverse content combinations. This significantly enhances the flexibility and personalization capabilities of ad delivery. As template libraries, tagging systems, and automated prompt recommendation tools continue to evolve, the integration of reinforcement learning and structured prompting has become a key driver of generative advertising platforms. This convergence is steering the industry toward a more controlled, predictable, and evaluative approach to content creation, marking a fundamental shift in the production and deployment of digital advertising content [3].

3.2. Multi Modal Fusion Generation Mechanism

With the rapid advancement of digital advertising, single-modality advertising content is increasingly insufficient to satisfy the growing perceptual and engagement demands of audiences across diverse media types — such as text, imagery, audio, and video. In this context, the multimodal integrative creative paradigm enabled by generative AI represents a transformative technological breakthrough. This approach eliminates the traditional silos between various content modalities, allowing for the synchronized and cohesive generation of rich, multisensory advertising experiences [4].

At the core of multimodal generative systems lies the principle of projecting inputs from heterogeneous modalities into a shared semantic latent space. This shared space enables the model to understand and generate coherent cross-modal outputs, where, for example, text can be used to generate corresponding images, or video narratives can be accompanied by stylistically aligned music and captions. The foundational modeling approach can be generally represented as:

$$\mathcal{Z} = f_T(x_T) = f_I(x_I) = f_A(x_A) \tag{4}$$

In this formulation, x_T, x_I and x_A represent text, image, and audio inputs, respectively. The functions f_T, f_I, and f_A are the corresponding modality-specific encoders. By projecting all input modalities into a shared semantic latent space z, generative models are able to achieve unified representation, enabling collaborative expression and understanding across multiple modalities. In the context of advertising content creation, a widely adopted implementation involves using text as the primary driver, where a model generates visual imagery conditioned on advertising copy. This process can be modeled as:

$$y_I = G_I(f_T(x_T)) \tag{5}$$

Here, x_T denotes the input textual content (e.g., an ad copy), f_T is the text encoder, and G_I is the image generator. By employing advanced image generation architectures such as Diffusion Models or Generative Adversarial Networks (GANs), it becomes feasible to generate coherent visual outputs that accurately reflect the semantic intent of the textual input. This enables every visual element within an advertisement to consistently convey the same promotional message [5].

To ensure semantic alignment between modalities and enhance the quality of multimodal outputs, recent models optimize the following objective function:

$$\mathcal{L} = \lambda_1 \cdot \mathcal{L}_{\text{text}} + \lambda_2 \cdot \mathcal{L}_{\text{visual}} + \lambda_3 \cdot \mathcal{L}_{\text{align}} \tag{6}$$

In this equation, $\mathcal{L}_{\text{text}}$ and $\mathcal{L}_{\text{visual}}$ are the loss functions for the text and image generation tasks, respectively, while $\mathcal{L}_{\text{align}}$ quantifies the semantic consistency between the textual and visual outputs. The parameters λ_1 , λ_2 and λ_3 are weights that balance each component of the loss. This optimization framework effectively mitigates issues such as mismatched styles or incoherence between generated text and imagery, thereby improving suitability for high-end brand campaigns and short-form video content.

With the continued evolution of cross-modal pre-trained models such as DALL·E, Imagen, and Runway Gen-2, the semantic fidelity and compositional integration of multimodal generation have significantly improved. On modern advertising platforms, users can now produce complete visual, audio, or video content — including interactive virtual interfaces — based solely on brief textual descriptions. This automation substantially reduces creative design costs, accelerates production timelines, and enriches the diversity and depth of advertising content.

Multimodal fusion technology not only facilitates the transformation of advertising content from static presentation to an immersive, experiential display mode, but also enables unified modeling, rapid iteration, and semantic alignment across content types. These capabilities provide comprehensive solutions for the generation and exchange of advertising information, making multimodal fusion a critical technological foundation for the development of intelligent advertising platforms.

3.3. RLHF Behavior Feedback Optimization Model

Reinforcement Learning from Human Feedback (RLHF) enhances the quality and effectiveness of advertising content by incorporating user behavior as an external reward signal. This approach enables the generative process to adapt dynamically to real-world publishing feedback, thereby optimizing both content quality and promotional outcomes. The core idea behind RLHF is to first train a foundational natural language processing (NLP) model and subsequently refine its output strategy using reinforcement learning techniques.

In the context of advertising, RLHF can be formalized by defining a reward function that maps user interaction data to content quality metrics. Specifically, for a given piece of generated advertising content y, the reward function can be defined as:

$$R(y) = \alpha \cdot CTR(y) + \beta \cdot CVR(y) - \gamma \cdot Bounce(y) \tag{7}$$

where $CTR(y)$ denotes the click-through rate, $CVR(y)$ the conversion rate, and $Bounce(y)$ the bounce rate associated with the content y. The coefficients α, β, and γ are weighting parameters that balance the importance of each user behavior metric. This evaluation system comprehensively considers three dimensions: user engagement, conversion effectiveness, and bounce rate.

The Policy Improvement Phase employs Proximal Policy Optimization (PPO) algorithms to iteratively refine the decision-making process of the generative model, with the primary objective of maximizing the expected reward:

$$\max_{\theta} \quad \mathbb{E}_{y \sim G_\theta(x)}[R(y)] \tag{8}$$

where x denotes the input, and $G_\theta(x)$ represents the advertising content generated by the current model parameterized by θ. The system dynamically updates θ by analyzing historical feedback data, thereby increasing the likelihood of producing high-quality content.

During the generation phase, distribution reconstruction and sampling optimization are conducted through behavior-feedback-driven sampling. The objective function incorporates Kullback–Leibler (KL) divergence to regulate policy updates and maintain stability:

$$\mathcal{L}_{RLHF} = -R(y) + \lambda \cdot D_{KL}(G_\theta \parallel G_{pre}) \tag{9}$$

where G_{pre} is a pre-trained baseline model, D_{KL} measures the divergence between the current and previous policies, and λ is a regularization coefficient preventing overfitting to feedback and preserving generalization.

By adopting RLHF, the system achieves a continuous "learning and optimization" loop in advertising content creation. Leveraging consumer behavior feedback as a reward signal, the iterative process aligns generated advertisements with user preferences and target consumer goals, facilitating semantic guidance and personalized content matching. In practical applications, RLHF has been integrated into automated ad delivery systems, creative A/B testing tools, and personalized creative generation engines. These integrations not only enhance the commercial value of advertising but also reduce manual trial-and-error costs and accelerate intelligent content deployment.

With ongoing advancements in online learning algorithms and feedback mechanisms, reinforcement learning-based feedback optimization (RLHF) is emerging as a pivotal technical solution driving the continuous iteration and enhancement of advertising creativity.

3.4. Intelligent Content Proofreading System

Leveraging natural language processing and image recognition technologies, an automated and scalable text inspection model — an intelligent content proofreading system — can be developed to ensure the legality and brand safety of advertising content. The system performs real-time evaluation of generated content y and outputs a binary decision $s \in \{0,1\}$, where s = 1 denotes content passing inspection, and s = 0 indicates the presence of violations or potentially objectionable material.

The text compliance assessment module constructs a content evaluation function $C_{text}(y_T)$ based on semantic classification models such as BERT, where y_T represents the generated textual content. The fundamental decision function is defined as:

$$s_T = C_{text}(y_T) = \sigma(W \cdot h_T + b) \quad (10)$$

Here, h_T denotes the semantic vector representation of the text, W is the weight matrix, and σ is the sigmoid activation function. The model is trained to identify risk factors in advertising copy by learning from labeled data encompassing abusive, false, and sensitive content.

The image and video content detection module filters visually generated advertising materials using image recognition models such as ResNet or Vision Transformer, combined with object detection and content review libraries. Its decision function is formulated as:

$$s_I = C_{image}(y_I) = Detect(Feature(y_I)) \in \{0,1\} \quad (11)$$

where y_I is the generated image or video frame, $Feature(y_I)$ denotes the visual feature extraction process, and $Detect()$ performs content verification tasks, including the identification of prohibited information and improper use of trademarks. The integration of frame-by-frame analysis with semantic understanding substantially enhances the accuracy of video content screening.

To provide a comprehensive safety assessment, multiple dimensions — including text, visual content, semantic consistency, and brand compliance — are integrated into a unified scoring function. A linear weighted model is employed as follows:

$$S_{total} = \lambda_1 \cdot s_T + \lambda_2 \cdot s_I + \lambda_3 \cdot s_R \quad (12)$$

where s_R quantifies the degree of compliance with brand guidelines, including brand attributes, terminology, and contextual relevance. The weights $\lambda_i \in [0,1]$ are adjustable according to specific operational requirements. Content is deemed compliant if $S_{total} \geq \theta$, where θ is a predefined threshold.

This architectural design enables the intelligent content screening system to establish a robust risk control loop for advertising content, effectively mitigating potential issues such as the dissemination of illegal or inappropriate materials.

4. Analysis of the Role of Generative AI in the Evolution of Digital Advertising Products

4.1. Refactoring the Paradigm of Advertising Content Production

In traditional digital advertising platforms, content production predominantly relies on manual division of labor, encompassing tasks such as copywriting, image design, and material organization. This approach often results in lengthy production cycles, low operational efficiency, and limited content scalability. By leveraging generative AI technologies — anchored in pre-trained language models, multimodal generative frameworks, and semantic control techniques — large volumes of advertising copy and imagery can be generated rapidly. Moreover, these outputs can be tailored to diverse platforms and individualized preferences, thereby significantly enhancing both the speed and diversity of creative content production (Table 1 and Figure 1).

Table 1. Comparison of Content Production Efficiency of a Brand Advertising Team before and after Adopting Generative AI.

Project	Traditional way	After introducing generative AI	Increase amplitude
Average copy writing output speed (articles/hour)	5	120	+2300%
Number of multi-platform material adaptation versions (types)	3	20	+566%
Average A/B testing cycle (days)	20	7	-65%
Proportion of manual intervention	100%	<30%	-70%

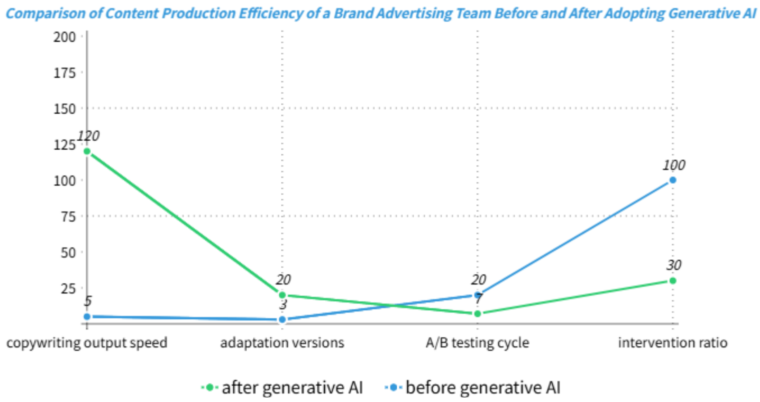


Figure 1. Comparison of Content Production Efficiency of a Brand Advertising Team before and after Adopting Generative AI.

The mentioned data demonstrate that the integration of generative AI not only enhances the scale and velocity of advertising content creation but also transforms the organizational structure of the creative process, propelling the advertising industry into an era of automated innovation.

4.2. Implementing Semantic Driven Real-Time Personalized Advertising

Traditional advertising primarily relies on broad demographic characteristics and static content libraries for information matching and content dissemination. This approach often falls short in accurately capturing consumer intent and adapting to real-time linguistic and contextual shifts. In contrast, under a semantic-driven paradigm, the system employs intent recognition technologies to construct semantic representations of user behavior. Utilizing prompt-based content generation tools, it can automatically produce advertising materials and complete the process of "dynamic matching → content creation → advertising presentation" prior to user interaction, thereby significantly enhancing personalization and responsiveness (Table 2 and Figure 2).

Table 2. Comparison of Changes in Key Indicators before and after Introducing Semantics-Driven Personalized Advertising on a Specific Advertising Platform.

Indicator project	Before introduction	After introduction	Increase amplitude
Ad click through rate (%)	23	48	+108.7%
User dwell time (seconds)	18.6	32.1	+72.5%
Content matching accuracy (%)	61	87	+42.6%
Average conversion rate (%)	6.5	14.3	+120%

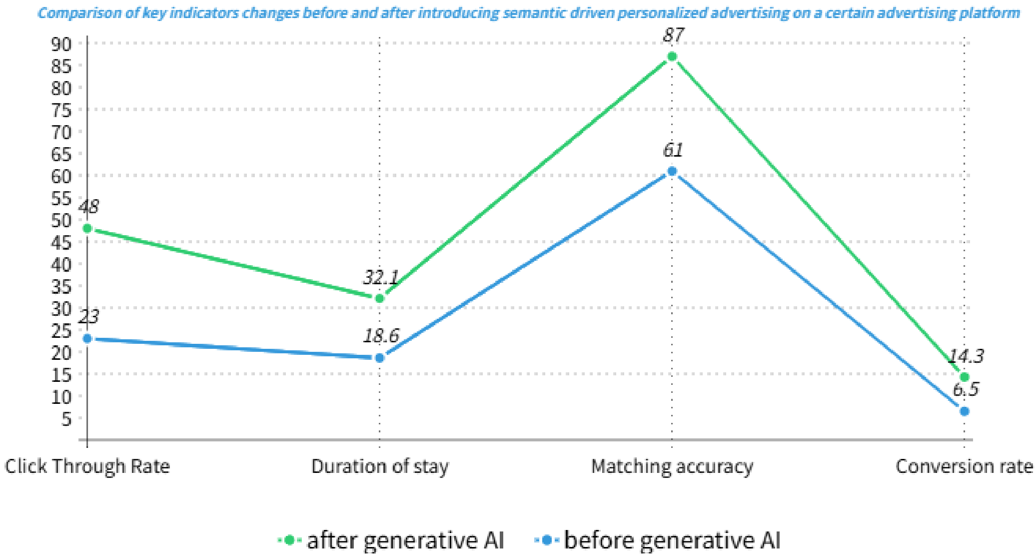


Figure 2. Comparison of Key Indicators Changes before and after Introducing Semantic Driven Personalized Advertising on a Certain Advertising Platform.

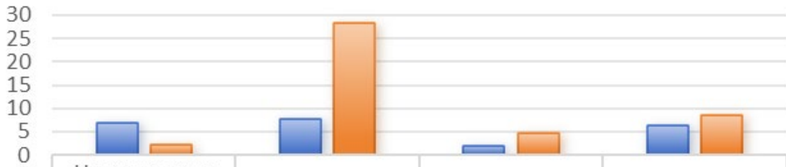
The introduction of generative AI technology enables advertising placement to better meet the current needs of consumers, significantly improving promotion rates and matching accuracy, laying a solid foundation for the development of an intelligent advertising placement model based on semantics-driven and feedback-based approaches.

4.3. Promote the Emergence of New Forms of Intelligent Interactive Advertising

As digital advertising shifts from "static display" to "intelligent interaction", artificial intelligence technology enables it to understand language and mimic the way people talk to each other, communicate with customers, and provide emotional feedback. New forms of advertising, such as voice chat ads, question and answer product introductions, and imitation of virtual characters, can dynamically adjust the path based on customers' real-time actions and behaviors, achieve personalized presentation of information, and greatly enhance customers' sense of immersion (Table 3 and Figure 3).

Table 3. Key Performance Indicators before and after Deploying Generative AI Interactive Advertising on a Certain Platform.

Indicator project	Traditional advertising	AI interactive advertising	Relative change
User response time (seconds)	6.8	2.1	↓69.1%
Dialogue trigger rate (%)	7.6	28.4	↑273.7%
Reply frequency	1.9	4.6	↑142.1%
User satisfaction (10-point scale)	6.3	8.5	↑34.9%



	User response time (seconds)	Dialogue trigger rate (%)	Reply frequency	User satisfaction
traditional advertising	6.8	7.6	1.9	6.3
AI interactive advertising	2.1	28.4	4.6	8.5

Figure 3. Key Performance Indicators before and after Deploying Generative AI Interactive Advertising on a Certain Platform.

By building an interactive, responsive, and guided advertising interaction system, generative AI enables advertising to finally gain the initiative to "understand consumers, guide consumer consumption, and meet consumer requirements", leading the advertising industry to a comprehensive upgrade from media display to intelligent communication.

5. Conclusion

Generative AI technology is fundamentally reshaping the production and dissemination paradigms within digital advertising. By leveraging innovative approaches such as structured prompts, cross-modal content synthesis, user behavior-driven optimization, and intelligent auditing, the industry has transitioned from manual to automated production processes — substantially enhancing both content generation efficiency and the precision of targeted advertising. Empowered by semantic parsing and interactive technologies, advertising formats are evolving beyond static templates toward greater customization, contextualization, and intelligence. This study explores the core technological roadmap and practical effectiveness of generative AI, demonstrating its impact across key dimensions such as advertising content creation and user experience. The findings underscore the transformative role of intelligent generation technologies, heralding a new era of advertising systems powered by natural language processing.

References

1. Z. Tian and C. Wang, "Rethinking Language Education in the Age of Generative AI," 2025. ISBN: 9781003426929.
2. X. Zhou, D. Teng, and H. Al-Samarraie, "The Mediating Role of Generative AI Self-Regulation on Students' Critical Thinking and Problem-Solving," *Educ. Sci.*, vol. 14, no. 12, p. 1302, 2024, doi: 10.3390/educsci14121302.
3. P. Gujar and S. Panyam, "Generative AI in Digital Advertising Campaigns," *Int. J. Comput. Trends Technol.*, vol. 72, no. 5, pp. 51–55, 2024, doi: 10.14445/22312803/IJCTT-V72I5P106.
4. G. L. Liu, R. Darvin, C. Ma, et al., "Unpacking the role of motivation and enjoyment in AI-mediated informal digital learning of English (AI-IDLE): A mixed-method investigation in the Chinese context," *Comput. Hum. Behav.*, vol. 160, p. 108362, 2024, doi: 10.1016/j.chb.2024.108362.
5. J. M. Higgs and A. Stornaiuolo, "Being human in the age of generative AI: Young people's ethical concerns about writing and living with machines," *Read. Res. Q.*, vol. 59, no. 4, pp. 632–650, 2024, doi: 10.1002/rrq.552.