

GRADES 6–8 · AGES 11–14 · 55 MINUTES · STRUCTURED DEBATE

Deepfakes: should they be illegal?

THE PROPOSITION

"Creating a deepfake of a real person without their consent should be against the law."

A structured debate on synthetic media that starts from consent and privacy and runs into free expression and enforcement. Students learn what a deepfake is technically, why detection is losing, and why 'just ban it' is harder than it sounds.

STUDENTS WILL BE ABLE TO

- Explain what a deepfake is and why they are getting harder to detect.
- Argue an assigned side on a proposed law with evidence and a rebuttal.
- Identify who a law would protect, who it would restrict, and how it would be enforced.

VOCABULARY FROM THE FIELD GUIDE

Deepfake · Consent (data use) · Privacy (when using AI) · AI regulation · Verification

WHAT'S IN THIS PACK

- Teacher guide: background, vocabulary, agenda, format, positions, questions, roles, rubric.
- Student handouts (photocopy one per student): position organizer, sentence starters, exit ticket.
- Standards: CCSS ELA SL.7.1 / SL.8.1; CCSS ELA W.7.1 / W.8.1 (argument); C3 Framework D2.Civ.12 (laws and rules); ISTE Students 2.a / 2.b (digital citizenship).

PRINT

US Letter, plain paper, duplex. Teacher pages once; student pages one per student. Free to photocopy for classroom and homeschool use. Not for resale.

Background for the teacher

Read this before class. Students get a version of the same facts on their handout; you get the framing and the traps.

A deepfake is synthetic audio, image, or video that shows a real person doing or saying something they did not. The models that make them are freely available and improving fast; detection tools lag behind and give false positives. The most common victims are not politicians but ordinary people – overwhelmingly women and girls – targeted with fake intimate images, including in middle and high schools.

Several U.S. states and other countries have passed laws targeting non-consensual intimate deepfakes and election deepfakes. The hard cases are satire, parody, art, and journalism – a law that bans all synthetic media of real people also bans a comedian's impression or a historical reconstruction. Students should feel that tension.

Enforcement is the second hard problem. The person who made it may be anonymous, in another country, or thirteen. Platforms can be required to remove content; that works better than prosecution but raises its own questions about who decides what's fake.

RUNNING DISCUSSION IN GRADES 6–8

Evidence enters the room. Students can hold a position they don't believe, and that's the point.

- Assign sides. Arguing a position you didn't choose is the fastest way to learn what the other side actually thinks.
- Require one piece of evidence per claim – a fact from the background sheet, a personal example, or a definition.
- Separate 'what is true' from 'what should we do.' Most middle-school debates tangle the two.
- Watch for the pile-on. If one side is winning on volume, pause and give the other side a 60-second huddle.

Vocabulary

Pulled from the Hi, Bot AI concepts field guide (hibot.space/learn/concepts). Read each card aloud, then have the class rewrite it in ten words or fewer.

Deepfake

Synthetic media that convincingly puts someone's face or voice where it wasn't.

Deepfakes use generative models to mimic a person's appearance or speech. They raise fraud, election, and harassment risks, especially when detection lags creation. Disclosure norms ("this is AI-generated") and provenance tools are part of the response; skepticism toward unverified media is the user-side defense.

Consent (data use)

Whether people agreed to their data being used to train or fine-tune systems.

Consent questions arise when personal writing, medical records, student work, or biometric data might enter a training or fine-tuning pipeline. Schools and companies often ban pasting identifiable student data into public models for this reason. "Public on the web" is not the same as "consented for AI training" in many ethical frameworks.

Privacy (when using AI)

What you paste into a model may be stored, logged, or used to improve products.

Privacy in AI use means knowing whether your prompts are retained, reviewed by humans, used for training, or shared with subprocessors. Enterprise tiers often offer "no training on customer data"; consumer free tiers may not. Treat prompts like email to a third party: don't send secrets you wouldn't want leaked.

AI regulation

Government rules on development, deployment, and accountability.

AI regulation spans safety testing, transparency, sector rules (health, finance, education), export controls on chips, and liability when systems cause harm. Examples include the EU AI Act, U.S. executive orders, and sector guidance from schools or hospitals. Regulation lags technology; compliance is often "best effort against moving targets."

Verification

Checking outputs against evidence – the habit that makes AI usable.

Verification means treating model output as a draft: cross-check claims, run code, read sources, ask what would falsify the answer. It's the skill that separates tool use from blind trust. Kids and adults need the same loop – models don't get a pass because they sound confident.

Warm-up

6 minutes.

Show (or describe) a harmless, obvious deepfake – a historical figure 'reading' a modern text. Ask: 'Should this be illegal?' Then describe the school case, without details: a student's face on a fake image, shared in a group chat. Ask again. Note that the technology is identical.

Agenda - 55 minutes

Minutes	Segment	What happens
0-6	Warm-up	Harmless deepfake vs. harmful one. Same tech, different answers?
6-14	Background + vocabulary	Read the sheet. Mark one fact for evidence. Class defines 'consent' in one sentence.
14-22	Team prep	Assigned sides. Position organizer: claim, three reasons, evidence, expected rebuttal. Both sides must address enforcement.
22-44	Structured debate	Opening 2 · huddle 3 · rebuttal 2 · cross-fire 6 · closing 1. Audience scores.
44-52	Draft the law	Mixed groups write ONE sentence of law: what is banned, what is exempt, who enforces. Read three.
52-55	Exit ticket	Handout.

How structured debate runs

Two teams, assigned sides. Opening statements (2 min each), rebuttal huddle (3 min), rebuttals (2 min each), open cross-fire with hands (5 min), closing statements (1 min each). The audience votes on which side argued better – not which side they agree with.

GROUND RULES – POST THESE

1. Attack ideas, never people. "That claim has a hole" – not "you're wrong."
2. One voice at a time. The listener's job is to be able to repeat what was said.
3. Evidence beats volume. A quiet fact outranks a loud opinion.
4. Changing your mind is winning. Say so out loud when it happens.
5. No naming real classmates or real local incidents. Hypotheticals only.
6. You argue the side you were assigned.

The two positions

Starter arguments are scaffolds, not scripts. Strong teams find their own. Assign sides where the format calls for it; students record their real view on the exit ticket.

Affirmative – ban it without consent

Your face and voice are yours. Using them to fake you is a harm the law should name.

- The victims are mostly ordinary people with no other recourse.
- Consent is the clean line: the person said yes, or they didn't.
- Laws against forgery and impersonation already exist; this is the same wrong with new tools.
- Without a law, platforms have no obligation to take it down.

Negative – a blanket ban goes too far

The harm is real, but the ban as written catches satire, art, and journalism, and can't be enforced.

- Parody and impressions of public figures are protected expression.
- Detection doesn't work; a law you can't enforce teaches contempt for the law.
- Target the harm (harassment, fraud, fake intimate images), not the technology.
- Makers are anonymous and global. Platform takedown rules would do more than criminal law.

Discussion questions

OPENING

1. What exactly is taken from someone when their face is faked?

CORE

1. Should the law treat a fake of a senator differently from a fake of a classmate?
2. If you can't reliably tell a deepfake from a real video, how would a court?
3. Is a comedian's impression a deepfake? Where's the line?

CLOSING

1. Write one sentence of law. Who does it protect, and who does it restrict?

Roles

Role	Job
First speaker	Opening: claim + strongest two reasons with evidence.

Role	Job
Rebuttal speaker	Restates the other side's best point, then answers it.
Cross-fire lead	Asks about enforcement and edge cases.
Closer	One reason to vote, stated plainly.
Audience judge	Scores on rubric. Votes on argument, not agreement.

Assessment rubric

Share this before the debate. Judges and self-assessors use the same sheet. 'Meets' is the target; note above or below in the margin.

Criterion	Meets
Claim	States a clear position that answers the question.
Evidence	Supports each claim with a fact, definition, or specific example.
Response	Restates the other side accurately before rebutting.
Vocabulary	Uses at least two lesson terms correctly.
Conduct	Follows ground rules; invites quieter voices in.

Exit ticket prompts

1. A deepfake is ____.
2. My assigned side: _____. My actual position: _____.
3. My one sentence of law: _____.

Extension

Find your state's law on deepfakes (or the nearest one). Summarize what it bans, what it exempts, and the penalty. Would it cover the school case?

SEND HOME

Ask at home: 'If you saw a video of a family member saying something awful, what would you do first – before reacting?'

STANDARDS ALIGNMENT

- CCSS ELA SL.7.1 / SL.8.1
- CCSS ELA W.7.1 / W.8.1 (argument)
- C3 Framework D2.Civ.12 (laws and rules)
- ISTE Students 2.a / 2.b (digital citizenship)

More lessons by grade band, and the full AI concepts field guide, at hibot.space/learn/debates. Tell us how it went: hello@hibot.space.

Deepfakes: should they be illegal?

"Creating a deepfake of a real person without their consent should be against the law."

WHAT YOU NEED TO KNOW

A deepfake is synthetic audio, image, or video that shows a real person doing or saying something they did not. The models that make them are freely available and improving fast; detection tools lag behind and give false positives. The most common victims are not politicians but ordinary people – overwhelmingly women and girls – targeted with fake intimate images, including in middle and high schools.

Several U.S. states and other countries have passed laws targeting non-consensual intimate deepfakes and election deepfakes. The hard cases are satire, parody, art, and journalism – a law that bans all synthetic media of real people also bans a comedian's impression or a historical reconstruction. Students should feel that tension.

Enforcement is the second hard problem. The person who made it may be anonymous, in another country, or thirteen. Platforms can be required to remove content; that works better than prosecution but raises its own questions about who decides what's fake.

WORDS WE'LL USE

Deepfake – Synthetic media that convincingly puts someone's face or voice where it wasn't.

Consent (data use) – Whether people agreed to their data being used to train or fine-tune systems.

Privacy (when using AI) – What you paste into a model may be stored, logged, or used to improve products.

AI regulation – Government rules on development, deployment, and accountability.

Verification – Checking outputs against evidence – the habit that makes AI usable.

MARK IT UP

Underline one sentence you agree with. Circle one you doubt. Write one question in the margin. Bring all three to the discussion.

Position organizer

Proposition: "Creating a deepfake of a real person without their consent should be against the law."

Name: _____ Side (assigned or chosen): _____

OUR CLAIM, IN ONE SENTENCE

THREE REASONS – AND THE EVIDENCE FOR EACH

Reason 1:

Evidence (a fact, a definition, or a specific example):

Reason 2:

Evidence (a fact, a definition, or a specific example):

Reason 3:

Evidence (a fact, a definition, or a specific example):

WHAT THE OTHER SIDE WILL SAY – AND OUR REPLY

Their best argument:

Our reply:

VOCABULARY I WILL USE CORRECTLY

☐ Deepfake ☐ Consent (data use) ☐ Privacy (when using AI) ☐ AI regulation ☐ Verification

Sentence starters

Use these until you don't need them. The goal is to respond to what was actually said.

- Our claim is ____, and the evidence is ____.
- The other side is right that ____, but that doesn't cover ____.
- How would your law handle the case of ____?
- The difference between those two cases is ____.
- Even if we can't enforce it perfectly, ____.

GROUND RULES

1. Attack ideas, never people. "That claim has a hole" – not "you're wrong."
2. One voice at a time. The listener's job is to be able to repeat what was said.
3. Evidence beats volume. A quiet fact outranks a loud opinion.
4. Changing your mind is winning. Say so out loud when it happens.
5. No naming real classmates or real local incidents. Hypotheticals only.
6. You argue the side you were assigned.

HOW STRUCTURED DEBATE WORKS

Two teams, assigned sides. Opening statements (2 min each), rebuttal huddle (3 min), rebuttals (2 min each), open cross-fire with hands (5 min), closing statements (1 min each). The audience votes on which side argued better – not which side they agree with.

YOUR ROLE TODAY

Role: _____

Role	Job
First speaker	Opening: claim + strongest two reasons with evidence.
Rebuttal speaker	Restates the other side's best point, then answers it.
Cross-fire lead	Asks about enforcement and edge cases.
Closer	One reason to vote, stated plainly.
Audience judge	Scores on rubric. Votes on argument, not agreement.

LISTENING TALLY

Keep count while others speak. Claims: ____ Evidence: ____ Questions: ____ Vocabulary used: ____
Minds changed: ____

Exit ticket

"Creating a deepfake of a real person without their consent should be against the law."

Name: _____ Date: _____

A deepfake is ____.

My assigned side: ____. **My actual position: ____.**

My one sentence of law: ____.

SELF-ASSESSMENT

Criterion	Meets looks like	Me today (above / meets / below)
Claim	States a clear position that answers the question.	
Evidence	Supports each claim with a fact, definition, or specific example.	
Response	Restates the other side accurately before rebutting.	
Vocabulary	Uses at least two lesson terms correctly.	
Conduct	Follows ground rules; invites quieter voices in.	

TAKE IT HOME

Ask at home: 'If you saw a video of a family member saying something awful, what would you do first – before reacting?'