Innovation Sandbox: Cybersecurity Investors Pivot to Safeguarding AI Training Models

News Analysis: If the winner of the RSA Innovation Sandbox says anything about the future of innovation and hype in cybersecurity, brace yourselves for a cottage industry of startups promising to protect AI machine learning models behind enterprise products.

At the annual RSA Conference shindig in San Francisco this week, a tiny Texas company called **HiddenLayer** won the 'Most Innovative Startup' prize for its technology that promises to monitor algorithms for adversarial ML attack techniques.

The HiddenLayer win signals an interesting shift in the startup ecosystem as venture capitalists pivot from hyping AI/ML security tools to investing in new companies to protect the code flowing in and out of AI training sets.

HiddenLayer's pitch is a future that includes MLMDR (machine learning detection and response) platforms that monitor the inputs and outputs of your machine learning algorithms for anomalous activity consistent with adversarial ML attack techniques. The company emerged from stealth in July 2022 with $6 million in funding.

What does winning the RSA Innovation Sandbox mean?

The RSA Innovation Sandbox, whether you take it seriously or not, provides a massive soapbox for investors and

entrepreneurs to pitch security wares, boost sales pipelines and validate new approaches to market categories.

Now in its 18th year, the top 10 Sandbox finalists have collectively seen over 75 acquisitions and raised more than $12.5 billion in investments since its inception. Previous winners include recognizable names like Imperva, Phantom, SECURITI.ai, Apiiro and Talon Cyber Security.

In previous years, the Sandbox finalists and pitches provided signs of investors rushing to fund startups in emerging categories like Data Security Posture Management (DSPM), API security, software supply chain security and intelligent identity and access management.

Now that HiddenLayer has captured the spotlight, look for a mad scramble to incubate and launch startups promising to protect the machine learning models and training sets behind tools like ChatGPT and other popular generative AI chatbots.

Consulting giant KPMG has already spun out a venture-backed startup building technology to secure AI (artificial intelligence) applications and deployments as organizations look to a future where AI models – and the data flowing through them – need to be secured.

KPMG's Cranium says it is working on "an end-to-end AI security and trust platform" capable of mapping AI pipelines, validating security, and monitoring for adversarial threats.

Big tech vendors Microsoft and Google have also started competing in the AI/ML space with Redmond first out of the gate with Microsoft Security Copilot, a ChatGPT-powered security analysis tool to automate incident response and threat hunting tasks.

Anti-malware vendor SentinelOne has followed suit with its own AI-powered threat hunting platform and Google's VirusTotal subsidiary rolled out a major generative AI feature upgrade.

In addition to security use cases for AI chatbots, the dramatic adoption of generative AI technology is sure to spur innovation among vendors helping with coming compliance and regulatory mandates.

Investors are seeing signs of revenue everywhere and the results of this year’s RSA Innovation Sandbox, a contest that includes VCs as judges, present a clear sign of what’s to come in cybersecurity innovation.

Related: RSA Conference 2023 Announcements: Day 1, Day 2, Day 3

Related: Microsoft Puts ChatGPT to Work on Automating Cybersecurity

Related: KPMG Tackles AI Security With Cranium Spinout

Related: ChatGPT Integrated Into Security Products as Industry Tests Capabilities

Innovation Sandbox: Cybersecurity Investors Pivot to Safeguarding AI Training Models