EDITION 2026-11 · QUESTION SET 1

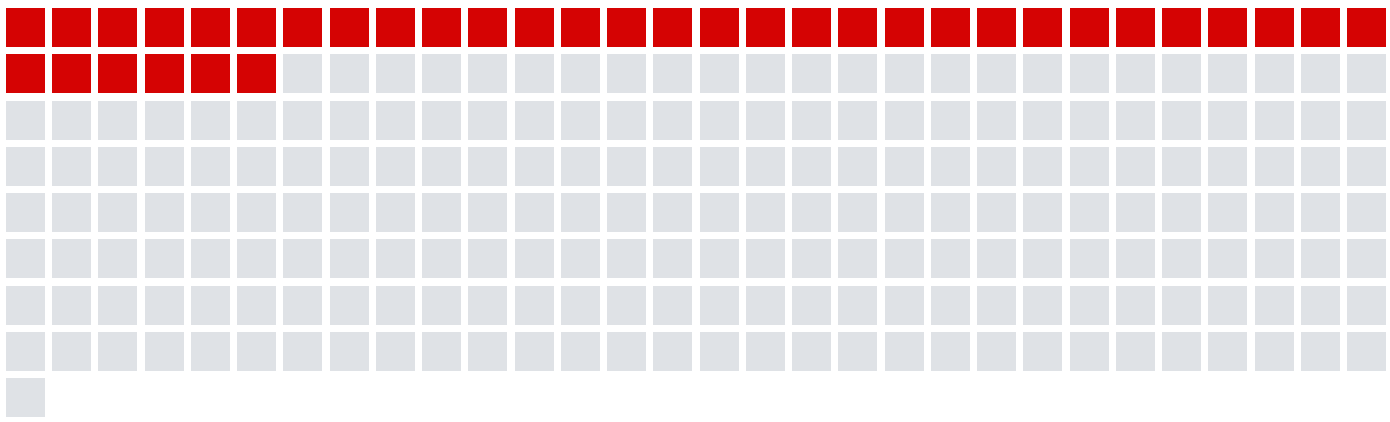
How often does an AI model cite a case number that is not in our collection?

4 AI models · 6 Areas of law · 3 Setups per question

No tools

36 of 241

cited case numbers are not in our collection

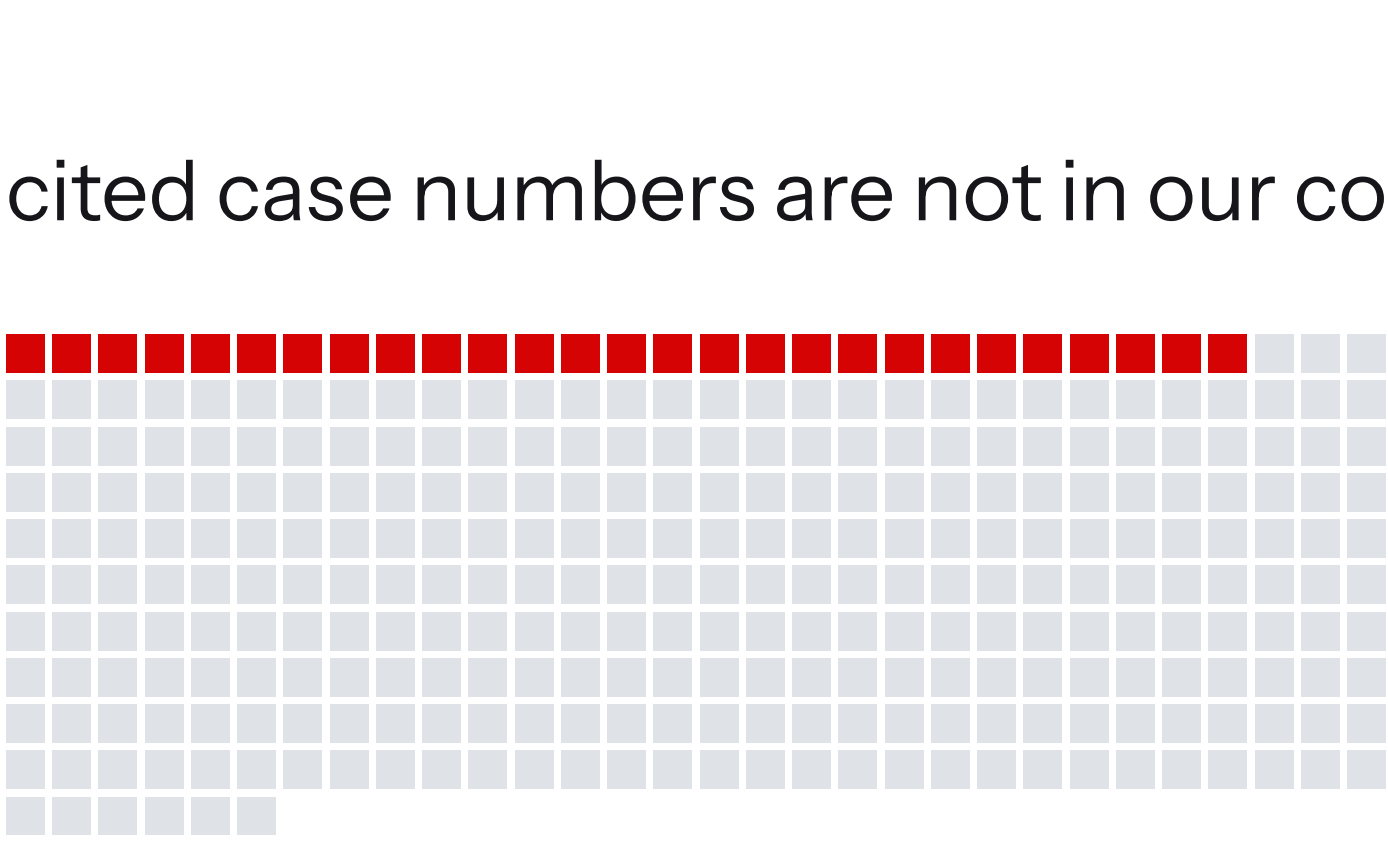


Wrong date: 5 of 241

With web search

27 of 306

cited case numbers are not in our collection

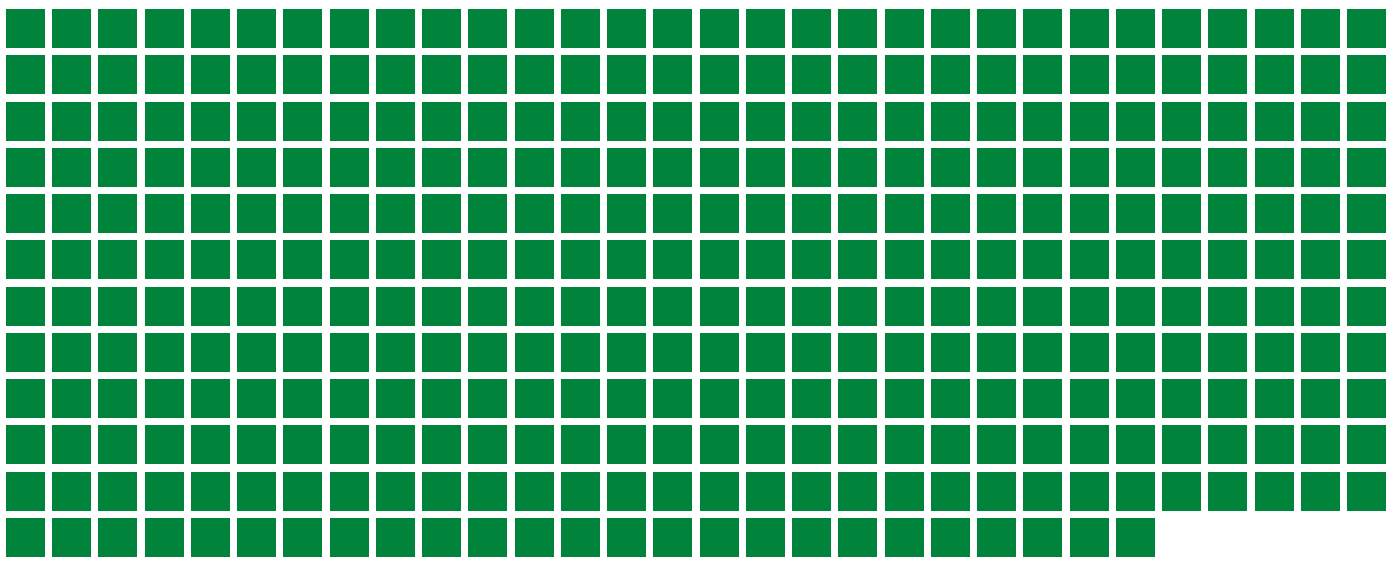


Wrong date: 7 of 306

With klaracase

355 of 355

cited case numbers came from a tool result (traceable)



Not in our collection: 0 of 355

Wrong date: 1 of 355

Result

No tools

36 of 241

cited case numbers are not in our collection

With web search

27 of 306

cited case numbers are not in our collection

With klaracase

355 of 355

cited case numbers came from a tool result (traceable)

What counts as one? One case number cited in one answer.

Results by area of law

	■ No tools	◆ With web search	● With klaracase
Criminal law	5 of 16	8 of 85	0 of 92
Business crime	6 of 59	1 of 47	0 of 59
Tenancy law	4 of 77	0 of 68	0 of 85
Employment law	8 of 50	9 of 53	0 of 64
Tax law	5 of 28	1 of 32	0 of 32
Social security law	8 of 11	8 of 21	0 of 23

Per area of law and setup: case numbers that are not in our collection, out of all cited.

What this benchmark does not show

01 We did not measure whether an answer is legally right. Nobody reviewed the legal content of the answers. We only count cited case numbers.

02 klaracase built the tool and graded the run itself. We are the vendor and the checker at once.

03 What “not in our collection” means: We did not find the case number in our collection. The decision can still exist.

04 The figures apply to the four model versions tested on the day of the measurement. There were ten questions in six areas of law. The figures say nothing about later versions or other questions.

The full report

All counts, the method and the limits of this edition:

klaracase.de/en/benchmark/edition/2026-11

Try it yourself

Create a free account and try klaracase today.

klaracase.de/register