The Transformation of Ad Engines

LLMs, Foundation Models, and Agentic Systems

Anjan Goswami

General Manager, SmartInfer.com

December 2025

1. Digital advertising enters a decisive transformation

The global digital advertising market is projected to exceed $870 billion by 2027, with AI-powered systems increasingly controlling every component of the advertising stack [1]. Large language models, multimodal foundation models, and autonomous agents are fundamentally reshaping how ads are created, targeted, and optimized—representing the most significant transformation since the advent of programmatic buying.

Meta's Generative Ads Model (GEM) achieved 5% higher conversions on Instagram and 3% on Facebook Feed in Q2 2025, with architectural improvements doubling efficiency in Q3 [2]. Google's AI Overviews now reach 1.5 billion users monthly across 200+ countries, with ads integrated directly into AI-generated summaries [3]. Meanwhile, 58% of consumers have replaced traditional search with generative AI tools, fundamentally altering discovery patterns [4].

The advertising infrastructure landscape is bifurcating. Classical ad stacks—built on keyword matching, gradient-boosted decision trees, and heuristic bidding rules—are giving way to AI-native architectures featuring semantic embeddings, foundation model rankers, and reinforcement learning agents. Meta plans to fully automate its advertising system by end of 2026, with AI generating ad creatives and determining target audiences based solely on a product image and budget [5].

Consumer attitudes present a countervailing force. Only one-third of consumers view AI in advertising positively, with 52% concerned about undisclosed AI-generated content [6]. Gartner forecasts that 50% of consumers will significantly limit social media usage due to platform quality perceptions [7]. The paradox: AI enables unprecedented advertising efficiency, but consumers are simultaneously becoming more resistant to being advertised to.

2. Classical ad architectures give way to AI-native systems

Traditional advertising platforms operate through a well-defined pipeline: candidate retrieval from ad indexes, ranking using predictive models (pCTR, pCVR), auction execution via generalized second-price mechanisms, and ad serving with logging for feedback. This architecture, refined over two decades, handles billions of impressions daily within sub-100ms latency requirements.

However, classical systems face fundamental limitations: reliance on sparse keyword and ID-based features, inability to understand content semantically beyond engineered features, and static decision policies that cannot adapt to long-term outcomes or changing contexts.

Table 1: Classical vs. AI-Native Ad System Components

Component	Classical Approach	AI-Native Approach
Retrieval	Keyword matching, inverted indexes, category filtering	Semantic embeddings, vector similarity, LLM intent parsing
Ranking	GBDT, logistic regression with engineered features	Foundation models (GEM), transformer rankers, multi-objective optimization
Creatives	Advertiser-provided static assets, manual A/B testing	Generative AI (text, image, video), dynamic personalization
Bidding	Heuristic rules, PID controllers, daily budget smoothing	RL agents (DQN, DDPG, actor-critic), real-time adaptive policies
Optimization	Offline batch training, periodic model updates	Continuous learning, streaming updates, multi-agent coordination

The transition is not binary. Major platforms are implementing hybrid architectures where AI components augment rather than replace classical systems, maintaining legacy backstops during transition. Infrastructure is evolving from feature stores to representation stores, with unified embedding pipelines replacing siloed feature engineering.

3. Foundation models are reshaping advertising at scale

Meta's GEM represents the clearest example of an advertising-specific foundation model. Announced in November 2025, GEM is described as the "central brain" for Meta's ads recommendation system—an LLM-scale model trained on thousands of GPUs across trillions of user-ad interactions [2].

3.1 Key Technical Innovations

- **Heterogeneous feature processing:** GEM handles sequence features (activity history) and non-sequence features (demographics, ad format, creative representation) through customized attention mechanisms with cross-feature learning.
- **4× efficiency gains:** Novel architecture delivers four times the ad performance improvement per unit of data and compute compared to previous-generation models.
- **Knowledge distillation at scale:** GEM transfers its learnings to hundreds of specialized "vertical models" through hierarchical knowledge transfer, achieving 2× the effectiveness of standard distillation methods.
- **23× training throughput:** Infrastructure innovations including multi-dimensional parallelism and custom GPU kernels enable efficient training at massive scale.

Google's integration of generative AI into search represents a parallel transformation. AI Overviews now serve 1.5+ billion users monthly, with ads appearing within AI-generated summaries on desktop and expanding to mobile globally [8]. AI Max for Search delivers 27% more conversions at similar CPA/ROAS for adopters versus traditional exact/phrase match campaigns [9].

The market impact is profound. Zero-click searches have risen to approximately 69% of queries (from 56%), and first-position organic CTR dropped from 7.3% to 2.6% between March 2024 and March 2025 [4]. This is driving more advertisers into paid ecosystems, elevating CPCs across industries.

4. Agentic systems enable autonomous optimization

AI agents—autonomous systems that receive high-level goals and independently execute multi-step operations—are rapidly becoming central to advertising operations. Unlike traditional automation that follows predefined rules, agentic systems learn, adapt, and make decisions with minimal human intervention.

Meta plans to fully automate its advertising system by end of 2026, with AI generating ad creatives and determining target audiences based solely on a product image and budget [5]. Google's new agentic capabilities announced at Google Marketing Live 2025 give marketers "the power of their very own Google AI," with agents that help small businesses adapt campaigns with less manual effort [3].

4.1 Reinforcement Learning for Bidding Optimization

RL-based bidding has matured from research concept to production reality. These systems treat bidding as a sequential decision problem, learning optimal policies that balance immediate rewards against long-term campaign objectives under budget constraints.

Table 2: Reinforcement Learning Performance in Advertising

Metric	Performance vs. Traditional Methods
Overall advertising performance	+23.4% improvement
Bid outcome prediction accuracy	85.7% (deep learning models)
Campaign ROI improvement	+21.56% via dynamic optimization
Conversion rate improvement	+16.7% consistent improvement
Small advertiser auction win rates	+31% increase with AI bidding

Industry observers predict media buying will shift toward agents negotiating with each other rather than humans. Gartner projects 90% of B2B buying will be AI-agent intermediated by 2028 [7]. Systems will communicate through AI agents in addition to traditional APIs, fundamentally changing how advertising inventory is transacted.

5. Consumer behavior shifts threaten traditional advertising models

The migration from traditional search to AI-powered discovery represents a fundamental shift in how consumers find information and products. This is not incremental evolution—it is a paradigm change that threatens the foundation of search advertising.

5.1 The Death of Traditional Search

- **Search replacement accelerating:** 58% of consumers have replaced traditional search with generative AI tools. ChatGPT preference over Google grew 8× in less than half a year (from 1% to 8% of respondents in Evercore survey) [4].
- **AI platforms exploding:** ChatGPT saw 44% traffic boost in late 2024 and now processes 2.5+ billion prompts daily. Perplexity reached 15 million monthly users with 191.9% year-over-year growth. Semrush predicts AI Search visitors will surpass traditional search visitors by 2028 [10].
- **Zero-click dominance:** 80% of consumers rely on zero-click results for at least 40% of their searches. Organic web traffic declining 15–25% as AI satisfies queries directly.
- **Shopping behavior changing:** Shopping queries now comprise nearly 10% of ChatGPT's prompts, growing 25%+ since early 2025. 71% of consumers want AI integrated into shopping experiences.

The implications are profound. As one industry observer noted: "Once, marketers cared where algorithms ranked them on a SERP. But search will be redefined, and reputation—what real people say about you on recommendation sites—will matter more than ever. The battle for online presence will shift decisively to the LLM's latent space" [5].

6. Social media advertising faces an existential reckoning

While AI is transforming search advertising, social media advertising faces a different but equally challenging disruption: growing consumer fatigue and rejection of advertising in social contexts. The platforms that have built trillion-dollar businesses on ad-supported models may face fundamental pressure.

6.1 Signs of Consumer Rejection

- **Platform fatigue rising:** Gartner forecasts that by 2025, 50% of consumers will significantly limit their social media usage due to their perception of platform quality [7]. Users report feeds have "started to look like TV, with endless short videos starring people they don't know."
- **Migration to private spaces:** Around half of all global social media users plan to increase time on emerging, community-driven platforms [6]. Reddit, Bluesky, Discord, and private groups are gaining at the expense of algorithmic feeds. Consumers are "craving closed spaces" where advertising is minimal or absent.
- **Ad saturation backlash:** 84% say ads that rotate or update feel more relevant, while nearly 40% feel uncomfortable when the same ad follows them after being ignored. Instagram CPMs reached $9.46 per thousand impressions in Q2 2025—higher spend for lower engagement [11].
- **Authenticity demands:** 55% of social users are more likely to trust brands that publish human-generated content (rising to 66% for Gen Z and Millennials). The #1 effort consumers want brands to prioritize in 2026 is "crafting human-generated content" [6].

6.2 Consumer Attitudes Toward AI-Generated Advertising

As platforms deploy more AI-generated content and targeting, consumer attitudes present a complex picture. Roughly one-third of consumers like AI in advertising, one-third are neutral, and one-third dislike it [12]. Men and younger consumers are more positive; women and older consumers more negative.

Only a small fraction fully trust ads created entirely by AI. Most prefer ads co-created by humans and AI. Nearly half say disclosure is essential to building trust [13]. Research shows AI-salient advertising generates greater online engagement but evokes more negative emotions that harm ad attitude—a disclosure dilemma with no easy resolution [14].

The convergence of these trends—AI replacing search, social fatigue, and AI-content skepticism—suggests the advertising industry may face a fundamental reckoning. As AI agents increasingly make or influence purchase decisions, advertising shifts from persuading humans to ensuring AI systems have accurate, complete information about products.

7. Platform-specific strategies diverge

7.1 Meta (Facebook, Instagram)

Meta is pursuing the most aggressive AI transformation. GEM foundation model drives 3–5% conversion lifts across surfaces. The Advantage+ suite automates creative, targeting, and bidding optimization. AI Sandbox provides generative tools for ad copy and image variations. Full automation planned by end of 2026 means advertisers will provide only a product image and budget—AI handles everything else [2].

7.2 Google (Search, YouTube)

Google is integrating AI throughout the search advertising experience. AI Overviews reach 1.5B+ monthly users with integrated advertising. AI Max delivers 27% more conversions for adopters. Veo (video) and Imagen 4 (image) models are now available in Google Ads for creative generation. Performance Max campaigns offer full cross-channel AI optimization [9].

7.3 TikTok/ByteDance

TikTok operates multimodal LLMs processing billions of daily videos for content understanding. Symphony generative AI suite enables TikTok-native ad creation. Gross merchandise value is projected to reach $66B in 2025, potentially doubling in 2026. INT8 quantization achieves 50% cost reduction for inference [15].

7.4 Connected TV and Audio

CTV is stealing linear TV ad share: Linear projected at less than 15% share by 2026 (down from 30% in 2019). GenAI-enhanced video ads are growing from 22% (2024) to 30% (2025) to 39% (2026). A concerning trend: if 60% of marketers adopt LLM-based planning, broadcast revenue could be cut in half within three years due to LLM planning bias toward digital channels with machine-readable performance data [16].

8. Strategic implications for market participants

The AI-powered advertising market is entering a decisive 2025–2028 window.

8.1 For Ad Platforms

1. **Embrace incremental transformation.** Start with assistive LLM features before moving to in-loop control. Maintain legacy backstops during transition.
2. **Invest in unified embedding infrastructure.** Build robust pipelines for generating, updating, and serving embeddings. Plan for vector database scaling.
3. **Deploy RL agents cautiously.** Start with constrained scope, implement safety nets, and build robust offline evaluation before broad rollout.
4. **Prepare for agent-to-agent commerce.** Design APIs and data structures that serve machine decision-makers.
5. **Address consumer trust proactively.** Develop clear AI disclosure policies and hybrid human+AI creative workflows that maintain authenticity.

8.2 For Advertisers

1. **Optimize for AI discovery.** Implement structured data, schema markup, and context-rich content. Your site needs to serve AI agents, not just human browsers.
2. **Balance AI efficiency with human authenticity.** Use AI for production efficiency but maintain human creative direction. The authenticity premium is real.
3. **Diversify beyond social advertising.** As consumer fatigue grows and private spaces gain share, over-reliance on ad-supported social platforms becomes riskier.
4. **Rethink measurement frameworks.** Track share of voice in AI responses, citation frequency, and brand mentions—not just traditional CTR and rankings.

8.3 Key Risks and Watchpoints

- **Consumer trust erosion:** The more AI is deployed without building trust, the greater the risk of broader advertising rejection.
- **LLM planning bias:** LLM-based media planning tools favor channels with machine-readable performance data. Traditional media may be systematically underweighted.
- **Rising costs:** As organic visibility declines and more advertisers move to paid AI placements, CPCs are rising across industries.
- **Regulatory uncertainty:** AI disclosure requirements, automated decision-making rules, and privacy regulations continue to evolve.

9. Conclusion

The transformation of ad engines through LLMs, foundation models, and agentic systems is no longer speculative—it is measurably happening. Meta's GEM, Google's AI Overviews, and the rapid adoption of AI shopping assistants represent inflection points that will define the next era of digital advertising.

However, this technological transformation is occurring against a backdrop of shifting consumer attitudes that complicate the picture. As AI takes over search, consumers increasingly expect helpful assistance rather than advertising. As social platforms saturate with AI-generated content, consumers migrate to private spaces and demand authenticity. The paradox: AI enables unprecedented advertising efficiency and personalization, but consumers are simultaneously becoming more resistant to being advertised to.

For ad platforms, the path forward requires balancing technological capability with consumer trust. The most sophisticated AI systems will fail if they erode the consumer relationships that make advertising valuable in the first place.

For advertisers, the imperative is twofold: embrace AI-native strategies for efficiency and reach, while investing in authentic human connections that differentiate in an increasingly AI-mediated world.

The winners in 2026 and beyond will not simply be those with the most advanced AI—they will be those who deploy AI in ways that genuinely serve consumers, building trust rather than eroding it. The advertising industry faces not just a technological transformation, but a fundamental test of whether it can adapt to a world where consumers have more power than ever to tune out.

References

[1] eMarketer. Agencies embrace ai as consumer sentiment remains mixed. eMarketer, December 2025. URL https://www.emarketer.com/.

[2] Meta Engineering. Meta's generative ads model (gem): The central brain accelerating ads recommendation ai innovation. Meta Engineering Blog, November 2025. URL https://engineering.fb.com/.

[3] Vidhya Srinivasan. The future of ai-powered search marketing. Think with Google, May 2025. URL https://www.thinkwithgoogle.com/.

[4] Amplitude. 2026 ai playbook. Amplitude, 2025. URL https://amplitude.com/.

[5] AdWeek. 2026 predictions: Search takes a backseat as brands battle inside ai's black box. AdWeek, December 2025. URL https://www.adweek.com/.

[6] Sprout Social. The future of social media: 2026 predictions. Sprout Social Insights, December 2025. URL https://sproutsocial.com/insights/. Q3-Q4 2025 Pulse Surveys.

[7] Gartner. Digital marketing hype cycle 2025. Gartner Research, 2025. URL https://www.gartner.com/.

[8] Search Engine Land. Ads arrive in google ai overviews. Search Engine Land, May 2025. URL https://searchengineland.com/.

[9] Google Ads Help. Google ads highlights of 2025. Google Ads Help Center, December 2025. URL https://support.google.com/google-ads/.

[10] Clarifai. Top llms and ai trends for 2026. Clarifai Blog, November 2025. URL https://www.clarifai.com/.

[11] SF Gate Marketing. Ad fatigue is real—how to keep your campaigns fresh and engaging. SF Gate, October 2025. URL https://marketing.sfgate.com/.

[12] Zappi. How consumers feel about the use of ai in advertising. Zappi Blog, 2025. URL https://www.zappi.io/.

[13] Smartly.io. Ai and advertising in 2025: Consumer expectations and research insights. Smartly Resources, 2025. URL https://www.smartly.io/.

[14] S. Hanson et al. The differential impact of ai salience on advertising engagement and attitude: Scary good ai advertising. *Journal of Advertising Research*, 65(2), 2025.

[15] AWS Machine Learning Blog. Bytedance processes billions of daily videos using their multimodal video understanding models on aws inferentia2. AWS Blog, 2025. URL https://aws.amazon.com/blogs/machine-learning/.

[16] WideOrbit. Ai and optimizing advertising in 2025: From hype to hard results. WideOrbit Blog, October 2025. URL https://www.wideorbit.com/.

Author's Note

This paper reflects applied analysis informed by work across multiple enterprise AI and advertising technology systems. The views expressed are intended to support strategic decision-making rather than prescribe specific vendors, products, or implementations.

About the Author

Anjan Goswami works on applied AI systems spanning evaluation, inference efficiency, and production-scale deployment. His experience focuses on the intersection of model behavior, infrastructure economics, and long-term system maintainability in real-world environments.

Executive Contact

Senior technical leaders or executives with questions related to AI-powered advertising systems, consumer behavior shifts, or large-scale deployment considerations may contact the author for further discussion.