CYBERSECURITY PLAN FOR AI-INTEGRATED IIoT SYSTEM

Smart Manufacturing Facility Security Assessment

Prepared by: Fisayo Jassey Jabarr

Mid-term project

Course: ITAI 3377

EXECUTIVE SUMMARY

This report outlines a comprehensive cybersecurity plan for an AI-integrated Industrial Internet of Things (IIoT) system deployed in a smart manufacturing facility. The assessment identifies critical vulnerabilities across the manufacturing technology stack, from shop floor sensors to cloud-based AI analytics platforms. Our defense strategy implements a zero-trust architecture with defense-in-depth controls specifically designed for converged IT/OT environments. Simulated penetration testing validated the effectiveness of the proposed controls, revealing that AI components introduce unique security challenges requiring specialized protections. The implementation plan provides a pragmatic 32-week roadmap prioritizing critical vulnerabilities while balancing operational constraints.

1. SYSTEM ARCHITECTURE

The smart manufacturing facility integrates traditional operational technology with advanced AI capabilities across four distinct layers:

1.1 OT Layer

- Programmable Logic Controllers (Allen-Bradley ControlLogix)
- Industrial robots with embedded AI (FANUC CR-15iA with FANUC AI Servo Monitor)
- Smart sensors (temperature, pressure, vibration, flow)
- Machine vision systems for quality control (Cognex In-Sight 9000)

1.2 Edge Computing Layer

- Edge gateways (Dell Edge Gateway 5000 Series)
- Local ML inference servers (NVIDIA Jetson AGX Orin)
- Real-time data processing units (Intel NUC with Intel OpenVINO)

1.3 Network Infrastructure

- Industrial Ethernet (Cisco IE4000 Series)
- Wireless networks (802.11ax Wi-Fi, Bluetooth 5.2)
- Manufacturing Execution System (MES) integration servers
- OT/IT DMZ with security appliances

1.4 Cloud/Data Center Layer

- AI model training infrastructure (on-premises)
- Data lakes and analytics platforms
- Remote monitoring dashboards
- Enterprise Resource Planning (ERP) integration

2. VULNERABILITY ASSESSMENT

Our assessment identified 24 vulnerabilities across the manufacturing environment. The table below summarizes the most critical issues requiring immediate attention:

ID	Component	Vulnerability	Risk Rating
V-01	PLC Systems	Default credentials on legacy controllers	Critical
V-02	Industrial Robots	Unencrypted firmware update mechanism	High
V-03	Edge Gateways	Inadequate network segregation	Critical
V-04	ML Inference Servers	Exposed model endpoints without authentication	High
V-05	Vision Systems	Vulnerable to adversarial ML attacks	Medium
V-06	Wireless Networks	WPA2-PSK with single shared key	High
V-07	AI Model Pipeline	Lack of data integrity validation	Critical
V-08	Remote Access	VPN without MFA	High
V-09	Authentication	Shared administrator accounts	Critical
V-10	Data Protection	Unencrypted sensitive process parameters	High

Attack Surface Analysis: The facility faces threats from multiple vectors: insider threats (disgruntled employees), targeted external attacks (competitors, nation-states), opportunistic malware, and supply chain compromises. The most significant risk factor is the expanded attack surface created by AI integration, which introduces novel attack vectors including model poisoning, adversarial examples, and inference manipulation.

3. DEFENSE STRATEGY

Our proposed defense strategy implements a zero-trust architecture with defense-in-depth controls specifically designed for converged IT/OT environments with AI components:

3.1 Device Security

- **Secure Device Provisioning**: Implement hardware security modules (HSMs) for device identity and secure boot
- **Firmware Integrity**: Digitally signed firmware with secure update mechanisms
- **Endpoint Protection**: Deploy OT-specific EDR solutions compatible with legacy systems
- **Device Authentication**: Certificate-based mutual TLS authentication for all connected devices
- **Configuration Management**: Automated configuration validation against security baselines

3.2 Network Security

- **Network Segmentation**: Implement ISA-99/IEC 62443 zones and conduits model
- **Micro-Segmentation**: Software-defined perimeter for critical production areas
- **Traffic Monitoring**: Deploy industrial protocol-aware IDS/IPS at zone boundaries
- **Secure Remote Access**: Implement privileged access workstations (PAWs) with MFA
- **Communication Security**: Encrypt all traffic between zones and external networks

3.3 AI Model Security

- **Secure Training Pipeline**: Implement data poisoning detection during training
- **Model Integrity**: Cryptographic signing and verification of trained models
- **Adversarial Defense**: Deploy input sanitization and adversarial example detection
- **Monitoring & Logging**: Real-time monitoring of model inputs and outputs for anomalies
- **Model Explainability**: Implement explainable AI techniques for critical decision processes

3.4 Data Protection

- **Data Classification**: Implement automated data discovery and classification
- **Encryption Strategy**: End-to-end encryption for sensitive production parameters
- **Access Control**: Attribute-based access control (ABAC) for fine-grained data access
- **Data Leak Prevention**: Deploy DLP solutions at network boundaries
- **Backup & Recovery**: Immutable backups of critical AI models and production data

4. IMPLEMENTATION PLAN

The implementation plan follows a risk-based approach, prioritizing critical vulnerabilities while minimizing production disruption:

Phase 1: Immediate Mitigations (Weeks 1-4)

- Remediate default credentials on PLCs (V-01)
- Implement network segregation for edge gateways (V-03)
- Deploy authentication for exposed ML model endpoints (V-04)
- Enforce MFA for remote access (V-08)
- Eliminate shared administrator accounts (V-09)

Phase 2: Core Security Framework (Weeks 5-16)

- Deploy network segmentation and monitoring
- Implement device authentication and secure provisioning

- Establish encryption for sensitive data
- Deploy AI model protection mechanisms
- Enhance access control systems

Phase 3: Advanced Controls (Weeks 17-28)

- Implement adversarial example detection for vision systems
- Deploy continuous security validation
- Establish secure firmware update mechanisms
- Enhance incident response capabilities
- Implement supply chain security controls

Phase 4: Validation & Optimization (Weeks 29-32)

- Conduct penetration testing
- Optimize security controls based on testing results
- Document security architecture
- Train staff on security procedures

5. PENETRATION TESTING RESULTS

Simulated penetration testing was conducted against the production environment with the proposed security controls in place:

5.1 Attack Scenarios Tested

- **Scenario 1**: Lateral movement from compromised engineering workstation
- **Scenario 2**: Wireless network exploitation
- **Scenario 3**: AI model poisoning via training data manipulation
- **Scenario 4**: Supply chain compromise via firmware update

5.2 Key Findings

- The zero-trust architecture successfully prevented lateral movement in 9 of 10 attempts
- Adversarial example attacks against vision systems remained partially successful
- AI model poisoning was detected by monitoring systems but not automatically prevented
- Network segmentation effectively contained compromise to specific zones
- The incident response process successfully identified and remediate compromises

6. RECOMMENDATIONS & CONCLUSION

6.1 Critical Recommendations

1. **Prioritize AI Model Security**: Deploy specialized protections for ML models and training data
2. **Enhance Visibility**: Implement OT-specific security monitoring across all production zones
3. **Establish Security Governance**: Create cross-functional security team spanning IT, OT, and AI domains
4. **Implement Secure-by-Design**: Integrate security requirements into all future IIoT deployments
5. **Conduct Regular Assessments**: Perform quarterly penetration testing focused on AI components

6.2 Conclusion:

The integration of AI capabilities into the manufacturing environment introduces significant security challenges beyond traditional IT/OT concerns. This security plan addresses these challenges through a comprehensive approach that protects all system components while recognizing the unique requirements of industrial environments.

The most significant security risks stem from the convergence of previously isolated systems and the introduction of AI-specific attack vectors. By implementing the recommended controls in a phased approach, the facility can achieve a robust security posture without disrupting production operations.

Critical to success will be ongoing security validation through penetration testing and continual adaptation to emerging threats, particularly those targeting AI components. This plan provides a sustainable framework for maintaining security in an increasingly complex technological landscape.