

AI-АГЕНТЫ В ПЕНТЕСТЕ: ЧТО РЕАЛЬНО РАБОТАЕТ, А ЧТО СЖИГАЕТ ТОКЕНЫ

БЕЗ МИФА ПРО “КНОПКУ БАБЛО”:
ПУТЬ ЧЕРЕЗ ОШИБКИ, ТЕСТЫ

РЕАЛЬНОСТЬ



Компания X продаёт автономного AI-хакера. Реальность продаёт вам счёт за токены, мусорные отчёты и риск бана.



Bugcrowd усиливает политики против AI slop.



HackerOne фиксирует рост отчётов с уязвимости, написанных AI.



VS



[\$(WHOAMI)]

СЕРГЕЙ ЗЫБНЕВ

CyberED

- Тимлид пентестеров в «Бастиян» с 5+ лет опыта в ИБ
- Автор Telegram-канала «Похек» (16к+ подписчиков)
- Опыт в AI/LLM:
 - \$6 000+ инвестировано в эксперименты с AI-агентами
 - 1,5 года R&D в AI для пентеста
- 5 OSS инструментов, ни один не остался в production



КЕЙС:

«БОЮСЬ, ЧТО МЕНЯ ЗАМЕНИТ ИИ»



Кого точно заменит



1. тех, кто держится только за ручные процессы;



2. тех, кто не умеет делегировать AI и проверять его ответы;



3. тех, кто игнорирует галлюцинации;



4. тех, чьё преимущество легко автоматизируется;



5. тех, кто не успевает перестроиться.



Кого вряд ли заменят



1. тех, кто умеет встроить AI в работу;



2. тех, кто быстрее разбирает вывод инструментов;



3. тех, кто формулирует гипотезы и валидирует их;



4. тех, кто собирает доказательства, а не просто генерирует отчёты;



5. тех, кто использует AI как усилитель, а не как «кнопку бабло».

НЕ «ИИ ПРОТИВ ЧЕЛОВЕКА»,
А ЧЕЛОВЕК С ИИ ПРОТИВ ЧЕЛОВЕКА БЕЗ ИИ.

ПОЧЕМУ НЕ ПОКАЗЫВАЕМ “ГОТОВУЮ СИСТЕМУ”

1



Готовая система = время,
методология, pipeline,
промтты, guardrails,
оркестрация.

2



На вебинаре
показываем демо

3



Полная система —
предмет курса



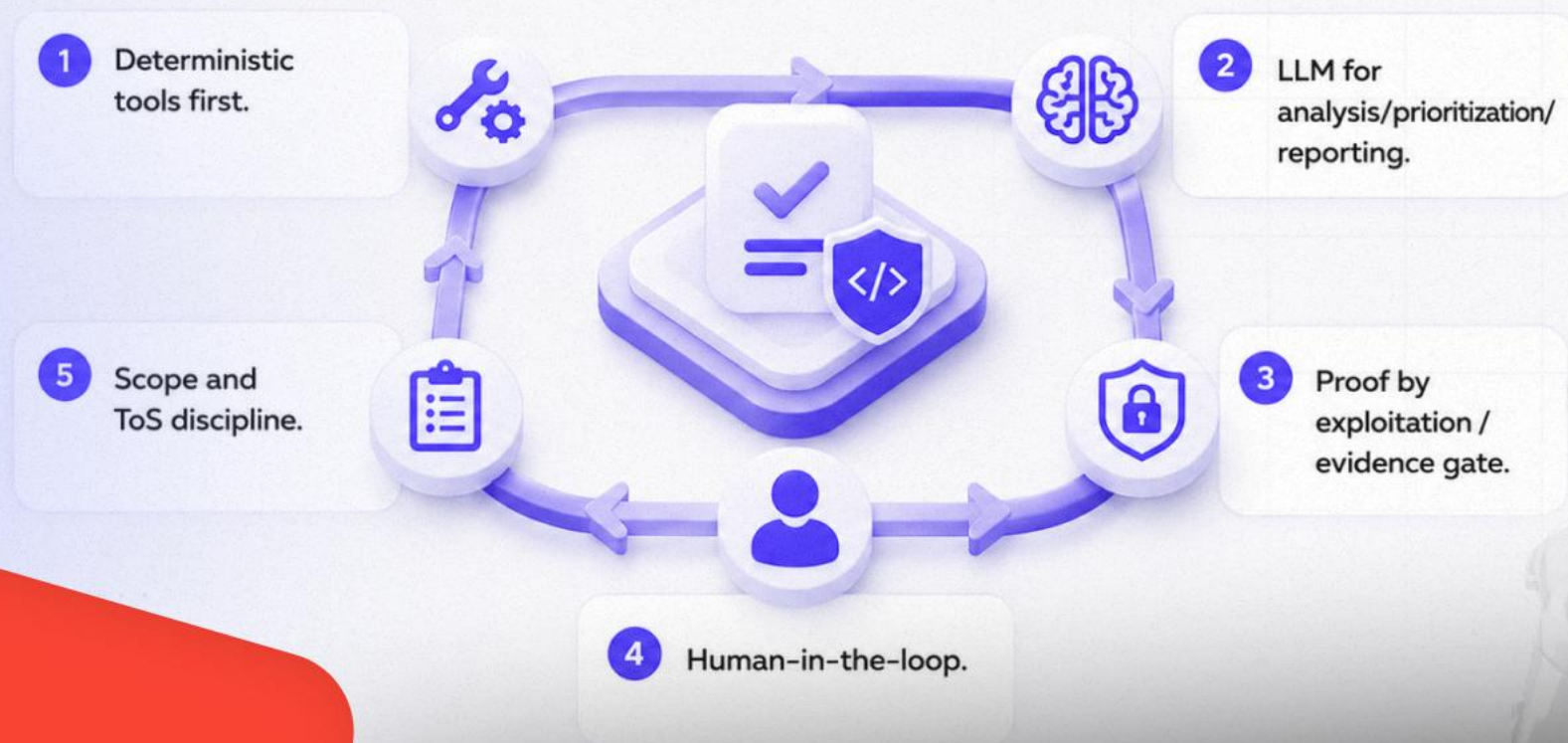
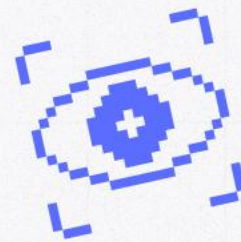
КАРТА ИНСТРУМЕНТОВ, КОТОРЫЕ “НЕ ВЗЛЕТЕЛИ”

Инструмент	 CAI	 PentAGI	 Strix	 Shannon	 Vulnhuntr
Тип	Universal AI security framework	Autonomous / multi-agent pentest framework	AppSec / CI/CD	White-box web pentest	LLM-assisted Python static analysis
Стоимость	~\$100 + локальный inference	до ~\$100 + локальный inference	\$50–500 за прогон / Claude subscription	~\$60–70	~\$200
Почему не основная система	На реальных задачах результат скромнее, чем на CTF; страшно запускать на боевую инфраструктуру	Учебные задачи решает, но для боевого применения рискованно; агент может заикливаться	Стоимость не оправдала результат; риск бана за offensive-промпты	Не подходит для black-box; слабое соотношение hype/result	Недетерминизм, только Python, большинство находок невалидны
Что вынести с вебинара	HITL важнее полной автономии	Архитектура красивая, но production-readiness не гарантирована	Validation-by-exploitation снижает шум, но дорого	Ограничения white-box подхода	Лучший кейс для тезиса «гибридный подход лучше автономного агента»

**НЕУДАЧА ИНСТРУМЕНТА — ЭТО ЦЕННОСТЬ:
ОНА ЭКОНОМИТ УЧАСТНИКУ МЕСЯЦЫ И ДЕНЬГИ.**

ЧТО РЕАЛЬНО РАБОТАЕТ

[FRAMEWORK:]



БЕЗОПАСНОЕ ДЕМО

☒ Target: OWASP Juice Shop.



ПОКАЗАТЬ:

- ✓ recon summary;
- ✓ candidate vuln hypothesis;
- ✓ manual validation;
- ✓ report skeleton.



НЕ ПОКАЗЫВАТЬ:

- ✗ весь pipeline;
- ✗ production targets;
- ✗ секретные prompts;
- ✗ agent orchestration целиком;
- ✗ обходы защит или ToS.



ДАВАЙТЕ РАЗБЕРЕМ
ВАШИ 
ВОПРОСЫ



ПОЧЕМУ АВТОНОМНЫЙ SWARM НЕ СПАСАЕТ



АВТОНОМНОСТЬ БЕЗ КОНТРОЛЯ БЫСТРО ПРЕВРАЩАЕТСЯ В РАСХОД ТОКЕНОВ, ШУМ И ОПЕРАЦИОННЫЙ РИСК.

РЕАЛЬНЫЕ ОГРАНИЧЕНИЯ



hallucinations



недоверие
к результатам



pipeline breaks



низкокачественные
reports



ban risks



privacy/data
leakage



дорогой
inference



shortage of clean
training/eval data



benchmark
contamination



gap между CTF
и real pentest

AI - МОЩНЫЙ ИНСТРУМЕНТ, НО НЕ ЗАМЕНА
ЭКСПЕРТИЗЕ И ПРОЦЕССУ.

BUG BOUNTY ~~IN~~ TOS

- ☐ SCOPE
- ☐ RATE LIMITS
- ☐ AUTOMATION ALLOWED
- ☐ AI-GENERATED CONTENT ALLOWED
- ☐ PROVIDER POLICY
- ☐ DATA HANDLING
- ☐ EVIDENCE
- ☐ HUMAN VALIDATION



AI УСИЛИВАЕТ КАРЬЕРНЫЙ И ФИНАНСОВЫЙ РОСТ

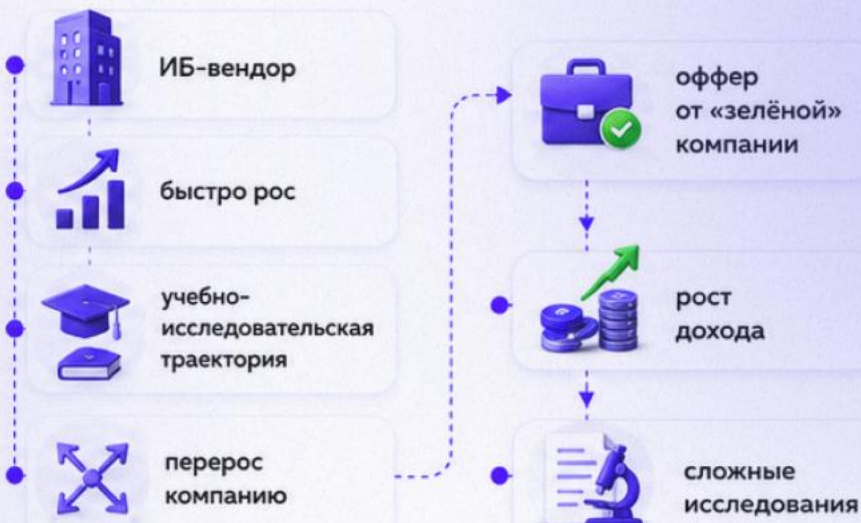
1 Мини-кейс 1: вымышленный Ваня



Средний рост дохода по такой траектории:

x1.5

2 Мини-кейс 2: вымышленный Петя



AI НЕ ОТМЕНЯЕТ УСИЛИЯ. ОН **УСИЛИВАЕТ** ТЕХ, КТО УЖЕ ГОТОВ РАСТИ, **УЧИТЬСЯ** И БРАТЬ **БОЛЕЕ СЛОЖНЫЕ ЗАДАЧИ**.

ЧТО ВАС ВОЛНУЕТ

1



Как работать
из России?

2



Запрещён ли AI?

3



Забанят ли на
bug bounty?

4



Что делать
с ToS?

5



Можно ли
локально?

6



Нужно ли
дообучать модель?



ЧТО ВЫ ПОКУПАЕТЕ



[НЕ «СЕКРЕТНУЮ КНОПКУ», А:]

1



Карта
проверенных
подходов

2



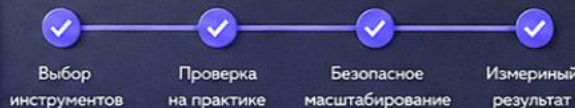
Готовые лабы
для проверки
своих AI-агентов

3



Reusable
workflow
и пайплайны

Проверенный путь



4



Критерии
выбора
инструментов /
моделей

5

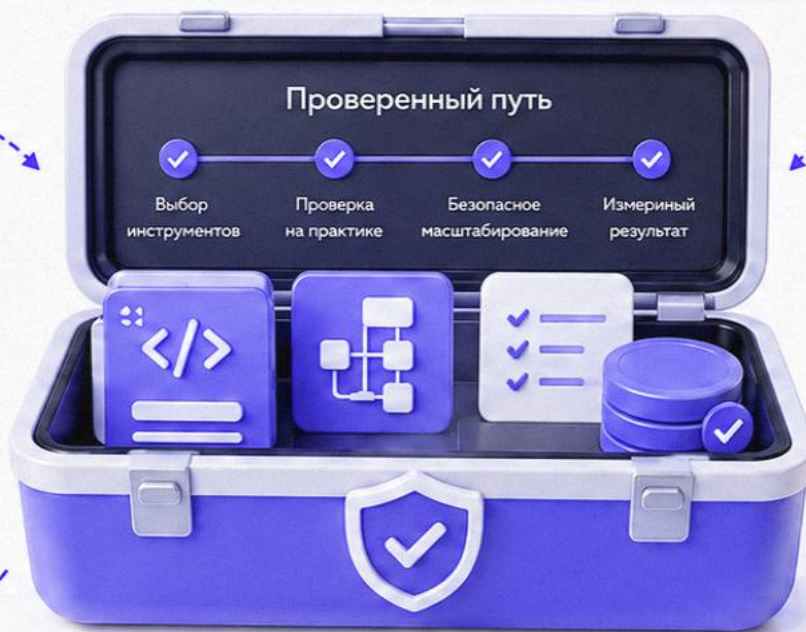


Ошибки,
которые уже
оплачены **мной**

6



Практика
с фидбеком +
понимание legal /
scope / ToS
boundaries



ЗАБЕРИТЕ WEB PENTEST SKILLS: СТАРТОВЫЙ НАБОР ПРОМПТОВ ПО ПЕНТЕСТУ



Небольшая часть — в открытом доступе.
Полная система работы с AI-агентами — в курсе APA.



СКАНИРУЙТЕ QR-КОД,
ЧТОБЫ ЗАБРАТЬ МАТЕРИАЛЫ



СПЕЦИАЛЬНОЕ ПРЕДЛОЖЕНИЕ



ОСТАВЛЯЙ ЗАЯВКУ!

НОВЫЙ КУРС: AI-усиленный пентест

AI-усиленный пентест за ~~60 000~~ рублей

Для всех: до 1 июня — 45 000 ₽

+ 10% скидка для тех, кто посетил вебинар

для участников
вебинара

ИТОГО: 40 500 рублей

УСПЕЙ КУПИТЬ ПО ВЫГОДНОЙ ЦЕНЕ ДО КОНЦА МАЯ

