REGIONAL CYBERSECURITY SUMMIT FOR AFRICA

Evolving Security Strategies for Artificial intelligence

Princess M Zawu
Liberia Telecommunication Corporation

20-23 November 2023
Kampala, Uganda

CONTENTS

- Security strategies for artificial intelligence (Al)
- AI and SOC Enhancement
- Challenges and Opportunities
- Comparing Traditional Security Measures and Al- specific strategies
- AI security strategies to safeguard Al systems
- preventing bias in Al models
- Summary

SECURITY STRATEGIES FOR ARTIFICIAL INTELLIGENCE (AI)

- AI proliferation across domains necessitates safeguarding its deployment.
- Key considerations for AI security strategies:

1. Complex Threat Landscape:

Evolving, sophisticated, and well-funded cybersecurity threats.

New attack vectors introduced by AI-powered tools and technologies. Rise of customized, high-impact cyberattacks.

2. The Role of Security Operations Centers (SOCs):

Facility housing cybersecurity professionals.

Responsible for real-time monitoring, investigation, and response to cyberthreats.

Protect assets such as intellectual property, confidential data, critical infrastructure, and brand reputation.

SECURITY STRATEGIES FOR ARTIFICIAL INTELLIGENCE (AI) contd….

3. SOC Evolution: -

- Initially implemented for government and defense organizations.
- Adoption by large enterprises and banks.
- Key milestones:

✓ 2005: Introduction of compliance objectives.

✓ 2007-2013: Emergence of critical security solutions (e.g., data leakage prevention, SIEM).

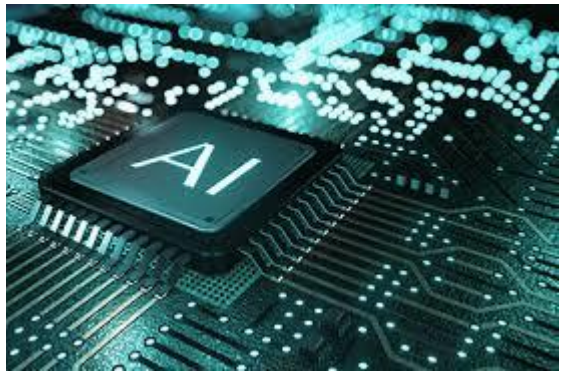
✓ 81% increase in advanced persistent threats (APTs) during this period

AI AND SOC ENHANCEMENT

- AI can significantly enhance SOC capabilities:
  - Threat detection: AI algorithms can identify patterns and anomalies.
  - Automation: AI-driven automation streamlines incident response.
  - Predictive analytics: AI predicts potential threats.
  - Behavioral analysis: AI detects abnormal activities

CHALLENGES AND OPPORTUNITIES

- Challenges in securing AI systems:
  - Bias: AI models must recognize bias in others without being biased themselves.
  - Resilience: AI systems need to withstand attacks and adapt. –
  - Discretion: Balancing transparency with security.



COMPARING TRADITIONAL SECURITY MEASURES AND AI- SPECIFIC STRATEGIES

Traditional security Measures

- Traditional security measures, such as firewalls, encryption and access controls, have been effective in safeguarding traditional IT systems.
- While traditional security measures are essential, they have limitations when applied to Al systems.
- Al systems often deal with vast amounts of complex and dynamic data, making traditional static security measures insufficient. Traditional measures focus on protecting the infrastructure and data, but may not adequately address Al specific vulnerabilities in the models themselves.

Al-Specific Security Strategies

- Al- specific strategies are required to address the unique risks and challenges posed by Al-systems.
- These strategies complement traditional security measures and focus on securing the Al algorithms, training data , and decision making processes.
- Al –specific security strategies are emphasize robust data governance practices.
- Al- specific strategies prioritize algorithm, transparency and explainability.
- Al-specific strategies employ adversarial training and defence mechanisms to enhance model robustness against attacks.

AI SECURITY STRATEGIES TO SAFEGUARD AI SYSTEMS

1. Open Worldwide Application Security Project (OWASP) AI Security and Privacy Guide:
   - Comprehensive guide on AI security and privacy provided by OWASP Foundation.
   - Covers designing, creating, testing, and procuring secure and privacy-preserving AI systems.
   - Key areas include transparency, auditability, bias countermeasures, and oversight.
2. AI Security Risk Assessment Framework:
   - Microsoft's AI security risk assessment framework for auditing and improving AI system security.
   - Empowers organizations to track and mitigate risks associated with AI deployments.
3. Incorporating AI in Cybersecurity Strategies:
   - AI enhances threat detection and response times.
   - Techniques like differential privacy and federated learning protect data and improve system robustness.

AI SECURITY STRATEGIES TO SAFEGUARD AI SYSTEMS contd…..

4. Good Software Engineering Practices:

- Apply software engineering practices to AI activities:
  - Versioning, documentation, unit testing, and code quality.
  - Mix data scientist profiles with software engineering profiles for future-proof, maintainable code.

5. Awareness of Particular AI Security Risks:

- Understand AI-specific risks:
  - Protect model parameters and monitor model access.
  - Guard against data leaks, model poisoning, and supply-chain attacks.

PREVENTING BIAS IN AI MODELS

- Preventing bias in AI models is crucial for fairness, equity, and ethical use of AI.
- Strategies to mitigate bias:

1. Blind Taste Tests:
   - Apply a blind approach to AI models, similar to blind taste tests.
   - Deny the algorithm information suspected of biasing the outcome.
   - Break the cycle of bias to promote greater equality across contexts.
2. Human-in-the-Loop Processes:
   - Involve humans in the decision-making loop alongside AI systems.
   - Humans can identify and correct biases that algorithms might miss.
   - Collaborative approach to mitigate bias and ensure a balanced outcome.

PREVENTING BIAS IN AI MODELS contd....

- 3. Diverse Training Data:
  - Monitor for outliers using statistical techniques and data exploration.
  - Compare and validate different samples of training data for representativeness.
  - Ensure training data includes diverse perspectives, backgrounds, and demographics.

4. Awareness and Training:

  - Train machine learning teams on AI bias.
  - Be aware of implicit biases and their influence on model training.
  - Screen participants for potential bias and establish clear data collection and annotation guidelines.

- While eliminating bias entirely is challenging, these strategies contribute to building more robust and equitable AI models.

SUMMARY /TAKE HOME

- Effective SOC, bolstered by AI, is crucial for countering evolving cyber threats.
- AI security strategies are essential to safeguard AI systems from threats and vulnerabilities.
- Preventing bias in AI models ensures fairness, equity, and ethical use of AI.
- Holistic approach combining technical measures, organizational responsibility, and ongoing vigilance is necessary.

Thank you!