

PAUSE Before You Trust or Share

Five steps for suspicious images, audio, video, and messages

Use PAUSE when a post, clip, or message tries to make you react before you think.

Pause	Stop before you forward, pay, click, or respond while emotions are high.
Ask	What exactly is this asking me to believe, do, send, or share?
Use the source	Trace it to the original account, full recording, article, or official website.
Search elsewhere	Search the key claim independently. Do not rely on one repost.
Exit pressure	If someone wants speed, secrecy, money, or personal information, verify another way.

USE IT WHEN...

- shock or outrage
- money or login codes
- no clear source
- "share this now" pressure

THREE QUESTIONS THAT SLOW THE RUSH

1 What is the exact claim?	2 What evidence actually supports it?	3 What happens if I wait one minute?
----------------------------	---------------------------------------	--------------------------------------

Common mistake: treating a screenshot, repost, or clipped video as the original source.

STILL UNSURE?

Do not share yet Waiting costs less than spreading a false claim.	Keep the original Save the post, file, account name, and time.	Ask a trusted source Use an official route or someone who can verify it.
---	--	--



SPOT THE FAKES

Scan to practice PAUSE with real examples

spotthefakes.org

PAUSE in Practice

A worked example

Example claim: "This video proves school is closed tomorrow."

Pause	Do not forward it yet. Waiting prevents one rumor from becoming ten.
Ask	What does the video itself prove? Does it show an official announcement, or only a claim?
Use the source	Open the school or district website yourself. Do not use a link supplied by the repost.
Search elsewhere	Check email alerts, the school portal, local notices, or another independent source.
Exit pressure	"Share before it is deleted" is pressure. It is not evidence.

IF THE OFFICIAL SOURCE IS SILENT

1. Wait

Do not turn missing information into confirmation.

2. Use a second route

Check the portal, saved number, or official email.

3. Label it honestly

Say "unconfirmed" instead of guessing.

BEFORE YOU SHARE

- ✓ Original source found
- ✓ Independent confirmation found
- ✓ Date and location checked
- ✓ No money, secrecy, or login-code pressure

If confirmation never appears, leave the claim unshared or label it clearly as unconfirmed.

PAUSE is the first move. Source, context, provenance, or technical tools may still be needed.

A SAFE RESPONSE

"I have not found official confirmation, so I am not sharing this yet."

LABEL THE STATUS HONESTLY

Confirmed

An official or independent source supports the claim.

Unconfirmed

You found the claim, but not enough evidence to support it.

False

Reliable evidence directly contradicts the claim.

Check the Source and Context

Real media can still carry a false story

A genuine image can mislead when the date, place, crop, or caption is wrong.

SAME PHOTO



ACTUAL CONTEXT

Karmakol Festival crowd, Sudan,
December 2017



FALSE CONTEXT

"Breaking: protest happening right
now"

The file is identical. The story is not.

Wrong date

Old footage presented as current.

Wrong place

A real event labeled as another
location.

Missing context

A crop or clip removes meaning.

False caption

The media is real; the attached claim is
not.

Original

Where did this first
appear?

Date

When was it
recorded?

Place

Where was it
recorded?

Claim

What does it actually
prove?

Separate the file from the caption: first describe only what the media shows. Then list every detail added by the caption.

Next: find the earliest credible copy, search the exact claim, and look for independent reporting that does not merely repeat the post.



SPOT THE FAKES

Scan to practice source and context checks

spotthefakes.org

Trace the Claim

A one-minute source and context method

Trace the claim without relying on one branded tool.

A ONE-MINUTE TRACE

1

Find the earliest source

Search the exact caption, event name, or distinctive phrase in quotation marks.

2

Check the date

Compare the first credible upload with the date claimed in the post.

3

Check the place

Use landmarks, signs, language, weather, uniforms, and maps.

4

Find the full video

Search a clear frame for earlier uploads, then compare the surrounding footage.

5

Find independent confirmation

A second source should do its own reporting, not simply repeat the same post.

SEARCH RECIPE

EXACT PHRASE

"exact caption" + place + date

VIDEO FRAME

pause - save clear frame - search for earlier matches

EXAMPLE: A CLIP SHOWS SMOKE

IT PROVES

Smoke was recorded in the clip.

IT DOES NOT PROVE

Where, when, why, who caused it, or whether the caption is accurate.

Do not let repetition fool you. Ten reposts can all come from one mistaken or deceptive source.

WORKED TRACE: THE FESTIVAL PHOTOGRAPH



Earliest credible source:
Wikimedia Commons

Place: Karmakol, Sudan

Date: December 2017

False claim tested: current protest

Do not ask only, "Is this real?" Ask, "Is this being presented truthfully?"

Verify the Person Another Way

A familiar voice or face is not proof of identity

A familiar voice, face, account name, or live video is not proof of identity. Verify through a route that did not come from the suspicious contact.

THE INDEPENDENT-ROUTE METHOD

- 1 End contact**
Stop replying, hang up, or close the message. Do not let the sender control the pace.
- 2 Open a trusted route**
Use a saved number, official directory, known email, school portal, or organization website.
- 3 Verify the request**
Ask whether the person actually contacted you and whether the money, code, or information request is real.
- 4 Act only after confirmation**
Do not send money, passwords, personal information, or one-time codes before independent verification.

THE INDEPENDENT-ROUTE TEST

Could you reach the person or organization without using the number, link, account, or instructions in the suspicious message?

YES

Use that route to verify the request.

NO

Find an official website, saved contact, or another trusted person first.

USEFUL WORDS TO SAY

"I am going to call you back using the number I already have."

"I will confirm this through the official portal before I act."

COMMON WARNING SIGNS

Urgency

Secrecy

New number

Gift cards or crypto

Login codes

PREPARE BEFORE AN EMERGENCY

Prepare now: save trusted contacts, choose a backup person, and agree on a private phrase that is not posted publicly.

The more convincing the impersonation, the more important independent verification becomes.



SPOT THE FAKES

Scan to build an identity-verification plan

spotthefakes.org

Choose an Independent Route

Match the situation to the safest verification method

Situation	Best verification route
Family member	Hang up and call the number already saved in your phone.
Boss or school	Use the official directory, portal, workplace system, or known email.
Official social account	Check the exact handle and navigate from the organization's official website.
Live video call	Treat the face and voice as appearance, not independent identity verification.

IF THE PERSON DOES NOT ANSWER

1. Stop Do not act on the original request.	2. Ask another person Contact another relative, coworker, or office.	3. Use official contact Call through a published number.
---	--	--

DO / DO NOT

DO • Call a saved number • Use an official portal • Ask another trusted person	DO NOT • Trust caller ID alone • Use a link from the message • Share one-time codes
---	--

WHERE TO FIND THE TRUSTED ROUTE

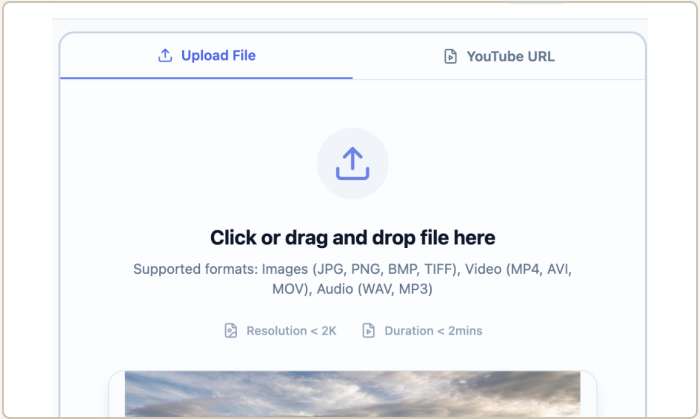
Bank or card issuer Use the number on the card or inside the official app.	School or employer Use the official site, portal, directory, or known internal system.
Government office Navigate to the agency's official website yourself.	Public figure or organization Start from the official website, not a link in the message.

A route is independent when you found it yourself, it existed before the suspicious message, and the sender cannot control it.

Deepfake Detectors

How to use a real multi-model detector

Deepfake-O-Meter lets you upload one file and run several research detectors. The screenshot below is the actual upload screen.



Actual upload screen: supported formats, resolution, duration, and file preview are visible.

USE IT CAREFULLY

1. Upload the original file

Keep the original image, video, or audio format whenever possible.

2. Choose suitable models

Check media type, manipulation type, and model year.

3. Wait for completion

A blank or "Processing" row is not a result.

4. Record every output

Save model name, year, label or score, and any error.

BEFORE YOU UPLOAD

File

Use the original, not a screenshot or repost.

Fit

Choose detectors built for your media and manipulation type.

Privacy

Review storage and sharing choices.

Privacy check: before uploading private media, review who operates the service and how files may be stored or shared.

A detector result is technical evidence, not a verdict about whether the entire story is true.



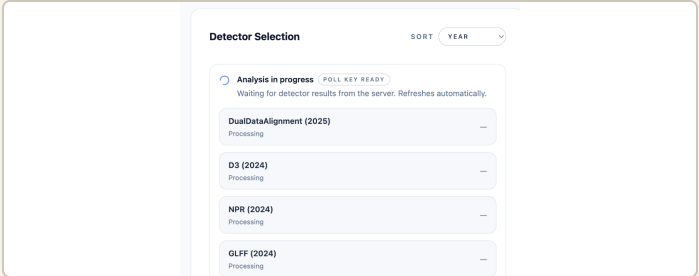
SPOT THE FAKES

Scan for the detector guide and current links

spotthefakes.org

Read the Results

What multi-model detector output does and does not mean



Actual detector-selection screen. "Processing" is not a result; wait for every selected model.

WHEN THE RESULTS ARRIVE

Result	What to think
Strong signal or high score	Useful warning evidence. Check whether several suitable detectors agree.
Weak or low signal	The model found little of the signal it was trained to detect. This does not prove authenticity.
Mixed outputs	Treat the case as unresolved. Model disagreement is information.
Error or processing	Unsupported, failed, or still running; wait or choose another model.

Do not compare scores as if standardized. Different models may use different scales, labels, and training targets.

WHY MODEL FIT MATTERS

Media type Image, video, and audio detectors are not interchangeable.	Manipulation type A face-swap model may not fit a generated landscape.
File quality Compression, cropping, and re-encoding can weaken signals.	Generator age Newer generators may fall outside training data.

USE THE RIGHT EVIDENCE CHAIN

Detector output	Source and context	Provenance	Independent confirmation
-----------------	--------------------	------------	--------------------------

"No AI detected" means no supported signal was found by that model. It does not mean "proven human-made."

Check DALL-E and Nano Banana

Actual results from OpenAI Verify and Gemini

Use the original file whenever possible. These are real results from the two provider-specific checkers.



OpenAI Verify

openai.com/verify

Use first for: supported OpenAI or ChatGPT images, including DALL-E media when signals are present.

Result shown: "Generated with OpenAI tools."



Gemini

gemini.google.com

Use first for: Nano Banana and other Google AI media.

Ask: "Was this image created or altered by Google AI?"

WHICH CHECKER SHOULD I TRY?

ChatGPT / DALL-E

Start with OpenAI Verify.

Nano Banana / Google AI

Start with Gemini's verification question.

Provider unknown

Try both, then continue with source, context, and provenance checks.

Signal found

Useful evidence that the matching provider generated or edited the file.

No signal found

Inconclusive. The file may be unsupported, changed, older, or from another provider.

Original file matters: screenshots, social-media downloads, and PDF conversion may remove or weaken provenance signals.



SPOT THE FAKES

Scan for OpenAI Verify, Gemini, and current instructions

spotthefakes.org

SPOT THE FAKES

Run the Two Checks

Step-by-step checks for OpenAI and Google AI media



Generated with OpenAI tools
This content was generated using OpenAI tools. The signals below provide supporting evidence for this result.
View details

OPENAI VERIFY - OPENAI / DALL-E MEDIA

- 1 Keep the original PNG, JPG, WEBP, or supported audio file.
- 2 Go to openai.com/verify, upload it, and save the exact result wording.



was this image generated by AI?

Based on our analysis, this image contains signals indicating it was generated or edited with Google AI.
If you have questions about this technology, you can find more information at support.google.com/gemini/verify_ai

GEMINI - NANO BANANA / GOOGLE AI MEDIA

- 1 Open Gemini and upload one image, video, or audio file.
- 2 Ask: "Was this image created or altered by Google AI?"

INTERPRET THE RESPONSE

Found

Save the exact wording and treat it as provider-origin evidence.

Not found

Do not call the file human-made. Try other evidence and consider file conversion.

COMMON MISTAKES

- ✓ Uploading a screenshot instead of the original
- ✓ Asking only "Is this AI?" instead of Google's verification question
- ✓ Using the wrong provider's checker
- ✓ Stopping after one missing signal

AFTER EITHER CHECKER

- ✓ Save the exact result wording
- ✓ Check source and context
- ✓ Keep the original file
- ✓ Do not treat a missing signal as proof

Scope matters: neither checker is universal. Provider signals answer an origin question, not whether the caption is true.

HOW TO DESCRIBE AN INCONCLUSIVE RESULT

"No supported provider signal was found. That does not prove the file is human-made."

Save a usable record: keep the original file, a screenshot of the exact result, and the date and claim you checked.

Content Credentials

Inspect a file's origin and edit history

Go to contentcredentials.org/verify. It opens the official verifier at verify.contentauthenticity.org.

1

Open the verifier

Use the official Content Credentials link.

2

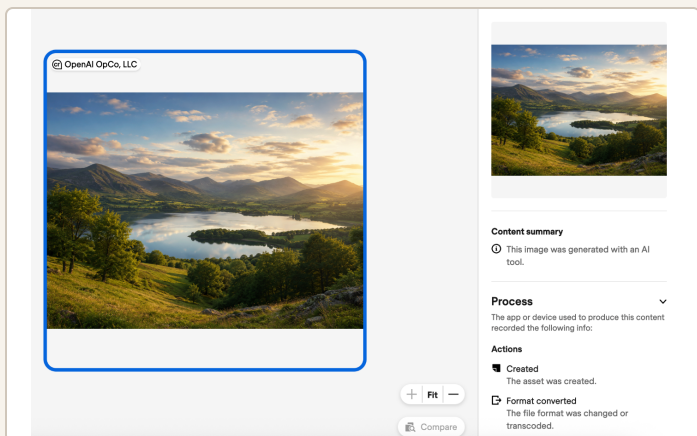
Upload the original file

Choose a device file or drag it onto the page.

3

Read the side panel

Inspect issuer, summary, process, and actions.



Actual verifier result: the main image and the full right-side provenance panel are visible.

1

Issuer

The badge says "OpenAI OpCo, LLC." This identifies who issued the credential.

2

Content summary

The panel states that the image was generated with an AI tool.

3

Process and actions

The record shows "Created" and "Format converted," so the history includes creation and a later format change.

The verifier shows provenance: who issued the record and what actions were recorded. It does not judge whether a caption is true.

NO CR PIN VISIBLE?

1. KEEP THE ORIGINAL FILE

A screenshot or converted copy may lose provenance information.

2. UPLOAD IT TO THE VERIFIER

If no credential appears, treat the result as inconclusive.



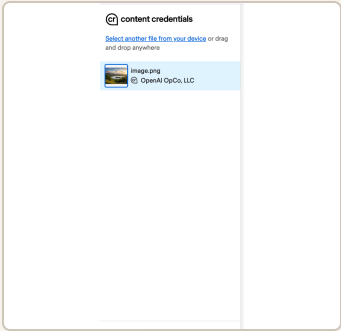
SPOT THE FAKES

Scan to open the Content Credentials verifier

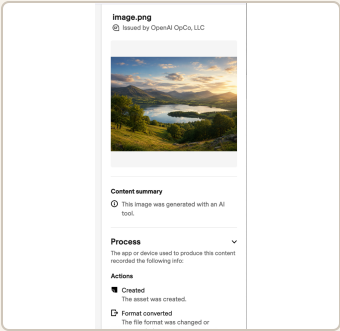
spotthefakes.org

How to Read the Verifier

What to click, what to inspect, and what not to conclude



1. Upload control and file list



2. Issuer, summary, process, and actions

FIVE THINGS TO DO

- 1

Upload the original file
Choose "Select another file from your device" or drag it in.
- 2

Check the issuer
Who signed the credential?
- 3

Read the content summary
Look for a plain-language statement about creation.
- 4

Expand Process
Review actions such as created, edited, or format converted.
- 5

Separate provenance from truth
A valid credential can support where the file came from. It cannot prove that a caption or interpretation is accurate.

WHAT YOU SEE AND WHAT YOU MAY CONCLUDE

What you see	Safe conclusion
Issued by OpenAI + AI-generated summary	Strong provenance evidence that OpenAI generated the file.
Created	The credential records creation of the asset.
Format converted	The file changed format after creation; that alone does not mean deception.
No credential found	Inconclusive. Many genuine and AI files have no available credential.

Provenance asks, "Where did this file come from?" Context asks, "Is the story about it accurate?" You often need both.

IF NO CREDENTIAL APPEARS

Do not label it fake Missing provenance is inconclusive.	Try a provider checker OpenAI Verify or Gemini may find provider-specific signals.	Continue with context Trace the source, date, place, and claim.
--	--	---