INTELLIGENT CYBER THREAT DETECTION AND RESPONSE SYSTEM USING COLLABORATIVE AI AGENTS

Demo Script and Presenter Notes

Purpose

Show a live, end-to-end cyber defense pipeline where multiple specialized AI agents collaborate to detect threats, enrich context, orchestrate decisions, and execute automated responses — with clear explainability at each step.

Audience Takeaways

- See real-time telemetry, alerts, and decisions on a modern dashboard
- Understand how detection, intel enrichment, orchestration, and response collaborate
- Observe explainable rationales and measurable outcomes in history and analytics

Pre-Demo Checklist

1) Launch: open Dashboard (Login/register if needed)
2) Ensure at least one device is visible on the dashboard
3) Be ready to trigger scenarios using the scenario buttons (Normal / BruteForce / Exfil / Stealth)

High-Level Architecture (1 slide or 20s verbal)

- Telemetry Stream: hosts/services/netflows generate continuous events
- DetectionAgent: rule+statistical (and model-based when enabled) anomaly scoring
- ThreatIntelAgent: classifies threat type, adds severity, IOCs, and recommended checks
- OrchestratorAgent: applies policy/playbooks, prioritizes, chooses response
- ResponseAgent: executes mitigation actions and logs the effect (e.g., device isolation)
- Forensics/History: timelines and evidence for review

Live Demo Flow (5–6 minutes)

0) Dashboard Overview (30s)
   - Point to KPIs (Devices, Alerts, Incidents, Isolated)
   - Show Telemetry, Alerts, and Agent Logs tables updating

1) Scenario A — Credential Brute Force (≈2 min)
   - Click BruteForce; mention: “We’re stressing authentication on a device.”
   - Telemetry shows surges in auth failures; Alerts table surfaces “BruteForce” with confidence
   - Open the alert’s details (explanation/evidence visible in the UI lists)
   - ThreatIntelAgent enriches: label=CredentialBruteForce, severity=High, IOC if applicable
   - OrchestratorAgent decides action: isolate the targeted device to prevent lateral movement
   - ResponseAgent executes: device status flips to “isolated”; Agent Logs record decision/result
   - Call out explainability: short human-readable reasons tied to the alert and response

2) Scenario B — Data Exfiltration Spike (≈2 min)
   - Click Exfil; mention: “Outbound transfer behavior increases beyond normal baseline.”
   - Telemetry highlights bytes_out anomalies; Alerts show ExfilSuspect with confidence

  - Enrichment assigns severity; Orchestrator selects containment action
  - Response completes; isolated device count increases; logs capture the full decision chain

3) Scenario C — Low‑and‑Slow Activity (≈1 min)
  - Click Stealth; mention: “Subtle deviations accumulate over time.”
  - Statistical signals raise alerts; when model is enabled, MLAnomaly may also appear
  - Orchestrator may choose isolation based on severity/confidence and policy

4) History & Analytics (≈1 min)
  - History page: show incident entries, start/end times, outcomes
  - Analytics page: charts for Alerts by Type, Severity distribution, and Confidence buckets
  - Tie back to KPIs: isolated devices reflect real containment actions

Key Talking Points
- Collaboration: dedicated agents handle detection, intel, decisioning, and action; logs provide a transparent chain of custody
- Explainability: each alert includes concise evidence and rationale; each response includes justification and impact
- Speed: seconds from detection to mitigation; dashboards update continuously
- Modularity: agents are swappable/upgradable (rule sets, statistical thresholds, models, and playbooks)

What to Say (Short Lines You Can Read Aloud)
- “As the telemetry flows in, the DetectionAgent evaluates windows of activity and emits alerts with confidence and evidence.”
- “ThreatIntelAgent classifies the alert and attaches severity and recommended checks, giving context for decisioning.”
- “The OrchestratorAgent applies policy and selects a safe containment action with a clear reason.”
- “ResponseAgent executes the action immediately and records the outcome; device state updates on the dashboard.”
- “Analytics summarize types, severity, and confidence so we can reason about risk and tuning.”

If Questions Arise
- Tuning thresholds: “Thresholds and policies are adjustable; analytics guide tuning for TPR/FPR balance.”
- Expanding responses: “Playbooks can include isolation, restart, or network controls; actions are recorded with outcomes.”
- Adding models: “Detectors can incorporate additional features or models without changing the orchestration contracts.”

Two‑Minute Mini Script (for quick presentations)
1) Login → Dashboard
2) Start BruteForce → point at Telemetry surge and new Alerts

3) Explain enrichment → show severity/IOCs → decision: isolate
4) Show device status changes to isolated; Agent Logs capture the chain
5) Open History to show the incident; open Analytics to show updated distributions

Closing Line
“You’ve just seen a live collaborative defense pipeline: rapid detection, contextual intelligence, policy-driven decisions, and automated response — all with clear explanations and measurable outcomes.”