Сработки сканеров и open-source триажер для Nuclei

RUSLAN SAFIULIN · 26.05.2026

— SLIDE 02 —

БОЛЬ

- >10k алертов на большом потоке
- Совпадение шаблона ≠ уязвимость
- Реальные находки тонут в шуме
- Триаж руками: дорого и медленно

— SLIDE 03 —

ПОЧЕМУ НЕ ПРАВИЛА

```
$ nuclei -id wordpress-detect -debug

[INF] matched wordpress-detect   https://blog.example/   [200]
[matcher] regex "/wp-content/uploads/(.*)" matched in body
[context] <link href="/wp-content/themes/twentytwentyfour/style.css">
          <meta name="generator" content="WordPress 6.4.2">
[check]   /wp-json/ → 200 (REST API WordPress)

[verdict] TRUE  это реальный WordPress
```

— SLIDE 04 —

ПОЧЕМУ НЕ ПРАВИЛА

```
$ nuclei -id wordpress-detect -debug

[INF] matched wordpress-detect   https://shop.example/   [200]
[matcher] regex "/wp-content/uploads/(.*)" matched in body
[context] <img src="https://cdn.partner.com/wp-content/uploads/2024/promo.png">
          ← хотлинк картинки с чужого WordPress
[check]   нет generator=WordPress, /wp-json → 404, Server: Tilda

[verdict] FALSE  сайт не на WordPress, лишь грузит картинку с него
```

— SLIDE 05 —

АРХИТЕКТУРА

— SLIDE 06 —

ПРЕПРОЦЕССОР

GET /actuator/env → 200 OK ~1.0 MB

```
HTTP/1.1 200 OK
Content-Type: application/vnd.spring-boot.actuator.v3+json
Content-Length: 6608

{"activeProfiles": ["prod"], "propertySources": [{"name": "systemEnvironment", "
... ещё ~1 010 000 байт тела ...
```

– SLIDE 07 –


```
{
  "summary": {
    "body_truncated": true,
    "match_count": 3,
    "status_code": 200,
    "injected_headers_checked": [
      "<HOST>"
    ]
  },
  "matcher_evidence": [
    {
      "pattern": "propertySources",
      "part": "body",
      "context": "{'activeProfiles': ['prod'], 'propertySour…"
    },
    {
      "pattern": "activeProfiles",
      "part": "body",
      "context": "{'activeProfiles': ['prod'], 'propertySour…"
    }
  ],
  "response_snippet": "{'activeProfiles': ['prod'], 'propertySource…",
  "extracted_results": [],
  "scan_target_mismatch": false
}
```

— SLIDE 08 —

ВЕРДИКТ ОТ ПРОТИВНОГО

```
{
  "verification_tag": true,
  "severity": "low"
  "verdict":
    Эндпоинт /actuator/env открыт и отдаёт propertySources с конфигурацией
    (200 + actuator JSON). FP сигналы не сработали: не WAF, продукт
    совпадает, условие matcher'а выполнено. Находка true; значения
    чувствительных ключей маскированы (******).
}
```

– SLIDE 09 –

ВЕРДИКТ ОТ ПРОТИВНОГО

FP сигналы:

waf_or_block_page	✗ не сработал	не страница блокировки
wrong_product	✗ не сработал	настоящий actuator JSON
matcher_contradiction	✗ не сработал	propertySources + activeProfiles в body
placeholder_value	✗ не сработал	реальная конфигурация, не заглушка

— SLIDE 10 —

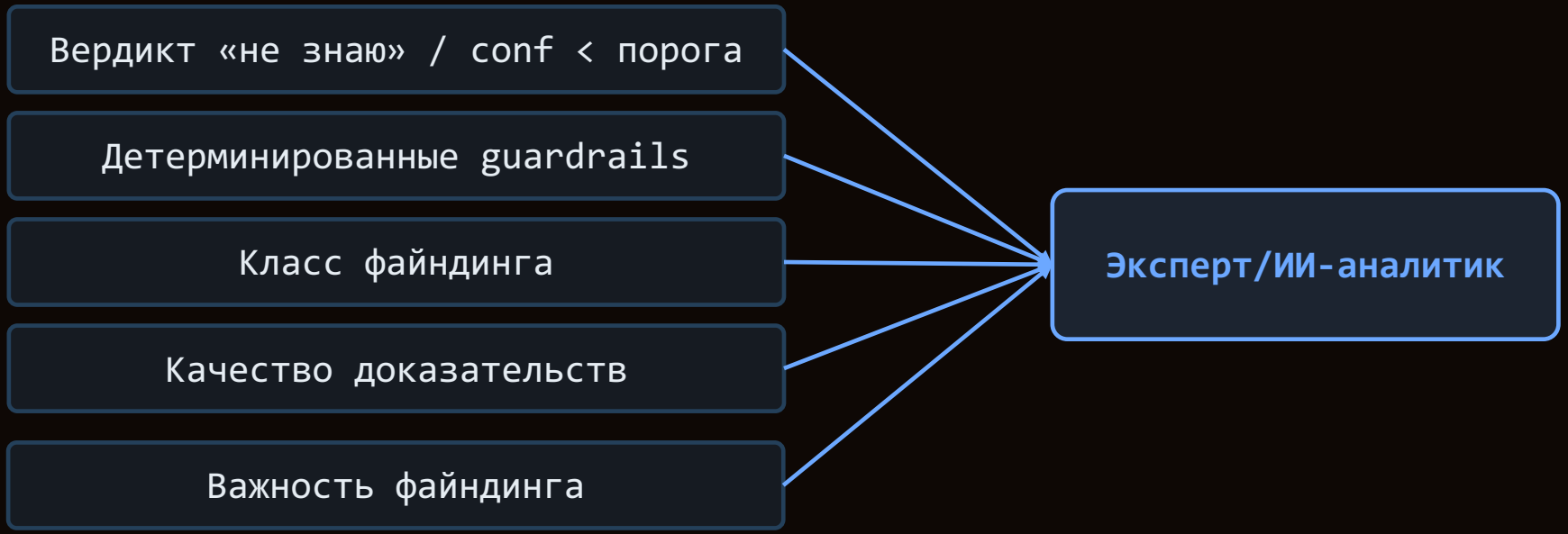
УВЕРЕННОСТЬ ЧЕРЕЗ АТАКУ

```
// атакуем вердикт → баллы из таблицы якорей

{
  "confidence_score": 0.75,
  "confidence_anchor": "F", // подтверждение по контенту
  "confidence_red_team_count": 0,
  "confidence_reason_code": "endpoint_exposed_values_masked"
}
```

– SLIDE 11 –

МАРШРУТИЗАЦИЯ



— SLIDE 12 —

РЕЗУЛЬТАТЫ

- Модель qwen36-fp8
- 844 размеченных вердиктов

Промпт	precision	recall	F1	TP	FP	FN
«докажи правду»	0.992	0.852	0.917	701	5	122
«ищи ложь»	0.984	**0.989**	0.945	814	13	9

— SLIDE 13 —

итоги

- Подготовка машиночитаемых доказательств
- Поиск FP-сигналов при выставлении вердикта
- Уверенность через атаку на вердикт
- Маршрутизация на эксперта/ИИ-аналитика

– SLIDE 14 –

OS NUCLEI-AUTOTRIAGE

https://github.com/cv3tomuzika/nuclei-autotriage