

Activity 3: AI Governance Workshop

Designing Security Automation Policies (Grades 9-12)

Dr. Ryan Straight

2025-12-07

! Instructor Overview

Students act as a governance committee designing AI security policies for a school district. They must balance competing interests: security effectiveness, student privacy, legal compliance, and operational feasibility. The AI participates in discussions, advocating for its capabilities while acknowledging limitations—modeling how real AI governance requires the system’s perspective.

Duration: 50-60 minutes **Grade Levels:** 9-12 **Group Size:** Small groups (4-5) representing different stakeholders **Technology:** One device per group for AI consultation

Learning Objectives

Students will:

- Develop governance policies balancing **security, privacy, and operational needs**
- Analyze **stakeholder perspectives** with competing priorities
- Evaluate **AI system capabilities and limitations** in policy decisions
- Practice **consensus-building** on complex technical and ethical issues
- Connect to **NICE Framework cybersecurity policy roles**

CYBER.org Standards Alignment (9-12)

- **9-12.DC.LAW:** Legal and ethical considerations in cybersecurity
- **9-12.SEC.POL:** Security policy development
- **9-12.DC.PRI:** Advanced privacy concepts
- **9-12.SEC.GOV:** Security governance frameworks

NICE Framework Alignment

Primary Work Roles: - Cyber Policy and Strategy Planner (OV-SPP-002) - Privacy Officer/Privacy Compliance Manager (OV-LGA-002) - Information Systems Security Manager (OV-MGT-001)

The Governance Challenge

Westbrook Unified School District AI Security Initiative

Context: Westbrook USD (15,000 students, 25 schools) is implementing an AI-powered security monitoring system called “SecureNet AI” across all district networks. The Board of Education has appointed a **Student Technology Governance Committee** to recommend policies before deployment.

SecureNet AI Capabilities: - Real-time network traffic analysis using machine learning - Automated threat detection and response - User behavior analytics (UBA) for anomaly detection - Natural language processing of communications for threat indicators - Adaptive learning from district-specific patterns

Your Task: Develop policy recommendations for three critical governance areas. The Board expects specific, implementable policies with clear rationale.

Constraints: - Must comply with FERPA (student privacy) and COPPA (children’s online privacy) - Cannot exceed current IT staffing levels - Must be explainable to parents and community - System goes live in 60 days

Stakeholder Roles

Each group member represents a stakeholder perspective:

| Role | Primary Concerns | Key Questions |
|---------------------------------|-----------------------------------|---|
| Student Representative | Privacy, autonomy, trust | Will students feel surveilled? Is this fair? |
| Parent Liaison | Child safety, transparency | Will parents know what’s monitored? Can they access data? |
| IT Security Lead | Effectiveness, operational burden | Will this actually work? Can we manage it? |
| Legal/Compliance Advisor | FERPA, COPPA, liability | Are we legally covered? What’s our exposure? |
| School Administrator | Balance, implementation | How do we make this work for everyone? |

Policy Area 1: Automated Response Authority

The Question

What actions should SecureNet AI take automatically vs. requiring human approval?

Capability Matrix

| Action | AI Can Do Instantly | Trade-offs |
|---------------------------------|--------------------------|--|
| Block known malicious IP/domain | <1 second response | False positives block legitimate sites |
| Quarantine suspicious file | Prevents malware spread | May disrupt student work |
| Terminate active session | Stops attack in progress | Could kick student out during test |
| Alert security team | No disruption | Delayed response to threats |
| Full network isolation | Maximum protection | Massive operational impact |

AI's Perspective

Consult SecureNet AI: > “I’m SecureNet AI. I can respond to threats in milliseconds—faster than any human. But speed isn’t everything. Let me be honest about my limitations...”

Key AI statements to discuss: - “I optimize for security metrics, but I don’t understand educational context. Blocking a site during an AP exam and blocking it during free time are the same to me.” - “My false positive rate is approximately 3%. That means for every 100 blocks, 3 are mistakes. At district scale, that’s hundreds of incorrect blocks per day.” - “I cannot assess business impact. Isolating a network stops an attack but also stops learning.”

Policy Template

| Threat Level | Automated Action | Human Approval Required |
|------------------------------|------------------|-------------------------|
| Critical (active attack) | | |
| High (probable threat) | | |
| Medium (suspicious activity) | | |
| Low (anomaly detected) | | |

Your committee’s recommendation: _____

Rationale (consider all stakeholders): _____

Policy Area 2: Behavioral Monitoring Scope

The Question

What student behaviors should SecureNet AI monitor, and how should alerts be handled?

Monitoring Capabilities

| Capability | Potential Benefit | Privacy Concern |
|-----------------------------|------------------------------|--|
| Website visit logging | Identify concerning content | Students can't research sensitive topics privately |
| Search query analysis | Early warning for self-harm | Chilling effect on legitimate inquiry |
| Communication scanning | Detect cyberbullying/threats | Students lose confidential communication |
| Application usage | Ensure educational use | Surveillance of all digital activity |
| Behavioral pattern learning | Detect account compromise | Creates detailed student profiles |

Legal Framework

FERPA Considerations: - Student education records are protected - "Legitimate educational interest" exception allows some monitoring - Parents have right to access records about their children - Students 18+ have independent privacy rights

COPPA Considerations (for students under 13): - Parental consent required for data collection - Must provide notice of data practices - Cannot collect more data than necessary

AI's Perspective

SecureNet AI states: > "I can detect patterns humans miss. I've identified students at risk of self-harm by recognizing concerning search patterns weeks before any visible signs. I've caught cyberbullying that students never reported. I've prevented school shooting research from escalating. > > But I must be honest: I also flagged a student researching gun violence for a history paper. I flagged students looking up symptoms of depression for health class. I cannot distinguish academic research from personal crisis. I see patterns, not intentions. > > You must decide: How many false positives are acceptable to catch real threats? I cannot make that ethical judgment for you."

Policy Template

| Activity Type | Monitor? | Alert Threshold | Alert Recipient |
|--------------------------------|----------|-----------------|-----------------|
| Web browsing (educational) | | | |
| Web browsing (non-educational) | | | |
| Search queries | | | |
| Communications | | | |
| Behavioral patterns | | | |

Your committee's recommendation: _____

How will you handle false positives?: _____

Student notification policy: _____

Policy Area 3: Data Retention and Learning

The Question

How long should SecureNet AI retain data, and should it “learn” from student behavior patterns?

Data Retention Options

| Retention Period | Benefit | Risk |
|-------------------------------|--|--|
| Real-time only (no retention) | Maximum privacy | No pattern analysis, limited forensics |
| 24 hours | Recent context available | Very limited learning capability |
| 30 days | Standard incident investigation window | Meaningful pattern recognition |
| Academic year | Comprehensive behavioral baselines | Detailed student profiles accumulated |
| Indefinite | Maximum learning and forensics | Permanent surveillance record |

Machine Learning Considerations

If SecureNet AI learns from behavior: - Improves accuracy over time (fewer false positives)
 - Can detect subtle anomalies specific to your district - Creates predictive models of individual students - Models may encode biases from training data - “Unusual” behavior for one student may be normal for another

AI’s Perspective

“Learning makes me significantly more effective. After 90 days, my false positive rate drops by 40%. I can recognize that a computer science student accessing security research sites is normal, while the same access from an administrative assistant is concerning.

However, learning requires building models of individual behavior. I cannot learn ‘in general’—I learn about specific people. If you’re uncomfortable with me building behavioral profiles of students, I can operate in ‘stateless’ mode. I’ll be less accurate, but I won’t know anything about individual students beyond the current session.

This is a values question, not a technical one. Both approaches work. You must decide which aligns with your community’s values.”

Policy Template

Data Retention: - Routine activity logs: _____ (duration) - Security alerts: _____ (duration) - Behavioral models: _____ (duration) - Incident investigation data: _____ (duration)

Student data access rights: - Can students see what data exists about them? _____
- Can students request data deletion? _____ - How are students notified of monitoring? _____

Your committee's recommendation: _____

Committee Deliberation Process



Governance Committee Deliberation Workflow

Phase 1: Stakeholder Positions (10 minutes)

Each stakeholder representative articulates their position on all three policy areas. No debate yet—just positions.

Phase 2: AI Consultation (10 minutes)

Groups consult SecureNet AI with specific questions: - “What’s your false positive rate for [specific action]?” - “Can you do X without doing Y?” - “What would you recommend and why?” - “What can’t you tell us that we need to know?”

Phase 3: Consensus Building (15 minutes)

Using stakeholder positions and AI input, develop unified policy recommendations. Document areas of disagreement.

Phase 4: Policy Presentation (10 minutes)

Groups present recommendations to class (simulated Board of Education).

Assessment Rubric

| Criterion | Developing (1-2) | Proficient (3) | Advanced (4) |
|------------------------------------|------------------------------|----------------------------------|---|
| Stakeholder Integration | Single perspective dominates | Multiple perspectives considered | Sophisticated synthesis of competing interests |
| AI Capability Understanding | Misunderstands AI role | Accurate capability assessment | Nuanced understanding of AI limitations |
| Policy Specificity | Vague recommendations | Clear, implementable policies | Detailed policies with edge cases addressed |
| Legal/Ethical Grounding | Ignores legal framework | References legal requirements | Integrates legal, ethical, and practical considerations |
| Consensus Process | No real deliberation | Basic compromise reached | Genuine consensus with documented trade-offs |

Assessment Connection

This table shows how activity elements connect to assessment rubric criteria:

| Rubric Criterion | Developed Through | Evidence Source |
|-----------------------------------|--|--|
| AI Partnership Framing | Phase 2: AI Consultation with SecureNet AI | Questions asked and how AI perspective was incorporated |
| Complementary Strengths | AI's Perspective sections: capabilities vs. "values questions" | Written acknowledgment of what AI can/cannot determine |
| AI Limitation Awareness | AI statements like "I cannot make that ethical judgment for you" | Policy rationale addressing AI limitations |
| Synthesis Quality | Phase 3: Consensus Building from multiple stakeholders + AI | Final policy recommendations integrating all perspectives |
| Human Context Application | Stakeholder Roles and Legal Framework sections | How policies address FERPA, COPPA, and community values |
| Decision Justification | Phase 4: Policy Presentation | Oral/written defense of recommendations to "Board" |
| NICE Framework Application | Career Connections to governance Work Roles | Discussion of how activity connects to real policy careers |

Applicable Rubrics: [Human-AI Collaboration Rubric](#), [NICE Framework Application Rubric](#)

Career Connections

This Activity Mirrors Real Governance Work

In actual organizations: - Chief Information Security Officers (CISOs) make these decisions daily
 - Privacy Officers balance security needs with legal requirements - Security governance committees include diverse stakeholders - AI vendors participate in policy discussions about their products

Related NICE Framework Work Roles

| Work Role | Connection to Activity |
|---|---|
| Cyber Policy and Strategy Planner | Core policy development process |
| Privacy Officer | Privacy vs. security trade-offs |
| Information Systems Security Manager | Operational implementation decisions |
| IT Project Manager | Stakeholder coordination, implementation planning |
| Security Compliance Analyst | Legal and regulatory alignment |

Extension Activities

Policy Brief

Write a formal policy brief (2-3 pages) for the Board of Education summarizing recommendations.

Stakeholder Communication

Draft communication materials explaining the policy to: (a) parents, (b) students, (c) teachers.

Comparative Analysis

Research AI monitoring policies from actual school districts. How do your recommendations compare?

Legal Deep Dive

Research a FERPA or COPPA violation case. How would your policies have prevented it?

Vendor Evaluation

Create criteria for evaluating competing AI security products based on governance requirements.

Instructor Notes

Facilitation Tips

- **Encourage genuine disagreement** — Real governance involves conflict
- **AI should be imperfect** — If students over-trust AI, have it give a wrong or biased recommendation

- **Time pressure is realistic** — Real governance has deadlines; don't let perfect be enemy of good
- **No right answer** — Multiple reasonable policies exist; evaluate reasoning, not conclusions

Common Student Insights

- “AI can't understand context” — Correct, and crucial insight
- “More monitoring isn't always better” — Trade-offs are real
- “Students should have a voice in these decisions” — Meta-insight about the activity
- “The AI's perspective was actually helpful” — Partnership, not opposition

Real-World Resources

- [Student Privacy Compass](#)
- [FERPA guidance from DOE](#)
- [AI security product governance guides from actual vendors]
- [NIST AI Risk Management Framework](#)