AppSecAI

AppSecAI Expert Triage Automation

Benchmark Report

SAST Scanner: CodeQL

Report Date: March 7, 2025

Executive Summary

Static Application Security Testing (SAST) tools are integral to identifying vulnerabilities. However, results must be manually triaged to remove false positives, a slow and expensive process that impedes application security, development and delivery. AppSecAI Expert Triage Automation (ETA) automates triage to lower manual triage time, tedium, and costs. This report evaluates AppSecAI and CodeQL against the open source OWASP benchmark with over 2,700 vulnerabilities. ETA analyzed a total of 4,161 findings in this assessment.

98.2%

CodeQL + AppSecAI Triage Accuracy

99.8%

CodeQL + AppSecAI True Positive Accuracy

1.7%

AppSecAI False Positive Rate

Metric	CodeQL	CodeQL + AppSecAI
Accuracy	61.9%	98.2%
True Positive Accuracy	90.2%	99.9%
False Positive Rate	43.3%	1.7%
False Positive Findings	1,801	72

AppSecAI Benefit

96.0%

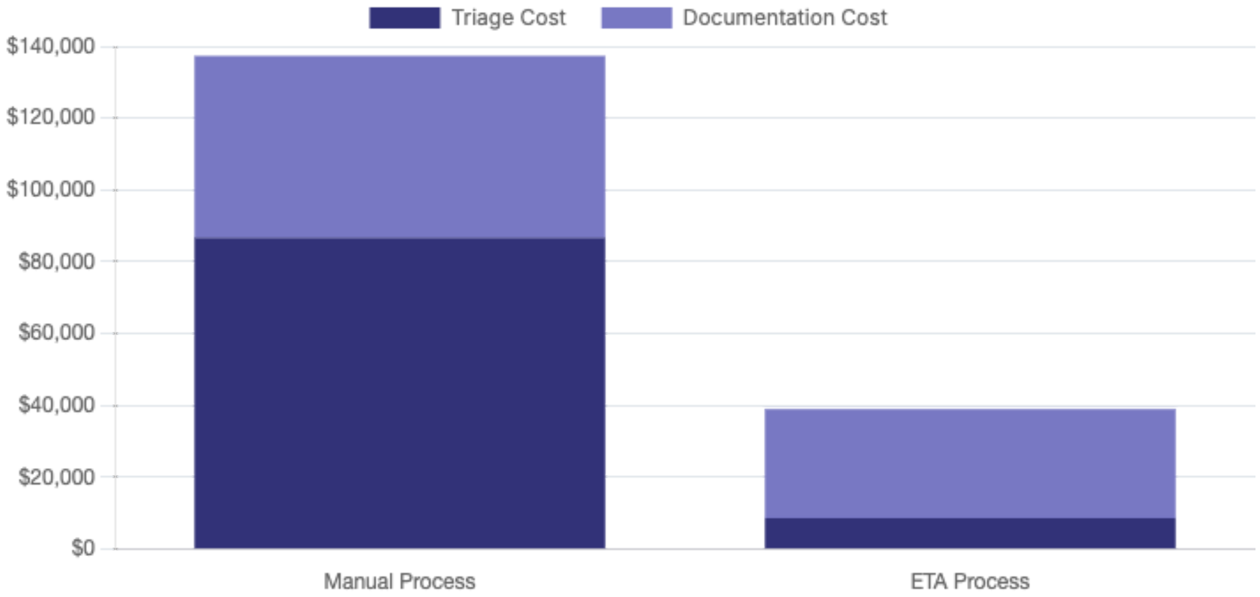
False Positive Reduction

1,729

False Positives Reduced

Financial Analysis

Cost Analysis Summary: Implementation of CodeQL + AppSecAI demonstrates significant cost savings across all security assessment processes. The total cost reduction is $98,268.75 (71.6%), with notable improvements in both triage and documentation costs. This substantial decrease in operational expenses enables security teams to process vulnerability assessments more efficiently while maintaining high accuracy.



Cost Rates

Developer Cost: $175/hour

AppSec Analyst Cost: $250/hour

Time Metrics

Manual Triage: 5 min/vuln

Documentation: 5 min/vuln

ETA Triage: 0.5 min/vuln

Appendix: Glossary of Terms

This glossary provides definitions for key terms used throughout the report to ensure clear understanding of the report metrics.

AppSecAI ETA (Expert Triage Automation)

An AI-powered solution that automates and accelerates vulnerability triage by removing false positives and documenting true positives.

Triage

The process of reviewing, categorizing, and prioritizing security findings to determine their validity and impact.

Accuracy

The percentage of findings that are reported correctly (whether in the OWASP Benchmark or not).

True Positives

Real Vulnerabilities (whether in the OWASP Benchmark or not) that were reported as vulnerable.

False Positives

Not Vulnerable Test Cases (whether in the OWASP Benchmark or not) that were reported as vulnerable.

True Negatives

Not Vulnerable Test Cases in the OWASP Benchmark that were also not Reported by the SAST Tool.

False Negatives

Not Vulnerable Test Cases (whether in the OWASP Benchmark or not) that were reported as vulnerable by the SAST tool.