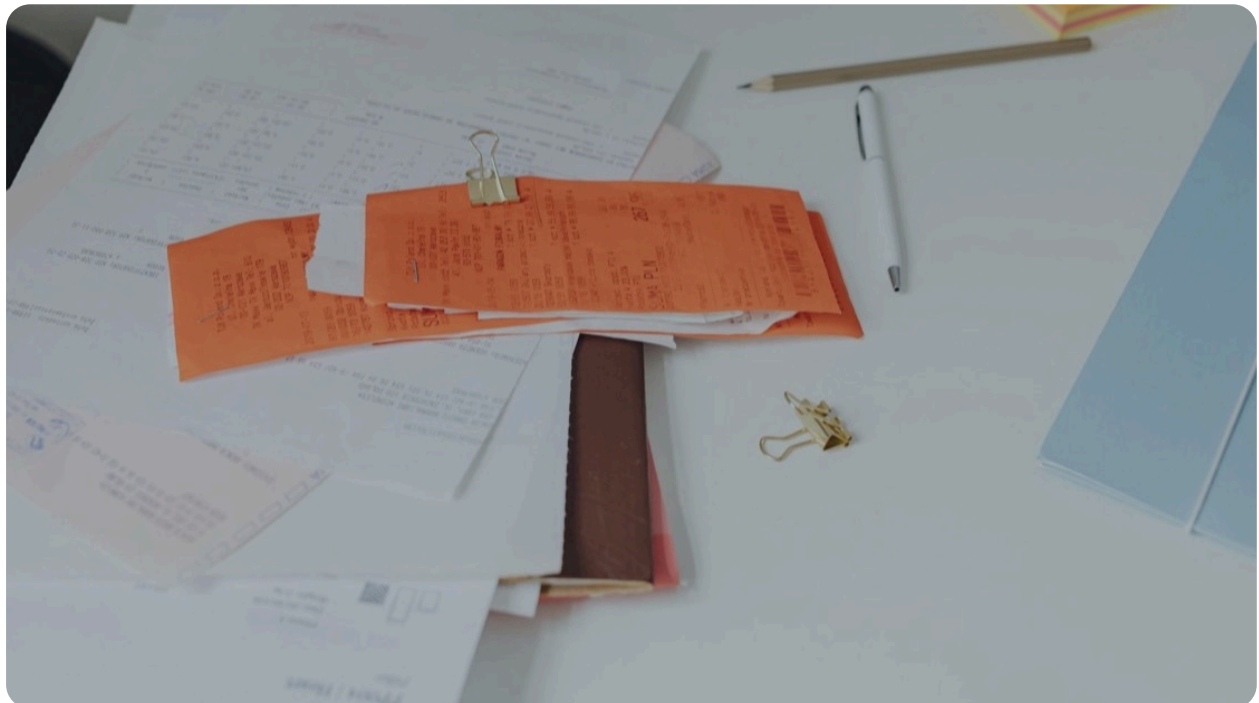


MAY 31, 2026

Structural hallucinations: why AI accuracy is a data problem, not a prompt problem

AI doesn't invent because the model is weak. It invents because the material you hand it is a mess. Fix the workspace, not the wording.

By Dave Taylor



AI doesn't hallucinate because the model is weak. It hallucinates because the material you hand it is a mess. The most dangerous version isn't a single made-up fact you can spot — it's a *structural hallucination*: a confident, well-formatted answer that's wrong underneath. And you cannot fix it with a sharper prompt.

The error you won't catch in time

In February 2025, a US federal court sanctioned lawyers from Morgan & Morgan — one of the largest law firms in America — after they filed motions citing cases that artificial intelligence had invented. Eight of the nine cases cited did not exist. The citations were

formatted correctly. They read as authoritative. Nobody caught them until the court did, and the firm ended up reimbursing the other side's legal fees and rolling out new internal safeguards ([ABA Journal](#)).

That is the shape of the threat. Not an obvious fabrication you'd catch on a quick read, but a made-up *structure* – the right format, the right tone, the wrong substance – that slips through precisely because everything about it looks correct. And it's not a one-off: courts are now sanctioning lawyers for AI-invented citations on a near-monthly basis ([National Law Review](#)).

Why a better prompt won't save you

When you tell a language model "don't make anything up," there is no internal truth-check for that instruction to hook into. The model has no separate pass where it verifies its own claims against reality. It produces the most plausible continuation of the text it's been given – and plausible is not the same as true. A sterner instruction just produces a more confident plausible answer.

This is why prompt engineering as the headline skill is fading. The leverage has moved off the wording and onto the inputs. The teams getting reliable output aren't writing cleverer prompts – they're doing the unglamorous work of organizing the material first. We've argued before that [prompt engineering is dead](#); structural hallucinations are the clearest evidence of why.

The leverage is the workspace, not the wording

Most people point a model at a pile of disorganised inputs – stale PDFs, half-finished strategy docs, conflicting Slack threads – and ask for a polished result. That's two jobs at once: work out what the mess actually is, then draft a clean artifact from it. Asked to do both in a single pass, the model does both poorly, and the gaps get filled with invention.

Give a model a dirty workspace and the dirt ends up in the final document. Give it a clean, structured one – the right files gathered, the originals preserved, the conflicts and gaps made visible before it drafts – and the same model produces work you can stand behind. The bottleneck was never the model's intelligence. It's usually [the knowledge still trapped in your team's heads](#), never written down where the model can see it.

What this means for your team

Next time an AI output is wrong in a way that's hard to catch, don't reach for a better prompt. Look at what you fed it. If you wouldn't hand that same pile to a new hire and expect a clean result, the model won't manage it either.

The accuracy problem you think you have with AI is usually an organization problem with your data – and that one you can actually fix. Get the workspace right and the hallucinations don't need a clever prompt to disappear. They just have less room to hide.