

[START HERE](#)

How to recognise AI-generated content.

Practical clues, sensible checks and the limits of AI detection.

[~/GUIDES/002](#)

Notice patterns. Check sources. Never mistake a clue for proof.



00 / INTRODUCTION

Clues are not proof.

AI-generated content is everywhere, but identifying it is more difficult than spotting a certain punctuation mark.

You might recognise generic wording, strange imagery or a familiar writing style. Those observations can help you decide what to check next. They cannot, on their own, tell you who wrote something or which tools were used.

THE RIGHT QUESTION

Instead of “Can I prove this is AI?”, start with “What evidence supports this content, and what would I need to verify it?”

IN THIS GUIDE

- 01 **Writing patterns**
Common signs and innocent explanations.
- 02 **Style myths**
Why punctuation is not a fingerprint.
- 03 **Detection tools**
What their scores actually mean.
- 04 **Images & video**
Visual clues and provenance.
- 05 **Verification**
How to check sources and claims.
- 06 **Responsible conclusions**
When to ask, not accuse.

01 / WRITING

Patterns worth noticing.

These features may appear in machine-written text, but human writers use them too. Look at the whole piece and its context.

01 / STRUCTURE

Overly tidy paragraphs

Every paragraph follows the same rhythm: statement, explanation and neat little conclusion. It may simply be the writer's preferred style.

02 / LANGUAGE

Stock phrases

Repeated phrases such as "in today's fast-paced world" or "unlock your potential" can make writing feel generic. Templates produce similar effects.

03 / DETAIL

Lots of words, little substance

The piece talks confidently without naming dates, examples, evidence or practical details. This is a reason to verify its claims, not proof of AI.

04 / CONSISTENCY

An oddly uniform voice

Long sections may share the same sentence length and tone. Editing, formal writing and brand guidelines can create that consistency too.

THE LIMIT

A distinctive style is a clue to investigate, not enough evidence to label somebody's work.

02 / STYLE MYTHS

The em dash is not proof.

A popular online shortcut is to treat em dashes, tidy lists or certain vocabulary as signs of ChatGPT. It is not a reliable test.

THE QUICK ASSUMPTION

“It uses an em dash.”

The conclusion: “It must be AI-written.” The same accusation is sometimes made about bullet points, semicolons, perfectly balanced paragraphs or British versus American spelling.

WHAT THE EVIDENCE SUPPORTS

“It has a particular style.”

An em dash has been normal punctuation for generations. A human might also write formally, use templates or ask for editing help. AI-assisted editing does not necessarily mean an AI created the whole piece.

BETTER APPROACH

Ask about the content’s accuracy, originality and sources rather than treating one punctuation mark as a smoking gun.

03 / AI DETECTORS

A score is not a verdict.

Automated text detectors look for statistical patterns. Their accuracy varies across writing types, models, languages and revisions.

WHAT THEY RETURN

An estimate, not a fact

A percentage in a detector report is not independent proof that a person used AI. It depends on the tool and the material it was tested against.

FALSE POSITIVES

Human work can be flagged

Research has shown some detectors mislabel human writing, particularly work by non-native English writers. Consequences can be serious if a score becomes an accusation.

FALSE NEGATIVES

Generated text can pass

Changing, editing or paraphrasing AI-generated text can weaken detection. A low score is not a certificate that no AI was involved.

RULE OF THUMB

Never use a detector score as the sole basis for accusing a student, colleague, customer or creator.

04 / VISUAL MEDIA

Look beyond the pixels.

Image models keep improving. Older visual giveaways are becoming less useful, and real photos can look unusual too.

01 / TEXT

Signs and labels

Garbled lettering or inconsistent typography may be worth checking. Modern generated images can also reproduce clean text.

02 / GEOMETRY

Physical details

Inspect reflections, repeated objects, shadows or architectural continuity. Lens distortion and editing can cause similar oddities.

03 / CONTEXT

Does the scene fit?

Is the claimed location real? Does the weather, event, date or background match independent reports? Context beats a single visual quirk.

04 / SOURCE

Where did it begin?

Find the earliest available upload, original photographer, event organiser or official source. Reposts often strip useful information.

IMPORTANT

An image can be genuine yet miscaptioned, altered or taken at another time. “Real photo” does not equal “true claim”.

05 / SOURCE CHECKS

Verify it systematically.

A practical method works better than guessing from the style of a post or the appearance of a picture.

01

Find the original

Look for the first publisher, photographer or document. A screenshot of a screenshot usually loses context.

02

Check the claim

Identify the precise claim: who, what, where and when. Search for corroboration from reliable independent sources.

03

Inspect the evidence

For media, check provenance details where available. Reverse-image search and surrounding context can help.

04

State the limits

If you cannot verify it, say “unverified”. Avoid making a stronger claim about AI authorship than the evidence supports.

ABOUT CONTENT CREDENTIALS

C2PA-style provenance can help show an asset’s recorded history, but missing metadata proves little and metadata alone cannot establish whether a claim is true.

06 / EXAMPLE

A local event rumour.

Imagine a viral Facebook post claiming an unusual incident happened on a street near you.

The post includes an impressive image, a dramatic description and no clear source. Here is a better way to assess it.

FIRST REACTION

Do not jump to “AI”.

The image looks strange and the caption feels generic. That might raise a question, but neither detail establishes what happened.

NEXT STEP

Identify a claim to check.

Search for the claimed place, date and event. Look for local reporting, public statements or another original recording.

FAIR CONCLUSION

Describe what you know.

“I could not verify where this image came from or whether it depicts this event.” That is stronger and fairer than an unsupported accusation.

If the post identifies a private person, avoid amplifying it or making allegations without reliable evidence.

07 / USING AI

Ask AI to help verify.

AI can organise questions and suggest checks. It can also invent citations or misread images, so its answer needs verification.

WEAK QUESTION

“Was this written by AI?”

A model might sound certain based on stylistic cues. It cannot reliably infer who wrote a passage merely from how it reads.

BETTER QUESTION

“What can we check?”

“List the factual claims in this post. Separate verifiable statements from opinion. Suggest independent sources and tell me what cannot be determined from the text alone.”

DO THE CHECKS YOURSELF

Open the suggested sources. Check dates, exact quotations and whether they support the claim. Do not copy invented citations.

08 / QUICK REFERENCE

A fair checking checklist.

Use these questions before you share a suspected AI-made post or challenge its author.

01 Original source

Can you identify where the material first appeared?

02 Clear claim

What exactly is being asserted, rather than implied?

03 Independent support

Do other reliable sources confirm the relevant details?

04 Media context

Could this picture or recording be edited, old or miscaptioned?

05 Provenance

Are credible creation records available? What do they actually show?

06 Detector limitations

Have you treated any detector output as fallible evidence?

07 Fair wording

Are you saying “possible” or “unverified” when certainty is unwarranted?

KEEP THIS PAGE

The aim is not to win a guessing game. It is to avoid being misled, or misleading somebody else.

09 / FAIRNESS

Be careful with labels.

Being observant is useful. Publicly accusing someone of using AI based on weak evidence is not.

PEOPLE

Ask before accusing

If AI use matters, ask whether tools were used and what role they played. An accusation based on style can damage trust and reputation.

EDUCATION

Respect assessment rules

Check the specific school or workplace policy. Grammar assistance, brainstorming and full generation may be treated differently.

PUBLISHING

Disclose where relevant

For your own work, explain meaningful AI assistance when the audience, client, policy or publisher expects that transparency.

PUBLIC CLAIMS

Separate authenticity from truth

Even verified human authorship cannot make an unsupported claim accurate. Evidence is still what matters.

THE BEST HABIT

Replace “That looks like AI” with “What evidence would actually settle this?”

10 / KEEP LEARNING

Look closer. Judge fairly.

Spotting AI use is less about catching a writing habit and more about checking what is genuine, sourced and accurate.

As AI tools become part of everyday writing, design and research, many pieces of work will combine human and machine input. There may be no useful single-word label. You will get further by examining the specific claim and the evidence behind it.

REMEMBER THESE THREE THINGS

01 **Style is not a fingerprint**

Generic phrases or punctuation are not proof of AI authorship.

02 **Detector scores have limits**

They can mislabel human writing or miss generated text.

03 **Verify the claim itself**

Source checks and fair wording matter more than a hunch.

FURTHER READING / REFERENCE

NAACL 2025 / Text detector evaluation



Liang et al. / False positives in AI detection



C2PA / Content Credentials explainer



Explore more at nathm.net