

GRADES 9–12 · AGES 14–18 · 60 MINUTES · PHILOSOPHICAL CHAIRS

Should you be watched to be kept safe?

THE PROPOSITION

"Schools should be allowed to use AI monitoring on students' devices and cameras to prevent harm."

Philosophical chairs on AI surveillance in schools, where students are the surveilled. Real trade-offs – the monitoring has flagged real crises and also false positives that outed students or punished creative writing – with privacy, bias, and consent as the vocabulary and seat-switching as the visible evidence of thinking.

STUDENTS WILL BE ABLE TO

- Describe what school monitoring software does, what it flags, and who sees the flags.
- Argue a position and switch when persuaded, stating what persuaded you.
- Propose limits (scope, retention, review, appeal) and evaluate whether they preserve the safety benefit.

VOCABULARY FROM THE FIELD GUIDE

Surveillance AI · Privacy (when using AI) · Bias (in data and models) · Consent (data use) · Explainability

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.9–10.1 / SL.11–12.1; CCSS ELA SL.11–12.3 (evaluate stance, premises, and links among ideas); C3 Framework D2.Civ.10 (civil rights and responsibilities); CSTA 3A-IC-29 / 3A-IC-30 (privacy and security trade-offs).

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.

Many districts run AI monitoring on school-issued devices and accounts, scanning documents, searches, and messages for signs of self-harm, violence, or bullying, and some use camera analytics. Vendors and some administrators credit the tools with intervening in real crises. Investigations have also found false positives that flagged literature assignments, outed LGBTQ students to families, and disproportionately flagged some students, along with data retained long after any risk.

The privacy question at this age is not abstract: students are the people being scanned, often without meaningful consent (a device agreement signed by a parent in August). Surveillance also changes behavior – students report not searching for help on sensitive topics because the search will be flagged. That chilling effect is a cost that never shows up in the vendor's numbers.

The most useful move is toward limits rather than a binary: what is scanned, who reviews a flag before anyone acts, how long data is kept, whether students are told, and whether there is an appeal. Both sides should be forced to say which limits they accept and what safety they give up.

RUNNING DISCUSSION IN GRADES 9–12

Stakes are real: college, jobs, creative work, civil rights. Students can steelman, cross-examine, and write policy.

- Insist on the steelman. Before rebutting, a speaker restates the other side's best version until they say 'yes, that's it.'
- Cross-examination is a question period, not a speech. Two minutes, questions only, no editorializing.
- Push past 'it depends.' Ask: depends on what, exactly – and who decides?
- Close with a written position. Speaking is the practice; the paragraph is the assessment.

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.

Surveillance AI

Automated watching – cameras, scoring, monitoring at scale.

Surveillance AI includes facial recognition, behavioral scoring in workplaces or schools, and mass metadata analysis. It raises civil-liberty questions separate from chatbot hype: power asymmetry, false positives, and chilling effects on speech. Not every camera has AI; not every AI deployment is surveillance – but the overlap is politically charged.

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.

Bias (in data and models)

Systematic skew – who's represented in training data and who gets hurt in outputs.

Bias in AI usually means unequal error rates or stereotypes baked in from historical data: hiring tools favoring certain résumés, image models defaulting to narrow demographics, language models mirroring toxic patterns from the web. Fixing bias requires dataset audits, evaluation across groups, and human review – not a single fairness slider.

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.

Explainability

Can we understand why a model produced a specific output?

Explainability (and interpretability) asks for reasons behind predictions – critical in lending, medicine, and discipline decisions. Deep neural networks are often black boxes; post-hoc explanations can be approximate or misleading. Regulated domains may require human-readable justification even when the model can't truly "explain" itself.

Warm-up

6 minutes.

Ask students to list, privately, three things they've typed on a school device this week they'd rather an administrator not read. No sharing. Then ask: 'How many of those would a threat-detection model flag?' Hands only.

Agenda - 60 minutes

Minutes	Segment	What happens
0-6	Warm-up	Private list. How many would be flagged?
6-16	Background + cases	Read the sheet: the real intervention, the flagged essay, the outing. Vocabulary cards.
16-42	Philosophical chairs	Agree / Disagree lines, Undecided row. Restate before responding. Switch freely; each switch is announced with what persuaded you. Teacher tallies switches.
42-52	Limits exercise	Mixed groups draft five limits (scope, review, retention, notice, appeal). Then state honestly what safety each limit trades away.
52-60	Written position	Exit ticket paragraph.

How philosophical chairs runs

Chairs face each other in two lines – Agree and Disagree – with a short Undecided row at the end. A speaker from one side speaks; the other side must summarize what was said before responding. Anyone may switch seats at any time, and switching is praised, never mocked.

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 personal disclosures required. 'A student I know of' is always fine.
6. Switching seats is applauded, never mocked.

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.

Agree – monitor, with limits

A student in crisis on a school device is the school's responsibility; a tool that catches even a few is worth real privacy costs.

- The interventions are real. One prevented suicide outweighs a hundred awkward flags.
- Students are minors on school property using school equipment.
- The fix for false positives is human review, not turning it off.
- Limits – scope, review, retention – can keep most of the benefit and cut most of the harm.

Disagree – the costs are hidden and fall on the wrong students

Constant scanning chills the exact searches a struggling student needs to make, flags the wrong kids, and normalizes being watched.

- Students stop searching for help when help is flagged. The tool prevents the cry for help it claims to catch.
- False positives have outed students and punished creative writing. Those harms are concentrated on the vulnerable.
- Consent signed by a parent in August is not consent from the student in March.
- A counselor who knows students catches more than a keyword model, and doesn't retain the data.

Discussion questions

OPENING

1. What does it change about a room when you know you're being watched?

CORE

1. One real intervention versus a hundred false flags – how do you weigh them? Does it matter who the hundred are?
2. Can a student consent to monitoring? Can a parent consent for them?
3. If the model can't explain why it flagged something, should anyone act on the flag?

CLOSING

1. Name the limits you'd accept, and the safety you'd give up to get them.

Roles

Role	Job
Restater	Every speaker restates the previous point before responding; a designated student checks.
Switch recorder	Logs each seat switch and the stated reason; reads the log at the end.
Undecided interrogator	Asks one hard question of each side before choosing a seat.
Limits scribe	Writes the group's five limits and the honest trade-off for each.

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
Argument	Claim, warrant, and evidence are explicit and connected.
Steelman	Presents the strongest version of the opposing view before attacking it.
Clash	Rebuttals respond to what was actually said, not a strawman.
Precision	Uses lesson vocabulary accurately and distinguishes is from ought.
Written position	Closing paragraph takes a defensible position and concedes one real cost.

Exit ticket prompts

1. My position: ____.
2. What moved me most today: ____.
3. Limits I'd accept: _____. Safety I'd give up: _____.

Extension

Find your district's device or acceptable-use policy. Identify what it says is monitored, who reviews flags, and how long data is kept. Write three questions you'd ask the school board.

SEND HOME

Ask at home: 'Were you ever watched at school or work in a way that changed how you acted? Did it make you safer?'

STANDARDS ALIGNMENT

- CCSS ELA SL.9–10.1 / SL.11–12.1
- CCSS ELA SL.11–12.3 (evaluate stance, premises, and links among ideas)

- C3 Framework D2.Civ.10 (civil rights and responsibilities)
- CSTA 3A-IC-29 / 3A-IC-30 (privacy and security trade-offs)

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.

Should you be watched to be kept safe?

"Schools should be allowed to use AI monitoring on students' devices and cameras to prevent harm."

WHAT YOU NEED TO KNOW

Many districts run AI monitoring on school-issued devices and accounts, scanning documents, searches, and messages for signs of self-harm, violence, or bullying, and some use camera analytics. Vendors and some administrators credit the tools with intervening in real crises. Investigations have also found false positives that flagged literature assignments, outed LGBTQ students to families, and disproportionately flagged some students, along with data retained long after any risk.

The privacy question at this age is not abstract: students are the people being scanned, often without meaningful consent (a device agreement signed by a parent in August). Surveillance also changes behavior – students report not searching for help on sensitive topics because the search will be flagged. That chilling effect is a cost that never shows up in the vendor's numbers.

The most useful move is toward limits rather than a binary: what is scanned, who reviews a flag before anyone acts, how long data is kept, whether students are told, and whether there is an appeal. Both sides should be forced to say which limits they accept and what safety they give up.

WORDS WE'LL USE

Surveillance AI – Automated watching – cameras, scoring, monitoring at scale.

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

Bias (in data and models) – Systematic skew – who's represented in training data and who gets hurt in outputs.

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

Explainability – Can we understand why a model produced a specific output?

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: "Schools should be allowed to use AI monitoring on students' devices and cameras to prevent harm."

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

☐ Surveillance AI ☐ Privacy (when using AI) ☐ Bias (in data and models) ☐ Consent (data use) ☐

Explainability

Sentence starters

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

- I heard ___ say ___. My response is ___.
- I'm switching because ___ persuaded me that ___.
- That benefit is real. The cost you're not counting is ___.
- With the limit of ___, would you still object?
- Who bears that cost, specifically?

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 personal disclosures required. 'A student I know of' is always fine.
6. Switching seats is applauded, never mocked.

HOW PHILOSOPHICAL CHAIRS WORKS

Chairs face each other in two lines – Agree and Disagree – with a short Undecided row at the end. A speaker from one side speaks; the other side must summarize what was said before responding. Anyone may switch seats at any time, and switching is praised, never mocked.

YOUR ROLE TODAY

Role: _____

Role	Job
Restater	Every speaker restates the previous point before responding; a designated student checks.
Switch recorder	Logs each seat switch and the stated reason; reads the log at the end.
Undecided interrogator	Asks one hard question of each side before choosing a seat.
Limits scribe	Writes the group's five limits and the honest trade-off for each.

LISTENING TALLY

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

Exit ticket

"Schools should be allowed to use AI monitoring on students' devices and cameras to prevent harm."

Name: _____ Date: _____

My position: ____.

What moved me most today: ____.

Limits I'd accept: ____. **Safety I'd give up: ____.**

SELF-ASSESSMENT

Criterion	Meets looks like	Me today (above / meets / below)
Argument	Claim, warrant, and evidence are explicit and connected.	
Steelman	Presents the strongest version of the opposing view before attacking it.	
Clash	Rebuttals respond to what was actually said, not a strawman.	
Precision	Uses lesson vocabulary accurately and distinguishes is from ought.	
Written position	Closing paragraph takes a defensible position and concedes one real cost.	

TAKE IT HOME

Ask at home: "Were you ever watched at school or work in a way that changed how you acted? Did it make you safer?"