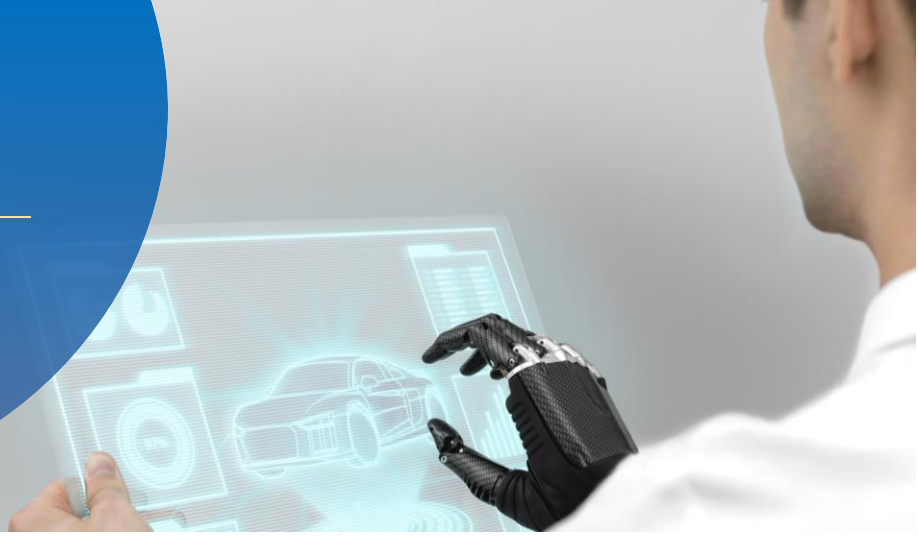
cogent infotech

Case Study

Ransomware Readiness Program: Immutable Backups + Recovery Drills for Public Sector



The Client

A global enterprise integrating Generative AI into internal productivity platforms and customer-facing digital services needed a structured security framework to scale AI adoption safely. As GenAI usage expanded, leadership identified growing risks related to prompt injection, the exposure of sensitive data, the misuse of AI models, and the need for stronger governance over AI access and use.

The Challenge

- Prompt injection attacks could manipulate AI behavior or bypass safety controls
- Sensitive enterprise and customer data faced potential exposure through AI interactions
- AI systems could be abused to generate harmful content or automate malicious activity
- Lack of standardized governance controls created policy and compliance risk
- Limited visibility into AI usage patterns, suspicious prompts, and model abuse attempts
- Need to scale GenAI securely without slowing innovation across internal and external platforms

About Cogent Infotech

Founded in 2003, Cogent Infotech is a trusted, award-winning firm with **23+ years** of experience**, 150+** government contracts, **10,000+** projects, and a **96% employee retention rate**. Recognized as an SBA Small Business and MBE-certified, we deliver excellence through diverse talent, AI-driven recruitment, and cooperative contracts like **NASPO Value Point and TIPS-USA.**

John Makely
Sr. Sales Manager
Business Development

John.makely@cogentinfo.com

323-573-2680

Solution and Process

- **AI Security Guardrails**
  Deployed prompt filtering and response moderation to detect and block malicious prompts and unsafe outputs
- **Data Protection Controls**
  Implemented data redaction, secure APIs, and encryption to prevent exposure of sensitive enterprise and customer information
- **Access and Usage Management**
  Established role-based access controls, multi-factor authentication, and API rate limits to prevent unauthorized access and model abuse
- **Monitoring and Threat Detection**
  Integrated AI interaction logs with SIEM and SOC monitoring tools to detect suspicious prompt activity and potential attack patterns
- **Secure AI Lifecycle and Governance**
  Conducted adversarial testing and implemented secure deployment pipelines to protect models as they evolved and support policy-compliant AI operations

Outcome

- Reduced risks associated with prompt injection and model manipulation
- Strengthened protection of sensitive enterprise and customer data
- Enabled secure and responsible scaling of GenAI applications across the organization
- Improved visibility into AI usage and emerging security threats
- Increased leadership confidence in the organization's AI security posture

Risk Analysis

- Prompt injection risk that could alter model behavior or bypass safeguards
- Data leakage risk involving confidential enterprise and customer information
- Model abuse risk through harmful content generation or malicious automation
- Access governance risk from unauthorized use of AI services and APIs
- Monitoring and compliance risk if AI activity is not centrally logged, reviewed, and controlled

Best Practices

- ✓ Applied prompt filtering and response moderation at every AI interaction point
- ✓ Used data redaction, encryption, and secure APIs for sensitive workloads
- ✓ Enforced RBAC, MFA, and API rate limiting to reduce unauthorized access and abuse
- ✓ Integrated AI activity monitoring with SIEM and SOC operations for threat visibility
- ✓ Performed adversarial testing and secure deployment validation throughout the AI lifecycle

TECHNOLOGIES

- Prompt Filtering and Response Moderation
- Secure APIs
- Encryption and Data Redaction Controls
- SIEM & SOC Monitoring Tools
- RBAC, MFA, and API Rate Limiting
- Adversarial Testing and Secure Deployment Pipelines

John Makely
Sr. Sales Manager
Business Development
John.makely@cogentinfo.com
323-573-2680