iterasec

Securing NDA Security company's AI-Enhanced Documentation Platform

Client:
NDA Security company

Industry:
Network Security and Infrastructure

Company Size:
300+ Employees

Services Used:
Marketing Website & Documentation Portal Penetration Testing, LLM Security Assessment

Client

An NDA Security company specialized in advanced solutions for network security, providing both software and hardware products designed to protect complex enterprise environments. By consistently innovating and adapting to emerging technological challenges, company's offerings helped a wide range of organizations maintain robust defenses against evolving cyber threats.

Background

Building on prior collaborations, Security company engaged Iterasec to conduct a comprehensive security review of their corporate websites, documentation portal, and customer communication platform. The standout feature in this project was the integration of Claude, a Large Language Model (LLM) chatbot, into the documentation portal. Security company aimed to enhance user experience by enabling customers to interact with detailed product instructions and quickly find answers to technical questions through an AI-driven interface.

The Challenge

- **New AI Integration**
  - With Claude integrated into the documentation portal, there were concerns regarding potential AI-specific vulnerabilities such as prompt manipulation or unauthorized data disclosure.
  - Ensuring that the chatbot adhered to privacy and compliance standards was vital to maintain customer trust.
- **External-Facing Web Properties**
  - Security company's websites and communication portals were accessible to a global audience, making them attractive targets for potential cyberattacks.
  - The portals handled sensitive user data, including support tickets and customer inquiries, necessitating strong access controls and data protection measures.
- **User Experience vs. Security Balance**
  While adding cutting-edge features, Security company needed to guarantee that new functionalities would not introduce exploitation paths or compromise their established security posture.

The Solution

Iterasec conducted a thorough penetration test focused on the websites, portals, and AI-powered documentation platform:

Website & Portal Security Testing

- **Since the documentation portal required authorization, Iterasec identified a 2FA bypass vulnerability.**
- **Assessed authentication and authorization flows, ensuring only legitimate users could access sensitive data and functionalities.**
- **The marketing websites were built on top of Wordpress, as a result, identified several minor WordPress misconfigurations and vulnerable plugins.**

AI Chatbot Security Assessment

- **Evaluated Claude's integration for prompt injection and data leakage vulnerabilities, simulating real-world attacks based on OWASP LLM Top 10 that could exploit AI-driven content generation.**
- **Collaborated closely with Skybox Security's development and data science teams to recommend guardrails that prevent malicious requests or unintended data access.**

Data and Privacy Compliance

- **Reviewed data handling and retention policies for stored chats and support interactions.**
- **Provided recommendations on encryption, data segmentation, and secure logging to reduce the risk of unauthorized data exposure.**

Ongoing Consultation

- **Established a clear communication channel for immediate reporting of critical findings.**
- **Advised on best practices for maintaining and updating AI-driven systems, including monitoring for emerging threats tied to LLM technology.**

The Outcome

- **Secured Web Portals and Marketing Websites**
  - Addressed multiple Wordpress misconfigurations and ensured the websites are protected according to the best practices.
- **Protected AI Integration**
  - Implemented robust safeguards within the chatbot, minimizing risk from prompt injection and ensuring strict adherence to user access controls.
  - Maintained data privacy and compliance through tailored logging and encryption mechanisms.
- **Improved Customer Experience**
  - Strengthened the documentation portal's overall functionality and reliability, allowing clients to engage confidently with both static content and the AI-driven interface.
  - Reinforced Security company's reputation for delivering user-friendly yet secure solutions.
- **Foundation for Future Enhancements**
  - Equipped Skybox Security's teams with methodologies and frameworks to continuously evaluate AI-based platforms.
  - Positioned the company to integrate additional conversational AI features while preserving a robust security posture.

Conclusion

By rigorously testing Security company's newly integrated AI chatbot and existing web portals, Iterasec helped ensure that these platforms delivered a seamless and secure user experience. The safeguards introduced not only mitigated immediate vulnerabilities but also established clear protocols for future enhancements. This collaborative effort underscored company's ongoing commitment to innovation and provided their customers with a trusted environment for both product exploration and real-time technical support.