발표요약문

발표자의 요청에 따라 발표자료 대신 요약문을 제공합니다.

위기에 처한 신뢰: 생성형 AI 시대의 미래를 보호하다

생성형 AI 시대가 도래하며 AI 활용이 보편화되고 있지만 이와 함께 반드시 해결해야 하는 중대한 위험도 동반된다. 이 중 가장 큰 위험은 AI 모델과 데이터의 보안이며, 이는 개발 과정뿐만 아니라 제품 공개 이후에도 중요한 문제다. 본 발표에서는 책임 있는 AI(Responsible AI)의 개념과 및 구글의 보안 AI 프레임워크(Secure AI Framework)를 소개하고, 최고경영진이 AI 개발 및 활용에 있어 적절한 거버넌스를 구현하는 방법을 조명한다. 또한 AI를 대상으로 발생할 수 있는 6가지 주요 공격 유형과 이에 대한 방어 전략을 소개하고, 경영진이 기업의 AI 자산과 데이터, 기반 인프라를 보호하기 위해 취해야 할 접근 방식에 대한 권고사항을 제시한다.

Trust at Risk: Securing the Future in the Age of Generative AI

In the Age of Generative AI, the use of AI is becoming prevalent, but this does not come without significant risks that must be addressed. Chief amongst these risks is the security of the AI models and data during their development, as well as after they are released into public use. This talk will focus on how the C-Suite can implement appropriate governance over the development and use of AI, by introducing the concepts of Responsible AI and Google's Secure AI Framework. This discussion will also cover 6 common types of attacks that can be made against AI, and how to defend against these attacks. Finally, the talk will provide recommendations for how the C-Suite should approach the defence of their corporation's AI assets, data and underlying infrastructure.