Group Project 3

Problem Statement

Target User

Our Approach (the 7 features list you just copied)

Human Impact Index (scoring model)

Mission Control Prototype Screens

Risk / Trust Model

Next Phase Plan (SOW)

Problem Statement

Small bio/health/food operators are flooded with cybersecurity and biosecurity alerts they don’t have time to interpret. They don’t know which alerts actually apply to them, how serious each one is in real-world terms, or what to do in the next hour to stay safe. Project Tuttwiler proposes a triage assistant that ingests signals, ranks them by human impact instead of just technical severity, and delivers one clear action at a time. The goal is fewer, smarter, safer decisions for under-resourced teams.

Target User

Primary user: the overworked technical lead at a small but critical operation (e.g., rural hospital IT admin, food safety plant supervisor, diagnostics lab manager). This person is responsible for uptime, safety, and compliance, but does not have a 24/7 SOC. They need:

1. to know if a new threat touches their environment,

2. to know how bad it is in *human* terms, and

3. to know the safe first step to take right now.

Our Approach (the 7 features)

Mission Control View — one priority card with a clear call: **GO / HOLD / ESCALATE**.

Human Impact Index — ranks by patient/food safety, accuracy risk, exploitability, patch status, dependency.

Safe Action Playbooks — short, plain-language steps safe for regulated environments.

Source Provenance / Trust Ledger — receipt of origin (CISA, vendor, OEM, etc.) + confidence.

"Is This Us?" Relevance Filter — quick local applicability check to suppress noise.

Tier-2 Escalation Workflow — one-click package to Bio-ISAC with context and audit.

Quiet Mode / Fatigue Guard — only human-safety (RED) alerts can break through.

Human Impact Index (scoring model)

Goal: Prioritize by *human and operational impact* for small operators.

Core factors (0–5) × weights

- **Human Safety Impact (×3.0):** none → *certain/ongoing harm*.
- **Clinical/Quality Accuracy (×2.0):** none → *proven error risk* (e.g., calibration drift).
- **Dependency & Substitutability (×2.0):** easy replacement → *sole-source/no backup*.
- **Exploitability / Activity (×2.0):** theoretical → *active in the wild*.
- **Patch / Remediation Status (×1.5):** patched → *no patch / won't fix*.
- **Operational Exposure (×1.5):** isolated → *internet-facing*.

Context modifiers

- **Scale/Reach:** −5…+5 (people/units affected).
- **Time Sensitivity:** −5…+5 (golden-hour, cold-chain, deadlines).
- **Source Confidence:** 0…−10 (vendor/CISA high → unverified low).

Formula

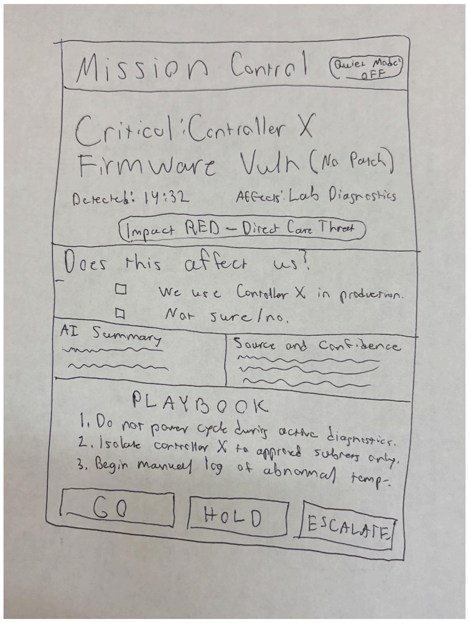
Total = Σ(score × weight) + Σ(modifiers)

Reference max (pre-modifiers) ≈ 60.

Thresholds

- **RED (Escalate):** Total ≥ 70 **OR** (Human Safety ≥ 3 **AND** (Exploitability ≥ 3 **OR** Patch ≥ 4)).
- **YELLOW (Monitor/Plan):** Total 40–69 and no RED trigger.
- **GREEN (Informational):** Total < 40.

Mission Control Prototype Screens



Mission Control

Critical: Controller X Firmware Vuln (No Patch)

Detected: 14:32 Affects: Lab Diagnostics

Impact RED – Direct Care Threat

Does this affect us?

□ We use Controller X in production.

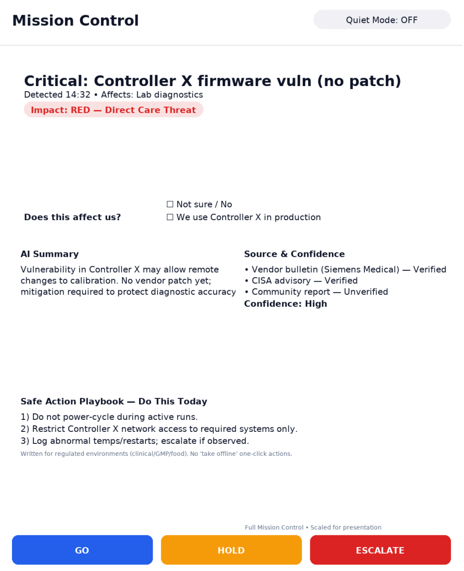
□ Not sure/no.

AI Summary

Source and Confidence

PLAYBOOK

1. Do not power cycle during active diagnostics.
2. Isolate controller X to approved subnets only.
3. Begin manual log of abnormal temp.



Mission Control

Critical: Controller X firmware vuln (no patch)

Detected 14:32 • Affects: Lab diagnostics

Impact: RED — Direct Care Threat

Does this affect us?

☐ Not sure / No

☐ We use Controller X in production

AI Summary

Vulnerability in Controller X may allow remote changes to calibration. No vendor patch yet; mitigation required to protect diagnostic accuracy

Source & Confidence

- Vendor bulletin (Siemens Medical) — Verified
- CISA advisory — Verified
- Community report — Unverified

Confidence: High

Safe Action Playbook — Do This Today

1) Do not power-cycle during active runs.
2) Restrict Controller X network access to required systems only.
3) Log abnormal temps/restarts; escalate if observed.

Written for regulated environments (clinical/GMP/food). No 'take offline' one-click actions.

Full Mission Control • Scaled for presentation

Intent: Reduce cognitive load; deliver one decisive card with safe actions.

1. **Mission Control — Full Screen**
   - Title + timestamp + affected function (e.g., “Lab diagnostics”).
   - **Impact pill** (RED/YELLOW/GREEN).
   - **AI Summary (proposed)** + **Source & Confidence** (CISA, vendor, community).
   - **Safe Action Playbook** (3–5 steps safe for clinical/GMP/food).
   - Decision row: **GO / HOLD / ESCALATE** (logs choice + packages context).
   - *Assets:* `Mission_Control_Full.pdf/.png` and `Mission_Control_Full_big.pdf/.png`.
2. **Mission Control — Simple Card (Lite)**
   - Single alert card with title, subline, impact pill, 2 bullets, **Schedule / Remind**.
   - Use for quick demos and mobile.
   - *Assets:* `Mission_Control_Simple.pdf/.png`.

Risk / Trust Model

Source provenance & integrity

- Maintain **allowlist** (CVE/NVD, CISA/CERT, vendor/OEM, FDA/EMA, CERT-EU) and a denylist for spoofers.
- Every alert includes origin URL, timestamp, **TLP**, content hash, and **confidence** tag.
- Unverified tips are **quarantined** for Tier-2 review; never auto-escalated.

User verification (lightweight)

- Org email + role attestation; referral/invite option; backchannel thumbs-up where available.

- Sensitive guidance is visible only to verified operators.

AI safety & human oversight

- AI drafts summaries and a **proposed** score; **humans confirm/adjust**.

- Guardrails block unsafe directives (e.g., “one-click shutdown”).

- All AI outputs and human overrides are **logged** for QA/tuning.

Alert-fatigue controls

- Single-card focus; dedupe by CVE/URL; suppress non-applicable items via **“Is this us?”**

- **Quiet Mode** lets only RED human-safety alerts break through protected windows.

Auditability & governance

- Every **GO/HOLD/ESCALATE** is timestamped with user, score breakdown, and provenance receipt.

- Tier-2 escalation bundles context and logs for follow-up; periodic review of false pos/neg.

Next Phase Plan (SOW)

Objective: Pilot a Tier-1 triage assistant with one willing operator site and Bio-ISAC Tier-2 support.

Scope & deliverables

1. **Ingest & Dedupe (MVP automation):**

    - Pull from CVE/NVD, CISA/CERT, selected vendor portals.

   - Normalize, hash-dedupe, tag TLP and confidence.
   - **Deliverable:** running intake + dedupe store; allowlist/denylist table.

2. **Human Impact Index v1.1 (tuned):**
   - Calibrate weights with Bio-ISAC; add examples and red/yellow/green bands.
   - **Deliverable:** rubric doc + worksheet; sample scored alerts.
3. **Prototype App (clickable + thin API):**
   - Mission Control screen; **Is this us?** checkbox; **GO/HOLD/ESCALATE** logging.
   - **Deliverable:** clickable prototype link + demo video (≤90s); minimal backend log.
4. **Safe Action Playbooks (starter set):**
   - Two scenarios: **No-patch critical (RED)**, **Patch within 72h (YELLOW)**.
   - **Deliverable:** PDF playbooks + blank template.
5. **Risk/Trust & AI Safety:**
   - Provenance receipts, verification concept, AI guardrails, logging.
   - **Deliverable:** Risk/Trust Model PDF; AI Usage Statement PDF.

Timeline (suggested)

- **Week 1:** finalize rubric, sources, and UX; build allowlist/denylist.
- **Week 2:** hook basic ingest/dedupe; produce scored sample alerts.
- **Week 3:** finalize Mission Control prototype; playbooks; demo video.
- **Week 4:** pilot walkthrough with Bio-ISAC; capture feedback; v1 SOW for Spring.

Success criteria

- Reviewers can answer in one glance: **“Does this affect us?”** and **“What’s the safe next step?”**

- At least **one** real-world alert scored and packaged for Tier-2 using the prototype.
- Evidence of **reduced alert fatigue** (deduped + suppressed non-applicable items).

Attachments / References (in Drive)