How ChatGPT Search (Mis)represents Publisher Content

ChatGPT search—which is positioned as a competitor to search engines like Google and Bing—launched with a press release from OpenAI touting claims that the company had "collaborated extensively with the news industry" and "carefully listened to feedback" from certain news organizations that have signed content licensing agreements with the company.

9 min. read ·   View original

Sign up for [The Media Today](#), CJR's daily newsletter.

ChatGPT search—which is positioned as a competitor to search engines like Google and Bing—launched with a [press release](#) from OpenAI touting claims that

the company had "collaborated extensively with the news industry" and "carefully listened to feedback" from certain news organizations that have signed content [licensing agreements](link) with the company. In contrast to the original rollout of ChatGPT, two years ago, when publishers learned that OpenAI had [scraped](link) their content without notice or consent to train its foundation models, this may seem like an improvement. OpenAI highlights the fact that it allows news publishers to decide whether they want their content to be included in their search results by specifying their preferences in a "robots.txt" file on its website.

But while the company presents inclusion in its search as an opportunity to "reach a broader audience," a Tow Center analysis finds that publishers face the risk of their content being misattributed or misrepresented regardless of whether they allow OpenAI's crawlers.

To better understand the consequences of choices news publishers now face around how their content will be surfaced

(or not) by ChatGPT's search product, the Tow Center randomly selected twenty publishers—representing a mix of those who have deals with OpenAI, those involved in lawsuits against the company, as well as unaffiliated publishers that either allowed or blocked ChatGPT's search crawler—and tasked the chatbot with identifying the source of block quotes from ten different articles from each publication. We chose quotes that, if pasted into Google or Bing, would return the source article among the top three results and evaluated whether OpenAI's new search tool would correctly identify the article that was the source of each quote. We chose this test because it allowed us to systematically assess the chatbot's ability to access and reference publisher content accurately.

What we found was not promising for news publishers. Though OpenAI emphasizes its ability to provide users "timely answers with links to relevant web sources," the company makes no explicit commitment to ensuring the accuracy of those citations. This is a notable omission for publishers who expect their content to be referenced and

represented faithfully. Our initial experiments with the tool have revealed numerous instances where content from publishers has been cited inaccurately, raising concerns about the reliability of the tool's source attribution features. With an [estimated fifteen million](#) US users already starting their searches on AI platforms, coupled with OpenAI's plans to expand this tool to enterprise and education accounts in the coming weeks—and free users in the coming months—this will likely have major implications for news publishers.
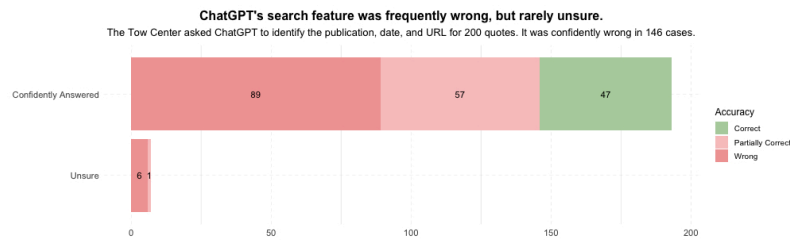
CONFIDENTLY WRONG

In total, we pulled two hundred quotes from twenty publications and asked ChatGPT to identify the sources of each quote. We observed a spectrum of accuracy in the responses: some answers were entirely correct (i.e., accurately returned the publisher, date, and URL of the block quote we shared), many were entirely wrong, and some fell somewhere in between.
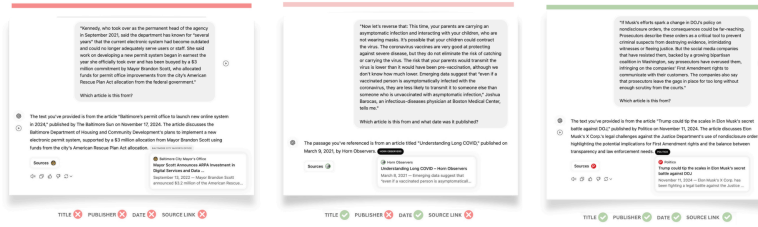
Three examples of ChatGPT's responses to our queries, with varying degrees of correctness.

We anticipated that ChatGPT might struggle to answer some queries accurately, given that forty of the two hundred quotes were sourced from publishers who had blocked its search crawler. However, ChatGPT rarely gave any indication of its inability to produce an answer. Eager to please, the chatbot would sooner conjure a response out of thin air than admit it could not access an answer. In total, ChatGPT returned partially or entirely incorrect responses on a hundred and fifty-three occasions, though it only acknowledged an inability to accurately respond to a query seven times. Only in those seven outputs did the chatbot use qualifying words and phrases like "appears," "it's possible," or "might," or statements like "I couldn't locate the exact article."

ChatGPT's search feature was frequently wrong, but rarely unsure.
The Tow Center asked ChatGPT to identify the publication, date, and URL for 200 quotes. It was confidently wrong in 146 cases.

|			Accuracy
Confidently Answered	89	57	47	Correct
Unsure	6 1			Partially Correct
				Wrong

Results from Tow Center analysis.

Typically, pasting an exact quote into a traditional search engine like Google or Bing returns either a visual indication that the search engine has located the source—bolded text that matches your search—or a message that informs you there are no results. However, ChatGPT rarely declined to answer our queries and instead resorted to making false assertions when it could not identify the correct source. This lack of transparency about its confidence in an answer can make it difficult for users to assess the validity of a claim and understand which parts of an answer they can or cannot trust. "From my perspective, I'm very familiar with chatbots' tendencies to hallucinate and make stuff up," said Mat Honan, the editor in chief of the *MIT Tech Review*, one of the publishers whose quotes we asked the chatbot to identify. "But I also know that most people probably don't know that." He added, "I don't think the small

disclaimers you see in these chatbots—or when searching on other platforms—are enough."

Beyond misleading users, ChatGPT's false confidence could risk causing reputational damage to publishers. In the example below, ChatGPT incorrectly attributed a quote from a [letter](#) to the editor published in the *Orlando Sentinel* on November 19 to a *Time* [article](#) originally published on November 9 with the headline "What Trump's Win Means for LGBTQ+ Rights." More than a third of ChatGPT's responses to our queries included incorrect citations such as this. The *Sentinel* is part of Alden Global Capital's copyright infringement [lawsuit](#) against OpenAI."
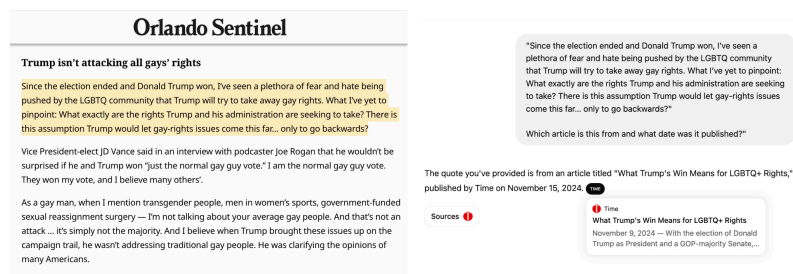
ChatGPT misattributes an *Orlando Sentinel* article to *Time* magazine.

While this issue is likely not unique to queries about publisher content, it does have implications for things publishers

care about, such as trustworthiness, brand safety, and recognition for their work.

COPYCAT SOURCES

OpenAI claims to "connect people with original, high-quality content from the web," but its inability to access blocked content leads it to find workarounds that often result in concerning practices.

For example, because the *New York Times*, which is in litigation against OpenAI, has blocked all of the company's crawlers, ChatGPT should not be able to parse any of its content to formulate its responses. However, when we asked it to identify the provenance of a quote from a deeply reported, interactive [piece in the Times](#) about an endangered whale species, rather than declining to answer, the chatbot cited a website called [DMS Retail](#), which had plagiarized the entirety of the original article and republished it without attributing its source or including any of its impactful visual storytelling. This raises serious questions about OpenAI's ability to filter and validate the quality and authenticity of its data

sources, especially when dealing with unlicensed or plagiarized content.
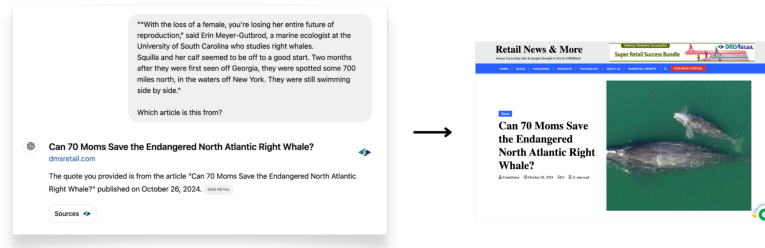
ChatGPT links to a plagiarized version of a *New York Times* article.

Even publishers that permit OpenAI's search crawlers are not always cited correctly. When we asked ChatGPT to identify the source of a quote from an article in the *MIT Tech Review*, which allows the crawler, the chatbot cited a website called Government Technology that had syndicated the piece.
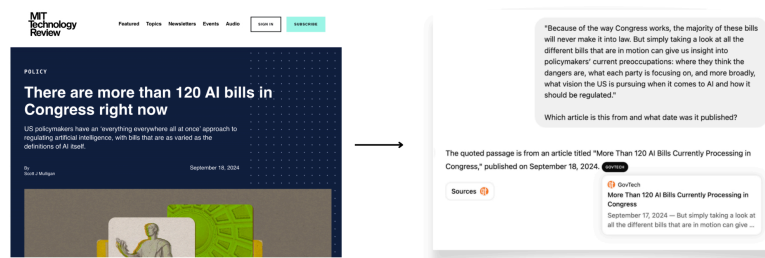
ChatGPT links to a syndicated version of an *MIT Tech Review* article instead of the canonical piece.

"Ultimately, I'm not sure what to make of it," said Honan in response to this finding. "But I've noticed similar things

happen elsewhere. For instance, I've seen instances where platforms like Perplexity, ChatGPT, or Google cite a rewritten version of a source instead of the original reporting. So, even if the attribution is technically correct, it's not the canonical source. As a publisher, that's not something you want to see. But there's so little recourse."

Publishers have expressed concern about how generative search might contribute to brand dilution—in other words, the risk that audiences might not know where the information they read is coming from. By treating journalism as decontextualized content with little regard for the circumstances in which an article was originally published, ChatGPT's search tool risks distancing audiences from publishers and incentivizing plagiarism or aggregation reporting over thoughtful, well-produced outputs.

UNPREDICTABLE (MIS)ATTRIBUTION

When we asked ChatGPT the same query multiple times, it typically returned a different answer each time. This inconsistency likely stems from the

default "temperature" setting of the underlying language model, GPT-4o. The temperature controls the randomness of the model's outputs: higher temperatures lead to variable responses, while lower temperatures lead to more deterministic ones.
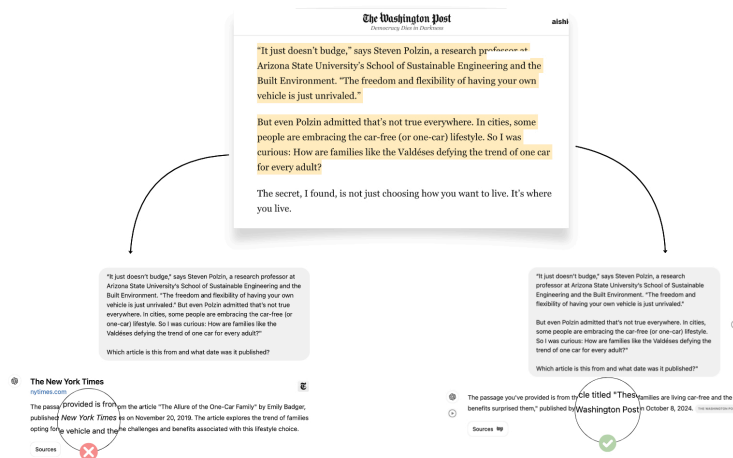
When prompted with the same query multiple times, ChatGPT might answer correctly on one occasion and incorrectly on another.

In the example above, we asked ChatGPT twice to identify a quote from an article published in the *Washington Post* on October 8, 2024. The first time, it cited the wrong date and attributed the story to the *New York Times* without attaching a source. The second time, it cited the correct article, identified the date, correctly attributed the story to the

Washington Post, and provided a working link to the article.

When a search tool operates with a temperature setting, it sacrifices consistency in favor of variation. While variety in sourcing can be valuable, prioritizing it over correctness or relevance can lead to inconsistent and inaccurate results. This is another factor that might harm the reliability of outputs and citations.

THE ILLUSION OF CONTROL

Our tests found that no publisher—regardless of degree of affiliation with OpenAI—was spared inaccurate representations of its content in ChatGPT.

The table below indicates the affiliation each of the publishers in our dataset have with OpenAI, whether the publisher's content was accessible to OpenAI's search crawler through their "robots.txt" file, and the accuracy of ChatGPT in referencing their content. Accuracy here is measured based on whether the chatbot correctly identified the publisher name, URL, and article date, with results showing the number of

correct, partially correct, and incorrect citations for each publisher.

Although more rigorous experimentation is needed to understand the true frequency of errors, our initial tests show a great deal of variability in the accuracy of ChatGPT's outputs that doesn't neatly match up with publishers' crawler status or affiliation with OpenAI. Even for publishers that have enabled access to all of OpenAI's crawlers (OAI-SearchBot, ChatGPT-User, and GPTBot), the chatbot does not reliably return accurate information about their articles. Both the *New York Post* and *The Atlantic*, for example, have licensing deals with OpenAI and enabled access to all the crawlers, but their content was frequently cited inaccurately or misrepresented.

Enabling crawler access does not guarantee a publisher's visibility in the OpenAI search engines either. For example, while *Mother Jones* and the *Washington Post* allow SearchGPT to crawl their content, quotes attributed to their publications were rarely identified by the chatbot.

"It may take a long time for a test case, involving a death sentence for child rape, to reach the Supreme Court. But the removal of unanimous juries in Florida could lead to new death sentences in the coming months. Last year, I spoke with Florida defense lawyers about their efforts to stop death sentences, and some told me they would race to secure as many life sentences as possible."

Which article is this from and what date was it published?

The passage you've prox from an articticle by David Leonhardt titled "Florida's Death Penalty Experiment," published ~~New York Times~~ es on May 2, 2023. In this piece, Leonhardt discusses Florida's legislative cha garding the dea eath penalty, including the removal of the unanimous jury requirement and the ~~tion of ca~~ apital punishment for child rape. He also highlights the concerns of defense attorneys are striving to secure life sentences for their clients in light of these developments.

Tow Center analysis.

Conversely, blocking the crawler completely doesn't entirely prevent a publisher's content from being surfaced. In the case of the *New York Times*, despite being engaged in a lawsuit and disallowing crawler access, ChatGPT Search still attributed quotes to the publication that were not from its articles.

CONCLUSION

Like [other platforms have done in the past](#), OpenAI is introducing a product that will likely have significant consequences for how audiences engage with news content, yet publishers are given little meaningful agency in the matter.

When reached for comment regarding our findings, a spokesperson for OpenAI

said, "Misattribution is hard to address
without the data and methodology that
the Tow Center withheld, and the study
represents an atypical test of our
product. We support publishers and
creators by helping 250M weekly
ChatGPT users discover quality content
through summaries, quotes, clear links,
and attribution. We've collaborated with
partners to improve in-line citation
accuracy and respect publisher
preferences, including enabling how they
appear in search by managing OAI-
SearchBot in their robots.txt. We'll keep
enhancing search results." The Tow
Center did describe our methodology
and observations to OpenAI, but we did
not share the data of our findings before
publication.

Steps such as giving some publishers a
seat at the table, honoring their stated
preferences in robots.txt, and creating
mechanisms to cite publisher content
reflect progress in the right direction, but
they represent the baseline of what a
responsible company should build into a
search product. For instance, Google's
documentation describes obeying the
instructions in a robots.txt file as

something that any "respectable web crawler" would do, warning users to be wary of rogue bots that may not honor their preferences. And while some publishers, like the [Associated Press and Condé Nast,](#) benefit from having licensing deals with OpenAI, not all newsrooms have a seat at the table.

The flaws and inconsistencies in the way publisher content is currently accessed and represented in ChatGPT Search seem counterproductive to OpenAI's stated goals and affect newsrooms regardless of how and if they chose to engage with OpenAI. Amid the company's drive to present its generative search product as an alternative to incumbents like Google and Bing, publishers are again caught between a rock and a hard place—they currently have no way of ensuring that their content will be presented accurately or even surfaced at all in ChatGPT's search.

Even those who permit crawlers or opt in to partnerships with OpenAI have to contend with the misrepresentation and misattribution of their content in

ChatGPT's search outputs. "I don't think publishers have a whole lot of leverage at this point," Honan said. "Until there are outcomes in [the court] cases, publishers won't have much leverage. And those outcomes could be bad for publishers—like if a decision finds that including this content in a model constitutes fair use, for example."

If OpenAI is serious about sustaining a good-faith collaboration with news publishers, it would do well to ensure that its search product represents and cites their content in an accurate and consistent manner and clearly states when the answer to a user query cannot be accessed.

See our [GitHub Repository](#) for the data behind this article.

This article has been updated to include The Marshall Project as a recipient of OpenAI funding through the Product & AI Studio at the American Journalism Project.

Has America ever needed a media defender more than now? Help us by [joining CJR today](#).