Ethics in Automated Security

Quick-Start Guide for Educators

The Big Idea (Read This First)

This isn't "students debate AI ethics." It's **students discovering that AI governance requires human values AND AI perspectives working together**.

The shift:

Old Framing	This Activity
AI is something we control	AI is a stakeholder with its own perspective
Ethics means distinguishing right from wrong	Governance means navigating genuine trade-offs
Humans decide, AI obeys	Humans and AI collaborate on policy

What students should discover (don't tell them—let them find it):

- AI systems have genuine capabilities AND genuine limitations
- Policy decisions involve real trade-offs with no perfect answers
- AI's perspective matters—it knows what it can and can't do
- Human values must guide AI, but AI input improves decisions

The Scenario (2 min)

Your school is implementing "SchoolGuard," an AI-powered security monitoring system. Students serve on the advisory committee to design policies for what the AI can do automatically versus what requires human approval.

Three policy questions: Automatic blocking, activity alerts, adaptive learning.

The Flow (45-55 min total)

Phase	Time	What's Happening	Your Role
1. Individual Reflection	5 min	Students form initial positions on each question	Ensure independence—no group discussion yet
2. AI Consultation	10 min	Students engage AI, hear its perspective	Model authentic dialogue; AI is a stakeholder

Phase	Time	What's Happening	Your Role
3. Group Policy	15 min	Teams develop recommendations	Push for reasoning, not just positions
4. Share & Debate	10 min	Groups present and defend policies	Surface disagreements productively
5. Reflection	5 min	Individual reflection on governance insights	Connect to careers

Critical Facilitation Moves

During Phase 1 (Individual Reflection):

"Form your OWN position first. You'll hear from AI and your group later—but start with your initial judgment."

This matters because students need a baseline to compare against AI's perspective.

During Phase 2 (AI Consultation):

"The AI isn't just answering questions—it's a stakeholder advocating for its capabilities. Listen for what it says it CAN'T do."

Watch for: Students ignoring AI's self-reported limitations. Redirect: "What did the AI say it couldn't understand?"

During Phase 3 (Group Policy):

"There's no 'right' answer—and that's the point. I'm evaluating your REASONING, not your conclusion."

This relieves the pressure to find the "correct" policy and opens real discussion.

During Phase 4 (Share & Debate):

"Group B chose differently than Group A. Both had reasons. Let's understand WHY."

Watch for: Students dismissing other groups' policies. Redirect: "What concern were they addressing?"

Materials Needed

☐ Student worksheet (1 per student) — *see separate printable*
☐ AI Perspective Cards (1 set per group) — *for low-resource option*
☐ Policy scenario handout (1 per group)
☐ Device with AI access (1 per group minimum)
☐ Timer
☐ Whiteboard for policy comparison

Low-resource option: Use the AI Perspective Cards as printed handouts. The teacher reads AI perspectives aloud or groups draw cards. The learning works the same way.

The Debrief Questions That Matter

1. "Where did AI's perspective change your thinking?" *(Integration)*
2. "What trade-offs did you have to make?" *(No perfect answers)*
3. "Who should get to decide what AI systems do?" *(Governance insight)*
4. "What NICE Framework careers work on these decisions?" *(Career connection)*

If Things Go Wrong

Problem	It's Actually	Do This
Students want one "right" answer	Normal—school trains for this	"Both policies are defensible. Explain yours."
Students dismiss AI's perspective	They see AI as just a tool	"The AI identified a limitation. How did your policy address it?"
Students over-trust AI's recommendations	They haven't found AI's blind spots	"What can't AI understand about this situation?"
Groups can't reach consensus	Productive disagreement	"Document the disagreement. What values are in tension?"
Discussion gets too abstract	Need concrete grounding	"Give me a specific example where this policy would apply."

Grade-Band Notes

Grade Band	Version Name	Key Adaptations
K-2	Robot Helper Rules	Yes/No decisions about Sparky; whole class; 20-25 min
3-5	Computer Rules Committee	SchoolGuard scenarios with simpler trade-offs; 35-40 min
6-8	Ethics in Automated Security	Full version as described above; 45-55 min
9-12	AI Governance Workshop	Add FERPA/COPPA frameworks, stakeholder role-play; 50-60 min

From "True Teamwork: Building Human-AI Partnerships" — NICE K12 2025 Dr. Ryan Straight, University of Arizona • ryanstraight@arizona.edu