How a leading Global Financial Services organization exposed critical AI vulnerabilities in a vast third-party ecosystem

54
models identified

229
technologies identified

57
vulnerabilites identified

The Challenge

Unmanaged third-party AI risks amidst a rapidly growing vendor ecosystem.

Large organizations rely on third-party vendors for services like customer support and analytics, many of which now incorporate AI technologies. However, the opaque integration and deployment of AI by these vendors create significant risks.

This lack of transparency can lead to unmanaged AI systems operating outside governance frameworks, causing compliance gaps and exposing systems to security threats. Vendors may also misjudge the regulatory requirements of their AI, such as international standards, while security vulnerabilities further heighten risks.

To address these challenges and maintain AI compliance, the organization prioritized visibility and control over vendors' AI systems, partnering with a solution provider for a comprehensive approach.

The Solution

Developing an inventory of vendor AI systems with Cranium CodeSensor

The organization first used the Cranium platform to identify vulnerabilities and compliance gaps in AI systems across 15 vendors. They requested AI Bills of Materials (AI BOMs), which detailed technologies, models, datasets, and infrastructure. With Cranium CodeSensor, vendors uncovered **54** models, **229** technologies, and **57** vulnerabilities, ensuring quick and accurate system assessments.

"Set up an integration and generated a Bill of Materials with CodeSensor in just 30 minutes—no support needed."

Streamlining compliance check with Cranium's AI Card

Each vendors self-attested to a NIST AI Risk Management Framework Starter Pack, generating a Cranium AI Card — a "digital passport" for their AI systems, providing detailed insights into development, deployment, and compliance status. This card could then be shared with the organization, along with additional partners, customers, and stakeholders.

Conducting in-depth AI model penetration testing in the Cranium AI Arena

Advanced security penetration testing of vendor models in Cranium's AI Arena revealed **57** vulnerabilities. Cranium's insights helped the organization visualize, categorize, and prioritize risks, enabling targeted mitigation efforts.

For instance, a vendor's sentiment analysis model was vulnerable to inference attacks, risking predictive accuracy. Cranium addressed this with a suite of targeted recommendations:

- Implementing robust access control measures
- Applying rate limits to model queries
- Conducting input validation and adversarial training
- Using model ensembles to improve security and reliability

Collaborating to develop industry-wide AI risk standards

In partnership with internal teams and external vendors, the organization spearheaded the development of a tailored security standard for third-party AI systems within the financial services sector. This standard was designed to align with key industry regulations, including the EU AI Act, while also setting a benchmark for other organizations within the sector to follow. By collaborating closely with its vendors, the organization ensured that this security framework would serve as a model for improving the safety and compliance of third-party AI systems.

The Result

Improved vendor AI security, compliance, and trust

The organization identified high-risk AI solutions within its vendor ecosystem, highlighting the need to address visibility gaps, mitigate vulnerabilities, and ensure regulatory compliance. Vendors welcomed the initiative as a valuable framework to enhance AI visibility, streamline compliance, and align with broader organizational needs.

"We're thrilled to see [company] leading this essential work for the industry and are excited to contribute to and support their innovative efforts."