

Summarization report, cutoff Oct 26, 2026

As of Oct 26, 2026, aiforlawyers.app counts 286 lawyer voices on summarization; 253 voices name one of 12 tools, 8 of which have 10+ voices and are rated. Highest floor of the 90% band: Claude (net +59, band +33 to +75, 27 voices, Low confidence).

Online version, data and updates: <https://aiforlawyers.app/report/summarization>

CITE THIS

aiforlawyers.app (AI Vortex), Summarization report, cutoff Oct 26, 2026,
<https://aiforlawyers.app/report/summarization>, licensed CC BY 4.0.

For one figure: “As of Oct 26, 2026, aiforlawyers.app counts 286 lawyer voices on summarization; 253 voices name one of 12 tools, 8 of which have 10+ voices and are rated. Highest floor of the 90% band: Claude (net +59, band +33 to +75, 27 voices, Low confidence).”

Every tool, with its voices

12 tools named in 286 voices on summarization. Rated tools first, ordered by the floor of the 90% band; then the tools with few voices, with their count and no band. Net = like minus dislike per 100 voices.

#	Tool	Status	Voices	Net	90% band	Confidence	Top complaint (voices)
1	Claude	Rated	27	+59	+33 to +75	Low	Accuracy (4)
2	ChatGPT	Rated	114	+5	-8 to +18	Medium	Accuracy (24)
3	Gemini	Rated	11	+36	-10 to +68	Low	Accuracy (2)
4	Copilot	Rated	19	+11	-19 to +38	Low	Accuracy (9)
5	Harvey	Rated	19	+11	-24 to +41	Low	Accuracy (4)
6	Westlaw AI	Rated	11	-9	-43 to +29	Low	Accuracy (6)
7	CoCounsel	Rated	33	-30	-49 to -7	Low	Accuracy (20)
8	Lexis+ AI	Rated	10	-40	-70 to +5	Low	Accuracy (6)
-	Clio	Few voices	3	-	no band: 3 of 10 voices needed	not rated	Accuracy (3)
-	Legora	Few voices	3	-	no band: 3 of 10 voices needed	not rated	Accuracy (2)
-	Perplexity	Few voices	2	-	no band: 2 of 10 voices needed	not rated	none on record

#	Tool	Status	Voices	Net	90% band	Confidence	Top complaint (voices)
-	Spellbook	Few voices	1	-	no band: 1 of 10 voices needed	not rated	none on record

What lawyers complain about, in their words

The most named complaints per tool, with a verbatim quote each. Counts are voices, not opinions of the firm or of us.

Claude Rated · 27 voices

6 voices of 27 criticize it or are mixed about it with a named complaint: accuracy (4), trust, liability (1), price (1).

"I tried to use Claude to create an index for me recently to a tranche of documents and I had to hand feed it small batches and then paste together the small tables it was able to create per batch."

r/Lawyertalk · Sep 4, 2026 · Accuracy

"Maybe US statutes are different but if you are an Australian lawyer asking Claude to summarise an Act, you should have your practising certificate under review."

Bluesky · Sep 8, 2026 · Trust, liability

ChatGPT Rated · 114 voices

60 voices of 114 criticize it or are mixed about it with a named complaint: accuracy (24), hallucination (18), privacy (11), trust, liability (6), integration (1).

"Not one sentence in his ChatGPT summary applied to his case."

r/Lawyertalk · Jan 9, 2026 · Accuracy

"Both times it hallucinated testimony and even quoted testimony."

r/Lawyertalk · Jun 1, 2026 · Hallucination

Gemini Rated · 11 voices

4 voices of 11 criticize it or are mixed about it with a named complaint: accuracy (2), hallucination (1), privacy (1).

"Gemini is definitely not accurate. Take any deposition and check the pincites Gemini provides and you will find many inaccuracies."

r/LawFirm · Oct 1, 2025 · Accuracy

"Gemini hallucinates like hell so I cannot really work with a 66 page legal pdf efficiently"

Hacker News · Feb 17, 2024 · Hallucination

Copilot Rated · 19 voices

11 voices of 19 criticize it or are mixed about it with a named complaint: accuracy (9), privacy (2).

“They absolutely trashed most conversation threads with barren slop and mixed up convo threads.”

Hacker News · Jun 18, 2026 · Accuracy

“Microsoft Copilot read and summarized confidential emails for almost a month.”

Bluesky · Mar 3, 2026 · Privacy

Harvey Rated · 19 voices

10 voices of 19 criticize it or are mixed about it with a named complaint: accuracy (4), hallucination (3), trust, liability (1), price (1), integration (1).

“Even when I add docs to a vault and ask for a review table, getting Harvey to actually “read” the docs is like pulling teeth.”

r/biglaw · Oct 15, 2026 · Accuracy

“As a paralegal i quality checked many such summaries done via harvey and found that they were largely unreliable/they would hallucinate facts and arguments”

r/biglaw · Aug 5, 2025 · Hallucination

Westlaw AI Rated · 11 voices

8 voices of 11 criticize it or are mixed about it with a named complaint: accuracy (6), hallucination (1), price (1).

“The new ones that I suspect are AI read like generic ChatGPT-voice, and the substance is borderline inaccurate.”

r/biglaw · Sep 25, 2025 · Accuracy

“I uploaded a federal regulation into westlaw and it broke/hallucinated on me to no end.”

r/LawFirm · Apr 25, 2025 · Hallucination

CoCounsel Rated · 33 voices

26 voices of 33 criticize it or are mixed about it with a named complaint: accuracy (20), price (3), hallucination (2), privacy (1).

“Had a rep demo co-counsel 2.0 today and it does nothing that 365 Copilot for enterprise (which is fully sandboxed) doesn't do.”

r/biglaw · Mar 31, 2026 · Accuracy

“Co-Counsel is not true AI. It's a Mechanical Turk (look it up).”

r/LawFirm · Dec 10, 2024 · Price

Lexis+ AI Rated · 10 voices

8 voices of 10 criticize it or are mixed about it with a named complaint: accuracy (6), trust, liability (1), price (1).

“Our firm purchased Lexis AI, I think it’s garbage but granted I don’t know how to function it very well and try to avoid it at all costs.”

r/paralegal · May 29, 2026 · Accuracy

“No, Lexis, I don't want an AI summary of the regulation I looked up. I prefer to do my own legal research, tyvm.”

Bluesky · Sep 28, 2026 · Trust, liability

Clio Few voices · 3 of 10

3 voices of 3 criticize it or are mixed about it with a named complaint: accuracy (3).

“I had one case where my client was injured in a not so common area but Clio ignored it.”

r/paralegal · Sep 28, 2026 · Accuracy

Legora Few voices · 3 of 10

2 voices of 3 criticize it or are mixed about it with a named complaint: accuracy (2).

“I am a tax solicitor and in one of my case studies, we uploaded a 500-page court judgment, however, Legora was not able to process it.”

r/legaltech · Nov 7, 2025 · Accuracy

Perplexity Few voices · 2 of 10

No named complaint on record among its 2 voices.

Spellbook Few voices · 1 of 10

No named complaint on record among its 1 voice.

Outside tests,

next to the
record not our
score

Public tests that include summarization work, with their own publisher, date and caveat. They measure question sets, not lawyers’ experience, and are never added to the ranking above. Others are on the arena.

Four tools were tested on seven tasks against a lawyer baseline: CoCounsel, Vincent AI, Harvey Assistant and Oliver (Lexis+ AI withdrew from the sections studied). Harvey took the top score on five of the six tasks it entered, including Document Q&A at 94.8% and Chronology Generation at 80.2%; CoCounsel scored 89.6% on Document Q&A and 77.2% on Document Summarization, the top score there. The lawyer baseline beat the AI tools on EDGAR research and on Redlining (79.7%).

Caveat. Vendors opted in task by task (Harvey six of seven, CoCounsel four), so tools were not compared on the same tasks; 2025 versions, over a year old; not a ranking of all legal AI.

Names: Harvey Assistant, CoCounsel, Vincent AI (vLex), Oliver (Vecflow, now August), Lexis+ AI (withdrew).

<https://www.vals.ai/industry-reports/vlair>

Method and cutoff

- **What a voice is.** One public post by a lawyer or legal professional that states an experience or a verdict about one AI tool, read by a language model with one fixed schema (tool, task, sentiment, main complaint). Posts that only ask a question or repeat someone else are not counted. Sources: Reddit, Hacker News, Bluesky, Lemmy, Mastodon, Stack Exchange; posts since Jan 1, 2023.
- **Net and band.** Net = (like – dislike) per 100 voices. The 90% band comes from resampling the same voices 4,000 times; a tool is rated from 10 voices and ordered by the floor of its band, so a few happy voices cannot outrank hundreds. Confidence follows the number of voices: Very low, Low from 10, Medium from 60, High from 300.
- **Few voices.** A tool with fewer than 10 voices on summarization is listed with its count and no band. It is not a negative finding.
- **Complaints.** Voices whose sentiment is dislike or mixed and that name one main complaint, counted per complaint. The quotes are verbatim, with their source and date.
- **Cutoff.** Data through Oct 26, 2026; the latest voice on summarization is from Oct 18, 2026. The report is rebuilt with the daily refresh, so the date moves every day.
- **Not in the numbers.** Outside tests are shown apart and never mixed with our figures. Usage is not quality: a tool people use more collects more posts of every

kind. Public posts are not a survey of all lawyers. Independence policy: no vendor pays for a place or a number. Full method: how we count.

Use these numbers: the counts, rates and bands in this report are free to reuse under CC BY 4.0. Credit “aiforlawyers.app (AI Vortex)” with the date and a link to <https://aiforlawyers.app/report/summarization>. Quotes belong to the people who wrote them: quote them with the link to the original post.

Report generated from the public record at <https://aiforlawyers.app>, cutoff Oct 26, 2026. Numbers are aggregates.