Policy Snippet (final)

AI Use in Security Operations
This system uses AI tools to identify phishing attempts, suspicious logins, and unusual network activity. Its purpose is to improve security by detecting threats quickly while reducing manual review time. When an alert is generated, users see a clear explanation (e.g., “possible phishing link” or “login from unusual location”).

AI decisions are never final. A trained analyst can review or override alerts, and users may appeal if they believe an item was flagged in error. Appeals are reviewed within **24 hours**, and responses are documented.

The system processes only the minimum necessary data (e.g., sender, subject line, metadata) with personal details redacted where possible. Alerts are retained for **90 days**, then securely deleted.

We track false positives, false negatives, and appeal times to ensure accuracy and fairness. AI models are reviewed quarterly and retrained at least once per year.

Controls & Metrics

- **Human-in-the-loop review:** All critical alerts verified by an analyst.
- **Appeals path:** User challenges resolved within **24 hours**.
- **Data minimization:** Only required fields processed; retention capped at **90 days**.
- **Accuracy targets:** False positive and false negative rates ≤ **5%**.
- **Review cadence:** Metrics reviewed **quarterly**; retraining at least **annually**.

Justification

These controls address risks highlighted in **Chapter 11** of *Cyber Ethics & Cyber Law* by focusing on bias, over-reliance, and accountability in AI systems. Human review and appeal processes counter over-reliance, subgroup testing and metrics address bias, and data minimization protects privacy.

Evidence Links

- CYBR-2100 Cyber Ethics & Cyber Law Syllabus – AI Ethics focus
- Real-world case: Amazon hiring AI bias (reported in news/ethics discussions).
- Predictive policing controversies (Chicago case studies).

Reflection

One trade-off I would revisit is the **24-hour appeal response window**. While it protects user trust, meeting this target consistently may require more staff time than some organizations can dedicate. Balancing fairness with operational resources will remain an ongoing challenge. Overall, I learned that building transparency and human oversight into AI systems is as important as technical accuracy.

AI Use Note

This text was drafted with support from **SAGE (AI assistant)** to structure sections, ensure clarity, and align with course materials.