TÀI LIỆU HƯỚNG DẪN • AN NINH MẠNG & AI

CẨM NANG CÀI ĐẶT & SỬ DỤNGPENTAGI BY VXCONTROL

AI Agent Tự Động Hóa Pentest • Lập Kế Hoạch Đa Tầng • Tự Khai Thác & Báo Cáo Lỗ Hổng

⭐ **GitHub Repository:** https://github.com/vxcontrol/pentagi **(~3.800+ Stars)**
👑 **Tác giả / Tổ chức: vxcontrol (Cộng đồng An toàn thông tin)**
Giấy phép: Apache-2.0 (Mã nguồn mở hoàn toàn)
Công nghệ lõi: Docker Compose, Langfuse / Graphiti, Python Backend, React UI
Mô hình hỗ trợ: OpenAI, Anthropic Claude, Ollama (Local), vLLM

1. PENTAGI LÀ GÌ & TẠI SAO ĐÁNG CHÚ Ý?

PentAGI là hệ thống AI Agent mã nguồn mở đột phá, được xây dựng chuyên biệt cho tác vụ kiểm thử thâm nhập (Penetration Testing) tự hành. Thay vì chỉ chạy các bộ quét bảo mật thụ động như Nessus hay Nikto, PentAGI sở hữu khả năng tự lên chiến lược tấn công, phân tích bối cảnh, thử nghiệm khai thác (exploit) và tổng hợp báo cáo chi tiết.

- **Hai chế độ linh hoạt (Automation vs Assistant):** Chế độ Automation cho phép agent tự động thực thi toàn diện theo mục tiêu được giao; Chế độ Assistant cho phép chuyên viên an ninh duyệt từng lệnh hoặc can thiệp bước tiếp theo.
- **Tích hợp công cụ an ninh hàng đầu:** Nmap, SQLmap, Metasploit, Gobuster, Nikto, ffuf, hydra... đều được container hóa và điều khiển thông qua AI Tool-Use API.
- **Tự lập kế hoạch tấn công (Attack Graph):** Sử dụng đồ thị tri thức (Graphiti/Neo4j) để ghi nhớ đường đi mạng, dịch vụ mở và các lỗ hổng đã xác nhận, tránh lặp lại các thử nghiệm vô ích.
- **Báo cáo chuẩn quốc tế:** Tự động phân loại lỗ hổng theo thang điểm CVSS v3.1, trích xuất bằng chứng (PoC) và đề xuất phương án vá lỗi cho kỹ sư DevOps.

2. YÊU CẦU HỆ THỐNG & ĐIỀU KIỆN TIÊN QUYẾT

Thành phần	Yêu cầu tối thiểu	Khuyến nghị sản xuất
Hệ điều hành	Linux (Ubuntu 22.04+), WSL2, macOS	Ubuntu Server 24.04 LTS x64
CPU / RAM	4 Cores CPU / 8 GB RAM	8 Cores CPU / 16 GB - 32 GB RAM
Ổ cứng	30 GB dung lượng trống	50 GB SSD NVMe

Phần mềm bắt buộc	Docker Engine 24+, Docker Compose v2	Docker Engine 26+, Podman rootless (tùy chọn)
API Key / LLM	Tối thiểu 1 API key (OpenAI/Anthropic)	OpenAI GPT-4o / Claude 3.5 Sonnet hoặc Ollama 70B local

3. HƯỚNG DẪN CÀI ĐẶT CHI TIẾT QUA DOCKER COMPOSE

Phương pháp cài đặt chuẩn và an toàn nhất cho PentAGI là sử dụng Docker Compose, giúp cô lập môi trường thực thi và bảo đảm an toàn cho máy chủ gốc:

Bước 1: Clone kho mã nguồn chính thức

```
git clone https://github.com/vxcontrol/pentagi.git
cd pentagi
```

Bước 2: Cấu hình biến môi trường (.env)

```
cp .env.example .env
# Mở file .env và điền API key mô hình ngôn ngữ của bạn
nano .env
```

Mẫu cấu hình tối thiểu trong file .env

```
# Cấu hình cổng và domain
PENTAGI_LISTEN_IP=0.0.0.0
PENTAGI_PORT=8080
PUBLIC_URL=http://localhost:8080

# LLM Provider (Chọn 1 hoặc nhiều nhà cung cấp)
OPENAI_API_KEY=sk-...
ANTHROPIC_API_KEY=sk-ant-...

# Bảo mật hệ thống
PENTAGI_JWT_SECRET=thay_the_bang_chuoi_ngau_nhien_rat_dai_32_ky_tu
```

Bước 3: Khởi động toàn bộ cụm container

```
docker compose up -d

# Xem tiến trình khởi chạy và nhật ký container
docker compose logs -f
```

4. HƯỚNG DẪN THỰC THI KIỂM THỬ AN NINH ĐẦU TIÊN

Sau khi cụm container khởi động thành công, hãy truy cập giao diện quản trị tại địa chỉ: http://localhost:8080

- **Bước 1:** Tạo tài khoản quản trị Admin đầu tiên và đăng nhập vào Dashboard.

- **Bước 2:** Chọn 'New Pentest Flow' -> Đặt tên chiến dịch (ví dụ: 'Kiểm thử máy chủ thử nghiệm web nội bộ').
- **Bước 3:** Xác định rõ phạm vi (Scope): Nhập IP/Domain đích (ví dụ: 192.168.1.100 hoặc demo.testfire.net) và phạm vi bị cấm đụng đến.
- **Bước 4:** Chọn chế độ 'Assistant Mode' cho lần chạy đầu để quan sát từng lệnh AI đề xuất trước khi bấm 'Approve' thực thi.

5. LƯU Ý PHÁP LÝ & BẢO MẬT (QUAN TRỌNG)

> **CẢNH BÁO PHÁP LÝ NGHIÊM NGẶT**
>
> PentAGI chứa các công cụ khai thác lỗ hổng thật. TUYỆT ĐỐI KHÔNG thực hiện kiểm thử trên các hệ thống, máy chủ, mạng lưới mà bạn KHÔNG có văn bản ủy quyền hợp pháp bằng văn bản. Mọi hành vi tấn công trái phép đều vi phạm luật an ninh mạng.

> **Cách ly môi trường an toàn**
>
> Luôn chạy PentAGI trong một VLAN cô lập hoặc máy ảo riêng biệt để ngăn ngừa rủi ro rò rỉ dữ liệu hoặc lan truyền mã độc ngoài ý muốn.

Chúc bạn xây dựng giải pháp an ninh mạng tự động và nâng cao năng lực phòng thủ thành công cùng PentAGI!

Nội dung kỹ thuật tổng hợp từ tài liệu chính thức của repo https://github.com/vxcontrol/pentagi