Discussion of AI policy to be implemented in the OWASP Project Policy, or as a standalone policy

Reasoning.

AI tools are already used without any governance, but we want to avoid AI slop and low effort contributions, especially to projects like the Cheat Sheet series. Lastly, we want to avoid AI tools being used to submit mass submissions of low effort, low quality security reports to our new VDP. There are two different points, that may mean we want a separate policy to cover all contributors and not just project leaders.

DRAFT Policy

Use of Artificial Intelligence (AI) Tools in projects

OWASP projects may use AI-assisted or generative AI tools to support the creation of project materials, including but not limited to documentation, code examples, graphics, and written content. The use of such tools is optional and at the discretion of the project contributors.

Regardless of whether AI tools are used, **the Project Leader and Contributors bear full responsibility for ensuring the accuracy, quality, security, and appropriateness of all AI-generated or AI-assisted content prior to its inclusion in any OWASP project deliverable**. This includes verifying factual correctness, removing unintended bias, ensuring compliance with OWASP standards, and confirming that no copyright-protected material has been inadvertently generated or incorporated.

AI-generated content **must not** be included in OWASP projects without human review and validation. The Project Leader is accountable for maintaining the integrity and quality of the project's outputs when AI tools are utilized. Where provenance of ideas is required, such as international standards or where a guarantee of original content is required, AI generated content may not be appropriate as it doesn't necessarily disclose its data training set where the AI generated the content.

–

The use of AI generated or discovered security vulnerabilities

Some projects, including curl, have withdrawn support for their vulnerability disclosure program due to the waste of time vetting and validating low effort AI submitted bugs. We need to have something that doesn't abuse both Foundation staff and volunteer time to project against such submissions

DRAFT policy on VDP submissions

AI-Generated Vulnerability Submissions Clause

OWASP expressly prohibits the use of AI tools to generate low-quality, hallucinated, or unverified security vulnerability reports for submission to OWASP's Vulnerability Disclosure Program (VDP). Submissions generated wholly or partially by AI **must be thoroughly validated by the submitter** to ensure the issues reported are real, reproducible, and technically sound.

OWASP has observed industry-wide evidence that AI-generated "slop" vulnerabilities place an unsustainable burden on open-source maintainers and security teams. Recent reporting shows that the cURL project cancelled its bug bounty program after being overwhelmed by large volumes of low-quality AI-generated submissions, many of which reported non-existent vulnerabilities and consumed significant maintainer time. To avoid similar harm, OWASP may reject, suspend, or permanently ban contributors who submit fabricated, low-effort, or AI-generated vulnerability reports that waste volunteer and Foundation resources. [bleepingcomputer.com], [arstechnica.com]

The burden of verification lies entirely with the submitter.

3. Events and Training AI policy

A common theme we are starting to see is low effort AI generated content being submitted. Weeding out these unoriginal and low effort talks takes a fair amount of time and effort by the Events team, CFP volunteers, and event organizers.

DRAFT policy for discussion

AI-Assisted Training, Talks, and Educational Content Clause

OWASP prides itself on running high value, research-led, original talks and training in its events and training days. Speakers and trainers can use AI tools to assist them in creating images, videos, music, decks, training materials, and labs, but they remain wholly responsible for the content. OWASP has always had a very high bar for submissions. AI-led talks and training are very unlikely to be accepted. In recent years, we have seen the rise of many low quality submissions. To avoid more being submitted, which wastes CFP and CFT volunteer time, and if accepted, audience time, OWASP is implementing an AI policy around events and training.

Contributors who use AI tools to generate any portion of OWASP talks, training materials, workshops, slide decks, demonstrations, or other educational content **retain full responsibility for the accuracy, validity, and quality of all claims and instructional materials** included in their submissions.

All AI-assisted content must be carefully reviewed, corrected, and validated by the submitter before being provided to OWASP. The OWASP Foundation is under **no obligation** to accept training sessions, talks, or presentations that appear to be AI-generated, low-quality, misleading, or factually inaccurate. OWASP may decline such submissions at any stage - initial review, planning, or post-submission - without further obligation to the submitter.

This requirement protects the integrity and credibility of OWASP events, prevents the spread of inaccurate security guidance, and ensures that content maintains OWASP’s standards of technical rigor and educational value.