BRIEFING NO. 4

Treat everything AI writes as a first draft

AI doesn't know when it's wrong. You can work with that — once you know why.

Use AI for a few weeks and it will happen to you: a confident, polished answer with a wrong date in it. A statute that doesn't exist. A quote nobody said. The first time, it feels like a betrayal — the tool *sounded* so sure. It isn't broken, and you didn't do anything wrong. This is just how these tools work, and once you know why, it stops being dangerous.

Why AI makes things up

An AI model doesn't look things up. It **predicts what comes next** — it has read an enormous amount of text, and it produces the words most likely to follow yours. When it knows a subject cold, prediction lands on the truth. When it doesn't, prediction lands on something truth-*shaped*: a plausible date, a realistic-sounding case name, a citation formatted perfectly for a paper that was never written.

The industry calls this a *hallucination*. Plain English: **a confident guess.** And because the guess arrives in the same fluent voice as everything else, nothing about it looks different. That's the trap — wrong answers don't come with a warning tone.

So here's the working rule, and it covers almost everything: treat what AI hands you the way you'd treat a first draft from a bright new hire. Impressive, fast, genuinely useful — and not going out the door until someone who knows the business has read it.

Four things to check, every time

Names and titles. People, companies, products, job titles. AI will cheerfully promote someone to CEO or misspell a client's name in an otherwise flawless letter.

Numbers and dates. Totals, percentages, deadlines, dollar figures. If a number matters, recalculate it or trace it back to the source you provided.

Quotes and citations. If AI quotes a person, a study, a law, or an article, confirm the source exists and says what it's claimed to say. Invented citations are the single most common serious error.

Anything you'd hate to be wrong about. The higher the stakes — a client letter, a filing, a price — the slower the read. Low-stakes drafts can be skimmed; commitments get the full pass.

The send-sign-bill rule

One sentence, worth putting on a sticky note: **before anything AI wrote gets sent, signed, billed from, or acted on, a person reads it.** Not skims — reads. That single habit converts AI from a risk into what it actually is: the fastest first-draft writer you'll ever hire.

One trick that prevents most of it

Errors multiply when AI is guessing from memory, and collapse when it's working from material you gave it. So paste in the source — the contract, the numbers, the original email — and ask it to work from that, not from what it "knows." If you read our first briefing on writing better prompts, this is the same lesson wearing different clothes: the more you give it, the less it invents.

And one more: put the peer review in the prompt

You can make the AI run its own first check before you ever see the draft. Add one line to the end of your prompt: "*When you're done, review your work like a skeptical colleague – double-check every name, number, date, and citation, and flag anything you're not certain of.*" It won't catch everything, and it doesn't replace your read — but it catches plenty, and the parts it flags tell you exactly where to look first. Free proofreading from the same temp who wrote the draft. Take it.

> Want to learn more about keeping your data private? Email **PrivateAI@sterling.net**