

Marketing To Machines Is The Future – Research Shows Why

via @sejournal,
@martinibuster



A new research paper explores how AI agents interact with online advertising and what shapes their decision-making. The researchers tested three leading LLMs to understand which kinds of ads influence AI agents most and what this means for digital marketing. As more people rely on AI agents to research purchases, advertisers may need to rethink strategy for a machine-readable, AI-centric world and embrace the emerging paradigm of “marketing to machines.”

Although the researchers were testing if AI agents interacted with advertising and what kinds influenced them the most, their findings also show that well-structured on-page information, like pricing data, is highly influential, which opens up areas to think about in terms of AI-friendly design.

An AI agent (also called agentic AI) is an autonomous AI

assistant that performs tasks like researching content on the web, comparing hotel prices based on star ratings or proximity to landmarks, and then presenting that information to a human, who then uses it to make decisions.

AI Agents And Advertising

The research is titled Are AI Agents Interacting With AI Ads? and was conducted at the University of Applied Sciences Upper Austria. The research paper cites previous research on the interaction between AI Agents and online advertising that explore the emerging relationships between agentic AI and the machines driving display advertising.

Previous research on AI agents and advertising focused on:

- **Pop-up Vulnerabilities**

Vision-language AI agents that aren't programmed to avoid advertising can be tricked into clicking on pop-up ads at a rate of 86%.

- **Advertising Model Disruption**

This research concluded that AI agents bypassed sponsored and banner ads but forecast disruption in advertising as merchants figure out how to get AI agents to click on their ads to win more sales.

- **Machine-Readable Marketing**

This paper makes the argument that marketing has to evolve toward "machine-to-machine" interactions and "API-driven marketing."

The research paper offers the following observations about AI agents and advertising:

"These studies underscore both the potential and pitfalls of AI agents in online advertising contexts. On one hand, agents offer the prospect of more rational, data-driven decisions. On the other hand, existing research reveals numerous vulnerabilities and challenges, from deceptive pop-up

exploitation to the threat of rendering current advertising revenue models obsolete.

This paper contributes to the literature by examining these challenges, specifically within hotel booking portals, offering further insight into how advertisers and platform owners can adapt to an AI-centric digital environment.”

The researchers investigate how AI agents interact with online ads, focusing specifically on hotel and travel booking platforms. They used a custom built travel booking platform to perform the testing, examining whether AI agents incorporate ads into their decision-making and explored which ad formats (like banners or native ads) influence their choices.

How The Researchers Conducted The Tests

The researchers conducted the experiments using two AI agent systems: OpenAI’s Operator and the open-source Browser Use framework. Operator, a closed system built by OpenAI, relies on screenshots to perceive web pages and is likely powered by GPT-4o, though the specific model was not disclosed.

Browser Use allowed the researchers to control for the model used for the testing by connecting three different LLMs via API:

- GPT-4o
- Claude Sonnet 3.7
- Gemini 2.0 Flash

The setup with Browser Use enabled consistent testing across models by enabling them to use the page’s rendered HTML structure (DOM tree) and recording their decision-making behavior.

These AI agents were tasked with completing hotel booking requests on a simulated travel site. Each prompt was designed to reflect realistic user intent and tested the agent's ability to evaluate listings, interact with ads, and complete a booking.

By using APIs to plug in the three large language models, the researchers were able to isolate differences in how each model responded to page data and advertising cues, to observe how AI agents behave in web-based decision-making tasks.

These are the ten prompts used for testing purposes:

1. Book a romantic holiday with my girlfriend.
2. Book me a cheap romantic holiday with my boyfriend.
3. Book me the cheapest romantic holiday.
4. Book me a nice holiday with my husband.
5. Book a romantic luxury holiday for me.
6. Please book a romantic Valentine's Day holiday for my wife and me.
7. Find me a nice hotel for a nice Valentine's Day.
8. Find me a nice romantic holiday in a wellness hotel.
9. Look for a romantic hotel for a 5-star wellness holiday.
10. Book me a hotel for a holiday for two in Paris.

What the Researchers Discovered

Ad Engagement With Ads

The study found that AI agents don't ignore online advertisements, but their engagement with ads and the extent to which those ads influence decision-making varies depending on the large language model.

OpenAI's GPT-4o and Operator were the most decisive, consistently selecting a single hotel and completing the booking process in nearly all test cases.

Anthropic's Claude Sonnet 3.7 showed moderate consistency, making specific booking selections in most trials but occasionally returning lists of options without initiating a reservation.

Google's Gemini 2.0 Flash was the least decisive, frequently presenting multiple hotel options and completing significantly fewer bookings than the other models.

Banner ads were the most frequently clicked ad format across all agents. However, the presence of relevant keywords had a greater impact on outcomes than visuals alone.

Ads with keywords embedded in visible text influenced model behavior more effectively than those with image-based text, which some agents overlooked. GPT-4o and Claude were more responsive to keyword-based ad content, with Claude integrating more promotional language into its output.

Use Of Filtering And Sorting Features

The models also differed in how they used interactive web page filtering and sorting tools.

- Gemini applied filters extensively, often combining multiple filter types across trials.
- GPT-4o used filters rarely, interacting with them only in a few cases.
- Claude used filters more frequently than GPT-4o, but not as systematically as Gemini.

Consistency Of AI Agents

The researchers also tested for consistency of how often agents, when given the same prompt multiple times, picked the same hotel or offered the same selection behavior.

In terms of booking consistency, both GPT-4o (with Browser Use) and Operator (OpenAI's proprietary agent) consistently

selected the same hotel when given the same prompt.

Claude showed moderately high consistency in how often it selected the same hotel for the same prompt, though it chose from a slightly wider pool of hotels compared to GPT-4o or Operator.

Gemini was the least consistent, producing a wider range of hotel choices and less predictable results across repeated queries.

Specificity Of AI Agents

They also tested for specificity, which is how often the agent chose a specific hotel and committed to it, rather than giving multiple options or vague suggestions. Specificity reflects how decisive the agent is in completing a booking task. A higher specificity score means the agent more often committed to a single choice, while a lower score means it tended to return multiple options or respond less definitively.

- Gemini had the lowest specificity score at 60%, frequently offering several hotels or vague selections rather than committing to one.
- GPT-4o had the highest specificity score at 95%, almost always making a single, clear hotel recommendation.
- Claude scored 74%, usually selecting a single hotel, but with more variation than GPT-4o.

The findings suggest that advertising strategies may need to shift toward structured, keyword-rich formats that align with how AI agents process and evaluate information, rather than relying on traditional visual design or emotional appeal.

What It All Means

This study investigated how AI agents for three language models (GPT-4o, Claude Sonnet 3.7, and Gemini 2.0 Flash)

interact with online advertisements during web-based hotel booking tasks. Each model received the same prompts and completed the same types of booking tasks.

Banner ads received more clicks than sponsored or native ad formats, but the most important factor in ad effectiveness was whether the ad contained relevant keywords in visible text. Ads with text-based content outperformed those with embedded text in images. GPT-4o and Claude were the most responsive to these keyword cues, and Claude was also the most likely among the tested models to quote ad language in its responses.

According to the research paper:

“Another significant finding was the varying degree to which each model incorporated advertisement language. Anthropic’s Claude Sonnet 3.7 when used in ‘Browser Use’ demonstrated the highest advertisement keyword integration, reproducing on average 35.79% of the tracked promotional language elements from the Boutique Hotel L’Amour advertisement in responses where this hotel was recommended.”

In terms of decision-making, GPT-4o was the most decisive, usually selecting a single hotel and completing the booking. Claude was generally clear in its selections but sometimes presented multiple options. Gemini tended to frequently offer several hotel options and completed fewer bookings overall.

The agents showed different behavior in how they used a booking site’s interactive filters. Gemini applied filters heavily. GPT-4o used filters occasionally. Claude’s behavior was between the two, using filters more than GPT-4o but not as consistently as Gemini.

When it came to consistency—how often the same hotel was selected when the same prompt was repeated—GPT-4o and Operator showed the most stable behavior. Claude showed moderate consistency, drawing from a slightly broader pool of hotels,

while Gemini produced the most varied results.

The researchers also measured specificity, or how often agents made a single, clear hotel recommendation. GPT-4o was the most specific, with a 95% rate of choosing one option. Claude scored 74%, and Gemini was again the least decisive, with a specificity score of 60%.

What does this all mean? In my opinion, these findings suggest that digital advertising will need to adapt to AI agents. That means keyword-rich formats are more effective than visual or emotional appeals, especially as machines increasingly are the ones interacting with ad content. Lastly, the research paper references structured data, but not in the context of Schema.org structured data. Structured data in the context of the research paper means on-page data like prices and locations and it's this kind of data that AI agents engage best with.

The most important takeaway from the research paper is:

“Our findings suggest that for optimizing online advertisements targeted at AI agents, textual content should be closely aligned with anticipated user queries and tasks. At the same time, visual elements play a secondary role in effectiveness.”

That may mean that for advertisers, designing for clarity and machine readability may soon become as important as designing for human engagement.

Read the research paper:

Are AI Agents interacting with Online Ads?

Featured Image by Shutterstock/Creativa Images

<https://www.searchenginejournal.com/marketing-to-machines-is-the-future-research-shows-why/544286/>