Consumer Protection in the Age of Chatbot Advertising

Erik Hermann[1*], Stefano Puntoni[2], David A. Schweidel[3], Kenneth C. Wilbur[4]

[1] European University Viadrina, Frankfurt (Oder), Germany.

[2] Wharton School of Business, University of Pennsylvania, Philadelphia, PA, USA.

[3] Goizueta Business School, Emory University, Atlanta, GA, USA.

[4] Rady School of Management, University of California, San Diego, CA, USA.

[*]Corresponding author: hermann@europa-uni.de

AI providers face increasing pressure to monetize their chatbots through advertising. *Contextual native ads* - personalized suggestions woven seamlessly into chatbot conversations - blur the line between assistance and persuasion, raising risks of covert influence, privacy violations, and consumer harm that demand urgent policy attention.

Chatbots based on large language models (LLMs)—also termed *conversational AI*—are becoming a dominant interface for information search, commerce, companionship, and other everyday tasks. Like earlier technologies such as search engines and social media, there is increasing pressure for AI providers to monetize their chatbots through advertising[1]. Several AI providers have begun implementing advertising or plan to do so. For example, Bing Chat integrates sponsored links. Perplexity's ads are text-based ads that appear as sponsored questions ("related questions") at the bottom of the chatbot response or as suggestions at the sidebar to the right of the response. Unlike other ads, clicking these ads continues the conversation rather than redirecting the user to an external website[2]. The adtech startup Nexad is building an "AI-native ads system" network with advertisers, affiliates, and third parties, allowing AI app providers to embed personalized text and image ads directly into AI chatbot responses (see Fig. 1). Engagement data are then used for continuous refinement of future ads[3].

Chatbot advertising can take different forms. Ads may resemble clearly labeled display ads placed alongside chatbot responses or appear as sponsored calls-to-action tied to query results. We focus on a new format that warrants scrutiny, *contextual native ads*: personalized suggestions embedded within conversations that are presented in natural language. Such chatbot ads could thus unfold within the conversational flow[1]. A user seeking financial advice might receive a credit card recommendation. While responses to his queries may appear like neutral assistance, they come with an important difference: their placements are sold to the highest bidders often without consumers' knowledge. The blurring of lines between AI assistance and commercial interests, delivered via persuasive content, raises urgent questions and challenges for public policy. This ad format combines the opacity of native advertising, the targeting of social media, and the authenticity of human-like conversations, making them appealing to advertisers. Given existing concerns about how advertising-driven business models may impact consumer wellbeing, it is crucial to understand the consumer welfare implications of sponsored chatbot ads.

TARGETING AND OBTRUSIVENESS

Contextual native ads take the form of narrative language embedded seamlessly within a conversational flow. That resembles an ad type that the marketing industry calls "native ads," designed to match the form and function of their surrounding content and the platform (e.g., news publishers or social media) on which they are distributed[4]. Unlike search engines that return a list of results, with sponsored results demarcated, chatbots provide a single response that may blend organic and sponsored content, blurring the line between them[5]. This would make chatbot advertising less transparent and harder to recognize for consumers, and more valuable to advertisers. Even when digital ads (including native ads) are labeled, consumers often fail to recognize the paid origin of the content[4]. Similarly, early research shows that users struggle to detect chatbot ads—even when they are labeled as sponsored—and that once they are made aware of them, they perceive chatbots including ads as manipulative[1].

Chatbot advertising formats can be described using two dimensions: targeting (i.e., how tailored are ads to individual users or queries) and obtrusiveness (i.e., how interruptive or noticeable are ads within user conversations)[6,7]. *Generic native ads* (low targeting, low obtrusiveness) are general brand mentions casually embedded in chatbot answers and conversations. *Generic obtrusive ads* (low targeting, high obtrusiveness) are explicit, clearly labeled promotional interruptions without contextual relevance. *Contextual native ads* (high targeting, low obtrusiveness) are personalized suggestions interwoven into chatbot conversations

without clear sponsorship labels. *Contextual obtrusive ads* (high targeting, high obtrusiveness) are explicit promotions that directly respond to user queries such as direct sponsored suggestions.

Of these four overarching categories of chatbot advertising, *contextual native ads* would likely have the greatest effect on consumer behavior[7] and simultaneously pose the greatest risk to consumers[1]. Because ads align closely with user queries and appear in the natural conversational flow, they are difficult to distinguish from non-sponsored (organic) responses. As they do not interrupt user engagement, they are appealing to chatbot operators. *Contextual native ads* would likely amplify advertiser conversion rates, as they would not trigger consumers' natural defenses to persuasion attempts. Unlike obtrusive formats that disclose their commercial intent, *contextual native ads* can deliver persuasion in the guise of neutral assistance.

CONTEXTUAL NATIVE ADS AND CONSUMERS

With simple prompts and queries, chatbots provide a comprehensive answer with information compiled from various sources. In web searches, the user synthesizes the results; in chats, the chatbot synthesizes on the user's behalf. Further, chatbots can generate personalized syntheses based on past user interactions, increasing their ability to generate personalized advertising that maximizes user response[7]. Moreover, flexibility in the communication and narrative style, tone, and modality (text vs. voice) can enable chatbots to appeal simultaneously to both factual and emotional modes of processing. Chatbots can be equipped with human-like features reflected by names, personae, and communication styles, enabling interactions to feel interpersonal. By adopting relational framings that mimic peer-to-peer recommendations, chatbots can present information as sincere advice rather than using generic persuasive appeals, lowering consumer defenses and enhancing influence[8]. This combination of features poses new risks to consumer welfare absent in earlier forms of digital marketing.

One risk is covert persuasion and deception. Because *contextual native ads* can be seamlessly embedded into otherwise neutral answers, consumers are unlikely to identify them as sponsored content[5], making them less likely to process the persuasion attempt critically and resist it. AI providers and/or advertisers could conceal the commercial intent of the content delivered by the chatbot. The inclusion of explicit disclosures that clearly indicate whether a response has been influenced by paid sponsorship is required for truthful advertising. Otherwise, the blurring of organic and sponsored content risks confusing and misleading consumers. It is also possible for LLMs to be calibrated to maximize persuasiveness, to the potential detriment of consumers. In extreme cases, autonomous systems calibrated to maximize sales may deceive consumers. Such covert forms of persuasion and potential deception are particularly problematic because they may exploit consumer sensitivities, such as by using guilt or fear appeals to increase compliance among those with emotional sensitivity or conflict aversion. Vulnerable populations (e.g., those with lower digital literacy, children) may therefore be disproportionately influenced by chatbot ads.

Privacy is another concern. Traditional search engines are designed for one-shot interactions; each set of search results has limited dependence on the user's prior searches. Conversational AI has long-term memory features that may enable ad designs that leverage previously disclosed sensitive data, without the consumer being aware of it. Chatbots' use of reciprocity strategies (e.g., empathy and emotional support) both engage consumers and encourage personal and sensitive data disclosure[9]. This can result in the unprecedented collection and use of personal data for targeted advertising, with concerns that could be exacerbated if the

chatbot provider shares or uses the information for advertising across the Internet (i.e., retargeting).

Contextual native ads can also shape how choices are presented and evaluated, thereby undermining consumer autonomy. Consumers may delegate key aspects of the decision-making process to chatbots, allowing (or even encouraging) chatbots to dictate the choice architecture. Whereas search engines provide results that users can compare, part of chatbots' appeal is the simplicity with which results are presented, often providing as a single answer or a few suggestions, without requiring the consumer to process the full set of alternatives.

POLICY CHALLENGES

The opportunities and risks arising from *contextual native ads* require proactive policies. With disclosure being the foundation of advertising policy, transparency is paramount. The US Federal Trade Commission (FTC) emphasizes that disclosure must be "clear and conspicuous," meaning easily noticeable and easily understandable by ordinary consumers. In its Enforcement Policy Statement on Deceptively Formatted Advertisements, the FTC states that an ad is deceptive when its promotional nature is not apparent to consumers and that even a factually accurate ad may be deceptive if the ad format encourages consumer misunderstanding of ad source or ad intent[10]. Prior research on native advertising shows that consumers often fail to recognize disclosures, and compliance among publishers and advertisers is weak[4]. Conversational AI that can intermingle paid content with what appears to be neutral advice exacerbates this concern, as consumers may not know which chat content is sponsored, even when sponsorship is disclosed[1]. Regulators must update disclosure principles to reflect disclosure of specifically how sponsorship influences machine-generated chat responses. Standardized, updated disclosure policies would also help to prevent harmful or excessive competition between chatbot providers in advertising sales, and help to protect user trust of the chatbot industry.

Another important area is privacy. Chatbots do not merely track queries, but they can elicit personal disclosures in real time, often without consumers fully realizing it[9]. Regulators might consider imposing privacy-by-design obligations. These could include (i) real-time notifications when a chatbot elicits or uses personal data, (ii) rules about storing (and using) sensitive data for commercial purposes without explicit opt-in consent, (iii) safeguards against manipulative elicitation techniques, (iv) the ability to delete or export chatbot data, and (v) restrictions on data sharing with third parties.

Consumer protection may require larger steps to protect vulnerable populations and contexts. Existing advertising rules already distinguish by context and consumer population, such as in advertising to children, for financial services, healthcare and pharmaceuticals, and tobacco products. As vulnerable consumers can be particularly susceptible to the influence of undisclosed advertising and to unintentionally providing sensitive information, more research is needed to understand how to best update existing regulations to protect against the risks posed by c*ontextual native ads*.

LOOKING FORWARD

The future of chatbot advertising is not yet written. Along one trajectory, advertising integrated into conversational AI may deliver relevant, well-targeted suggestions that improve decision making and subsidize access to powerful tools. With clear disclosures and enforced protection,

consumers could benefit from ads that are more timely, less obtrusive, and more contextually useful than other digital ad formats. Ads could also subsidize access to chatbots. Conversely, the same features that make chatbot ads persuasive create inherent risks. If sponsorship is covert and disclosures ineffective, consumers may be deceived and trust in AI systems could erode over time.

Ultimately, consumer protection must co-evolve with advertising in conversational AI. Well-regulated practices could make advertising more useful for consumers while enabling AI providers to offer high-quality chatbots without charging consumers directly for use. However, covert and deceptive formats ads could amplify the shortcomings and failures observed in earlier digital advertising formats. The challenge for policymakers, advertisers, and AI developers is to shape this emerging market in ways that promote trust, fairness, and long-term consumer welfare without hindering innovation.

References

1. B. J. Tang, K. Sun, N. T. Curran, F. Schaub, K. G. Shin, in *Proc. IMWUT* (2025).
2. S. Berry, https://www.webfx.com/blog/ppc/perplexity-advertising/ (WebFX, 2025).
3. S. Bradley, M. Meaney, https://www.businessinsider.com/adtech-startup-nexad-raises-seed-ai-native-ads-pitch-deck-2025-4 (Business Insider, 2025).
4. C. Campbell, P. Grimm, *J. Public Policy Mark.* 38, 110-123 (2019).
5. I. Zelch, M. Hagen, M. Potthast, in *Proc. CHIIR* (2024).
6. H. Goldfarb, C. Tucker, *Mark. Sci.* 30, 389–404 (2011).
7. S. C. Matz, J. D. Teeny, S. S. Vaid, H. Peters, G. M. Harari, M. Cerf, *Sci. Rep.* 14, 4692 (2024).
8. M. A. Tuk, P. W. J. Verlegh, A. Smidts, D. H. J. Wigboldus, *J. Consum. Psychol.* 19, 38–47 (2009).
9. X. Zhan, J. C. Carrillo, W. Seymour, J. Such, in *Proc. SEC* (2025).
10. Federal Trade Commission, https://www.ftc.gov/system/files/documents/public_statements/896923/151222deceptiveenforcement.pdf (2015).

Competing Interests

The authors declare no competing interests.