How We Write About AI

By Mandy Hathaway — AI Ethics Specialist & Technical Writer · mandyhathaway.com

A plain-language guide for writers, editors, and product teams communicating about AI systems to non-specialist audiences.

This guide covers how we talk about artificial intelligence in external-facing content — product copy, help documentation, blog posts, and press materials. The goal is accuracy without jargon, and clarity without oversimplification.

These are not rules about what AI can do. They are rules about how we describe what it does — because the language we use shapes what readers believe, and what readers believe about AI has real consequences.

01 — PREFERRED TERMINOLOGY

Use these terms consistently across all external content. When in doubt, choose the simpler option.

Use this →	Not this →
AI system	'AI' alone when referring to a specific product or tool
model	'brain,' 'mind,' or 'intelligence' (avoid anthropomorphism)
generates / produces	'thinks,' 'decides,' 'understands,' 'knows'
trained on	'learned from' (acceptable), 'studied' (avoid)
confidence score	'certainty,' 'accuracy,' 'how sure it is'
output	'answer,' 'response,' 'opinion' (context-dependent)
limitation	'flaw,' 'bug,' 'mistake' (unless a genuine error occurred)
large language model	'LLM' on first use — spell out first, abbreviate after

Anthropomorphism — attributing human qualities to AI systems — is one of the most common and consequential writing errors in AI communication. It misleads readers about what these systems actually are and how they work, and it inflates trust in ways that can cause real harm.

The problem

When we write that an AI 'understands' your question or 'knows' the answer, we imply a kind of awareness the system does not have. Current AI systems are sophisticated pattern-matchers. They do not comprehend meaning — they identify statistical relationships in data and produce outputs accordingly. Language that suggests otherwise is inaccurate, regardless of how natural it feels to write.

In practice

✓	The model generates a response based on patterns in its training data.
✗	The AI understands your question and knows the right answer.
✓	The system produces output that may not reflect current information.
✗	The AI thinks carefully before giving you its opinion.
✓	Results may vary. The model was trained on data up to [date].
✗	The AI will always give you accurate, up-to-date information.
✓	If the output seems wrong, verify with a primary source.
✗	Trust the AI — it has processed millions of examples.

Every AI system has limitations. How we write about those limitations matters as much as acknowledging them. Vague disclaimers do not serve readers — specific, actionable language does.

Be specific about what the system cannot do

✓	This model was trained on data up to March 2024 and may not reflect recent events.
✗	Results may not always be accurate.
✓	The system performs best on straightforward factual queries. Complex reasoning tasks may produce unreliable output.
✗	Like all AI, this tool has some limitations.
✓	Outputs should be reviewed by a qualified professional before use in medical, legal, or financial decisions.
✗	Always use your own judgment when reviewing AI-generated content.

Do not bury limitations

Limitation disclosures belong near the relevant feature — not in a footnote, a separate FAQ page, or a legal disclaimer users will never read. If a system has a known failure mode that affects a specific use case, say so where that use case is described.

04 — WRITING ABOUT BIAS & FAIRNESS

AI systems can produce biased outputs. This is not a hypothetical risk — it is a documented property of systems trained on human-generated data. Our communications should reflect this honestly without either dismissing the issue or overstating our ability to resolve it.

What to say

✓	This system was trained on data that may reflect historical biases. We actively monitor outputs and update training to address these issues.
✗	Our AI is unbiased and treats all users equally.

✓	Performance may vary across demographic groups. We are committed to ongoing evaluation and improvement.
✗	We have eliminated bias from our AI system.
✓	If you receive an output that seems discriminatory or unfair, please report it using [link].
✗	AI is inherently objective because it is based on data, not human opinion.

05 — QUICK REFERENCE: WORDS TO AVOID

These words and phrases routinely appear in AI writing and routinely mislead readers. Flag them in review.

Avoid	Why
understands	Implies comprehension. The system processes — it does not understand.
knows / believes	Implies awareness or conviction. Use 'produces' or 'generates.'
thinks / decides	Implies deliberation. These systems execute — they do not deliberate.
always / never	No AI system performs with absolute consistency. Avoid absolutes.
intelligent	Contested and misleading. Prefer 'capable' or describe the specific capability.
unbiased / fair	No system trained on human data is fully unbiased. Do not claim otherwise.
accurate	Accuracy varies by task and context. Qualify or avoid.
revolutionary	Marketing language. Describe what the system actually does.