Trust Signal: Corporate Cybersecurity Awareness Game

Game Overview

Trust Signal is a team-based detection game where corporate employees learn to identify AI-generated deepfakes and social engineering attacks through realistic workplace scenarios. Teams analyze communications and make trust/verify decisions under time pressure, building essential cybersecurity awareness skills.

Core Game Mechanics

Simple Detection Challenge

- **Teams of 4-6** review workplace communications (voice messages, emails, video calls)
- **Binary decisions**: "Trust" (authentic) or "Signal" (suspicious/fake)
- **30-60 seconds per scenario** to maintain engagement
- **Immediate feedback** with explanation of detection clues
- **Progressive difficulty** increases 25% each session

Scoring System

- **+2 points** for correct detection
- **-1 point** for false positives (flagging authentic content)
- **-3 points** for missing actual threats
- **Bonus points** for identifying specific AI artifacts or social engineering tactics

Role Rotation (Each Session)

- **Employee**: Focus on personal threat recognition
- **Manager**: Evaluate team member requests and communications
- **IT Support**: Technical analysis perspective
- **Executive**: High-stakes decision making under pressure

Four-Session Progressive Structure

Session 1: Foundation (1 hour)

Learning Objective: Recognize basic deepfake indicators

Format:

- 5 min: Welcome + team formation
- 10 min: Threat landscape overview (recent real cases)
- 30 min: Game rounds (6 scenarios, increasing difficulty)
- 10 min: Team debrief + ethics discussion
- 5 min: Preview next session

Scenarios:

1. **Obvious voice clone** - CEO with wrong accent/speech patterns
2. **Suspicious email** - AI-generated but with subtle grammar errors
3. **Authentic urgent request** - Real executive communication for contrast
4. **Basic video deepfake** - Face mapping with obvious artifacts
5. **Voice + email combination** - Multi-channel but inconsistent details
6. **Pressure scenario** - Time-sensitive request with emotional manipulation

Ethics Component: Discussion of appropriate verification methods vs. insulting colleagues

Session 2: Escalation (1 hour)

Learning Objective: Handle multi-channel sophisticated attacks

Format: Same structure, new roles, harder scenarios

Scenarios:

1. **High-quality voice clone** - CEO's voice requesting wire transfer
2. **AI-enhanced BEC** - Perfect grammar, personalized details from social media
3. **Deepfake video call** - Short clip claiming technical difficulties
4. **Coordinated attack** - Voice + email + text message with consistent story
5. **Insider threat simulation** - Colleague acting suspiciously (social engineering)
6. **Time pressure scenario** - "Board meeting in 10 minutes" type urgency

Ethics Component: Balancing security with workplace relationships and trust

Session 3: Advanced (1 hour)

Learning Objective: Detect sophisticated AI-generated content and verify properly

Format: Same structure, new roles, expert-level scenarios

Scenarios:

1. **Perfect voice clone** - Requires verification through secondary channels
2. **AI-generated documents** - Fake contracts, invoices with subtle inconsistencies
3. **Deepfake video meeting** - Full conversation requiring behavioral analysis
4. **Advanced social engineering** - Long-term relationship building attempt
5. **Supply chain compromise** - Fake vendor communications
6. **Crisis exploitation** - Using current events to create urgency

Ethics Component: Corporate responsibility for AI use and deepfake prevention

Session 4: Mastery (1 hour)

Learning Objective: Coordinate team response to complex threats

Format:

- 5 min: Welcome + final role assignments
- 10 min: Real-world case study analysis
- 35 min: Complex team scenario requiring coordination
- 10 min: Action planning for workplace implementation

Team Challenge: Multi-stage attack simulation requiring different team members to:

- **Employee**: Receives initial suspicious contact
- **Manager**: Evaluates escalation and team consultation
- **IT Support**: Provides technical verification methods
- **Executive**: Makes final decision on high-stakes request

Ethics Component: Developing organizational policies for AI governance

Implementation Details

CISO-Led Facilitation

- **Pre-session**: CISO reviews current threat intelligence and customizes scenarios
- **During session**: CISO provides real-world context and organizational policy guidance
- **Post-session**: CISO tracks metrics and adjusts future content

Technology Requirements

- **Laptop/tablet per team** for scenario viewing
- **Audio playback capability** for voice samples
- **Simple scoring app** or paper-based tracking
- **Timer for each scenario** (30-60 seconds)

Scenario Content Sources

- **Real attack vectors** adapted from current threat intelligence
- **Anonymized internal examples** (with permission)
- **Industry-specific threats** relevant to organization type
- **Regulatory compliance scenarios** for ethics training

Measurable Outcomes

- **Pre/post assessment**: Deepfake detection accuracy
- **Behavioral metrics**: Verification protocol usage
- **Incident reporting**: Increase in legitimate suspicious activity reports
- **Culture indicators**: Team confidence in threat recognition

Spaced Repetition Support

- **Monthly mini-sessions** (15 minutes) with new scenarios
- **Quarterly refreshers** with updated threat landscape
- **Annual full program** with advanced scenarios
- **Real-time alerts** when new attack types emerge

Success Metrics

Immediate (Session-Level)

- **95%+ participation** across all sessions
- **85%+ detection accuracy** by Session 4
- **Consistent engagement** measured by active participation

Short-term (3 months)

- **60% improvement** in deepfake detection on assessments
- **40% increase** in appropriate verification behaviors
- **25% increase** in security incident reporting

Long-term (1 year)

- **Zero successful** deepfake/social engineering attacks
- **Organization-wide culture** of verification before action
- **Regulatory compliance** with AI governance requirements

Competitive Advantages

Professional Focus

- **Workplace-relevant scenarios** rather than abstract exercises
- **Realistic time pressures** matching actual work environments
- **Business impact context** for all scenarios

Scalability

- **Facilitator guide** enables CISO-led implementation
- **Modular content** allows customization by industry/role
- **Progressive difficulty** accommodates diverse skill levels

Regulatory Alignment

- **Built-in ethics training** meets compliance requirements
- **Documented learning outcomes** support audit needs
- **Policy integration** helps organizations develop governance frameworks

Commercial Positioning

Target Market: Mid-market organizations (500-5,000 employees) with dedicated IT but limited security specialization

Value Proposition: "Turn your monthly security meeting into an engaging team-building exercise that actually reduces cyber risk"

Pricing Model: $2-4 per employee per month, including facilitator materials, scenario updates, and outcome tracking

This game design balances professional credibility with engaging mechanics, providing measurable cybersecurity awareness improvement while respecting the time constraints and learning preferences of corporate employees.