

THE TELL

The Tell

The working framework I use to keep AI-assisted writing reading like a person wrote it.

Christopher Jan Benitez · christopherjanb.com

The method

AI Writing Tells

A field guide for spotting and removing the patterns that make text read as AI-generated. The catalog is adapted from Wikipedia's community-maintained "Signs of AI writing" guide (WikiProject AI Cleanup), generalized so it applies to blog posts, marketing copy, articles and emails, not just encyclopedia entries.

The one rule that governs everything else

These are signals, not proof. Large language models learned to write from human text, so every pattern below also appears in human writing: editorials, blogs, fan fiction, press releases. A single transition word or one em dash means nothing. What matters is clustering and density: AI-generated text tends to stack many of these tells together, and to smooth specific facts into generic statements that could apply to almost any topic.

So the job is never "ban these words." It's "notice when enough signals pile up that the text reads as machine-smoothed, then either flag it (detection) or fix the underlying flatness (revision)." Treat a finding as a question worth asking, not a verdict. Avoid confident accusations; false positives are easy and costly.

Two modes

Decide which one the request calls for. Often it's both.

1. Detect: "Was this written by AI?" / "Does this read as AI?" Produce a calibrated read with the specific evidence, not a yes/no with false certainty.
2. Scrub: "Clean the tells out of this draft" / "Humanize this." Produce a revised version that fixes the patterns *at the root* (add specifics, vary rhythm, take a stance), not a thesaurus swap that just hides the surface words.

Workflow

1. Load the catalog. Read [references/signals-catalog.md](#) for the full annotated list of signals with examples and fixes. Do this before auditing anything substantial. It's the substance of the skill.
2. Read the text against each category. Go category by category. For every hit, capture: the exact snippet, which signal it matches, and (for scrub mode) a concrete rewrite. Note *where* signals cluster: three tells in one paragraph is far more telling than three spread across 2,000 words.
3. Weigh the signals. Not all are equal:
 - Strong (rarely appears in careful human writing, especially in clusters): chatbot leftovers ("Certainly! Here's...", "As an AI language model", knowledge-cutoff disclaimers), fabricated or mismatched citations, the "significance" participial tail used repeatedly, bold-lead-in bullet lists where the bold term is just restated, leftover prompt text or meta-commentary.
 - Moderate: inflated-significance puffery, promotional/travel-brochure tone, "not only... but also" / "it's not just X, it's Y" contrasts, "It's important to note that", rule-of-three stacking, transition-word spam, "In conclusion"/"Overall" wrap-ups.
 - Weak on their own (only meaningful in a cluster): em dashes, title-case headings, vague attributions ("studies show", "experts say"), uniform sentence/paragraph length, generic statements with no lived specifics.

Em dashes deserve a special note: plenty of skilled humans use them. They are at most a faint signal and never proof on their own. Don't lead with them.
4. Calibrate the read. Base your confidence on density and the presence of strong signals, and state it honestly. Account for base rates: if you would tag ten pieces as AI, expect at least one to be a false positive even when you're good at this. Remember human writing is increasingly LLM-influenced, so the signals are converging over time. A clean text today is weaker evidence of human authorship than it was a year ago.
5. Produce the output in the format below for the mode you're in.

Output format: Detect mode

```
## Read: [Strong / Moderate / Weak / Inconclusive] signal of AI generation

[1-2 sentences: the overall impression and what's driving it, stated with
appropriate humility.]

## What's driving it
- **[Signal name]** ([strength]): "[exact snippet]" – [why it reads as AI]
- **[Signal name]** ([strength]): "[exact snippet]" – [why]
[... grouped strongest-first ...]

## What cuts against it
- [Anything that reads as genuinely human: specific lived detail, an unusual
stance, idiosyncratic voice, a real verifiable source. Always include this section
– it's the honesty check.]
```

Caveat

These are probabilistic signals, not proof. [Note the strongest alternative explanation – e.g. "a human writer trained on web copy could produce all of this," or "this looks like a human draft lightly polished by AI."]

Output format: Scrub mode

Lead with the revised text, then a short changelog. Fix root causes:

- Replace puffery with specifics. "Played a pivotal role in the region's development" → name the actual thing it did, with a number, date, or named example.
- Vary rhythm. Break up uniform sentence lengths. Let some sentences run; cut others to three words.
- Take a stance. AI hedges and balances everything. Say the thing the writer actually believes.
- Add lived detail. A concrete example, a real source, a specific that only someone who knows the subject would include.
- Cut, don't reword, the empty connectors. Delete "It's important to note that"; don't swap it for "It's worth highlighting that."

[Revised text]

Changelog

- [Signal] → [what you changed and why], e.g. "Cut three 'significance' tails; replaced with the specific 1989 founding detail."

When the source context matters

A few signals in the catalog are context-dependent and you should weight them by where the text lives:

- Citation/sourcing tells (fabricated cites, echoing reliability-guideline language like "significant independent coverage") are strong in encyclopedic or journalistic contexts, weaker in a personal blog.
- Markup leakage (stray Markdown asterisks in a context that doesn't render Markdown, broken wiki markup) is a strong tell wherever it appears, because it means text was pasted from a chatbot without cleanup.
- Promotional tone is a tell in neutral/encyclopedic writing but expected in actual marketing copy, so in marketing contexts, downweight it and look harder at the other categories.

A note on the source

This catalog reflects Wikipedia's "Signs of AI writing" guide as of early 2026. That page is living and updated as models change their habits, and newer models actively try to disguise these

patterns. Treat the catalog as current best-known tells, not a permanent list. The deeper signal, generic and fact-smoothed prose with no specific point of view, outlasts any individual phrase, so anchor on that when the surface tells have been scrubbed.

The signals

Signals Catalog

The full annotated list of AI-writing tells, adapted from Wikipedia's "Signs of AI writing." Each entry has: what it is, why it reads as AI, and a bad → better example. Read SKILL.md first for how to weigh these: clustering matters more than any single hit, and none of these is proof on its own.

Table of contents

- 1. Language and tone
- 2. Sentence-level constructions
- 3. Style and formatting
- 4. Communication intended for the user (chatbot leftovers)
- 5. Citations and sourcing
- 6. Markup leakage
- 7. The deepest tell: regression to the mean

1. Language and tone

1.1 Inflated significance / puffery

- What: Statements that puff up the importance of the subject by tying it to a broader theme, usually in generic terms.
- Why it's a tell: Models pad importance because the training data (bios, press releases, promo copy) does. The phrases recur in a small, predictable repertoire.
- Watch for: "stands as," "serves as a testament to," "plays a vital/significant/crucial role," "a pivotal moment," "part of a broader movement," "rich cultural heritage," "rich tapestry," "enduring legacy," "leaves a lasting impact," "continues to captivate," "deeply rooted," "indelible mark," "watershed moment."

READS AS AI	READS AS HUMAN
The festival stands as a testament to the region's rich cultural heritage and continues to captivate visitors from around the world.	The festival has run every August since 1974 and drew about 12,000 people last year, most from the surrounding three counties.

1.2 Promotional / travel-brochure tone

- What: Marketing admiration where neutral description belongs. Landscapes are always scenic, views breathtaking, everything clean and modern.
- Why it's a tell: Reads like ad copy, not reporting. Strong tell in encyclopedic/neutral writing; expected (so downweight it) in actual marketing copy.
- Watch for: "breathtaking," "stunning," "must-see," "nestled in the heart of," "vibrant," "picturesque," "boasts," "a hidden gem," "state-of-the-art," "seamless."

READS AS AI	READS AS HUMAN
Nestled in the heart of the valley, this picturesque town boasts breathtaking views and a vibrant local scene.	The town sits at the north end of the valley, population 4,200, known locally for its Saturday produce market.

1.3 The "significance" participial tail

- What: A trailing clause, usually a present participle, that gestures at importance without adding a fact.
- Why it's a tell: One of the most identifiable AI habits. Once you see it you see it everywhere. Repeated use is a *strong* signal.
- Watch for: sentence-final "...emphasizing the significance of...," "...reflecting its enduring relevance...," "...highlighting its role in...," "...underscoring its importance...," "...solidifying its status as...," "...cementing its place as...," "...marking a turning point in..."

READS AS AI	READS AS HUMAN
The bridge opened in 1932, reflecting the city's growing ambition and underscoring its emergence as a regional hub.	The bridge opened in 1932. It cut the river crossing from forty minutes by ferry to under five.

1.4 Editorializing and superficial analysis

- What: Interpretation or opinion inserted where neutral statement was asked for; throat-clearing that announces importance instead of demonstrating it.
- Why it's a tell: Models add a layer of "meaning-making" reflexively, even when told to stay neutral.
- Watch for: "It's important to note that," "It's worth noting that," "This highlights," "This underscores," "This serves to," "Importantly," "Notably," "Crucially."

READS AS AI	READS AS HUMAN
It's important to note that the policy had wide-reaching effects on the community.	The policy cut bus fares by half; ridership rose 30% within a year. (Show the effect; drop the announcement.)

1.5 Vague attribution

- What: Claims sourced to an unnamed authority.

- Why it's a tell: Models reach for the *shape* of evidence without a real source. A bias toward rebutting with unspecified "critics" is a noted variant.
- Watch for: "studies have shown," "experts say," "research suggests," "industry reports indicate," "it is widely believed," "critics argue," "some say."

READS AS AI	READS AS HUMAN
Studies have shown that the technique improves outcomes.	A 2023 trial of 400 patients found a 12% improvement — or, if no real source exists, cut the claim entirely.

2. Sentence-level constructions

2.1 Negative parallelism / false contrast

- What: A construction that sets up a contrast to sound insightful: "not X, but Y." Often used to "correct" a misconception the reader never held.
- Why it's a tell: Stereotypical AI move. Negation framing is one of the most-cited tells.
- Watch for: "It's not just X, it's Y," "More than just X," "Not only... but also...," "rather than merely... it...," "isn't about X — it's about Y."

READS AS AI	READS AS HUMAN
This isn't just a kitchen tool; it's a statement about how you live.	The knife holds an edge well and balances at the bolster. It costs \$90.

2.2 Rule of three

- What: Compulsive triads: three adjectives, three clauses, three examples, whether or not three things actually exist.
- Why it's a tell: Models lean on the cadence of triads as a default rhythm.
- Watch for: lists of exactly three that feel padded; "X, Y, and Z" where Z is a synonym of Y.

READS AS AI	READS AS HUMAN
The program is fast, efficient, and streamlined.	The program runs in under a second. (Three near-synonyms collapse to one real claim.)

2.3 Section-summary padding / "this article explores"

- What: Meta-commentary that narrates the structure instead of delivering content.
- Why it's a tell: Models announce what they're about to do.
- Watch for: "This article explores," "In this section, we will examine," "Let's dive into," "This guide covers."

READS AS AI	READS AS HUMAN

In this section, we will explore the history of the company and its many achievements.

The company was founded in 1998 in a Helsinki garage. (Just start.)

3. Style and formatting

3.1 Transition-word spam

- What: Overuse of a small set of connectors at the start of sentences.
- Why it's a tell: Models over-signpost logical flow. Density is the giveaway.
- Watch for: "Moreover," "Furthermore," "Additionally," "However," "Consequently," "Nevertheless," "In addition," "On the other hand," opening many consecutive sentences.

READS AS AI

Moreover, the design is sleek. Furthermore, it's durable. Additionally, it's affordable.

READS AS HUMAN

The design is sleek and durable, and it's cheap.

3.2 Conclusion tells

- What: A wrap-up paragraph that restates everything, often flagged with a connector.
- Why it's a tell: Models default to a summary close even when none is needed.
- Watch for: "In conclusion," "Overall," "In summary," "To sum up," "All in all," "Ultimately," especially leading a final paragraph that adds nothing new.

READS AS AI

In conclusion, the product offers strong value for its many users.

READS AS HUMAN

(Cut it, or end on the most useful concrete takeaway.)

3.3 Bold-lead-in bullet lists

- What: Bullets that open with a bolded phrase, a colon, then a sentence that just restates the bold phrase.
- Why it's a tell: A signature ChatGPT structure, rare in most edited prose. The bold term adding nothing the following sentence doesn't is a *strong* tell.

Bad:

- Scalability: The system is designed to scale across use cases.
- Reliability: The system is reliable and dependable.

Better: prose, or bullets where each adds a distinct fact: *Handles up to 10,000 concurrent users; survived a full data-center failover in March with no downtime.*

3.4 Excessive boldface and title case

- What: Bolding scattered through prose for emphasis; Headings In Title Case where sentence case is the norm.

- Why it's a tell: Models over-format. Title Case Headings and frequent mid-paragraph bold are uncommon in most human long-form.

3.5 Punctuation and characters

- Em dashes (—) and en dashes: a *weak* signal. Many skilled writers use them. Never lead with this or treat it as proof.
- Emoji: in professional/encyclopedic contexts, a tell.
- Curly "smart" quotes and apostrophes appearing inconsistently can indicate paste-from-chatbot, though many editors produce them too.

3.6 List overuse

- What: Breaking content into bullets where flowing prose is expected.
- Why it's a tell: Models default to lists. In contexts that call for prose (encyclopedia, essay, narrative), heavy listing reads as machine-structured.

4. Communication intended for the user (chatbot leftovers)

These are the strongest signals: text a careless editor pasted without removing the chatbot's framing. Their mere presence is close to dispositive.

- Watch for:
 - Openers: "Certainly!," "Sure! Here's...," "Of course!," "Absolutely!"
 - Self-reference: "As an AI language model," "As a large language model, I..."
 - Knowledge-cutoff disclaimers: "As of my last knowledge update," "As of [date], I don't have information on..."
 - Sign-offs: "I hope this helps!," "Let me know if you need anything else," "Feel free to ask."
 - Sycophancy: "Great question!," "You're absolutely right," "What a fascinating topic."
 - Offers to continue: "Would you like me to expand on...," "I can also provide..."
 - Refusal/hedge boilerplate: "I'm sorry, but I can't...," "It's important to approach this topic responsibly."
 - Leftover prompt echo: the article restating the instruction it was given ("Here is a 500-word encyclopedia-style article about...").

Any of these in published text is a near-certain sign the text came straight from a chatbot. Weight accordingly.

5. Citations and sourcing

Strong in encyclopedic/journalistic contexts; weaker on personal blogs.

- Watch for:
 - Fabricated citations: plausible-looking references (author, title, journal, year, DOI) that don't exist or don't match anything real.
 - Mismatched citations: a real source attached to a claim it doesn't actually support.
 - Echoing reliability-guideline language: text that parrots the *criteria* for good sourcing rather than just sourcing. Phrases like "significant, independent, reliable coverage," "substantial secondary coverage, not trivial mentions," "widely-read outlets." These echo Wikipedia's notability wording and suggest a model was prompted to write a Wikipedia article.
 - Over-listing minor media spots to manufacture notability, the way a personal bio would, rather than citing independent reporting.
 - Dead, redirected, or hallucinated URLs; fake ISBNs/DOIs.

READS AS AI	READS AS HUMAN
Profiled in multiple high-quality, independent, and widely-read outlets, providing significant secondary coverage rather than trivial mentions.	Covered by the Mesabi Daily News (March 2024) and SBS News (April 2024). (Name them; drop the notability-criteria language.)

6. Markup leakage

A *strong* tell wherever it appears, because it means text was pasted from a chatbot without cleanup.

- Watch for:
 - Stray Markdown in a place that doesn't render it: literal `**asterisks**`, `#` headers, or `-` bullets sitting as plain characters.
 - Broken or mismatched markup (a context expecting one markup language receiving another, e.g. Markdown pasted into a wiki).
 - Inconsistent heading styles within one document, as if assembled from a chatbot's default output.

7. The deepest tell: regression to the mean

This is the signal that survives after every surface phrase has been scrubbed, and it's the one to anchor on.

- What: AI-generated text smooths specific facts into generic statements that could apply to almost any subject. It explains, generalizes, and balances, but rarely sounds like a particular person who has actually been somewhere, built something, or holds a real opinion.
- Symptoms:

- Uniformity: sentences of near-identical length; paragraphs of near-identical size. Lifeless rhythm.
- No lived specifics: no number, name, date, or example that only someone close to the subject would know.
- No stance: every point hedged and counter-balanced; nothing the writer is willing to claim plainly.
- Interchangeable content: swap the topic noun and the paragraph would still "work," because it isn't really about anything.
- The fix (and the human counter-signal): specificity and voice. A real example, a concrete number, an unhedged claim, an idiosyncratic aside, varied pacing. When you're auditing, genuine lived specifics are the strongest evidence *against* AI authorship, so always look for them before concluding.

READS AS AI	READS AS HUMAN
The company has had a significant impact on the industry and continues to be a major player, known for its commitment to quality and innovation.	The company shipped the first sub-\$50 model in 2019, which is why three of its competitors now copy its hinge design.

The signal catalog is adapted from Wikipedia's "Signs of AI writing" (WikiProject AI Cleanup), used under CC BY-SA 4.0, and this adaptation is shared under the same license. Wikipedia and its contributors are not affiliated with, and do not endorse, this work. Applied and shared by Christopher Jan Benitez, christopherjanb.com.