

CLASSROOM LESSON · GRADES 9-12

Real or fake?

Verifying voices, videos, images and offers in the age of AI

Students learn how AI-made audio, video, images and polished scam messages are used in fake emergencies, fake job and scholarship offers and online rumors, practice lateral reading and verification by a known route, sort real-world style messages into Verify first, Red flag scam or Likely legitimate, and write a personal verification routine. The lesson also covers responsible behavior with AI media of real people and, briefly and carefully, where to get help if someone threatens to share an image.



GRADE BAND
9-12

TIME
50 minutes

FORMAT
Whole class + groups

PREP
Print & cut, ~10 min

BIG IDEA



You cannot always tell real from fake by looking or listening, but you can always check by a route you already trust, and you can choose never to make or spread fake media of a real person.



Students will be able to

- ☐ Describe how AI-made voices, videos, images and messages are used in scams and rumors, including fake family emergencies and fake job or scholarship offers aimed at teens
- ☐ Use lateral reading (leaving the source and checking what trusted sources say) and verification by a known route to judge a claim or message
- ☐ Sort messages and claims into Verify first, Red flag scam or Likely legitimate and justify each decision
- ☐ Explain why creating or sharing deceptive media of a real person causes harm, can break school rules and in some cases the law, and name where to get help if someone threatens to share an image



Materials

- Projector or board for the slide outline
- Worksheet 1: Verify first, Red flag scam or Likely legitimate (1 per student; contains 12 written messages and claims)
- Worksheet 2: My verification routine (1 per student)
- Scenario cards (1 set per group of 3-4, cut apart)
- Optional: one device per group for a 3-minute lateral reading practice on a teacher-chosen public claim
- Half sheets for the exit ticket

Before class

- Print both worksheets and one card set per group.
- Write on the board: the school counselor's name, how students report a safety concern at your school, and the two help resources: takeitdown.ncmec.org and the NCMEC CyberTipline (report.cybertip.org or 1-800-843-5678).
- Review your district's student conduct policy on harassment, impersonation and sharing images of others so you can describe consequences in one or two general sentences.
- Talk briefly with your school counselor before teaching. If a student discloses that they are being threatened or harmed, follow your school's reporting procedures and connect the student with the counselor the same day.
- If you use the optional lateral reading practice, choose a public, school-appropriate claim in advance (for example, a viral 'new law' rumor) and confirm the trusted sources you expect students to find.



Vocabulary

- **Deepfake** — Audio, video or an image made or changed with AI so that a real person seems to say or do something they never said or did.

Vocabulary (continued)

- **Voice cloning** — Using AI and a short recording of someone's voice to make new audio that sounds like them.
- **Synthetic media** — Any image, audio, video or text created by AI rather than captured from real life. It is not always harmful, but it can be used to deceive.
- **Lateral reading** — A fact-checker habit: leave the unfamiliar source, open new tabs and search what trusted sources say about the person, site or claim.
- **Verify by a known route** — Checking a message by contacting the person or organization through a phone number, app or website you already had, never the one in the message.
- **Social engineering** — Tricking people into giving up money, information or access by playing on trust, fear, urgency or excitement.
- **Safe word** — A private word or phrase a family agrees on in advance to confirm it is really them during an emergency call.
- **Sextortion** — When someone threatens to share a real or fake sexual image of a person unless they pay or send more. It is a crime, and the person targeted is not at fault.

Standards alignment

- **ISTE Standards for Students 1.3.b** — Evaluate information (paraphrased): students judge the accuracy, validity, bias, origin and relevance of digital content before they trust or share it.
- **ISTE Standards for Students 1.2.a** — Digital footprint (paraphrased): students manage their digital identity, understand the lasting effects of their online actions on themselves and others, and make safe, legal and ethical choices.
- **ISTE Standards for Students 1.2.b** — Online interactions (paraphrased): students interact with empathy and responsibility online, which includes not creating or spreading deceptive media of other people.
- **CSTA PK-12 Computer Science Standards (2026) HS-ALG-PS-05** — Paraphrased: evaluate AI-generated output such as text and images for accuracy, bias and potential harm.
- **CSTA PK-12 Computer Science Standards (2026) HS-SYS-SE-32** — Paraphrased: classify the causes and impacts of security breaches and social engineering attacks for individuals and other groups.
- **CSTA PK-12 Computer Science Standards (2026) HS-SOC-HU-43** — Paraphrased: evaluate how human choices in using and deploying computing technologies create risks, benefits and long-term impacts.
- **CSTA K-12 Computer Science Standards (2017) 3A-IC-24** — Paraphrased: evaluate how computing affects personal, ethical, social, economic and cultural practices.

Aligns with commonly used digital citizenship and media literacy frameworks (ISTE Standards for Students) and with both the 2026 CSTA PK-12 and the 2017 CSTA K-12 computer science standards; Georgia districts can map it to their state computer science and digital literacy expectations and to local student conduct, acceptable-use and student safety policies.

Lesson sequence · 50 minutes

5' Warm-up

- 1 Show slide 1. Ask: Have you seen a video or image online this month that you later found out was fake? No names or links needed.
- 2 Ask students to vote with fingers (1 to 5) on how confident they are that they could spot an AI-made voice on a phone call. Note the spread.
- 3 Explain the goal: today is not about becoming a deepfake detector. It is about building a routine that works even when you cannot tell.

Say: AI fakes are getting good enough that looking harder will not always work. So we are going to stop asking 'does it look real?' and start asking 'how can I check this by a route I already trust?'

12' Mini-lesson

- 1 Slide 2 (2 minutes): What AI fakes are used for. Cloned voices in fake family emergencies, fake video calls and profile photos, polished scam texts with no spelling mistakes, and fake job and scholarship offers that target teens.
- 2 Slide 3 (2 minutes): Why the old clues fail. Good grammar no longer means a message is safe. Visual glitches like odd hands, teeth or shadows can help, but newer fakes may have none.
- 3 Slide 4 (3 minutes): Lateral reading. Leave the source, open a new tab and search the person, site or claim. Check what trusted outlets, the official organization's own site or fact-checkers say. Three posts that all copy one source count as one source.
- 4 Slide 5 (2 minutes): Verify by a known route. Call the number you already have, open the official app or site yourself, ask in person, and agree on a family safe word.
- 5 Slide 6 (3 minutes): Responsible behavior and help. Do not create or share fake images, audio or video of real people. If someone threatens to share an image of you, real or fake: do not pay, do not send more, save the messages, block, and tell a trusted adult. Point to the help resources on the board.

Say: *If anyone ever threatens to share an image of you, real or fake, it is not your fault, you are not in trouble, and you do not have to handle it alone. Don't pay and don't send more. Tell a trusted adult or the counselor, and there are free services on the board that help take images down.*

18' Activity

- 1 Groups of 3-4 read the 12 messages and claims on Worksheet 1, sort each into Verify first, Red flag scam or Likely legitimate and write a one-line why (about 9 minutes).
- 2 For every Verify first item, groups write the exact known route they would use to check (for example, 'the district website typed in myself' or 'call Mom's saved number') (about 3 minutes).
- 3 Optional lateral reading practice: groups search a teacher-chosen public claim in a new tab and report what two trusted sources say (about 3 minutes). Without devices, groups instead discuss two scenario cards.
- 4 Quick share: each group explains one item they disagreed about and how they settled it (about 3 minutes).

Say: *For Verify first, 'look it up' is not enough. Tell me exactly where: which number, which app, which website you would type in yourself.*

8' Discussion

- 1 Use two or three discussion questions, starting with the fake principal video, because it shows how sharing a fake can cause harm even without a scam.
- 2 Discuss responsibility: What happens to a real classmate when a fake clip of them spreads? Describe your school's conduct policy in general terms.
- 3 Restate the help routes on the board. Nobody gets in trouble for asking for help.

7' Exit ticket

- 1 Students start Worksheet 2 (My verification routine) and complete at least steps 1 to 4 in class (about 4 minutes).
- 2 Students answer the three exit ticket questions on a half sheet (about 3 minutes). The rest of Worksheet 2 and the family sheet go home.



Discussion questions

1. **A video of the principal 'canceling school' gets shared thousands of times, but there is nothing on the district website. Who is harmed if people believe it, even though nobody asked for money?**

Listen for: Families who miss work or keep students home, the principal whose identity was used, the school's credibility. Students connect sharing a fake to real harm and name the fix: check the district's own channels before sharing.

2. **Why does a polished, perfectly written message no longer prove it is legitimate?**

Listen for: AI tools write clean, personalized messages and translate well, so spelling and grammar are weak clues. The stronger clues are the request (money, codes, personal information), the pressure and whether it checks out by a known route.

3. **Your 'brother' calls crying from an unknown number, asks for money now and says not to tell your parents. What makes this a red flag even if the voice sounds exactly right?**

Listen for: Urgency, secrecy, an unknown number and a money request. Voices can be cloned from short clips. Hang up and call your brother or a parent on a number you already have; use a family safe word.

4. **Someone makes an AI 'joke' video of a classmate saying something embarrassing. Is it harmless if it is labeled as fake?**

Listen for: Labels get cropped out, clips spread beyond the intended audience, and the classmate's reputation and well-being are affected. It can break school conduct rules on harassment or impersonation. The responsible move is not to make it, not to share it and to report it.

5. **How is lateral reading different from reading the 'About' page of a website?**

Listen for: The About page is the source describing itself. Lateral reading checks what independent, trusted sources say about it, which is how professional fact-checkers work.



Exit ticket

Answer on a half sheet. No names needed.

1. What does 'verify by a known route' mean? Give one example. → **Check by contacting the person or organization through a number, app or website you already had, not the one in the message. Example: hanging up and calling a parent's saved number, or typing the district website address yourself.**
2. Name two red flags of a fake job or scholarship offer. → **Any two of: you must pay a fee, it is guaranteed, they want bank or Social Security details before any interview, they send a check and ask you to send money back, they contacted you out of nowhere by text, the deadline is extreme.**
3. If someone threatens to share an image of you unless you pay or send more, what are your first moves? → **Do not pay and do not send more, stop replying, save the messages, block and tell a trusted adult or the counselor. It is not your fault. Take It Down and the NCMEC CyberTipline can help.**



Support

- Start with four clear items (scholarship fee, brother voice call, coach text, recruiter text) before the harder Verify first items.
- Provide a three-question card: What do they want me to do? Is there pressure? How can I check by a route I already trust?
- Give sentence frames for the why: 'This is a red flag because ____.' 'I would verify by ____.'



Extend

- Students research one public example of a debunked viral image or clip from a reputable fact-checking outlet and explain which verification step exposed it.
- Students write a one-page proposal for how the school should respond when a deepfake of a student or staff member spreads, balancing support for the person harmed, investigation and student privacy.
- Debate: Should apps be required to label AI-generated media? Students consider what labels can and cannot solve.



Multilingual learners

- Pre-teach deepfake, verify, scholarship, fee, guaranteed and safe word with examples and icons.
- Point out that scams also arrive in Spanish and other languages, and AI translation makes them look fluent. Invite students to share how they would verify in their home language.
- Allow group discussion in home languages with an English report-out. Share the Spanish family sheet with Spanish-speaking families.



Accessibility

- All messages and claims are written out in words, so they work with screen readers, large print and audio. No item depends on viewing an actual image or video.
- Allow students to complete the verification routine orally or with speech-to-text.
- Offer a quiet option for students who find the sextortion topic difficult, and let students know they can talk privately with the counselor.



Family extension

Real or fake? A family plan for checking voices, videos and offers.

Today your student learned that AI can copy voices, create realistic videos and images and write very convincing scam messages, including fake job and scholarship offers aimed at teens. The best defense is a routine: pause, check what trusted sources say and verify by a phone number, app or website you already know. Please take 10 minutes to set up your family's plan.

Send home the family sheet (English and Spanish) included in this lesson pack.



Follow-up: free 2-minute lesson

k12cyberforward.com/learn/ai-tools-check

Project it for a quick review or share the link with families.
Related Cyber Academy course: academy-students-9-12.

Sources and further reading.

- FBI IC3 – Criminals Use Generative Artificial Intelligence to Facilitate Financial Fraud (PSA, December 2024): <https://www.ic3.gov/PSA/2024/PSA241203>
- FBI IC3 – Child Sexual Abuse Material Created by Generative AI and Similar Online Tools is Illegal (PSA, March 2024): <https://www.ic3.gov/PSA/2024/PSA240329>
- FTC – Scammers use AI to enhance their family emergency schemes: <https://consumer.ftc.gov/consumer-alerts/2023/03/scammers-use-ai-enhance-their-family-emergency-schemes>
- FTC – That job offer text is probably a scam: <https://consumer.ftc.gov/consumer-alerts/2026/04/job-offer-text-probably-scam>
- FTC – How To Avoid Scholarship and Financial Aid Scams: <https://consumer.ftc.gov/articles/how-avoid-scholarship-and-financial-aid-scams>
- FTC – Avoiding and Reporting Gift Card Scams: <https://consumer.ftc.gov/articles/avoiding-and-reporting-gift-card-scams>
- CISA Secure Our World – Recognize and Report Phishing (AI can produce polished messages): <https://www.cisa.gov/secure-our-world/recognize-and-report-phishing>
- NCMEC – Take It Down: <https://takeitdown.ncmec.org/>
- NCMEC – CyberTipline: <https://report.cybertip.org/>
- NCMEC NetSmartz – Sextortion: <https://www.missingkids.org/netsmartz/topics/sextortion>
- Digital Inquiry Group – A Civic Online Reasoning Module for High School (lateral reading): <https://cor.inquirygroup.org/blog/module-for-high-school>
- ISTE Standards for Students: <https://iste.org/standards/iste-standards-for-students>
- CSTA – 2026 PK-12 Computer Science Standards: <https://csteachers.org/pk12standards/>