AppSecAI Contributes Python Benchmark to OWASP - Advances Metric-Based Security Testing

Evaluates Python SAST, DAST, IAST and LLM-based security tools that power AI development and vibe coding

AppSecAI - Application Security Fix Automation

LOS ALTOS, CA, UNITED STATES, November 6, 2025 /EINPresswire.com/ -- AppSecAI today announced its contribution of a Python version of the Open Web Application Security Project (OWASP) Benchmark to the open-source community, developed in collaboration with David Wichers, creator of the original Java OWASP Benchmark and 15-year co-leader of the OWASP Top 10 project. This contribution brings to the Python ecosystem the same gold-standard framework that has defined objective measurement of Static, Dynamic, and Interactive Application Security Testing (SAST/DAST/IAST) tools since 2015—empowering developers and security teams to accurately evaluate the accuracy, coverage, speed, and effectiveness of their vulnerability detection and triage automation tools.

"We created the original OWASP Benchmark because the security industry needed to move beyond subjective claims to objective measurement," said David Wichers, creator of the OWASP Benchmark and former OWASP Top 10 project leader. “Python's role in AI development and modern applications makes this standardized testing framework critical. Security teams deserve the transparency to understand what their software vulnerability detection tools actually deliver."

Addressing Critical Gap as Python Powers AI/ML, Vibe Coding, and Enterprise Applications

With Python powering critical applications across AI/ML, web development, vibe coding

environments, and enterprise systems, the lack of a standardized vulnerability detection benchmark has made it difficult to objectively compare software vulnerability detection tool performance. This contribution gives organizations data-driven insights to make informed decisions about their Python security tools.

Application security testing has become more critical as AI-accelerated development creates new challenges. A 2024 study published on arXiv found that 24.7% of AI-generated code contains security vulnerabilities. Meanwhile, BlackDuck's 2025 Global State of DevSecOps report found that 57% of organizations say AI coding assistants have introduced new security risks or made it harder to detect issues. In the same study, 16.28% of organizations cited "AI introducing vulnerabilities at scale and speeds that exceed AppSec capacity" as their primary security concern.

Thousands of Exploitable Test Cases Validate Both Traditional and AI-Powered Security Tools

The Python Benchmark includes over 1,000 comprehensive test cases spanning major vulnerability categories including SQL Injection, Cross Site Scripting (XSS), Command Injection, Path Traversal, Weak Cryptography, and more—each mapped to specific Common Weakness Enumeration (CWEs). Every vulnerability is actually exploitable, ensuring fair evaluation of any SAST, DAST, IAST, or LLM-based security product.

The benchmark also includes false positives that often fool security tools. This dual focus measures tools' ability to accurately identify real vulnerabilities while filtering out false alarms - a critical capability as organizations face vulnerability alerts at volumes that exceed traditional manual triage capacity.

In the coming weeks, AppSecAI will contribute more test cases that cover additional CWEs.

Measuring AI Solutions' Effectiveness at Eliminating False Positives

Beyond evaluating traditional security scanners, the Python Benchmark provides a standardized framework for measuring how effectively AI-powered security solutions eliminate false positives. As organizations increasingly deploy AI-based triage and remediation tools, objective benchmarks enable data-driven comparison of these solutions' accuracy in distinguishing true vulnerabilities from false alarms.

This capability addresses a critical need in the application security industry, where BlackDuck's 2025 Future of Application Security report found that 71% of organizations say a significant portion of their security alerts are noise: false positives or duplicate findings from different tools. The OWASP Benchmark enables organizations to validate vendor claims about AI-powered false positive reduction and new automated triage capabilities with reproducible, objective data.

Evidence-Based Security: Objective Metrics Replace Vendor Marketing Claims

The Python Benchmark follows OWASP's mission to make application security visible, giving security teams the transparency needed to understand tool strengths and weaknesses rather than relying solely on vendor marketing claims. This approach enables security teams to validate tool effectiveness, justify investments, and demonstrate improvement over time.

Without objective measurement, security teams cannot make informed decisions about their security tools. This benchmark enables data-driven decision-making backed by reproducible results that any organization can verify.

The Python OWASP Benchmark is available as open-source software at https://owasp.org/www-project-benchmark/About AppSecAI

About AppSecAI

AppSecAI transforms application security through AI-powered automation, enabling organizations to secure applications at portfolio scale through automated vulnerability triage and remediation. AppSecAI's results-based pricing model charges per vulnerability actually fixed versus $5,000-$20,000 for manual remediation processes. Founded by industry veterans and backed by application security experts, we combine existing application security tools with advanced AI technologies to deliver efficient, accurate, and scalable application security solutions. Learn more at www.appsecai.io.

About OWASP Benchmark

The OWASP Benchmark Project is a test suite designed to evaluate the accuracy, coverage, and speed of automated software vulnerability detection tools. Created by Dave Wichers, the benchmark provides thousands of test cases with known vulnerabilities to enable objective comparison of security testing tools. https://owasp.org/www-project-benchmark/.
Media Contact

Kira Wojack
Merritt & Rose Communications
Kira@MerrittAndRose.com
+1 415 419-4062

Kira Wojack
Merritt & Rose Communications
+1 415-419-4062
email us here

This press release can be viewed online at: https://www.einpresswire.com/article/864658653

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors

try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.
© 1995-2025 Newsmatics Inc. All Right Reserved.