Using AI Safely: What Every Author Should Know

I teach AI as a thinking partner, and I mean it. Used well, it sharpens your ideas and saves you hours. But the ways it can hurt an author are quiet ones. Your voice slowly stops sounding like yours. A quotation turns out never to have existed. A chapter you pasted in for a quick opinion sits on a company's servers under terms you never read. None of it announces itself, which is why I'd like you to know about it before it happens to you.

Your voice can disappear

AI writing has habits: tidy symmetry, grand abstract words, neat little closing lines. If you let it draft or "polish" your prose, your pages slowly start to sound like everyone else's. Use it to question your writing, not to replace it.

It will tell you what you want to hear

These tools tend to agree with you and to praise ideas that haven't earned it. Ask it to argue against your positioning, to name the weakest claim in a chapter, to say what a sceptical reader would object to.

It invents things

Quotes, statistics, sources, even whole studies can be made up and stated with complete confidence. Check anything factual against the original before it goes near your manuscript.

Your thinking is your asset

If the AI comes up with your position, you end up with the average of what's already been written. Do your own thinking first, then bring it in to be tested.

Protect what's yours, and what isn't

Start with your own work. An unpublished manuscript, a framework you've spent years developing, a book idea nobody has seen yet: once these are out of your hands, you can't get them back. Before you paste any of it into an

AI tool, ask yourself whether you'd be comfortable if it turned up somewhere you didn't expect. If not, keep it out, or share only the part you need help with. The same goes for your own personal details, such as your address, finances, health, and family. There is rarely a good reason to put them in. And keep your own dated copies of your drafts in your own files, so the record of what you wrote, and when, belongs to you.

Then there's what was never yours to begin with: anything confidential about clients, interviewees, or colleagues, and anything shared with you in trust. Their permission matters more than your convenience.

In ChatGPT, the setting that governs training is called Improve the model for everyone. Go to Settings, then Data controls, and switch it off. On a personal account it's on by default, so please check it today. It applies to your whole account on every device, but it's per account, so check each one you use.

Switching it off stops your new conversations being used for training. It doesn't delete anything, and your chats stay in your history. One catch: if you give a thumbs up or thumbs down on a reply, the whole conversation attached to it may still be used for training. So don't rate answers on anything sensitive. For a conversation you'd rather not keep, use Temporary Chat.

Other tools have a similar setting under a different name. In Claude it's Help improve Claude, under Settings, then Privacy. Menu names change from time to time, so if you can't find it, search the tool's help pages for 'training' or

improve. If you're on a business or team plan, training is normally off by default, but check with whoever administers it.

The setting helps, but it doesn't replace your own judgement about what goes in.

Check your contracts

Publishers increasingly ask authors to state whether AI was used. Read what you've signed, and be ready to answer honestly.

Keep your own muscles working

If you reach for the AI before you've tried to structure a problem yourself, that skill weakens. Try first, then ask.

None of this means avoiding AI. If anything here is unclear, bring it to me and we'll go through it together.