FIVE BLOCKS RESEARCH

REPUTATION REPORT · OCTOBER 2026

Which News Sites Can AI Read, and Which Does It Cite?

Part 1: which of 44 news outlets let eight AI engines in. Part 2: what those engines actually cited in 14,392 answers about the Fortune 100 and their CEOs.



EXECUTIVE SUMMARY

Most big outlets let some AI engines in and keep others out. Whether that matters depends on the engine and the question

This report keeps two things apart. **Access** is what each outlet's rules allow each engine to read. **Use** is what the engines actually showed as sources when we asked them about the Fortune 100. Every page is labeled.

ACCESS WHAT THE RULES ALLOW

Most outlets pick and choose. Of our 44 outlets, 25 block Google's Gemini app, 25 block Perplexity's crawler, 18 block Claude, 14 tag their pages to keep them out of Copilot and 12 block ChatGPT's search bot. None block Google's AI Overviews or AI Mode.

Twelve block no engine, four of them press release wires. The New York Times, the Wall Street Journal, Barron's, MarketWatch, the Economist and NPR leave only three of the eight unblocked.

Each engine reaches the news its own way. Google's Gemini app and its search answers follow different rules. Grok publishes no crawler name to block. Perplexity says its live fetcher generally ignores the rules.

Deals and paywalls add a layer. 12 outlets have content deals with OpenAI. Of the five paywalled sites we could test, four still send the full article in the page; The Information sends only a teaser.

USE WHAT THE ENGINES CITED

The question decides whether news shows up. Asked "tell me about" a company, Copilot cited one of our 44 outlets in 1% of answers and ChatGPT in 34%. Asked for the latest news, 68% and 87%.

A few outlets carry most of it. On news questions, Yahoo Finance, Reuters and CNBC were cited most. Stock and investing sites such as MarketBeat appeared in 40% to 60% of news answers on every engine.

The Gemini app barely cites the big outlets, 7% to 13% of answers on news questions, and leans on stock sites instead.

Answers move. Only 30% to 60% of the links in a news answer came back the next day.

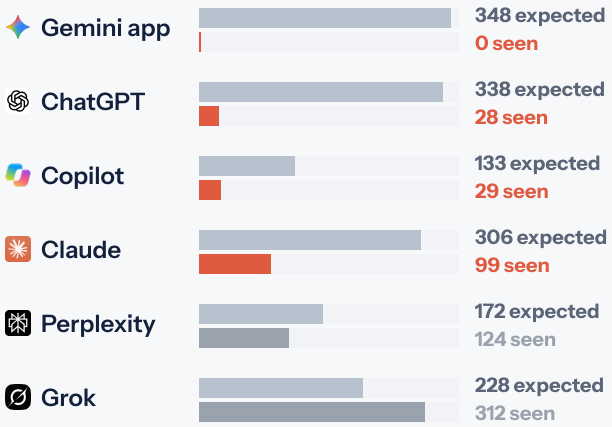
ACCESS vs USE DO THE RULES HOLD?

We compared what the rules allow with what the engines cited. The **Gemini app** cited no articles from the 25 outlets that block it, where Google's AI Mode cited them in 69% of news answers. **ChatGPT** cited outlets that block it in 28 news answers, against about 338 if the block made no difference. **Copilot** and **Claude** cite blocked outlets a fraction as often as expected. **Perplexity** and **Grok** cite them anyway.

ChatGPT's count rose each day of the test (1, 7, then 20), so its block is not airtight.

ANSWERS CITING OUTLETS THAT BLOCK THE ENGINE

Four news questions, three days: expected if the block made no difference vs seen



Rules from each site's robots.txt file and page tags read Oct 23, 2026; 36 of the 44 robots files were read again on Oct 26 and had not changed (8 sites refused the second request). Deals from public announcements. Source: Five Blocks AIQ, Oct 24-26, 2026. Eight AI engines, six questions about each Fortune 100 company and its CEO, asked daily (14,392 answers). Article pages only. A citation is a link the engine showed with its answer.

HOW TO READ THIS REPORT

Three labels, so you always know which question a number answers

ACCESS

WHAT THE RULES ALLOW

Can the engine read the outlet? The publisher's rules: robots.txt files, page tags, paywalls and content deals, read on October 23, 2026. Access is a fact about the outlet, the same whatever anyone asks.

Part 1, pages 4 to 7

USE

WHAT THE ENGINES CITED

Did the engine show the outlet as a source? The links engines showed with their answers, from 14,392 answers to six questions about the Fortune 100, October 24 to 26, 2026. Use depends on the question and the day.

Part 2, pages 9, 13 and 14

ACCESS vs USE

DO THE RULES HOLD?

Do the rules match the behavior? We put the two side by side: if an outlet blocks an engine, does that engine stop citing it? These pages say which numbers are rules and which are citations.

Part 2, pages 10 to 12

1 · THE QUESTION DECIDES THE SOURCES

We asked two kinds of questions. Answers to "Tell me about [company]" and "Tell me about [company]'s CEO" lean on the company's own site and Wikipedia, so they say little about news. Four more questions call for news: the latest news about the company and about its CEO, what analysts and investors are saying, and whether the company has been involved in controversies.

2 · ALLOWED IS NOT THE SAME AS CITED, AND A LINK SHOWN IS NOT THE SAME AS A SOURCE USED

An open outlet may still never be cited, and a blocked outlet can still appear, through copies on other sites or fetchers that ignore the rules. And we can only count the links an engine shows. An engine can read a page it never shows, such as a press release behind a news story. Nobody outside the AI companies can see what an engine read behind the scenes.

WHY THIS MATTERS NOW

Fewer people click through from AI than from search

4% vs 19%

of all adults surveyed in 27 markets say they always or often click through to sources from AI chatbots, vs from search

Reuters Institute Digital News Report 2026

79%

of the top 100 US and UK news sites block at least one AI training crawler

BuzzStream, December 2025

2.5B

monthly users of Google AI Overviews, the AI answer at the top of Google search

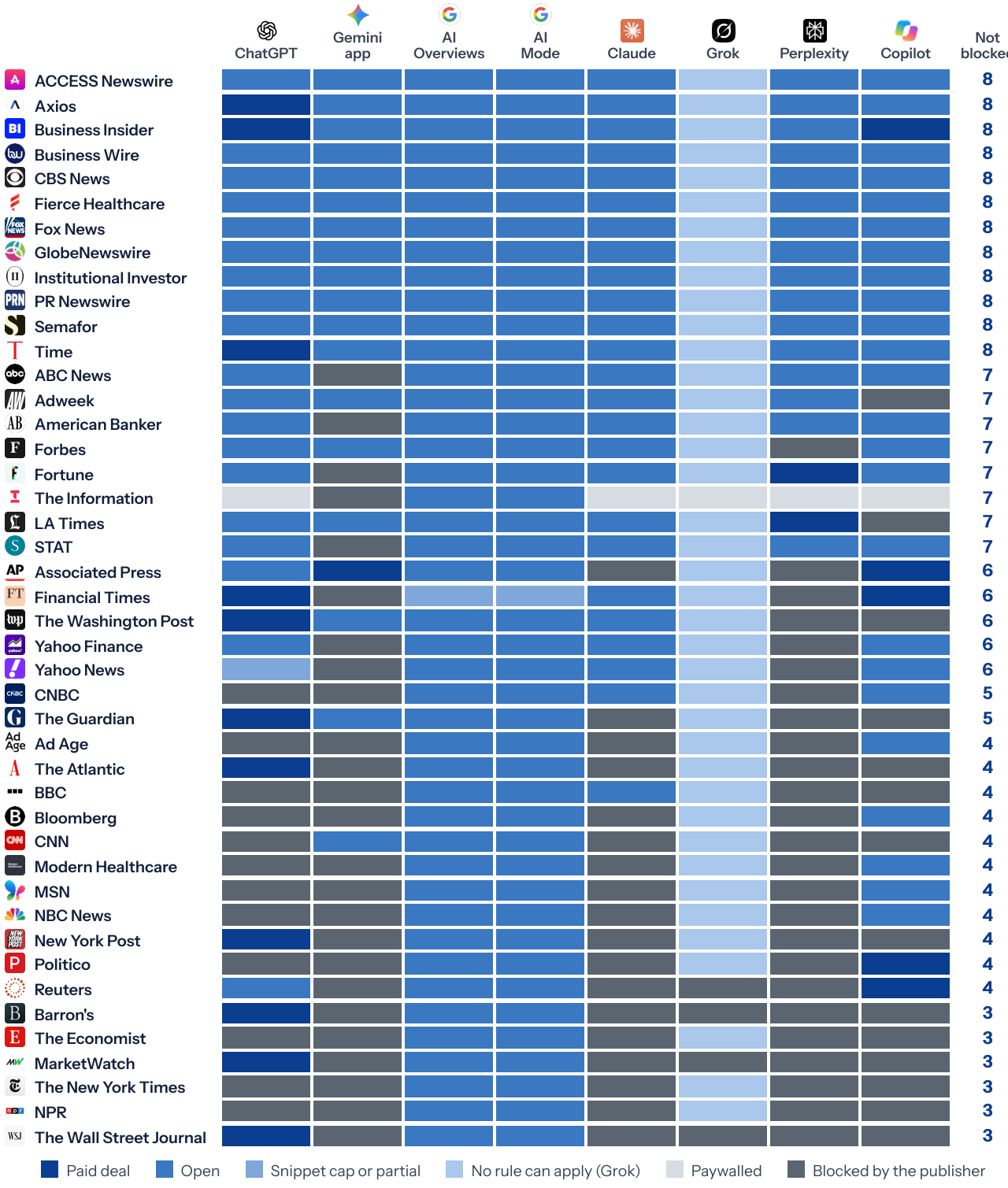
Google, June 2026

PART 1 · ACCESS · THE REPORT CARD

ACCESS WHAT THE RULES ALLOW

Each outlet's rule for each of eight engines

Color = the publisher's rule for that engine as of October 23, 2026. This page shows rules only, not what the engines cited (see Part 2). Last column = engines the rules do not block.



Rules from each site's robots.txt file and page tags read Oct 23, 2026; 36 of the 44 robots files were read again on Oct 26 and had not changed (8 sites refused the second request). Deals from public announcements. A block in robots.txt is shown even where a deal exists (Politico for ChatGPT). Grok counts as not blocked wherever no catch-all rule applies, because it publishes no crawler name. For Perplexity, gray means the site blocks its index crawler, PerplexityBot. Yahoo News blocks ChatGPT's search bot only on pages under /articles/, so it shows as partial.

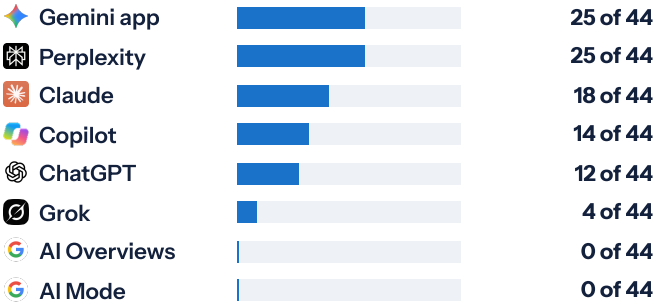
PART 1 · ACCESS · READING THE REPORT CARD

ACCESS WHAT THE RULES ALLOW

Twelve outlets block no engine. The Times and the Journal leave only three unblocked

OUTLETS THAT BLOCK EACH ENGINE

By robots.txt, page tag or catch-all rule, of our 44 outlets



AI Overviews and AI Mode use Google's regular search crawler. Blocking it would mean leaving Google Search, which no outlet here does. One outlet, the FT, limits how much Google may quote.

BLOCK NO ENGINE

ACCESS Newswire · Axios · Business Insider · Business Wire · CBS News · Fierce Healthcare · Fox News · GlobeNewswire · Institutional Investor · PR Newswire · Semafor · Time

NOT BLOCKED FOR ONLY 3 OF 8

Barron's · The Economist · MarketWatch · The New York Times · NPR · The Wall Street Journal

For these six, AI Overviews and AI Mode are open. The third is Grok for the NYT, the Economist and NPR (Grok has no crawler name to block), and ChatGPT, through News Corp's deal, for the WSJ, Barron's and MarketWatch.

Trade and TV split. Fierce Healthcare and Institutional Investor block no engine; Modern Healthcare and Ad Age block four. CBS and Fox News block none; NBC and CNN, four.

WHAT EACH BLOCK DOES

Different rules for different jobs

Search bots

Crawlers that fetch pages for AI answers, such as OAI-SearchBot (ChatGPT) and Claude-SearchBot. Blocking these is what keeps an outlet out of answers.

Training bots

Crawlers that collect text to train models, such as GPTBot. Many outlets block these but leave the search bots open.

Live fetchers

Fetches made while a person waits for an answer. OpenAI, Google and Perplexity say these may not follow robots.txt.

Page tags

Instructions inside each page. Bing says "noarchive" keeps a page out of Copilot answers.

Rules from each site's robots.txt file and page tags read Oct 23, 2026; 36 of the 44 robots files were read again on Oct 26 and had not changed (8 sites refused the second request). Deals from public announcements.

PART 1 · ACCESS · HOW EACH ENGINE REACHES THE NEWS

ACCESS WHAT THE RULES ALLOW

Each engine gets content its own way, so publishers have a different lever for each

ENGINE	HOW IT GETS CONTENT	WHAT A PUBLISHER CAN DO	DO LIVE FETCHES FOLLOW ROBOTS.TXT?
ChatGPT	Own search crawler (OAI-SearchBot), training crawler (GPTBot), live fetcher; also third-party search "such as Bing"	Block OAI-SearchBot (search answers) or GPTBot (training)	OpenAI: "robots.txt rules may not apply"
Gemini app	Google's index, if the site allows the Google-Extended token	Block Google-Extended (Google Search is not affected)	Google: user fetchers "generally ignore robots.txt rules"
AI Overviews and AI Mode	Google's regular search index (Googlebot); the same for AI Overviews	nosnippet or max-snippet tags limit what can be quoted	Answers come from the search index
Claude	Three Anthropic bots; web search through Brave and, since Oct 15, 2026, Parallel Web Systems	Block Anthropic's bots; Brave has no separate crawler name	Yes: the only vendor that says its live fetcher honors robots.txt
Grok	No crawler name in xAI's documentation	Only a catch-all rule or network blocking	Not documented
Perplexity	Index crawler (PerplexityBot) and live fetcher (Perplexity-User)	Block PerplexityBot	Perplexity: user fetcher "generally ignores robots.txt rules"
Copilot	Bing's index	Block bingbot (leaves Bing) or use the noarchive tag	No separate Copilot fetcher documented

GOOGLE · CRAWLER DOCUMENTATION

"Google-Extended does not impact a site's inclusion in Google Search"

So an outlet can keep the Gemini app out and still appear in Google's AI Overviews and AI Mode. Google runs two AIs on two sets of rules.

BING · WEBMASTER GUIDELINES

"NOARCHIVE prevents content from being used in Copilot responses and grounding results."

The tag is on the NYT, WSJ, Washington Post, Barron's, MarketWatch, LA Times, NPR, CNN, NY Post, the Atlantic, the Guardian, the BBC, the Economist and Adweek. AP and the FT carry it too but have Microsoft deals.

PERPLEXITY · CRAWLER DOCUMENTATION

"Perplexity-User... generally ignores robots.txt rules"

OPENAI · CRAWLER DOCUMENTATION

"robots.txt rules may not apply"

OpenAI says this about fetches made for a user. Its search bot is the one publishers block.

Source: each vendor's crawler documentation, fetched October 2026. Tags checked on cited article pages, Oct 23, 2026. We tested the Copilot app; all Copilots ground answers in Bing.

PART 1 · ACCESS · DEALS AND PAYWALLS

ACCESS WHAT THE RULES ALLOW

Deals and paywalls are a second layer of access

CONTENT DEALS WITH AI COMPANIES AMONG OUR 44 OUTLETS

AI COMPANY	OUTLETS WITH A CURRENT DEAL
OpenAI (ChatGPT)	The Wall Street Journal, The Washington Post, Financial Times, The Atlantic, Business Insider, Politico, Time, The Guardian, New York Post, Barron's, MarketWatch, Axios. Politico's deal stands, but its robots file still blocks OpenAI's search bot.
Microsoft (Copilot)	Financial Times, Associated Press, Reuters, Business Insider, Politico
Perplexity	LA Times, Fortune. Revenue-sharing publisher program.
Google (Gemini app)	Associated Press

Deals usually cover display in answers and sometimes training. Their terms are rarely public, so a deal tells you an engine may use an outlet, not how often it will show it. Google's paid News AI pilot (more than 400 publications, including the FT, CNN, Axios and Forbes) is not tied to any one AI engine, so we do not count it above.

PAYWALL TYPE, SELECTED OUTLETS

OUTLET	PAYWALL
The Wall Street Journal	Hard paywall
Barron's	Hard paywall
Financial Times	Mostly subscriber-only
The Information	Hard paywall, teaser only
The New York Times	Metered
Bloomberg	Metered
Fortune	Metered
The Washington Post	Metered
Reuters	Registration wall
CNBC	Free
Forbes	Free
Business Insider	Free

A paywall stops readers, not always crawlers. Of the five paywalled sites we could test, four (Fortune, CNN, the LA Times and American Banker) send the full article in the page and hide it from readers with a script. The Information sends only a teaser. The WSJ, NYT, FT and Bloomberg blocked our test requests, so we could not check them.

Deals from public announcements and the publishers' and AI companies' sites, October 2026; lawsuits are not counted as rules. Paywall type from each outlet's pages and published policies, October 2026.

PART 2 · USE · WE PUT IT TO THE TEST

USE WHAT THE ENGINES CITED

Six questions about every Fortune 100 company, eight engines, three days

What we asked. Five Blocks AIQ asked eight AI engines six questions about each Fortune 100 company every day from October 24 to 26, 2026. Two are general questions, the AI version of googling a company. Four call for news. That makes 14,392 answers (the Walmart CEO question ran on two of the three days).

What we counted. The links each engine showed with its answer. We count an outlet as cited when the engine linked to one of its article pages, not a stock quote or company profile page. We compare engines and questions within these three days.

What it can and cannot show. This is one sample: large US companies, six questions, three days. It shows how the engines behave on questions people really ask about big companies. It cannot show what an engine read but did not link, and three days is too short to call a trend.

THE EIGHT ENGINES

- ChatGPT
- Gemini app
- AI Overviews
- AI Mode
- Claude
- Grok
- Perplexity
- Copilot

ANSWERS PER QUESTION

All eight engines, three days

	Question	Type	Answers
1	Tell me about [company].	General	2,400
2	Tell me about [company]'s CEO.	General	2,392
3	What is the latest news about [company]?	News	2,400
4	What is the latest news about [company]'s CEO?	News	2,400
5	What are analysts and investors saying about [company]?	News	2,400
6	Has [company] been involved in any controversies?	News	2,400

WHY TWO KINDS OF QUESTIONS

The two general questions show how AI describes a company. The four news questions show which outlets AI turns to for current events. Asking both on the same days, about the same companies, shows how much the question itself decides which sources appear.

Source: Five Blocks AIQ, Oct 24-26, 2026. Eight AI engines, six questions about each Fortune 100 company and its CEO, asked daily (14,392 answers). Article pages only. A citation is a link the engine showed with its answer.

USE WHAT THE ENGINES CITED U1 · THE QUESTION DECIDES

Ask about a company and the engines describe it from its own pages. Ask for news and the big outlets appear

1% → 68%

of Copilot answers cited one of our 44 outlets when asked "tell me about" a company vs asked for its latest news.

Every engine except the Gemini app cites the big outlets far more often on news questions. ChatGPT goes from 34% to 87%, AI Mode from 17% to 59%. Questions about the CEO sit in between: once the question is about a person, the press comes in.

SHARE OF ANSWERS CITING AN ARTICLE FROM ONE OF THE 44 OUTLETS

	Tell me about the company	Tell me about the CEO	Latest news: company	Latest news: CEO	Analysts and investors	Controversies
ChatGPT	34%	63%	87%	84%	85%	76%
Gemini app	2%	3%	7%	10%	1%	13%
AI Overviews	9%	38%	54%	78%	90%	78%
AI Mode	17%	51%	59%	80%	81%	74%
Claude	7%	39%	22%	63%	55%	50%
Grok	6%	44%	62%	44%	35%	77%
Perplexity	48%	59%	73%	81%	69%	81%
Copilot	1%	28%	68%	61%	24%	29%

WHAT TO DO Before judging how AI treats your coverage, check which question you are testing. "Tell me about" answers come from your own pages and Wikipedia; news questions come from the press.

Source: Five Blocks AIQ, Oct 24-26, 2026. Eight AI engines, six questions about each Fortune 100 company and its CEO, asked daily (14,392 answers). Article pages only. A citation is a link the engine showed with its answer.

WHAT THE ENGINES CITED INSTEAD: SHARE OF ANSWERS CITING EACH TYPE OF SOURCE

	"Tell me about [company]"			"What is the latest news about [company]?"		
	Own website	Wikipedia	News articles	Own website	Wikipedia	News articles
ChatGPT	99%	5%	44%	96%	2%	95%
Gemini app	65%	54%	11%	77%	0%	64%
AI Overviews	96%	96%	10%	90%	1%	70%
AI Mode	89%	89%	22%	80%	1%	73%
Claude	32%	16%	10%	44%	0%	56%
Grok	87%	80%	18%	70%	0%	92%
Perplexity	98%	26%	65%	94%	0%	90%
Copilot	83%	81%	2%	77%	0%	80%

News articles here = article pages from any site AIQ classes as news, not only our 44 outlets, which is why these shares are higher. Own website matched by company domain list.

ACCESS vs USE DO THE RULES HOLD? U2 · DO THE BLOCKS HOLD?

If an outlet blocks the Gemini app, ChatGPT or Copilot, those engines rarely cite it. Perplexity and Grok cite blocked outlets anyway

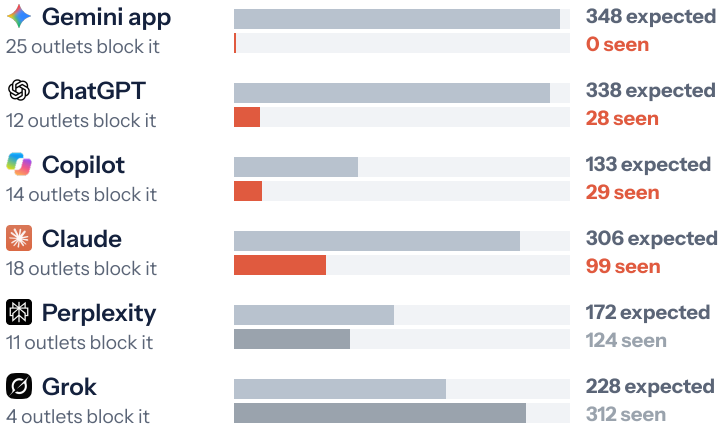
0 vs 348

Gemini app answers citing the 25 outlets that block it: seen vs expected if the block made no difference. ChatGPT: 28 vs about 338.

Access: which outlets block each engine (Part 1). **Use:** how often each engine cited those outlets anyway. The blocked outlets are cited a fraction as often as expected by the Gemini app, ChatGPT, Copilot and Claude. Blocks do much less against Perplexity and nothing against Grok.

ANSWERS CITING OUTLETS THAT BLOCK THE ENGINE, FOUR NEWS QUESTIONS

Expected (rules not followed): the engine's rate for outlets that let it in, scaled by how much more or less Google AI Mode cites the blocked outlets than the open ones. **Seen:** what the engine actually cited in three days.



WHAT TO DO Check which engines can use each outlet on your media list. A story in an outlet that blocks an engine will rarely reach that engine's users as a link.

Source: Five Blocks AIQ, Oct 24-26, 2026. Eight AI engines, six questions about each Fortune 100 company and its CEO, asked daily (14,392 answers). Article pages only. A citation is a link the engine showed with its answer. Rules from each site's robots.txt file and page tags read Oct 23, 2026; 36 of the 44 robots files were read again on Oct 26 and had not changed (8 sites refused the second request). Deals from public announcements. Zero links does not prove an outlet's reporting never shaped an answer.

Gemini app

Zero articles from the 25 outlets that block it, on every question and every day. It cites few big outlets at all.

ChatGPT

Seen rose each day: 1, 7 and 20 answers, mostly Bloomberg and CNBC. Their robots files had not changed.

Copilot

The 14 outlets with the noarchive tag and no Microsoft deal. The WSJ, Washington Post and Guardian still appear now and then.

Claude

Seen is about a third of expected, mostly Bloomberg, NPR, NBC and CNN. Reuters, which blocks Claude, never appeared.

Perplexity

NYT articles appeared in 59 answers, the BBC in 28. Its live fetcher says it generally ignores robots.txt.

Grok

Grok publishes no crawler name, so only 4 outlets block it, with a catch-all rule. Reuters alone appeared in 280 answers.

ACCESS vs USE DO THE RULES HOLD? U3 · GOOGLE AND COPILOT

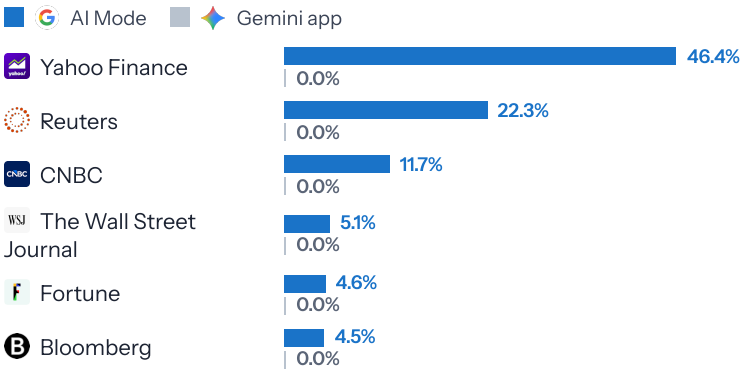
Google's two AIs follow different rules, and the results show it

69% vs 0%

of AI Mode vs Gemini app news answers cited articles from the 25 outlets that block the Gemini app.

Access: the Gemini app respects the Google-Extended block; AI Mode uses Google's search crawler, which these outlets leave open. **Use:** AI Mode cites them constantly, the Gemini app never.

SHARE OF NEWS ANSWERS CITING THE OUTLET



All six block the Gemini app's token (access). None block AI Mode.

WHAT TO DO Treat the Gemini app and Google's search answers as two audiences. An outlet that blocks Gemini can still reach AI Mode and AI Overviews.

Source: Five Blocks AIQ, Oct 24-26, 2026. Eight AI engines, six questions about each Fortune 100 company and its CEO, asked daily (14,392 answers). Article pages only. A citation is a link the engine showed with its answer. Rules from each site's robots.txt file and page tags read Oct 23, 2026; 36 of the 44 robots files were read again on Oct 26 and had not changed (8 sites refused the second request). Deals from public announcements.

ACCESS vs USE DO THE RULES HOLD? U4 · COPILOT

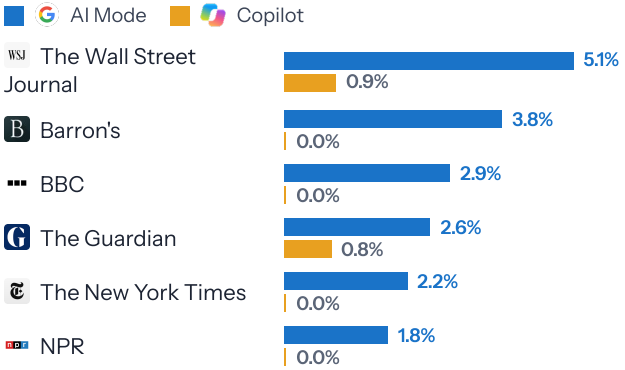
A one-word page tag keeps much of the national press out of Copilot

2.4% vs 17.5%

of Copilot vs AI Mode news answers cited the 14 outlets with the noarchive tag and no Microsoft deal.

Access: Bing says the tag keeps pages out of Copilot answers. **Use:** Copilot cited those outlets about a seventh as often as AI Mode, mostly on controversy questions.

SHARE OF NEWS ANSWERS CITING THE OUTLET



All six carry the noarchive tag and have no Microsoft deal (access).

WHAT TO DO For audiences that use Copilot, a Journal or Times story will rarely show up as a link. Plan on outlets without the tag.

Source: Five Blocks AIQ, Oct 24-26, 2026. Eight AI engines, six questions about each Fortune 100 company and its CEO, asked daily (14,392 answers). Article pages only. A citation is a link the engine showed with its answer. Tags checked on cited articles, Oct 23, 2026.

ACCESS vs USE DO THE RULES HOLD? U5 · DEALS

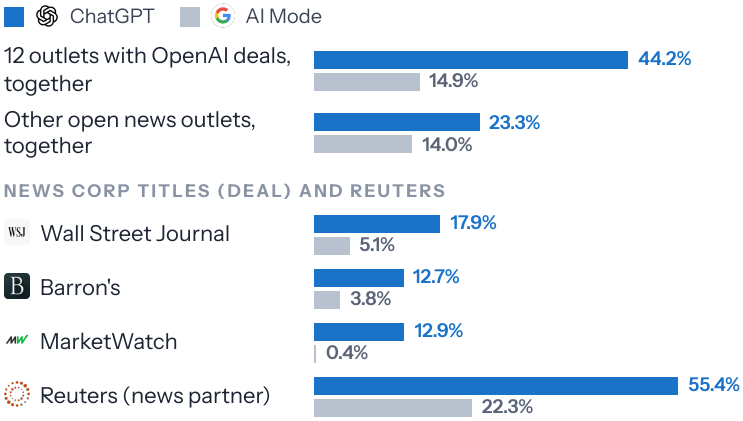
ChatGPT cites outlets it has deals with about twice as often as other open outlets

44% vs 23%

of ChatGPT news answers cite the 12 outlets with OpenAI deals vs other open news outlets. Google's AI Mode, with no such deals, shows almost no gap.

Access: 12 of our outlets have OpenAI content deals (Part 1). **Use:** ChatGPT cites them far more than AI Mode does; the News Corp titles stand out most. Reuters, which OpenAI named as a news partner when ChatGPT search launched, is still ChatGPT's single biggest source.

CHATGPT VS AI MODE (YARDSTICK), FOUR NEWS QUESTIONS



WHAT TO DO Count a deal as a real plus for ChatGPT on news questions, though not a guarantee: Reuters, a named news partner, is cited more than all the deal outlets together.

Source: Five Blocks AIQ, Oct 24-26, 2026. Eight AI engines, six questions about each Fortune 100 company and its CEO, asked daily (14,392 answers). Article pages only. A citation is a link the engine showed with its answer. Other open news outlets let OpenAI in and have no current OpenAI content deal, excluding Yahoo Finance and the press release wires.

ACCESS vs USE DO THE RULES HOLD? U6 · PAYWALLS

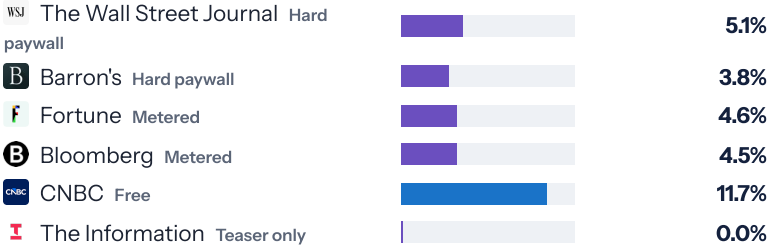
A hard paywall did not keep the Wall Street Journal out of Google's AI answers

5.1%

of AI Mode news answers cited the WSJ, behind a hard paywall. The Information, which sends crawlers only a teaser, was almost never cited by any engine.

Access: most paywalled sites still let crawlers read the article. **Use:** paywalled outlets are cited alongside free ones. What a crawler can read matters more than what a reader pays.

AI MODE NEWS ANSWERS CITING THE OUTLET



WHAT TO DO Do not treat a paywalled outlet as invisible to AI. Check the engine's crawler rules for that outlet as well as the paywall.

Source: Five Blocks AIQ, Oct 24-26, 2026. Eight AI engines, six questions about each Fortune 100 company and its CEO, asked daily (14,392 answers). Article pages only. A citation is a link the engine showed with its answer. Paywall type from each outlet's pages and published policies, October 2026.

USE WHAT THE ENGINES CITED U7 · WHO GETS CITED

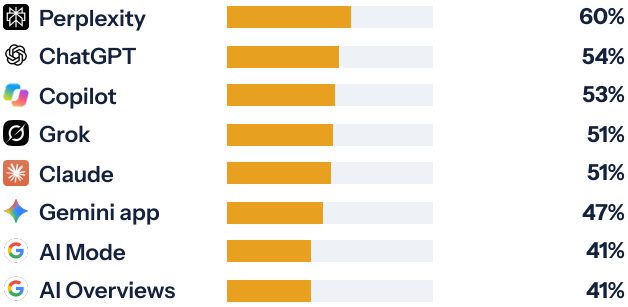
On news questions, a handful of outlets and a crowd of stock sites fill the answers

Yahoo Finance

was the most-cited of our 44 outlets on news questions, followed by Reuters and CNBC.

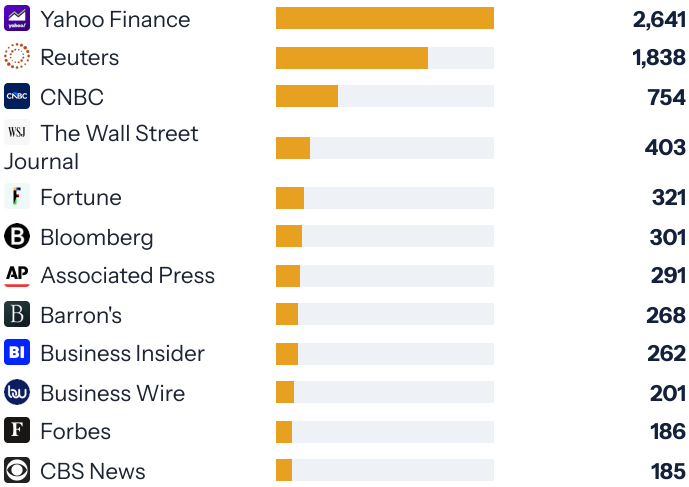
Stock and investing sites, such as MarketBeat, Investing.com, Seeking Alpha and StockAnalysis, appear in 40% to 60% of news answers on every engine. They are not among our 44 outlets, but they compete with them for the same answers. The Gemini app, which 25 of our outlets block, leans on them most: its top sources were marketbeat.com (21%), investing.com (9%), en.wikipedia.org (9%).

NEWS ANSWERS CITING A STOCK OR INVESTING SITE



MOST-CITED OF THE 44 OUTLETS, FOUR NEWS QUESTIONS

Answers citing the outlet's articles, all eight engines, three days



Open does not mean cited: Fox News, which blocks no engine, appeared in 15 of 9,600 news answers. The Economist, which blocks five engines, appeared in none.

WHAT TO DO Ask which sites actually show up for your company, not only which outlets ran the story. A stock site may be telling your news story to more AI users than the outlet that broke it.

Source: Five Blocks AIQ, Oct 24-26, 2026. Eight AI engines, six questions about each Fortune 100 company and its CEO, asked daily (14,392 answers). Article pages only. A citation is a link the engine showed with its answer. Stock and investing sites: a list of 23 domains, including MarketBeat, Investing.com, Seeking Alpha, StockAnalysis, TipRanks, Simply Wall St, Zacks and Benzinga.

USE WHAT THE ENGINES CITED U8 · CHANGE

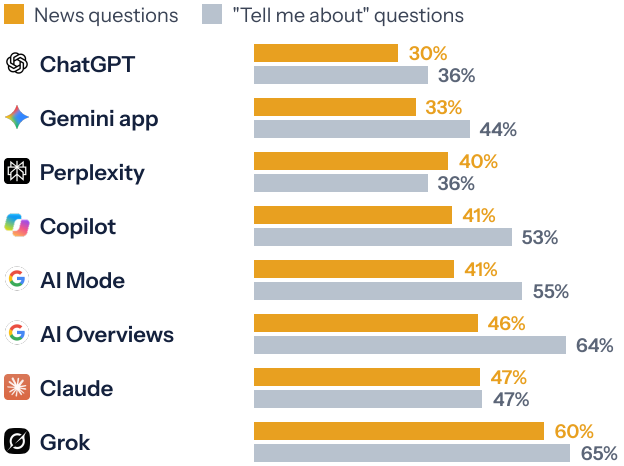
Ask the same question tomorrow and about half the links change

30% to 60%

of the links in an engine's news answer came back the next day for the same question. ChatGPT is lowest; Grok is highest.

General "tell me about" answers are steadier on most engines, because they rest on the company's own site and Wikipedia. One screenshot of an AI answer tells you little; measure over weeks.

LINKS BACK THE NEXT DAY, SAME ENGINE AND QUESTION



WHAT TO DO Track the questions your stakeholders ask, engine by engine, over weeks.

Source: Five Blocks AIQ, Oct 24-26, 2026. Eight AI engines, six questions about each Fortune 100 company and its CEO, asked daily (14,392 answers). Article pages only. A citation is a link the engine showed with its answer. Two day pairs per question (Oct 24-25, Oct 25-26). All links, not just news. Links compared by full page address.

USE WHAT THE ENGINES CITED U9 · PRESS RELEASES

Every engine is allowed to read the release wires. Few show release pages as sources

3.4%

of Perplexity news answers link a Business Wire release page, the highest of any engine. None of the four wires blocks any engine (access).

Low counts do not mean the engines ignore releases. An engine can take facts from a release and show a news story or the company's own site instead.

NEWS ANSWERS LINKING A RELEASE PAGE ON THE WIRE'S SITE

	ChatGPT	Gemini app	AI Overviews	AI Mode	Claude	Grok	Perplexity	Copilot
Business Wire	3.1%	1.6%	2.1%	2.2%	0.9%	0.8%	3.4%	2.6%
PR Newswire	2.9%	1.2%	1.0%	1.8%	1.0%	1.8%	2.2%	1.9%
GlobeNewswire	1.3%	0.1%	0.0%	0.3%	0.3%	0.8%	1.3%	0.0%
ACCESS Newswire	0.0%	0.2%	0.2%	0.2%	0.0%	0.0%	0.1%	0.0%

Source: Five Blocks AIQ, Oct 24-26, 2026. Eight AI engines, six questions about each Fortune 100 company and its CEO, asked daily (14,392 answers). Article pages only. A citation is a link the engine showed with its answer.

ENGINE BY ENGINE

Eight engines, eight ways of reaching and using the news

	ACCESS · THE RULES		USE · WHAT THEY CITED (OCT 24-26)					
ENGINE	OUTLETS THAT BLOCK IT (OF 44)	HOW IT REACHES THE NEWS	CITES ONE OF THE 44: "TELL ME ABOUT" COMPANY	CITES ONE OF THE 44: NEWS QUESTIONS (AVG)	BLOCKED OUTLETS: SEEN OF EXPECTED	TOP OUTLET ON NEWS QUESTIONS	STOCK OR INVESTING SITE, NEWS QUESTIONS	LINKS BACK NEXT DAY, NEWS
ChatGPT	12	Own search bot; deals with 12	34%	83%	28 of ~338	Reuters 55%	54%	30%
Gemini app	25	Google index if Google-Extended allowed	2%	8%	0 of ~348	CBS News 2%	47%	33%
AI Overviews	0	Google search index	9%	75%	n/a	Yahoo Finance 44%	41%	46%
AI Mode	0	Google search index	17%	74%	n/a	Yahoo Finance 46%	41%	41%
Claude	18	Own bots; Brave search	7%	47%	99 of ~306	Yahoo Finance 24%	51%	47%
Grok	4	No crawler name	6%	55%	312 of ~228	Reuters 23%	51%	60%
Perplexity	25	Own crawler; fetcher ignores rules	48%	76%	124 of ~172	Yahoo Finance 41%	60%	40%
Copilot	14	Bing index; noarchive tag	1%	46%	29 of ~133	Yahoo Finance 27%	53%	41%

Rules from each site's robots.txt file and page tags read Oct 23, 2026; 36 of the 44 robots files were read again on Oct 26 and had not changed (8 sites refused the second request). Deals from public announcements. Source: Five Blocks AIQ, Oct 24-26, 2026. Eight AI engines, six questions about each Fortune 100 company and its CEO, asked daily (14,392 answers). Article pages only. A citation is a link the engine showed with its answer. AI Overviews and AI Mode are blocked by none of the 44 in robots files or page tags, so there is nothing to compare. Perplexity's blocked count is for its index crawler; the seen-vs-expected test uses the 11 outlets that block both its crawler and its live fetcher by name.

WHAT WE ARE TESTING NEXT

Questions this data cannot answer yet

Is ChatGPT's block weakening?

Its citations of blocked outlets rose each day of this test. We will keep measuring.

Exclusive or wire?

Which reaches more AI answers, and how fast each engine picks up a new story.

Do corrections reach AI?

Corrected stories, tracked for 30 days.

Beyond the Fortune 100

The same questions for mid-size companies.

WHAT THIS MEANS FOR COMMUNICATORS

Five implications, each labeled with the evidence behind it

01 ACCESS **Add one column to every media list: which engines can use the outlet.** Read it off Part 1. The Times, the Journal and the BBC block several engines; Fox News, CBS and Axios block none.

02 ACCESS vs USE **Expect blocks to matter for the Gemini app, ChatGPT and Copilot, and not much for Perplexity or Grok.** Those three rarely cited outlets that block them; Perplexity and Grok cited them anyway.

03 USE **Make sure the core facts live where every engine can read them.** On "tell me about" questions the engines lean on your own site and Wikipedia. Keep them current, and check that your own robots.txt does not block the AI search bots.

04 USE **Plan for the stock sites.** On news questions, sites like MarketBeat appear in about half of answers on every engine. Know what they say about you.

05 USE **Measure by question and over weeks.** News answers change daily. Test the questions your stakeholders ask, engine by engine, before relying on one screenshot.

WHAT IT TAKES

Earned media still works. It now needs a second job done

A placement is only the start. What matters is whether the facts reach the answers your stakeholders read, engine by engine, and stay there.

An open record AI can cite

Company pages built around the questions people ask AI, refined as the answers change.

Placements planned for AI reach

Media lists that show which engines can use each outlet, alongside audience size.

Disclosed work on Wikipedia

Through Wikipedia's approved process, with an expert firm, on the articles AI reads.

Steady tracking

How each engine answers, measured every day, so shifts in sources are caught early.

Method. Access: robots.txt files for 44 outlets read October 23, 2026 and scored against each engine's published crawler names, with page tags on cited articles, paywall checks and announced deals; 36 robots files were read again on October 26 and had not changed. Use: Five Blocks AIQ asked eight engines six questions about each Fortune 100 company daily from October 24 to 26, 2026 (14,392 answers); only article pages count. For blocking, expected answers assume the engine treats blocked outlets like the outlets that let it in, adjusted by how often Google AI Mode cites each group. The findings were checked by AI agents acting as a fact-checker and a skeptic, and by Five Blocks.

SOURCES AND NOTES

Where every number comes from

Vendor documentation

OpenAI crawlers (developers.openai.com/api/docs/bots); OpenAI ChatGPT search help (help.openai.com/en/articles/9237897); Google crawlers and user-triggered fetchers (developers.google.com/crawling/docs/crawlers-fetchers); Google AI features (developers.google.com/search/docs/appearance/ai-features); Google robots meta tag (developers.google.com/search/docs/crawling-indexing/robots-meta-tag); Anthropic crawler help (support.claude.com/en/articles/8896518); Anthropic subprocessors (trust.anthropic.com/subprocessors); Perplexity crawlers (docs.perplexity.ai/guides/bots); Brave Search crawler (search.brave.com/help/brave-search-crawler); Bing Webmaster Guidelines (bing.com/webmasters/help/webmaster-guidelines-30fba23a); Bing robots meta tags (bing.com/webmasters/help/robots-meta-tags-and-attributes-that-bing-supports-5198d240). All fetched October 2026.

Deals

OpenAI partnership announcements (openai.com); Microsoft Publisher Content Marketplace, February 2026 (about.ads.microsoft.com); Perplexity Publisher Program (perplexity.ai); Google News AI pilot update, October 15, 2026, and publisher partnerships, December 10, 2025 (blog.google).

Third-party studies

Reuters Institute, Digital News Report 2026 (reutersinstitute.politics.ox.ac.uk); BuzzStream, "Which News Sites Block AI Crawlers," December 2025 (buzzstream.com/blog); Google blog, June 2026.

Five Blocks data

AIQ export of 14,392 answers to six questions, October 24-26, 2026; robots.txt files fetched October 23 and 26, 2026; page tags checked October 23, 2026; paywall checks October 2026.

Limits that matter

Fortune 100 companies only, six questions, three days; links shown, not hidden sources. Rules dated October 23, 2026; eight sites (CNBC, Politico, Axios, Business Wire, GlobeNewswire, Modern Healthcare, PRWeek and Ad Age) refused the October 26 recheck. Page tags could not be checked for GlobeNewswire, ACCESS Newswire, Time, Axios or Fierce Healthcare, so their Copilot status rests on their robots files. PRWeek's robots file could not be retrieved, so it is not scored. Fortune, which publishes the Fortune 100 list, is among the most-cited outlets, which may partly reflect the companies we chose.

This research was done with the help of Claude and other AI tools, with human review throughout.

ABOUT FIVE BLOCKS

Five Blocks tracks, analyzes and improves the digital reputations of companies and their leaders around the world

Since 2003, Five Blocks has worked on how organizations are seen on Google, on Wikipedia and now in AI answers. Our IMPACT and AIQ platforms follow tens of thousands of brands every day across search results and the leading AI engines.

This research shows which news outlets each AI engine is allowed to read and which it actually cites. It is meant to help communications teams plan coverage for the way people now look companies up. For more information, please contact us.

CONTACT

Sam Michelson

Founder and CEO, Five Blocks
samm@fiveblocks.com
+1.646.801.8669

WEBSITE

fiveblocks.com

RELATED RESEARCH

Reputation Report: How AI Describes Data Centers (October 2026)
Reputation Report: How AI Describes 25 of the Largest AI Data Center Projects (October 2026)