**The 19 tell-tale signs an article was written by AI**

Avoid these AI cliches (or use them to detect AI writing in the wild)

[[](https://medium.com/@JimTheAIWhisperer?source=post_page-----cf4ef7b4fe6d--------------------------------)](https://medium.com/@JimTheAIWhisperer?source=post_page-----cf4ef7b4fe6d--------------------------------)

[[The Generator](https://medium.com/the-generator?source=post_page-----cf4ef7b4fe6d--------------------------------)](https://medium.com/the-generator?source=post_page-----cf4ef7b4fe6d--------------------------------)

[Jim the AI Whisperer](https://medium.com/@JimTheAIWhisperer?source=post_page-----cf4ef7b4fe6d--------------------------------)

·

Follow

Published in

[The Generator](https://medium.com/the-generator?source=post_page-----cf4ef7b4fe6d--------------------------------)

·

16 min read

·

Sep 29, 2023

7.9K

138

**“Did AI write this sentence?”** That’s become *the* question of this year. It’s probably our biggest conundrum since Hamlet got all indecisive while chatting to a hollowed out skull. It’s rolled together all our 21st century doubts and insecurities of Fake News, The Matrix, and misinformation.

If you’re **reading** **content** **online in 2023**, it’s a question you want answered. And if you’re **writing content online in 2023**with the assistance of AI, it’s a question you perhaps *don’t* want answered. (For the record: I do write this blog. Last I checked, I’m a real human. Jiminy Cricket will vouch for me 🧚🏻‍♂️).

A child with a long nose and a long nose

Description automatically generated

I call it the Pinocchio Paradox: Are writers “real humans”, or long-nosed liars?

AI writers are engaged in a cloak-and-dagger routine not dissimilar to Cold War espionage. Think about it: during those times, spies would pass coded messages, employ deception, and use the technology of their era to get the upper hand. Now, our “espionage” is the realm of information and content creation. It’s a Turing Test where some copywriters are assuming false identities and trying to conceal whether AI had a hand in their content.

This digital shadow realm is like a spy-game where creators and consumers need to be discerning, questioning, and aware. With AI capabilities at our fingertips, the cards we hold are more powerful than ever, and knowing when and how to play them is our best defence, and the art of the game.

Just want the cold hard tips, no background explanation? I get you. Just scroll to the numbered list “The 19 Tell-Tale Signs of AI Content” below!

**Why you need to be able to spot AI content yourself**

**Readers** shouldn’t be relying on content detectors (or assuming Google will protect them from AI content), and **AI assisted writers** shouldn’t be fretting (or thinking big tech is out to get them). There’s always been chatter that AI will harm SEO, and that search engine algorithms will buffer readers from non-human written content. That isn’t the case, and would be impossible to implement. Language is language, at the end of the day. SEO does prioritize relevance, engagement, and authenticity, so there is some truth to it, but it doesn’t discern AI from human text. Frankly, not all low-quality content is AI. And in the right hands, AI generated content can appear spectacular.

**From AI with Love: The double agents of digital content**

I’m pretty good at deciphering AI content. I’m like James Bond at the poker table. I do this for a living; I can spot a good bluff. I don’t go for the obvious tells, like “AI writing is emotionless”. Emotion is a red herring (and cuts out many fine, objective, human writers who craft with precision and subtlety). Some of the most moving writing I’ve read in the last year is [written by AI](https://www.theguardian.com/technology/2023/jul/18/ai-chatbots-grief-chatgpt).

No, I look for patterns, repetitions, or an overuse of particular phrases or references. I analyze the flow of the content, the choice of examples, the depth and breadth of topics covered in a piece. I use skills I honed across two decades as a copywriter, an assistant editor, academic and examiner.

(There’s scientific evidence for this, [by the way](https://arxiv.org/pdf/2212.12672.pdf). Some people *are* better at spotting AI generated content: people with experience in generative AI, immersed in its ebbs and flows, and people trained as close readers).

**What you’ll learn from this article**

In this article, I’m going to reveal how to use critical analysis — yes, the type of close reading they teach you at college — to expose AI. Use it for good, use it for evil. I’m not the boss of you. It can help you detect content yourself, or to edit and occlude telltale signs. I will say, *if* you use it to hide AI content, these tips will improve your writing (that can only be a good thing, right?).

Ultimately, it’s the *quality* and *readability* of the content that will matter in the future, and not whether it was a collaboration with machine learning.

Now, let’s take our sunglasses off, and lay our cards face up on the table…

A person in a suit playing poker

Description automatically generated

**Why not just use an AI detector?**

Why not use a lie detector while we’re at it? Both are notoriously unreliable. But I think most people are a little squeamish about the use of lie detectors.

In fact, AI content detectors are less valid than polygraphs (which get things wrong about 10% to 30% of the time, [depending on the study](https://pubmed.ncbi.nlm.nih.gov/30284848/)). In July 2023, OpenAI closed down its AI Classifier. It had an embarrassing ***74% fail rate.***

Wharton professor and tech expert Ethan Mollick advises “AI detectors have high false positive rates, and they should not be used as a result.” However, detectors continue to be used, and to offer false confidence (or false fears, depending which side of the spy-war over AI content you’re on).

(The whole thing is a bit like **Marvel’s Secret Invasion**, where everyone is a suspect. Nobody knows quite who is human behind our screens anymore).

**[Anger over AI art in Marvel’s “Secret Invasion” misses the point](https://medium.com/the-generator/marvels-secret-invasion-ai-art-sparks-controversy-in-opening-credits-49967e0c7277?source=post_page-----cf4ef7b4fe6d--------------------------------" \t "_blank)**

[MCU conspiracy thriller taps into fears of artists being replaced with AI](https://medium.com/the-generator/marvels-secret-invasion-ai-art-sparks-controversy-in-opening-credits-49967e0c7277?source=post_page-----cf4ef7b4fe6d--------------------------------" \t "_blank)

[medium.com](https://medium.com/the-generator/marvels-secret-invasion-ai-art-sparks-controversy-in-opening-credits-49967e0c7277?source=post_page-----cf4ef7b4fe6d--------------------------------" \t "_blank)

**The shortcomings of AI content detectors**

Turnitin was used to evaluate 38 million student essays for AI writing in 3 months alone (April-June 2023), but the company admits their algorithm incorrectly red flags 4% on a sentence-by-sentence level. That’s potentially 1.8 million essays [falsely labelled ‘AI content’](https://www.reddit.com/r/college/comments/159rexs/falsely_accused_of_using_ai/). An informal experiment by [**The Washington Post**](https://www.washingtonpost.com/technology/2023/06/02/turnitin-ai-cheating-detector-accuracy/) found Turnitin got it wrong >50% of the time (*n*=16).

In an [Educator FAQ](https://help.openai.com/en/collections/5929286-educator-faq) last week, OpenAI addressed the elephant in the room:

*“Do AI detectors work? In short, no. While some (including OpenAI) have released tools that purport to detect AI-generated content, none of these have proven to reliably distinguish between AI-generated and human-content.”*

Whether or not they’re good at catching AI, the fact is detectors are in wide use, and if you’re writing with AI you’d be wise to avoid getting red flagged.

**“I don’t use AI, why should I care?”**

Ironically, we might have to start using AI content detector evasion tactics for original *human* work (almost like a writing version of a CAPTCHA, but baked into writing styles and words and syntax to avoid). Again, this might  
not be a bad thing, and encourage better rhetoric and writing techniques.

This is why my list of tell-tale signs of AI to watch out for is useful whether you’re a genuine human writer, or an AI assisted one. It’s about avoiding cliches and adding value.

**What are the best AI content detectors?**

Some of the most popular AI content detector programs out there include:

* GPTZERO
* OPENAIWRITER
* CROSSPLAG
* COPYLEAKS
* SAPLING
* CONTENTATSCALE
* ZEROGPT

**Using an AI content detection remover tool (Undetectable.ai)**

One way to avoid AI content detectors is to use a generator specifically trained to evade them. The leading tool in this niche is **Undetectable.ai**.

**[The Truly Undetectable AI Content Writing Tool](https://undetectable.ai/?fpr=jimtheaiwhisperer&source=post_page-----cf4ef7b4fe6d--------------------------------" \t "_blank)**

[We transform your flagged AI-generated content into premium quality writing that aligns with your brand voice and makes…](https://undetectable.ai/?fpr=jimtheaiwhisperer&source=post_page-----cf4ef7b4fe6d--------------------------------" \t "_blank)

[undetectable.ai](https://undetectable.ai/?fpr=jimtheaiwhisperer&source=post_page-----cf4ef7b4fe6d--------------------------------" \t "_blank)

People can run their AI generated content from any other AI writing tool through **Undetectable.ai**. (Do what you will with this information! I’m merely letting you know it exists in the ever-escalating AI arms race).

[**Undetectable.ai**](https://undetectable.ai/?fpr=jimtheaiwhisperer) uses a federated multi-model architecture with custom and fine-tuned models to identify text that might trigger alarm bells, removes it, then automatically rephrases it for you. Whether this paraphrasing is more human-like (or simply passing under the AI’s radar) is up to you to decide.

It does allow you to see how the major free and paid AI detectors (including all those listed above) perceive your text. Even though we’ve established that content detectors aren’t successful most of the time (correctly identifying AI content in perhaps only 20–30% of cases), an AI content detection remover like **Undetectable.ai** can help to keep it above suspicion.

**How AI content detectors work (Or don’t, as the case may be!)**

There are two main metrics that the AI detection tool kits measure: **perplexity**and**burstiness**. I’ve discussed these in detail previously.

**[The Dummy Guide to ‘Perplexity’ and ‘Burstiness’ in AI-generated content](https://medium.com/the-generator/the-dummy-guide-to-perplexity-and-burstiness-in-ai-generated-content-1b4cb31e5a81?source=post_page-----cf4ef7b4fe6d--------------------------------" \t "_blank)**

[Understanding Language Models: A Simplified Guide](https://medium.com/the-generator/the-dummy-guide-to-perplexity-and-burstiness-in-ai-generated-content-1b4cb31e5a81?source=post_page-----cf4ef7b4fe6d--------------------------------" \t "_blank)

[medium.com](https://medium.com/the-generator/the-dummy-guide-to-perplexity-and-burstiness-in-ai-generated-content-1b4cb31e5a81?source=post_page-----cf4ef7b4fe6d--------------------------------" \t "_blank)

Basically, “**Perplexity**” refers to how predictable the next word in a sentence is, and “**Burstiness**” refers to the uniformity of the sentences (humans write in fits and starts, with varied sentence lengths, while AI is more consistent). These same metrics are also used to measure performance levels of LLMs.

AI content detectors use a **generative adversarial approach** (don’t worry, I’ll break it down). The detector basically asks: “Would I have written this?” Or rather: “Would I have completed the next word in the sentence this way?”

It’s a game where the adversarial AI (the detector) attempts to predict the next conceivable word. This is actually how AI writes (procedurally, one-word-at-a-time, based on statistical outcomes. If it guesses correctly, it’s logical a similar AI model might have generated the text in the first place.

**The Red Queen: Why we may never be able to reliably detect AI**

That’s good as far as it goes, but relies on both AIs using the same lexicon. You can bypass (or “perplex”) any one detector by combining different models. Many AI text generators leverage multiple LLMs, and switch between whichever is most suitable for the use case (**Jasper AI**, for example, uses [OpenAI, Neo X, T5, Bloom, plus a proprietary model](https://medium.com/p/60536071ca6a)).

Additionally, returning to our Cold War analogy, it’s an arms race. The latest AI language models are increasingly fluent and human-like, which escalates the need for even more powerful means of AI detection (which are trained with the adversarial process). In turn, models like [**Undetectable.ai**](https://undetectable.ai/?fpr=jimtheaiwhisperer)are trained to bypass *that* technology. Ultimately, it echoes the Red Queen’s dilemma: “It takes all the running you can do, to keep in the same place.”

A red and white chess piece

Description automatically generated

AI image of the Red Queen from Alice “Through the Looking-Glass”. [Learn how to spot AI art HERE](https://medium.com/the-generator/15-tips-to-tell-if-an-image-is-ai-art-or-not-2ae703d0ac36).

Computer scientists warn that “[ultimately, there is nothing special about AI-written text that always distinguishes it from human-written, and detectors can be defeated](https://arstechnica.com/information-technology/2023/09/openai-admits-that-ai-writing-detectors-dont-work/#:~:text=Ultimately%2C%20there%20is%20nothing%20special,abysmal%2026%20percent%20accuracy%20rate.)”. But in my experience: there’s a rhetorical element that shouldn’t be overlooked. That’s where our human intuition comes in.

**The human touch: Why rely on humans for AI detection?**

When we’re dealing with language-based AI, a sound knowledge of figures of speech is our best defence. Critical reading is the antidote to AI content. While technology evolves at breakneck speeds, there remains an intrinsic value in the subtle nuances and imperfections of human communication.

In situations where AI models generate descriptive or connective content, they tend to lean on a default mechanism to link ideas or provide reference, even if it’s repetitive or not entirely contextually appropriate. Over-reliance on such devices lends the content an inauthentic air, more reminiscent of a machine echoing human speech patterns than genuine human discourse.

It’s not that AI text doesn’t look like human writing any more; it’s that AI emulates the structure and rhythm of human prose *too* perfectly, without capturing the spontaneity and unpredictability inherent in genuine human communication. While AI can replicate the framework, it sacrifices the subtle deviations and unique expressions — the irregularities — that make our words come alive, in favour of conformity and the technically correct.

It takes a human, attuned to the intricacies and idiosyncrasies of language, to detect the clunky clichés that algorithms inadvertently produce in their ceaseless pursuit of patterns, consistency, and machine-driven precision.

**The 19 Tell-Tale Signs of AI Content**

1. **Caught red handed.**This one is ridiculous, but [it snuck under the noses of the editors of the scientific journal **Nature**](https://futurism.com/the-byte/paper-retracted-authors-used-chatgpt), so bears repeating: look for the phrases “Regenerate response” or “[as an AI language model, I](https://www.theverge.com/2023/4/25/23697218/ai-generated-spam-fake-user-reviews-as-an-ai-language-model)”. That’s a sure sign of sloppy copy+paste straight from the chat window.
2. **Textbook perfect spelling.** Because it’s a Large Scale Language Model, AI knows how virtually every word is spelled. Additionally, AI doesn’t make typos, because, well: no fingers. Humans on the other hand…
3. **Middle-of-the-road opinions.**AI-generated content is non-committal or overly neutral. This tendency is because AI models are designed to avoid showing bias, or causing offense, and to cater to a broad audience. The content often adopts an “on the one hand, on the other hand” approach. This is different from a balanced view, which weighs pros and cons but may arrive at a conclusion. AI’s approach is more about hedging its bets.
4. **Overuse of “crucial”.** Once you recognize its prevalence in AI-generated content, the word “crucial” stands out like a sore thumb. A tell-tale sign is the repetitive usage in formulas like, “When you’re [insert activity; often ‘learning about *x*’ or ‘starting out in *x*’], it’s crucial to…”. While the term can be, well, *crucial* in certain contexts, keeping an eye out for repetitive patterns can help differentiate AI text from human writing.
5. **“Delve, Dive, Discover”.**AI has a penchant for particular word patterns. Among these are the use of the verbs “Delve,” “Dive,” and “Discover” in Calls-to-Action (CTAs). Especially recognizable when followed by “…into the exciting world of *x*”. It’s a copywriting technique, but after noticing its frequency in AI-generated content, I’ve personally refrained from it.
6. **“Unlock”**. Another word to eschew. Especially common in AI marketing.
7. **“Ensure”**. AI uses this word more than any reasonable human ever does.
8. **“A dash of *x*”.** This is a colloquialism that I literally only see AI use now.
9. **Overuse of demonstrative pronouns.**High frequency of words like “the, this, these, that, those,” can be an indicator of AI . Daphne Ippolito from **Google Brain** says “[a really easy cue is the word “the” occurs too many times](https://solvid.co.uk/detect-ai-writing/#:~:text=%E2%80%9CIf%20you%20have%20enough%20text,scanning%20a%20piece%20of%20text.)”. Demonstrative pronouns are among the most common words in English. Thus, when generating content, AI models tend to use these words frequently as they are statistically common in training data.
10. **Starting sentences with “Remember,…”.** When you come across a piece of content that constantly “reminds” the reader, it could be a sign of AI. The rhetorical device where the writer breaks off to address an abstract concept, asking them to pause and take note of the subsequent point, is called an **Apostrophe**. For example, starting with the word “Remember” followed by a comma, is a favorite apostropheof AI models. Of course, it’s worth noting that human writers *also* use the same rhetorical devices. The difference lies in the frequency with which they are applied. An AI might overuse or apply these devices where a human wouldn’t. Remember, it’s all about spotting patterns! (See what I did?)
11. **Metaphor + Tricolon construction.** This pattern typically goes “[Topic] is more than just [basic description or literal interpretation]; it’s a [deeper or metaphorical interpretation] [verb] [a broader context or significance, expressed as a tricolon].” Example: “*Reading is more than just a hobby; it’s a bridge connecting our society’s history, imagination, and collective consciousness****.”***When you start to notice this rhetorical device — where a metaphor is immediately followed by a tricolon to reinforce its significance — you’ll start to see it everywhere in AI-generated content.
12. **Inability to construct anastrophes**. This one’s a bit harder to explain. An anastrophe is a rhetorical device where the expected order of words in a sentence is reversed. It’s used to create a particular emphasis or poetic impact. At its most obvious, it’s classic Yoda-speak (“Powerful you have become”). Anastrophe also can be spotted in ordinary speech; it’s also used in advertising to make a simple phrase more memorable or achieve immediate recognition. E.g. **“Like a good neighbor, State Farm is there”**.  
    I’ve written before about how hard it is to get [ChatGPT to reliably write in reverse](https://medium.com/the-generator/ai-cant-write-sentences-backwards-and-what-this-tells-us-about-llms-4a9ee42c81e2), and what this means for how AI approaches language. This is one of those implications. Although you won’t encounter anastrophe in human writing often, I hope this highlights how straightforward AI is in getting from point A to point B. The text AI generates is procedural, i.e. one word after another, so it’s not good at the forethought required to form an anastrophe. AI tends to avoid doing backflips and linguistic gymnastics. So: “A sentence AI would be unlikely to write is this one”.
13. **AI can’t commit catachresis**. This is another rhetorical device that’s hard to describe, but once you get it, *IYKYK*. AI never writes anything wrong, but humans do, sometimes intentionally. When something’s so wrong it’s right, that’s a catachresis. It’s behind many memes and song lyrics, and it’s catchy. E.g. “Because reasons”, “They don’t think it be like it is, but it do,” and “Set fire to the rain”. Human writers use unconventional grammar and phrasing that deviates from the norm to convey humor, emotion, and originality. AI is so focused on getting things right, it doesn’t know when to break the rules. *Much wow, such grammar.*
14. **“In today’s world”.**Even if it’s not an AI artefact, please flush this turd.
15. **“Take your *x* to the next level”.**Nobody talks like that outside of the Red Bull Cinematic Universe. Seriously, hyped up jargon is rampant in online marketing and advertising. But most real-life conversations steer clear of such hyperbolic statements. AI models, trained on vast amounts of data from the web, have a tendency to over-inflate trendy buzzwords.
16. **“Master”**. Just no. Marketing material might refer to “mastering” a new skill, platform, or tool. But in 2023 it’s outdated. AI didn’t get the memo.
17. **Antithetical Parallelism.** This common rhetorical formula is where two contrasting ideas are juxtaposed in parallel structures. Think of phrases like “From the oldest traditions to the latest trends” or “Whether you’re a seasoned pro or a newcomer to the game”. It lends itself to making content sound balanced and comprehensive (see #4). While human writers use it too, overuse of this structure can be a giveaway of AI.
18. **Limited Use of Parentheses and Em Dashes**: While parentheses and em dashes are common in human writing for creating nuance or offering asides — writers *love* to drop an em dash for emphasis or squirrel away an extra thought in parentheses (guilty as charged) — AI usually favors more straightforward syntactic constructions. The absence or rarity of punctuation marks can serve as a spoiler that you’re reading AI content.
19. **“If you’re anything like me”**. This is the AI equivalent of the Steve Buscemi meme “How do you do fellow kids”. *Hello* fellow humans.

**So there you have it!**(Actually, wait, add that cliche too while you’re at it).

The signoff “So there you have it!” is a common copout when AI wraps up articles or blog posts. AI often puts an overly neat bow on complex issues or discussions when tasked to conclude a piece of content, usually followed by a brief restatement of the main points, suggesting an artificial sense of closure that doesn’t do justice to the subject matter. It doesn’t have the contextual understanding or ability to discern nuance like humans do.

A bird made of metal parts

Description automatically generated

Overuse of such phrases can feel more like AI parroting popular speech than genuine human communication.

**What can I do with this knowledge? Or, Spy vs. Spy**

Hopefully the above will help you identify who-is-who. As we’ve discussed, an unnatural frequency of stock phrases, especially in contexts where they don’t quite fit, can serve as a telltale sign that the content was AI-generated.

Yet, this isn’t just about pinpointing who or what wrote the content, but about assessing the added value of the piece. Does it enlighten? Does it engage? Or does it merely parrot recognizable phrases? In the future, it won’t matter whether a text is AI or human written, but if it adds value.

Armed with this understanding, one becomes a content detective, peeling back layers of rhetoric to reveal the true nature of a text. It’s a skill, much like the art of counterintelligence. One learns to pick out allies from foes. The future battlefield isn’t AI vs. human. It’s meaningful vs. meaningless.

**Casino Royale of Clichés: Why AI loves copywriting tropes**

Drawing parallels to the world of poker, AI plays its hand based on the cards (data) it’s been dealt. AI tends to gamble on the familiar, resorting to tried-and-true strategies. But why? It comes down to both its training data, and the predictability of procedural generation. AI learns from vast amounts of information, and in the world of writing, clichés are abundant. AI picks up patterns in language from this data, and generates based on “what word comes next in the sentence”. Of course that’s going to repeat and amplify the cliches exisiting in the corpus. It regresses toward a mean. To an AI, these phrases are not tired or overused; they are statistically safe bets.

**[Is ChatGPT getting dumber? Let’s talk about ‘AI Drift’](https://medium.com/the-generator/is-chatgpt-getting-dumber-lets-talk-about-ai-drift-58f0fc5ebad2?source=post_page-----cf4ef7b4fe6d--------------------------------" \t "_blank)**

[Why is AI losing its edge? AI Drift explained 🚗💨🤖](https://medium.com/the-generator/is-chatgpt-getting-dumber-lets-talk-about-ai-drift-58f0fc5ebad2?source=post_page-----cf4ef7b4fe6d--------------------------------" \t "_blank)

[medium.com](https://medium.com/the-generator/is-chatgpt-getting-dumber-lets-talk-about-ai-drift-58f0fc5ebad2?source=post_page-----cf4ef7b4fe6d--------------------------------" \t "_blank)

**Generative Adversarial Network** (GAN) AI detectors focus on a procedural word-by-word basis (much like AI generators themselves). Consequently, they overlook the broader prose. This means they get lost in the prediction of forthcoming words in a sequence and miss challenging the overused tropes that a keen human eye, trained in critical reading, might catch.

My critical-reading method differs from the (fallible) AI content detectors. Unlike the fallible AI content detectors which focus on the likelihood of individual words, my approach zeroes in on common rhetorical figures.

**Hey, if any AI detector software companies are reading this and want to incorporate rhetoric detection,**[**reach out to me**](http://jim@jimtheaiwhisperer.com/)**for prompt engineering!**

When AI plays a royal flush of clichés, it’s the discerning reader, the astute player at the table, who decides whether to call its bluff or fold. In the end, it’s not about man or machine, but about authenticity, insight, and instinct.