DELL Technologies

5 GenAI 효과를 안전하게 극대화하기 위한 권고

1	2	3	4	5
GenAI 시스템의 레이어 보호	**제로 트러스트 (zero trust) 원칙 활용**	**거버넌스 및 인적 감독 유지**	**사용 가능한 GenAI 보안 툴 활용**	**확신을 가지고 혁신 달성**
인프라스트럭처	절대 신뢰하지 않고 항상 검증	주요 이해 관계자 참여 유도	내용	임무를 저해하지 않고 촉진하면서 사이버 보안을 목표로 함
OS 및 Kubernetes	최소 권한 액세스	윤리 및 규정 준수, 데이터 관리를 위한 정책 수립	위험 예측	성숙된 사이버 보안을 통해 조직이 혁신에 대한 자신감을 갖도록 함
GenAI 애플리케이션	시스템 강화	책임 소재 모니터링 및 적용	지식 및 자동화	
데이터	ID 관리	훈련 및 교육		
	세분화			
	로깅, 모니터링 및 감사			

Generative AI 기술은 혁신적인 기능을 약속하지만 특유의 보안 문제를 수반합니다.

Generative AI는 전에 없던 혁신을 주도하고 경쟁 우위를 제공하는 독보적인 이점을 통해 비즈니스에 혁신을 일으키고 있습니다. 이 기술은 혁신적인 잠재력을 가지고 있지만 자체적인 보안 문제를 수반하고 있습니다.

Dell SME(Subject Matter Expert)이자 Services Product Manager인 Steve Brodson 씨와과 Cybersecurity Consultant인 Eitan Lederman 씨는 APEX 및 AI 마케팅 팀의 Chris Cicotte 씨와 함께 이러한 문제를 해결하고 GenAI 효과를 안전하게 극대화하는 방법에 대해 논의했습니다. dell.com/cybersecuritymonth에서 대화 요약 및 주제에 대한 추가적인 인사이트를 읽어보고 전체 토론을 시청하십시오.

> 결국에는 사람을 학습시키는 것입니다. 사람들은 GenAI 시스템을 어떻게 사용해야 하는지 알아야 합니다. 무엇을 해야 하는지 뿐만 아니라 무엇을 하지 말아야 하는지도 배워야 합니다."
>
> **Eitan Lederman**
> Dell Cybersecurity Consultant

GenAI 시스템의 레이어 보호

GenAI는 비교적 새로운 기술이지만 대부분의 보안 프로토콜은 다른 워크로드를 보호하는 데 사용되는 기존의 사이버 보안 기술과 동일합니다.

인프라스트럭처 - 공격 노출 지점 최소화에 집중:

- 취약성 및 침투 테스트
- 패치
- 강화
- 강력한 비밀번호인 MFA(Multi-Factor Authentication)를 포함한 ID 관리
- 모니터링 및 감사
- 타사 공급망의 보안 보장

OS 및 Kubernetes - 역시 다음과 같은 공격 노출 지점을 줄이는 데 중점을 둠:

- 취약성 검사
- 정기적인 패치
- Kubernetes 구성 요소 업데이트
- ID 관리, RBAC(Roles-Based Access) 및 최소 권한 액세스를 기반으로 액세스 제어 제한
- API 서버, 기밀, kubelet 및 기타 구성 요소를 포함하는 제어 플레인을 안전하게 보호
- 네임스페이스 사용

GenAI 애플리케이션 - GenAI가 생성하는 새로운 공격 노출 지점을 대상으로 보안 조치 구현:

- 프롬프트 인젝션, 기밀 정보 공개, 모델 도난, 학습 데이터 오염 문제를 해결하기 위한 ID 관리
- 학습 데이터 오염, 모델 편향으로부터 보호하기 위한 데이터 소스 검증
- 모델 DOS, 모델 도난, 기밀 정보 공개, 이상 징후 감지, 포렌식을 식별 및 예방하기 위한 모니터링 및 감사

데이터 - 강력한 데이터 보호 수단을 통합하여 언어 모델 및 애플리케이션에서 데이터 보호:

- 에어 갭 처리된 사이버 볼트
- 암호화
- 인시던트 대응 계획
- 학습 데이터와 결과물의 모니터링 및 감사

학습 입력값, 모델 출력값 그리고 RAG(Retrieval Augmented Generation)와 관련된 모든 데이터 등에 데이터 보호 원칙이 적용되고 있는지 확인하십시오. 또한 모든 관련 데이터 보호 규정을 지속적으로 준수해야 합니다.

제로 트러스트(zero trust) 원칙 활용

이미 언급된 ID 관리, 최소 권한 액세스, 시스템 강화 및 패치 적용과 같은 몇 가지 제로 트러스트 원칙의 역할은 GenAI 워크로드를 보호하는 데 있어 제로 트러스트 원칙의 가치를 나타냅니다. 제로 트러스트 아키텍처 또한 네트워크 활동에 대한 지속적인 로깅, 모니터링 및 감사가 필요하며, 이는 결과 조작 및 데이터 오염과 같은 GenAI 관련 위험을 방지할 수 있습니다.

DELL Technologies

이외에도 제로 트러스트는 침해의 영향을 줄이는 마이크로 세분화를 촉진합니다. 또한 전송 중 데이터와 저장 상태 데이터 모두에 대해 데이터 암호화를 요구하는데, 이는 전체 데이터 보호 전략에서 중요한 부분입니다.

앞서 설명한 것들은 제로 트러스트를 통해 GenAI 워크로드를 보호하는 방법 중 몇 가지일 뿐이며, 제로 트러스트 원칙을 도입하는 것을 모범 사례로 고려해야 합니다.

거버넌스 및 인적 감독 유지

GenAI의 가치는 대부분 인간이 일반적으로 실행하는 작업을 자동화하는 데 있지만, 애플리케이션의 보안과 적절한 기능을 보장하기 위해서는 인간의 관리가 중요합니다. 거버넌스 모델을 통해 조직 전반에서 주요 이해 관계자들이 윤리 및 규정 준수, 데이터 관리 정책 및 절차에 대한 지침과 요구 사항을 설정하며 궁극적으로는 책임 소재를 명확히 할 수 있습니다.

적절한 거버넌스 및 감독은 지나친 모델 의존, 편향, 결과 조작, 기밀 정보 공개, 데이터 오염과 같은 문제를 해결하는 데 도움이 될 수 있습니다.

Lederman 씨는 학습의 중요성도 짚었습니다. "결국에는 사람을 학습시키는 것입니다. 사람들은 GenAI 시스템을 어떻게 사용해야 하는지 알아야 합니다. 무엇을 해야 하는지 뿐만 아니라 무엇을 하지 말아야 하는지도 배워야 합니다."

조직의 GenAI 애플리케이션이 야기하는 위험 외에도, GenAI 기반 사이버 공격도 확산되고 있으며 인간의 개입이 필요한 경우도 많습니다. 악의적인 공격자가 딥페이크를 사용하여 사람의 행동을 유도하는 경우나 사람의 글쓰기 또는 말하기 스타일을 보다 정확하게 모방하여 훨씬 더 속기 쉬워진 피싱 공격을 예로 들 수 있습니다. 지속적인 학습과 교육은 이러한 위험을 해결하는 가장 효과적인 방법 중 하나이며, 이에 따라 다시 인적 요소가 중요해집니다.

사용 가능한 GenAI 보안 툴 활용

대부분의 초점이 위험에 맞춰져 있지만 GenAI에는 보안 노력을 강화할 수 있는 잠재력도 있습니다. 이러한 기능은 현재 초기 단계이지만 세 가지 주요 영역에서 이점을 제공합니다.

- **콘텐츠:** 보안 정책 생성, 맞춤형 교육, 데이터 분류 및 보고
- **예측:** 위험 및 공격 활동 예측, 해결 조치 제안
- **지식:** 환경 쿼리(시스템과 대화), 포렌식, 자동화

GenAI를 보안 툴에 활용하는 것은 보안 팀의 역량을 극대화하고 비용을 절감하며 방어 체계를 강화하는 데 도움이 될 수 있습니다. 점점 더 증가하고 있고 성숙도가 높아지는 이러한 솔루션을 활용하십시오.

확신을 가지고 혁신 달성

무엇보다 중요한 것은 보안 위험 때문에 혁신 잠재력을 가진 기술을 활용하지 못하는 상황을 방지하는 것입니다. GenAI는 효율성, 자동화, 비용 절감, 문제 해결 및 창의성 유도 등 다양한 방식으로 비즈니스를 혁신할 수 있습니다.

GenAI는 강력하면서도 때로는 새로운 사이버 보안 조치를 요구하지만, 목표는 조직의 임무를 방해하는 것이 아니라 촉진하는 것입니다. 올바른 사이버 보안 전략을 개발하면 성장하고 혁신할 수 있다는 자신감을 갖게 됩니다.

dell.com/cybersecuritymonth에서 오늘날의 주요 사이버 보안 과제를 해결하는 방법을 알아보십시오.

DELL Technologies