AI Writing Fingerprints: Identify (& Fix) AI-Generated Content via @sejournal, @MattGSouthern

New research shows that ChatGPT, Claude, and other AI systems leave distinctive “fingerprints” in their writing.

Here’s how you can use this knowledge to identify AI content and improve your AI-assisted output.

The AI Fingerprint: What You Need to Know

Researchers have discovered that different AI writing systems produce text with unique, identifiable patterns.

Analyzing these patterns, researchers achieved 97.1% accuracy in determining which AI wrote a particular piece of content.

The study (PDF link) reads:

> *“We find that a classifier based upon simple fine-tuning text embedding models on LLM outputs is able to achieve remarkably high accuracy on this task. This indicates the clear presence of idiosyncrasies in LLMs.”*

This matters for two reasons:

- **For readers**: As the web becomes increasingly saturated with AI-generated content, knowing how to spot it helps you evaluate information sources.

- **For writers**: Understanding these patterns can help you better edit AI-generated drafts to sound more human and authentic.

How To Spot AI-Generated Content By Model

Each major AI system has specific writing habits that give it away.

The researchers discovered these patterns remain even in rewritten content:

> *"These patterns persist even when the texts are rewritten, translated, or summarized by an external LLM, suggesting that they are also encoded in the semantic content."*

1. ChatGPT

Characteristic Phrases

- Frequently uses transition words like "certainly," "such as," and "overall."
- Sometimes begins answers with phrases like "Below is…" or "Sure!"
- Periodically employs qualifiers (e.g., "typically," "various," "in-depth").

Formatting Habits

- Utilizes **bold** or *italic* styling, bullet points, and headings for clarity.
- Often includes explicit step-by-step or enumerated lists to organize information.

Semantic/Stylistic Tendencies

- Provides more detailed, explanatory, and context-rich

answers.
- Prefers a somewhat formal, “helpful explainer” tone, often giving thorough background details.

2. Claude

Characteristic Phrases

- Uses language like “according to the text,” “based on,” or “here is a summary.”
- Tends to include shorter transitions: “while,” “both,” “the text.”

Formatting Habits

- Relies on simple bullet points or minimal lists rather than elaborate markdown.
- Often includes direct references back to the prompt or text snippet.

Semantic/Stylistic Tendencies

- Offers concise and direct explanations, focusing on the key point rather than lengthy detail.
- Adopts a practical, succinct voice, prioritizing clarity over elaboration.

3. Grok

Characteristic Phrases

- May use words like “remember,” “might,” “but also,” or “helps in.”
- Occasionally starts with “which” or “where,” creating direct statements.

Formatting Habits

- Uses headings or enumerations but may do so sparingly.
- Less likely to embed rich markdown elements compared to

ChatGPT.

Semantic/Stylistic Tendencies

- Often thorough in explanations but uses a more "functional" style, mixing direct instructions with reminders.
- Doesn't rely heavily on nuance phrases like "certainly" or "overall," but rather more factual connectors.

4. Gemini

Characteristic Phrases

- Known to use "below," "example," "for instance," sometimes joined with "in summary."
- Might employ exclamation prompts like "certainly! below."

Formatting Habits

- Integrates short markdown-like structures, such as bullet points and occasional headers.
- Occasionally highlights key instructions in enumerated lists.

Semantic/Stylistic Tendencies

- Balances concise summaries with moderately detailed explanations.
- Prefers a clear, instructional tone, sometimes with direct language like "here is how…"

5. DeepSeek

Characteristic Phrases

- Uses words like "crucial," "key improvements," "here's a breakdown," "essentially," "etc."
- Sometimes includes transitional phrases like "at the

same time” or “also.”

Formatting Habits

- Frequently employs enumerations and bullet points for organization.
- May have inline emphasis (e.g., “**key** improvements”) but not always.

Semantic/Stylistic Tendencies

- Generally thorough responses that highlight the main takeaways or “breakdowns.”
- Maintains a relatively explanatory style but can be more succinct than ChatGPT.

6. Llama (Instruct Version)

Characteristic Phrases

- “Including,” “such as,” “explanation the,” “the following,” which signal examples or expansions.
- Sometimes references step-by-step guides or “how-tos” within text.

Formatting Habits

- Levels of markdown usage vary; often places important points in numbered lists or bullet points.
- Can include simple headers (e.g., “## Topic”) but less likely to use intricate formatting than ChatGPT.

Semantic/Stylistic Tendencies

- Maintains a somewhat formal, academic tone but can shift to more conversational for instructions.
- Sometimes offers deeper analysis or context (like definitions or background) embedded in the response.

7. Gemma (Instruct Version)

Characteristic Phrases

- Phrases like “let me,” “know if,” or “remember” often appear.
- Tends to include “below is,” “specific,” or “detailed” within clarifications.

Formatting Habits

- Similar to Llama, frequently uses bullet points, enumerations, and occasionally bold headings.
- May incorporate transitions (e.g., “## Key Points”) to segment content.

Semantic/Stylistic Tendencies

- Blends direct instructions with explanatory detail.
- Often partial to a more narrative approach, referencing how or why a task is done.

8. Qwen (Instruct Version)

Characteristic Phrases

- Includes “certainly,” “in summary,” or “title” for headings.
- May appear with transitions like “comprehensive,” “based,” or “example use.”

Formatting Habits

- Uses lists (sometimes nested) for clarity.
- Periodically includes short code blocks or snippet-like formatting for technical explanations.

Semantic/Stylistic Tendencies

- Detailed, with emphasis on step-by-step instructions or bullet-labeled points.

- **Paraphrase**-friendly structure, meaning it can rephrase or re-organize content extensively if prompted.

9. Mistral (Instruct Version)

Characteristic Phrases

- Words like “creating,” “absolutely,” “subject,” or “yes” can appear early in responses.
- Tends to rely on direct verbs for commands (e.g., “try,” “build,” “test”).

Formatting Habits

- Usually applies straightforward bullet points without heavy markdown.
- Occasionally includes headings but often keeps the structure minimal.

Semantic/Stylistic Tendencies

- Prefers concise, direct instructions or overviews.
- Focuses on brevity while still aiming to be thorough, giving core details in an organized manner.

How to Make AI-Generated Content More Human

The study revealed that word choice is a primary identifier of AI-generated text:

> *“After randomly shuffling words in the LLM-generated responses, we observe a minimal decline in classification accuracy. This suggests that a substantial portion of distinctive features is encoded in the word-level distribution.”*

If you’re using AI writing tools, here are practical steps to

reduce these telltale patterns:

- **Vary your beginnings**: The research found that first words are highly predictable in AI content. Edit opening sentences to avoid typical AI starters.
- **Replace characteristic phrases**: Watch for and replace model-specific phrases mentioned above.
- **Adjust formatting patterns**: Each AI has distinct formatting preferences. Modify these to break recognizable patterns.
- **Restructure content**: AI tends to follow predictable organization. Rearrange sections to create a more unique flow.
- **Add personal elements**: Incorporate your own experiences, opinions, and industry-specific insights that an AI couldn’t generate.

Top Takeaway

While this research focuses on distinguishing different AI models, it also demonstrates how AI-generated text differs from human writing.

As search engines improve their ability to spot AI content, heavily templated AI writing may lose value.

By understanding how to identify AI text, you can create content that rises above the average chatbot output, appealing to both readers and search engines.

Combining AI’s efficiency with human creativity and expertise is the best approach.

Featured Image: Pixel-Shot/Shutterstock

https://www.searchenginejournal.com/ai-writing-fingerprints-how-to-spot-fix-ai-generated-content/541613/