Cobalt

CASE STUDY

Zest AI secures its AI lending intelligence assistant with Cobalt

THE CHALLENGE

As a leader in credit risk assessment, Zest AI empowers financial institutions to expand access to credit for underserved populations. While Zest AI has had a longstanding relationship with Cobalt for web application and API pentests, a new and significant security challenge emerged with the development of LuLu, their lending intelligence assistant.

LuLu allows non-technical financial executives to analyze complex data through natural language queries. The novel architecture and potential vulnerabilities inherent in this LLM-powered application necessitated specialized pentesting. Although Zest AI possessed strong internal security expertise, they needed a security partner capable of assessing the unique attack surface of LuLu to ensure its resilience, maintain their high security standards, and continue providing assurance to their customers and auditors.

THE SOLUTION

Zest AI turned to their trusted pentesting partner, Cobalt, to test LuLu. Waylan Wong, Cloud Engineering, Sr. Manager at Zest AI said: “We’ve been working with Cobalt for years before they started offering LLM pentesting. We knew Cobalt was committed to staying at the forefront of industry best practices, so when we started developing LuLu, it was a no-brainer to continue to work with them.” Cobalt followed our proprietary methodology based on the OWASP Top 10 for LLM applications. The detailed testing methodology and actionable

ZEST AI

ABOUT CUSTOMER

Zest AI is a pioneer in AI lending innovation with a mission to broaden access to lending using smarter, more efficient AI. Zest AI is a proven partner to banks, credit unions, and specialty lenders, helping them decrease risk and boost growth opportunities throughout the lending process.

INDUSTRY

FinTech

SIZE

150+ employees

HEADQUARTERS

Burbank, California, US

COBALT SERVICES

AI and LLM application pentest

BY THE NUMBERS

Remediated 100% of findings in 2 week

> “We’ve been working with Cobalt for years before they started offering LLM pentesting. We knew Cobalt was committed to staying at the forefront of industry best practices, so when we started developing LuLu, it was a no-brainer to continue to work with them.”
>
> **WAYLAN WONG,** CLOUD ENGINEERING, SR. MANAGER AT ZEST AI

recommendations provided in the reports offered significant reassurance to Zest AI's executives and customers that LuLu was being rigorously evaluated against industry-leading security standards.

THE OUTCOME

The pentests for LuLu focused on a tailored subset of the OWASP Top 10 for LLMs relevant to its functionality. The pentest uncovered an indirect prompt injection which allowed data exfiltration via hidden markdown images, which would have allowed attackers to exfiltrate sensitive chat data to external domains if users were tricked into executing the malicious prompt. This vulnerability was quickly remediated by disabling automatic image rendering and implementing an image allowlist. Cobalt pentesters also validated a strong baseline of positive security controls in place for web applications, confirming the application's resilience against common attacks like XSS, SQL injection, and unauthorized API action, while noting strong input sanitization and session management.

The detailed findings provided valuable insights that led to broader improvements. Waylan noted: "The reports were incredibly clear, showing us exactly what was tested and the results. That level of detail made it easy to work with our data science team to quickly resolve vulnerabilities and help us understand our own application better." These insights contributed to UI adjustments and improved response generation methods for LuLu. Another significant outcome was an update to the customer contractual obligation which clarified the intended use and expectations for LuLu.

The overall speed and efficiency of the testing, coupled with the rapid resolution of findings and the understanding of existing security strengths, allowed Zest AI to confidently advance LuLu from its beta phase to market, reinforcing their commitment to secure and innovative financial technology.

> "The reports were incredibly clear, showing us exactly what was tested and the results. That level of detail made it easy to work with our data science team to quickly resolve vulnerabilities and help us understand our own application better."
>
> **WAYLAN WONG,** CLOUD ENGINEERING, SR. MANAGER AT ZEST AI

To learn more about what Cobalt can do for your organization, visit www.cobalt.io/get-started

Cobalt