**AI-Driven Penetration Testing (GHOSTCREW) Assessment Event**

**07 May - 09 May 2025**

**Submit NLT 09 April 2025 at or before 11:59 PM ET**

**U.S. Citizens Only**

Purpose  
The Cyber Fusion Innovation Center (CFIC), in collaboration with U.S. Army Cyber Command (ARCYBER) and Army Cyber Technology and Innovation Center (ArCTIC), invites qualified industry partners to submit proposals for the development of GHOSTCREW, an AI-driven platform designed to accelerate experience accumulation for Red Team/Pentest operators. This RFP outlines the requirements for a prototype system capable of accelerating operator experience accumulation by providing numerous recommended courses of action and simulating operator decisions prior to execution in the live environment thus exposing the operator to more learning during a single operation than would occur without the AI assistance and in-stride simulated outcomes.

Background/Synopsis  
This Request for Proposal (RFP) solicits innovative solutions via a whitepaper submission in Vulcan for the development of GHOSTCREW, an AI-driven, operational Red Team platform designed to significantly accelerate the experience accumulation of Red Team/Pentest operators. By providing real-time, AI-assisted decision support, simulated tool use outcomes and a centralized repository of lessons learned, GHOSTCREW aims to bridge the gap between foundational training and practical experience, ultimately enhancing the effectiveness and efficiency of Red Team operations using an AI assisted platform during real-world red team operations.

A significant challenge facing our organization is conferring experience to our Red Team/Pentest personnel. While foundational training is readily available, practical experience is primarily gained through live operational missions, which are costly and time-consuming. GHOSTCREW aims to address this challenge by providing an AI assisted platform where operators can conduct operational missions while receiving in-stride assistance, experiment with different attack vectors and learn from previous operations during live operations. The GHOSTCREW platform will accelerate experience accumulation using AI assisted decision making and simulated course of action outcome feedback built directly into the tool chain and user interface of the Red Team Operator.

Key Capabilities

The proposed GHOSTCREW platform should possess the following key capabilities within the context of Red Team pentesting operations within a conventional IT environment:

1. AI-Assisted Visualization and Decision Making: Provide operators with real-time, AI-driven visualization and corresponding recommendations for Courses of Action (COAs) selection during penetration testing operations. These recommendations should be context-aware and tailored to the operator's experience level, considering factors such as stealth, impact, and risk. They should also factor in previous operational results and relevant aspects of commercial threat intelligence (CTI).
2. Simulated Attack Outcomes: The platform should provide likely outcomes to specific tool/techniques/tactic use cases for the purpose of enabling higher level course of action selection by the operator. For example, if a given tool is not successful in achieving the intended outcome without generating adverse effects, the operator should be warned accordingly. The simulated outcomes can be the result of actual simulations where tools are used in a localized, virtual range or from analysis of previous operational outcomes, CTI and related data.
3. Threat Intel Integration: Ingest and integration with threat intelligence feeds to update the AI solution regularly. This will ensure that the platform remains current with the latest threats and attack methodologies plus improved COA recommendations.
4. Risk Assessment: Provide accurate risk assessments for tool/COA/action recommendations, considering the operator's level of experience and the potential impact of their actions.
5. Centralized Lessons Learned: Maintain a centralized repository of log events and lessons learned from past red team operations, enabling operators to leverage lessons learned through Retrieval-Augmented Generation, reinforcement learning or similar supported method while allowing operators to access and learn from the experiences of other operators. This should include maintaining up-to-date target knowledge pertaining to specific IT services or systems.
6. Efficient Resource Usage: The platform should be designed to minimize resource consumption and operate efficiently while leveraging the necessary amount of GPU/Compute support required to support the onboard AI subsystem. We desire as self-contained of a system as possible but look forward to understanding innovative solutions to supply sufficient GPU/Compute needed to effectively implement the capability in as offline a capacity as possible (not closing the door on reach back support for GPU compute).
7. Quick State Updates: The platform should be able to provide quick, granular, system and operational state updates to the user in an intuitive way that enables the operator to have a good situational understanding of the current operation.
8. Connectivity and Minimal Operational Impact: The platform should be designed to minimize any potential resource impact on the systems being red teamed and are expected to connect via standard RJ45 ethernet connectors or virtual networking interfaces and should possess limited physical dependances on the system being penetration tested.
9. Error Handling & Recovery: The platform must have robust error handling and recovery mechanisms to handle unexpected situations, recover from failures, maintain system stability, protect operational data, and ensure a consistent state.

Deliverables

The contractor shall provide the following deliverables:

1. A functional prototype of the GHOSTCREW platform.
2. Documentation, including user manual, installation/configuration guides, and relevant API documentation needed for extensibility.
3. A detailed report outlining the design and implementation of the AI algorithms, reinforcement learning methodology used in the platform (i.e. RAG, Vector Search, conventional indexed Search, etc.).
4. A demonstration of the platform's capabilities and performance.

Assessment Criteria

The proposed GHOSTCREW prototype will be evaluated based on the following criteria, with each criteria item scored on a scale of 1-5 where: 1 = Unsatisfactory; 2 = Marginal; 3 = Satisfactory; 4 = Good; 5 = Excellent

1. Functionality:
   1. Criterion: The platform successfully implements the key capabilities outlined in the RFP Section 3.0 (AI-Assisted Decision Making, Simulated Attack outcomes, Lessons Learned and threat Intel Integration, etc.).
   2. Passing Level: The prototype demonstrably performs most core functions with at least basic competency. Minor limitations are acceptable, but critical features must be operational.
2. AI Performance (Decision Support):
   1. Criterion: The AI component provides relevant and accurate recommendations for Courses of Action (COAs) during simulated attacks.
   2. Passing Level: In a standardized test scenario, the AI's suggested COAs are judged by experienced Red Teamers and AI Subject Matter Experts (SMEs) to be functionally/tactically sound and aligned with realistic attack methodologies at least 75% of the time.
3. Accuracy of Simulated Outcome information:
   1. Criterion: The platform can accurately indicate probable outcome of penetration test tool use given a specific target.
   2. Passing Level: Subject matter experts judge the simulated outcome to accurately reflect realistic penetration test results at least 75% of the time.
4. Threat Intelligence and Lessons Learned Integration:
   1. Criterion: The platform effectively integrates with lessons learned and threat intelligence feeds while incorporating up-to-date information on vulnerabilities, exploits, and TTPs.
   2. Passing Level: The prototype demonstrates the ability to ingest lessons learned and threat data from at least five lesson learned documents plus five standard commercial threat intelligence reports and use that information to inform operators in successive operations. Examples here are using some type of reinforcement learning method to inform the operator AI model and/or simulated attack scenarios.
5. Usability and User Interface:
   1. Criterion: The platform's user interface is intuitive, easy to navigate, and supports efficient interaction for both trainees and instructors.
   2. Passing Level: Novice users can complete basic tasks within the platform with minimal guidance. Experienced Red Teamers find the interface to be conducive to their workflow and provide positive feedback on its usability.
6. Scalability and Performance:
   1. Criterion: The platform demonstrates the ability to handle a realistic number of tasks such as attack simulations and log data volume without significant performance degradation.
   2. Passing Level: The prototype maintains acceptable responsiveness (defined by specific benchmarks) during a load test simulating a realistic number of users and operations.
7. Security of the Platform:
   1. Criterion: The GHOSTCREW platform itself is designed with robust security measures to prevent unauthorized access, data breaches, and misuse of the system.
   2. Passing Level: A third-party penetration testing effort finds limited critical vulnerabilities in the platform's security posture consistent with prototype level capabilities. Key standard security practices including user authentication, authorization and data encryption are implemented.
8. Lessons Learned & Knowledge Management:
   1. Criterion: The platform effectively captures, stores, and presents lessons learned from simulated attacks, commercial threat intel and other sources making this knowledge accessible to users for future operations and improvement.
   2. Passing Level: The prototype demonstrates a clear mechanism for recording key takeaways from user actions, simulated attacks (successful and unsuccessful) and AI recommendations. This information is easily retrievable and presented in a digestible format.
9. Documentation and Training Materials:
   1. Criterion: The vendor provides documentation (user manual, installation and configuration guide, API docs) to support the platform's deployment, use, and maintenance.
   2. Passing Level: Documentation is judged to be complete, accurate, well-organized, and easy to understand by specified users (system administrators, instructors, and red team operators).
10. Vendor Support and Collaboration:
    1. Criterion: The vendor demonstrates a willingness and ability to work collaboratively with the government team throughout the prototype development and assessment process.
    2. Passing Level: The vendor is responsive to inquiries, proactive in communication, and open to incorporating feedback from government stakeholders.
11. Functionality: The extent to which the platform meets the key capabilities outlined above.
    1. AI Performance: The accuracy, relevance, and effectiveness of the AI-assisted decision-making capabilities. This will include reinforcement learning, verification and validation regarding the non-determination of LLM data relevancy.
    2. Usability: The ease of use and intuitiveness of the platform's user interface.
    3. Scalability: The ability of the platform to perform sufficient number of the tasks listed above to complete a penetration test mission on a small to medium sized conventional IT network consisting of switches, routers, Windows and Linux based hosts and servers.
    4. Security: The security of the platform and its ability to protect sensitive data.
    5. Innovation: The novelty and originality of the proposed solution.
    6. Cost-Effectiveness: The overall cost of the proposed solution, considering its capabilities and benefits.

Whitepaper Submission Information

Interested vendors are requested to submit a proposal through Vulcan that includes the following information:

* A detailed description of the proposed GHOSTCREW platform
* A technical architecture diagram of the platform
* A description of the AI algorithms used in the platform
* A plan for connecting the platform to the penetration test target environment
* A proposed timeline for development and deployment
* A cost estimate
* Company qualifications and experience

What is a Virtual Assessment Event?  
ARCYBER submits specific problems to solve and CFIC utilizes its ecosystem and market research to attract best-of-breed submissions to solve the problem. ARCYBER subsequently reviews and selects participants with the highest value to present and/or demonstrate their capability in a private, virtual pitch / demo session hosted through an online platform with Government Stakeholders on a specific date.

Why Should You Participate?  
ARCYBER seeks to enter into non-FAR or FAR-based agreements with Industry, Academic, and National Lab partners whose solutions are favorably evaluated by ARCYBER Subject Matter Experts. As such, these events are considered competitive in the same manner as a Broad Agency Announcement (BAA) or Commercial Solutions Opening (CSO), and solutions will be evaluated independently of one another primarily for technical merit.

Timeline  
Phase 1: 12 February 2025 – Collaboration Event  
Industry, academia, and government partners met in-person and identified current limitations and ideated ways to overcome limiting factors. Problem statements were developed to frame AI-driven penetration testing software needs.

**Phase 2: 05 March - 09 April 2025 – Submission Window / Q&A Session**  
Industry, academia, and government partners MUST submit a 6-page submission response whitepaper document through this link – [**CLICK HERE**](https://www.vulcan-sof.com/login/ng2/submission?collectionUuid=acd7d0ed-109d-49c7-bbc0-f36f541383a6). Submit NLT (No Later Than) 11:59 p.m. Eastern on Wednesday, 09 April 2025.

**Q&A Response Session:** Fill out [**this form**](https://forms.monday.com/forms/7443d87379a8859dcb2dff0531e5c0e2?r=use1) to submit questions related to this release which can be addressed by the Principal Stakeholder. The form will be open until 12:00pm (noon) Eastern on 19 MAR. Thereafter, a virtual Response Session will be hosted on 20 MAR.

**Phase 3: 10 April - 24 April 2025 – Submissions Reviewed & Downselects**  
Submitted documents will be placed in review based on rubric parameters identified through the requirements and criteria included in this document. Submissions will be reviewed by the Assessment Team and subsequently graded for potential down selection. The Assessment Team will conduct a final meeting to review graded whitepaper documents for down-selects of those respondents/submissions they feel have the highest potential to satisfy their needs. Favorably evaluated submissions will receive an invitation to virtually present.

**Phase 4: 07 May - 09 May 2025 – Virtual Demonstration Presentations**    
Demonstration presentations will be held virtually through MS Teams. During the time allotted, each presenting entity will demo their solution and capabilities that should satisfy, to the best of their ability, the requirements and criteria noted in this release. Each group will be notified of their presentation time.

**Phase 5: Follow-On**  
The Government has several acquisition/contract vehicles (i.e., various IDIQ contracts, purchase orders, etc.) to potentially purchase the final solution if operationally viable. The Government may elect to purchase all, some, and/or none of the solutions from Phase 5 for further projects. Once awarded, contracting with the awardee company is slated to begin on or near 19 May 2025.

How You Can Participate  
Submissions will be accepted through the Vulcan platform – [**CLICK HERE TO SUBMIT**](https://www.vulcan-sof.com/login/ng2/submission?collectionUuid=acd7d0ed-109d-49c7-bbc0-f36f541383a6).  
Whitepaper submission format and template are on our website – [**CLICK HERE TO LEARN MORE & DOWNLOAD**](https://www.cyberfic.org/ghostcrew).  
Vulcan submission instruction sheet – [**DOWNLOAD THE PDF**](https://6541d37e-a8b0-42c8-9e53-5676469b2638.usrfiles.com/ugd/6541d3_cd8e737edd2f4aed946087cf30a02be3.pdf).

Interested in **teaming opportunities** with other companies? Fill out [**this form**](https://forms.monday.com/forms/9f22147bc985858a74e7bc764bf7a417?r=use1) NLT 5:00pm Eastern on 24 MAR to submit your capabilities and information will be collated and provided to you that was submitted by other companies interested, as well. All information provided will be sent out on 26 MAR.

To learn more about how to use Vulcan, please refer to this information on the CFIC website - https://www.cyberfic.org/joinvulcan.

Questions?  
For submission-related questions, please contact Brandon Sizemore at [bsizemore@cyberfic.org](mailto:bsizemore@cyberfic.org) and Amanda Green at [agreen@cyberfic.org](mailto:agreen@cyberfic.org).

DISCLAIMERS:An award under 10 U. S. Code, Section 2371b may result in award of a follow-on production in accordance with 10.U.S.C. 2371(f). Upon determination that the competitively awarded prototype project(s) has been successfully completed, and subject to the availability of funds, the prototype project(s) may result in the award of a follow-on production contract or transaction without the use of competitive procedures. Such awards may include multiple phases.

Non‐Government advisors may be used in the evaluation of submissions and will have signed Non‐Disclosure Agreements (NDAs) with the Government. The Government understands the information provided in this announcement is presented in confidence and may contain trade secret or commercial or financial information and agrees to protect such information from unauthorized disclosure to the maximum extent permitted and as required by law. An organization's participation in any part of the selection process under this announcement indicates concurrence with the aforementioned use of contractor support personnel.