

Verify the Person Another Way

Use this when a call or message may impersonate someone.

Trace the Source and Context

Check whether media is AI-made, real, or miscaptioned.

Deepfakes can look and sound real.

AI can imitate a person, voice, place, or event. That can make scams and false stories more convincing.

Which image is fake: A or B?

One is a photograph. One was generated with AI. Turn the brochure upside down for the answer.



Images: Ben Bowens, Iceland, CC0; VulcanSphere, FLUX 1.1 Pro Raw, CC0. Both via Wikimedia Commons.

Why use it: AI can copy a familiar face or voice. This check does not tell you whether the file is AI. It tells you whether the real person made the request.

A familiar face, voice, account name, or caller ID does not prove who is behind the message.

How to use an independent route

- End the suspicious contact.** Hang up, stop replying, or close the message. Do not click its links.
- Open a route you already trust.** Use a saved number, official app, known email, official website, portal, or directory.
- Verify the exact request.** Ask whether the person contacted you and whether the request for money, information, a password, or a code is real.
- Act only after confirmation.** Do not send money, personal information, passwords, or one-time login codes until the request is confirmed.

Choose a route the sender cannot control

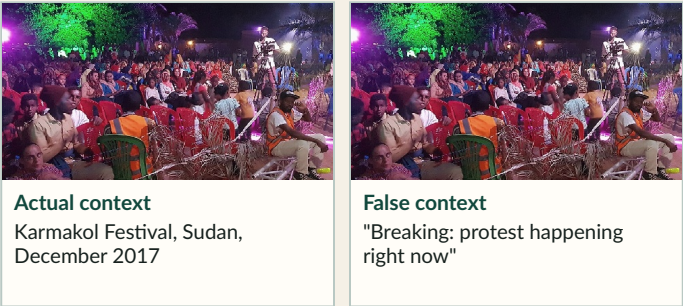
Family or friend	Call the number already saved in your phone.
Bank or card	Use the number on the card or inside the official app.
School or work	Use the official portal, directory, or known email.
Organization	Open its official website yourself. Do not use a link in the message.

If the person does not answer

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

Warning signs: urgency, secrecy, a new number, gift cards or crypto, login codes, or pressure not to contact anyone else.

Why use it: Finding the earliest reliable copy can show that a photo or clip is old, cropped, relabeled, or out of context. It can also show whether you still need a watermark or detector check.



How to use it

- Open the original.** Find the full post, recording, article, or official account. Do not rely on a screenshot or repost.
- Reverse-search it.** In the Google app, tap Lens and upload the image. For video, save and search a clear frame.
- Check the source, date, and place.** Compare the earliest reliable upload with the claim.
- Search the exact words.** Put the caption or a distinctive phrase in quotation marks. Add a place or date.
- Separate the file from the story.** Write what the media itself shows. Then list what the caption adds and check the claim with an independent source.

What your search can tell you

Earlier match found: compare how the date, place, crop, or caption changed.

Real file, false story: the media may be real even when the attached claim is misleading.

No reliable match: you still do not know. Try a watermark, Content Credentials, or detector check.

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

A deepfake is media made or changed with AI. It can copy people, invent evidence, or make a false story look convincing. Verify before you believe, share, pay, or respond.

22,000+

FBI complaints involved AI-enabled crime in 2025.

\$893M+

in reported losses were tied to those complaints.

Hard to spot: deepfakes were about 1 in 5 biometric fraud attempts in one report, and 53% of Americans lacked confidence in telling AI-made content from human-made content.

Why protect yourself

Impersonation	A copied face or voice can pretend to be someone you trust.
Fraud	A fake emergency can pressure you to send money, codes, or personal information.
Misinformation	AI media or false captions can make invented claims look real.

Four checks inside

1. Trace the source and context

3. Use a deepfake detector
2. Verify the person another way

4. Check watermarks and credentials

Statistics: FBI IC3 2025; Entrust 2026; Pew Research Center, Sept. 2025.

Answer: B is AI-generated.

Spot the Fakes

Deepfake defenses exist. People should know how to use them.

Spot the Fakes teaches people how to check suspicious images, videos, audio, and messages. The site gives practical steps, practice, and free materials. No technical background is needed.

Practice with real examples

Use the online quiz and guided examples to practice tracing a source, verifying a person, using a detector, and checking watermarks or Content Credentials.

Get free educational materials

Access free printable guides that explain what to do when an image, video, audio clip, or message seems suspicious.

What you will find online

- Step-by-step verification guides
- Current links to public tools
- Practice quizzes and worked examples
- Free printable educational materials

Why this exists: people often hear that deepfakes are dangerous without being shown what to do. Spot the Fakes gives them clear steps, practice, and free materials.



FOR MORE INFORMATION
spotthefakes.org

Practice the checks and access free materials.

More people should know these defenses exist and know how to use them.

Check Watermarks and Content Credentials

Check for provider watermarks or a signed file history.

Why use them: Some AI tools add invisible watermarks or Content Credentials. A match can show which system made or edited the file. No match does not prove the file is human-made.



OpenAI Verify
openai.com/verify

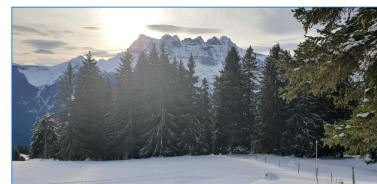
1. Upload the original supported file.
2. Save the exact result wording.
3. Try this first for supported OpenAI or ChatGPT media.



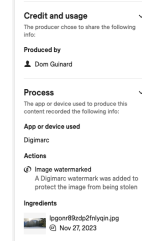
Gemini and SynthID
gemini.google.com

1. Upload the original image, video, or audio.
2. Ask: "Was this created or altered by Google AI?"
3. Save Gemini's exact response.

Check Content Credentials



File preview



Actual verifier example: the file preview is on the left and the signed details are on the right.

1. **Open the verifier.** Go to contentcredentials.org/verify.
2. **Upload the original.** Use the original file from your device.
3. **Read the details.** Check issuer, summary, process, and actions.

How to read any result

Match found: save the exact wording. It can show the provider, issuer, or recorded file history.

No match found: the result is inconclusive. The file may be unsupported, changed, older, or from another provider.

Keep the original: screenshots, social-media downloads, edits, and format changes can remove these signals.

These tools can show where a file came from. They cannot prove that its caption or story is true.

Use a Deepfake Detector

Check whether a file may be AI-made or edited.

Why use it: A detector looks for patterns linked to AI generation or editing. Use one when the source is unclear or no provider watermark is available. A detector gives a clue, not a final answer.

Detector	Result	Details
NoDown	0.0% AI-Generated Likelihood	Completed
GLFF	2.8% AI-Generated Likelihood	Completed
CLIP-ViT	0.1% AI-Generated Likelihood	Completed
NPR	Processing... (estimated time: 30s)	

Actual DeepFake-O-Meter results. Each model gives a separate score, and one model is still processing.

Short tutorial

1. **Use the original file and check privacy.** Avoid screenshots, reposts, and converted copies. Review how the service stores uploads.
2. **Open the detector and upload one file.** Use the current DeepFake-O-Meter link at spotthefakes.org and choose image, video, or audio.
3. **Choose suitable models and wait.** Match the model to the media type. "Processing," a blank row, or an error is not a result.
4. **Save every result.** Record the model name, year, score or label, and any error.

How to read the result

Several suitable models show a strong signal: useful warning evidence. Continue with source, context, and origin checks.

Low or no AI signal: the model found little of what it was trained to detect. This does not prove the file is human-made.

Mixed results: treat the case as unresolved. Different models can disagree or use different scales.

Error or processing: there is no usable result yet. Wait, retry, or choose another suitable model.

"No AI detected" does not prove the file is human-made.