F E R R A Z W E B

Investigação OSINT

Reconhecimento Passivo com theHarvester

Autor	João Pedro Ferraz da Costa Marques Pereira
Marca	FerrazWeb
Data de inicio	14 de Março de 2026
Data de conclusão	17 de Abril de 2026

Investigação conduzida exclusivamente com ferramentas e fontes públicas - OSINT passivo

Índice de Conteúdos

1. O que é OSINT 3
2. Para que Serve - Aplicações Reais 3
3. Setup do Ambiente de Trabalho 5
4. Ferramentas Utilizadas 6
4.1 *theHarvester* 6
4.1.1 Instalação - Problemas Encontrados e Soluções 6
4.1.2 Opções e Flags Disponíveis 8
4.1.3 Fontes Disponiveis (-b) 8
4.1.4 Exemplos de Comandos 9
5. Websites e Fontes Utilizadas na Investigação 10
6. Domínios Investigados e Razões de Escolha 10
7. Comandos Executados e Resultados 11
7.1 Investigação - nasa.gov 11
7.1.1 Resumo dos Resultados 11
7.1.2 Email Encontrado 11
7.1.3 IP's Identificados 12
7.1.4 Subdomínios de Maior Interesse 12
7.1.5 Interpretação e Significado 13
7.2 Investigação - lagplace.xyz 14
7.2.1 Resumo dos Resultados 14
7.2.2 *ASN's* e Infraestrutura 14
7.2.3 *IP's* Identificados 15
7.2.4 Subdomínios Encontrados e Interpretação 16
7.2.5 *URL's* Marcadas como Interessantes pelo theHarvester 16
7.2.6 Interpretação e Valor Defensivo 17
7.3 Investigação - rapid7.com 18
7.3.1 Resumo dos Resultados 18
7.3.2 ASNs e Infraestrutura Cloud 18
7.3.3 Descobertas de Maior Relevância 19
7.3.4 Interpretação e Significado 19
8. Análise Comparativa dos Resultados 20
9. Conclusões 21
10. Disclaimer Ético e Legal 22
11. Bibliografia 23

1. O que é OSINT

OSINT significa Open Source Intelligence - Inteligência de Fontes Abertas. É o processo sistemático de recolher, correlacionar e analisar informação que se encontra publicamente disponível, sem necessidade de qualquer acesso não autorizado a sistemas privados.

As fontes de informação utilizadas incluem:

- Motores de busca (Google, Bing, DuckDuckGo, Baidu)
- Redes sociais (LinkedIn, Twitter/X, Facebook)
- Registos DNS e WHOIS públicos
- Bases de dados de certificados SSL/TLS (Certificate Transparency Logs)
- Repositórios de código públicos (GitHub, GitLab)
- Documentos públicos e os seus metadados
- Bases de dados de brechas de segurança (HaveIBeenPwned, etc.)
- Motores de busca especializados para dispositivos expostos (Shodan, Censys)

> Nota importante: OSINT não é *hacking*. Toda a informação recolhida já estava publicamente disponível - as ferramentas apenas automatizam e correlacionam o que qualquer pessoa poderia encontrar manualmente, poupando horas de pesquisa.

2. Para que Serve - Aplicações Reais

O OSINT tem aplicações em múltiplas áreas profissionais e académicas. Compreender o que está publicamente exposto sobre uma organização é o primeiro passo tanto para a atacar eticamente como para a defender.

Área	Aplicação
Cibersegurança Ofensiva	Fase de reconhecimento passivo num pentest ético para mapear a superfície de ataque antes de qualquer interação.
Cibersegurança Defensiva	Perceber exatamente o que um atacante consegue ver da tua organização sem tocar em nenhum sistema.
SOC / Threat Intelligence	Monitorizar a exposição contínua de dados da empresa e detetar novas superfícies de ataque.
Bug Bounty	Mapear toda a infraestrutura pública antes de procurar vulnerabilidades especificas.
Jornalismo Investigativo	Investigar entidades, pessoas públicas ou organizações com base em fontes abertas verificáveis.
Recursos Humanos	Due diligence sobre candidatos, parceiros ou fornecedores.
Forense Digital	Rastrear atividade online e correlacionar identidades em investigações.

No contexto deste projeto, o OSINT é utilizado para desenvolver competências práticas de reconhecimento passivo, demonstrando capacidade de recolher, interpretar e documentar informação publicamente disponível sobre organizações. Estas competências são diretamente aplicáveis em funções de SOC Analyst, Penetration Tester, Threat Intelligence Analyst e Bug Bounty Hunter.

A grande utilidade do OSINT defensivo pode ser resumida simplesmente: antes de alguém mal-intencionado descobrir o que esta exposto sobre a tua organização, és tu a descobrir primeiro.

3. Setup do Ambiente de Trabalho

Todas as investigações descritas neste relatório foram conduzidas no seguinte ambiente:

Componente	Detalhe
Sistema Operativo	Linux Mint Cinnamon
RAM	16 GB
Ligação a Internet	Wi-Fi
Ferramenta Principal	theHarvester 4.10.1
Terminal	Bash / Terminal do Linux Mint
Editor de Texto	Notepad++
Data de inicio	14 de Março de 2026
Data de conclusão	16 de Março de 2026

O Linux Mint Cinnamon é uma distribuição baseada em Ubuntu/Debian, amplamente utilizada em cibersegurança pela sua estabilidade, compatibilidade com ferramentas da área e gestor de pacotes *apt* robusto. É uma escolha sólida para projetos *OSINT* e *pentesting*, sendo acessível a utilizadores que transitam do Windows para um ambiente Linux completo.

4. Ferramentas Utilizadas

4.1 *theHarvester*

O *theHarvester* é uma ferramenta de reconhecimento passivo *open-source*, desenvolvida em *Python* por Christian Martorella da *Edge-Security Research*. É incluída por padrão no *Kali Linux* é disponível para outras distribuições. O seu objetivo é recolher emails, subdomínios, *IP's*, *URL's* e nomes de funcionários associados a um domínio, consultando automaticamente dezenas de fontes públicas de forma simultânea.

O *theHarvester* funciona como um agregador inteligente: envia pedidos a motores de busca e bases de dados públicas exatamente como um utilizador humano faria, mas de forma automatizada. Analisa as respostas com técnicas de *scraping* para extrair padrões, sendo endereços de email, subdomínios, *IP's* e agrega tudo numa lista única sem duplicados. Para cada subdomínio encontrado, pode ainda realizar resolução *DNS* para obter os endereços *IP* reais dos servidores.

> Versão utilizada neste projeto: theHarvester 4.10.1, instalada via GitHub (repositório oficial: github.com/laramies/theHarvester)

4.1.1 Instalação - Problemas Encontrados e Soluções

O processo de instalação encontrou vários obstáculos que são documentados aqui por serem comuns e relevantes para qualquer utilizador que tente reproduzir este projeto.

Método 1 - apt (falhou): O método oficial documentado não funcionou nesta configuração do Linux Mint.

```
sudo apt install theharvester
```

Método 2 - GitHub (utilizado com sucesso): Instalação direta a partir do repositório oficial.

```
git clone https://github.com/laramies/theHarvester.git
cd theHarvester
pip3 install . --break-system-packages
```

> A flag --break-system-packages é necessária em distribuições modernas baseadas em Debian/Ubuntu (incluindo Linux Mint) que protegem o ambiente Python do sistema. Sem ela, o pip recusa a instalação.

Erro encontrado - ImportError (circular import): Tentativa de execução incorreta de dentro da pasta do módulo.

```
# ERRADO - executado de dentro da pasta do módulo
cd theHarvester && python3 theHarvester.py -d nasa.gov -b google
# ImportError: cannot import name '__main__' from partially initialized module
```

Solução - flag -m: Execução correta a partir da pasta raiz do projeto.

```
# CORRETO - executado com -m a partir da pasta raiz
python3 -m theHarvester -d nasa.gov -b google
```

O erro de circular import acontecia porque o ficheiro theHarvester.py tentava importar o módulo theHarvester, que é a própria pasta em que se encontrava - criando um ciclo. A flag -m instrui o Python a tratar a pasta como um módulo e resolver as importações corretamente.

Foi ainda necessário adicionar o diretório de scripts ao PATH do sistema:

```
echo 'export PATH="$HOME/.local/bin:$PATH"' >> ~/.bashrc
source ~/.bashrc
```

Problema	Causa	Solução
apt install falhou	Versão desatualizada/indisponível no Mint	Instalação via GitHub com git clone
ImportError: circular import	Execução do .py de dentro da pasta errada	Usar python3 -m theHarvester na pasta raiz
Scripts não reconhecidos no terminal	~/.local/bin não estava no PATH	Adicionado ao ~/.bashrc com export PATH
Google removido como fonte	theHarvester 4.10.1 removeu suporte ao Google	Usar bing, crtsh, dnsdumpster, all
Conflito websockets com yt-dlp	Dependência de versão de outro programa	Ignorado - não afeta o theHarvester

4.1.2 Opções e Flags Disponíveis

Flag	Opção longa	Descrição
-d	--domain	Obrigatório. Define o domínio alvo. Ex: -d nasa.gov
-b	--source	Obrigatório. Define as fontes de pesquisa. Ex: -b bing,crtsh
-l	--limit	Limita o número de resultados. Ex: -l 200 (padrão: 500)
-S	--start	Começa a partir do resultado número X. Útil para paginação
-f	--filename	Guarda os resultados em ficheiros JSON e XML. Ex: -f resultados
-v	--virtual-host	Ativa verificação de virtual hosts por resolução DNS inversa
-e	--dns-server	Define um servidor DNS específico para as pesquisas
-r	--dns-resolve	Resolve os subdomínios encontrados para obter os seus *IP's*
-n	--dns-lookup	Realiza pesquisa DNS completa nos subdomínios encontrados
-c	--dns-brute	Executa brute-force de subdominios usando dicionário interno
-t	--take-over	Verifica se subdomínios estão vulneráveis a *subdomain takeover*
-p	--proxies	Usa *proxies* definidos para anonimato ou contornar rate limits

4.1.3 Fontes Disponiveis (-b)

Na versão 4.10.1, o Google foi removido como fonte nativa. As fontes disponíveis incluem:

Fonte	O que recolhe	API Key?
bing	Emails e subdomínios via Bing Search	Não
crtsh	Subdomínios via Certificate Transparency Logs	Não
dnsdumpster	Subdomínios e mapeamento DNS	Não
urlscan	URLs e subdomínios via urlscan.io	Não
anubis	Subdomínios via Anubis DB	Não
hackertarget	Subdomínios e IPs via HackerTarget	Não
otx	Indicadores de ameaça via AlienVault OTX	Não
rapiddns	Subdomínios via RapidDNS	Não
shodan	IP's e serviços expostos via Shodan	Sim (obrigatória)
securityTrails	Histórico DNS e subdominios	Sim (obrigatória)
hunter	Emails profissionais via Hunter.io	Sim (obrigatória)
all	Usa todas as fontes disponíveis simultaneamente	-

4.1.4 Exemplos de Comandos

Pesquisa básica com Bing:

```
theHarvester -d nasa.gov -b bing
```

Múltiplas fontes simultâneas:

```
theHarvester -d nasa.gov -b bing,crtsh,dnsdumpster,urlscan,anubis
```

Todas as fontes com exportação para ficheiro:

```
theHarvester -d nasa.gov -b all -f resultados_nasa
```

Com resolução DNS e limite de resultados:

```
theHarvester -d up.pt -b crtsh,bing -r -l 100 -f resultados_up
```

Com brute-force de subdomínios:

```
theHarvester -d exemplo.com -b bing -c
```

Os resultados são guardados em dois formatos: JSON (estrutura de dados legível por máquina) e XML. Na versão 4.10.1, o formato HTML foi descontinuado em favor do JSON.

5. Websites e Fontes Utilizadas na Investigação

Para além das fontes consultadas automaticamente pelo *theHarvester*, foram utilizados os seguintes serviços web para complementar a investigação e contextualizar os resultados. Todos estes serviços são públicos, gratuitos e legais de consultar.

Ferramenta	URL	Porque foi escolhida
crt.sh	https://crt.sh	Base de dados de *Certificate Transparency Logs* - revela subdomínios via certificados SSL/TLS. Completamente passivo e público
DNSDumpster	https://dnsdumpster.com	Mapeamento DNS visual e gratuito. Mostra subdomínios, *IP's* e relações entre eles de forma clara
Shodan	https://shodan.io	Motor de busca para dispositivos e serviços expostos na internet. Mostra portas abertas, versões de software, etc.
WHOIS	https://who.is	Informação de registo de domínios: proprietário, datas de registo/expiração, *nameservers*
HackerTarget	https://hackertarget.com	Ferramentas *OSINT* gratuitas online - subdomínios, DNS, *traceroute*, informação de *IP*
VirusTotal	https://virustotal.com	Histórico de subdomínios, *IPs* e reputação de domínios. Útil para cruzar dados
URLScan.io	https://urlscan.io	*Screenshots*, tecnologias utilizadas e estrutura de URL's de domínios analisados

6. Domínios Investigados e Razões de Escolha

Foram investigados três domínios distintos, escolhidos para cobrir diferentes tipos de organizações e proporcionar resultados variados e comparáveis.

Domínio	Tipo	Razão da Escolha
nasa.gov	Agência governamental	Organização pública de grande dimensão com infraestrutura extensa. Ideal para demonstrar a quantidade de informação exposta por grandes organizações
lagplace.xyz	Servidor pessoal	Domínio de Guilherme Cruz, colega do curso de Redes e Sistemas Informáticos, que possui um servidor pessoal. A investigação foi realizada com a sua autorização total. Demonstra o que fica exposto num servidor pessoal e o valor do *OSINT* defensivo
rapid7.com	Empresa de cibersegurança	Criadores do Metasploit e do InsightVM. Investigar uma empresa de segurança com as suas próprias técnicas tem valor pedagógico e ironia técnica.

7. Comandos Executados e Resultados

7.1 Investigação - nasa.gov

A NASA (National Aeronautics and Space Administration) é a agência espacial e aeronáutica dos Estados Unidos, fundada em 1958. Com dezenas de centros de investigação distribuídos pelos EUA e infraestrutura tecnológica de grande escala, é um dos domínios públicos com maior superfície de exposição. O domínio nasa.gov é utilizado por investigadores, cientistas e sistemas críticos, tornando-o um caso de estudo exemplar para OSINT.

Comando executado:

```
theHarvester -d nasa.gov -b all -f resultados_nasa
```

O que este comando faz: Consulta simultaneamente todas as fontes disponíveis no theHarvester para o domínio nasa.gov e guarda os resultados em ficheiros JSON e XML.

7.1.1 Resumo dos Resultados

Categoria	Quantidade / Detalhe
Emails encontrados	1
Subdomínios/Hosts	Milhares (ficheiro truncado na analise)
IP’s identificados	7
Shodan	0 (requer API key)

7.1.2 Email Encontrado

```
support@cce.nasa.gov
```

O endereço support@cce.nasa.gov pertence ao Center for Climate Economics (CCE) da NASA. A identificação deste email revela o formato de nomenclatura utilizado pela organização e confirma a existência de um departamento interno dedicado a estudos climáticos económicos.

7.1.3 IP's Identificados

Endereço IP	Identificação
15.200.58.24	Amazon AWS - infraestrutura cloud
23.22.39.120	Amazon AWS - infraestrutura cloud
52.0.14.116	Amazon AWS - infraestrutura cloud
99.86.159.121	Amazon AWS - CloudFront CDN
192.0.66.108	Automattic / WordPress.com
192.149.130.241	NASA - infraestrutura própria
198.117.0.121	NASA - infraestrutura própria

7.1.4 Subdomínios de Maior Interesse

Infraestrutura interna exposta:

- vpn.arc.nasa.gov - servidor VPN do Ames Research Center
- vpn-faa.arc.nasa.gov - VPN especifica para a FAA (Federal Aviation Administration)
- accounts.arc.nasa.gov e accounts-dev.arc.nasa.gov - sistemas de gestão de contas
- agencytokens.nasa.gov (192.149.130.161) - Sistema de tokens de autenticação da agência

Ambientes de desenvolvimento e teste:

- airpermit-dev.arc.nasa.gov (128.102.23.51) - ambiente de desenvolvimento
- airpermit-stage.arc.nasa.gov (128.102.23.50) - ambiente de staging
- webdev.arc.nasa.gov - servidor de desenvolvimento web
- ucmdb-test.arc.nasa.gov - ambiente de teste de gestão de configurações

Bases de dados com IPs públicos:

- acesdb00.arc.nasa.gov (128.102.149.250)
- acesdb01.arc.nasa.gov (128.102.149.251)
- acesdb02.arc.nasa.gov (128.102.149.252)

Sistemas físicos mapeados:

- voip-*.arc.nasa.gov - dezenas de telefones VoIP com nomes resolúveis
- ups-eaton120-b10.arc.nasa.gov - sistema de energia ininterrupta
- visitor-badging3.arc.nasa.gov - sistema de controlo de visitas

Entradas curiosas:

- wtfhax.arc.nasa.gov - nome atribuído por alguém internamente, possivelmente servidor de testes informais
- whatismyip.arc.nasa.gov - serviço interno para verificação de IP
- worldwind20 a worldwind49.arc.nasa.gov - cluster do projeto NASA WorldWind

7.1.5 Interpretação e Significado

A investigação a NASA revela uma organização com infraestrutura massiva e distribuída. Os pontos mais relevantes do ponto de vista de segurança são os ambientes de desenvolvimento e *staging* expostos publicamente - estes sistemas frequentemente recebem menos atenção de segurança do que os ambientes de produção, e podem conter vulnerabilidades já corrigidas no site principal.

A visibilidade das bases de dados com *IP's* diretos e também notável. Mesmo que não estejam acessíveis externamente, o facto de os nomes DNS resolverem publicamente fornece informação detalhada sobre a estrutura interna da infraestrutura. Num cenário de *pentest* autorizado, estes seriam alvos prioritários de investigação.

7.2 Investigação - lagplace.xyz

O domínio lagplace.xyz pertence a Guilherme Cruz, colega do autor no curso de Redes e Sistemas Informáticos, que possui e gere um servidor pessoal próprio. Esta investigação foi conduzida com a sua autorização explícita e total conhecimento, no âmbito do projeto OSINT da FerrazWeb. O autor não é proprietário deste domínio.

> A investigação ao domínio lagplace.xyz foi realizada com a autorização expressa de Guilherme Cruz, proprietário do domínio e colega do curso de Redes e Sistemas Informáticos. Todos os resultados são partilhados com o seu conhecimento e consentimento total.

Comando executado:

```
theHarvester -d lagplace.xyz -b all -f resultados_lagplace
```

7.2.1 Resumo dos Resultados

Categoria	Quantidade / Detalhe
Subdomínios encontrados	14
IP's identificados	5 (+ 1 IPv6)
ASN's identificados	2 (Vodafone Portugal + Cloudflare)
URL's marcadas como interessantes	4
Emails encontrados	0

7.2.2 *ASN's* e Infraestrutura

ASN	Fornecedor	Significado
AS12353	Vodafone Portugal	O servidor físico está em Portugal, ligado a rede da Vodafone - servidor domestico ou VPS português.
AS13335	Cloudflare	O domínio usa Cloudflare como proxy/CDN - proteção DDoS e ocultação parcial do IP real

7.2.3 *IP's* Identificados

IP	Fornecedor	Observação
104.21.72.186	Cloudflare	IP publico do proxy Cloudflare - não é o servidor real
172.67.153.220	Cloudflare	IP publico do proxy Cloudflare - não é o servidor real
2606:4700:3035::...	Cloudflare	Endereço IPv6 do proxy Cloudflare
77.54.32.253	Vodafone Portugal	IP REAL do servidor - exposto via DNS de subdomínios não protegidos
77.54.55.120	Vodafone Portugal	IP REAL do servidor - exposto via DNS de subdomínios não protegidos

ATENÇÃO: Apesar de utilizar Cloudflare para proteger o domínio principal, os *IP's* reais da Vodafone Portugal (77.54.x.x) foram expostos através do DNS de subdomínios não protegidos. Esta é uma das falhas mais comuns em servidores com Cloudflare: O domínio raiz está protegido, mas subdomínios individuais apontam diretamente para o servidor real, revelando o IP verdadeiro.

7.2.4 Subdomínios Encontrados e Interpretação

Subdomínio	Serviço Identificado	Relevância de Segurança
casaos.lagplace.xyz	CasaOS - painel de gestão doméstico	Expõe o sistema operativo de gestão do servidor
code.lagplace.xyz	VS Code Server ou similar	Editor de código remoto acessível publicamente
crafty.lagplace.xyz	Crafty Controller - gestão Minecraft	Revela presença de servidor de jogos
dmra.lagplace.xyz	Projeto/serviço personalizado	Serviço não identificado publicamente
download.lagplace.xyz	Servidor de downloads	Servidor de ficheiros com acesso publico
escola.lagplace.xyz	Plataforma educativa/escolar	Projeto escolar exposto publicamente
filebrowser.lagplace.xyz	FileBrowser - gestor de ficheiros web	ALTO RISCO - interface de gestão de ficheiros publica
mail.lagplace.xyz	Servidor de email	Revela infraestrutura de email propria
map.lagplace.xyz	Dynmap ou similar	Mapa dinâmico - possivelmente do servidor Minecraft
myadmin.lagplace.xyz	phpMyAdmin	ALTO RISCO - interface de gestão de base de dados MySQL publica
status.lagplace.xyz	Página de status de serviços	Revela quais os serviços ativos e o seu estado
teste.lagplace.xyz	Ambiente de testes	Ambientes de teste raramente têm segurança adequada
upload.lagplace.xyz	Servidor de uploads	Interface de upload de ficheiros publicamente acessível
*.lagplace.xyz	Wildcard DNS ativo	Todos os subdomínios não definidos apontam para o mesmo servidor

7.2.5 *URL's* Marcadas como Interessantes pelo theHarvester

- https://filebrowser.lagplace.xyz/ - gestor de ficheiros web com interface publica
- https://lagplace.xyz/ - página principal do domínio
- https://myadmin.lagplace.xyz/ - interface phpMyAdmin de gestão de base de dados
- https://teste.lagplace.xyz/ - ambiente de testes

7.2.6 Interpretação e Valor Defensivo

Este é o exemplo mais valioso do projeto do ponto de vista do *OSINT* defensivo. Em menos de 5 minutos, sem qualquer acesso não autorizado, foi possível descobrir o sistema operativo de gestão do servidor (CasaOS), a presença de um servidor Minecraft com Crafty Controller, que o phpMyAdmin e o FileBrowser estão expostos publicamente com interfaces web acessíveis, o *IP* real da Vodafone Portugal apesar do uso de Cloudflare e a existência de ambientes de teste e upload ativos.

Esta investigação demonstra com clareza a proposta de valor do *OSINT* defensivo: antes que alguém mal-intencionado descubra esta informação, o proprietário descobre-a primeiro e tem a oportunidade de corrigir as exposições desnecessárias.

7.3 Investigação - rapid7.com

A Rapid7 e uma das maiores empresas de cibersegurança do mundo, fundada em 2000 e sediada em Boston, Massachusetts. São conhecidos principalmente por serem os criadores e mantenedores do Metasploit Framework, o framework de *pentesting* mais utilizado no mundo - e pela plataforma InsightVM de gestão de vulnerabilidades, utilizada por milhares de empresas globalmente. A escolha desta empresa tem um elemento de ironia técnica: estamos a usar tecnicas *OSINT* para investigar uma empresa que vende exatamente este tipo de serviço de segurança a outras organizações.

Comando executado:

```
theHarvester -d rapid7.com -b all -f resultados_rapid7
```

7.3.1 Resumo dos Resultados

Categoria	Quantidade / Detalhe
Subdomínios/Hosts	Centenas
IP's identificados	300+
ASN's identificados	7
URL's interessantes	Dezenas (incluindo endpoints de API e blog posts de CVEs)
Shodan	0 (requer API key)

7.3.2 ASNs e Infraestrutura Cloud

ASN	Fornecedor	Papel na Infraestrutura
AS14618	Amazon AWS	Infraestrutura cloud principal - a maioria dos serviços
AS16509	Amazon AWS	Infraestrutura cloud principal - serviços adicionais
AS13335	Cloudflare	CDN e proteção DDoS para serviços públicos
AS209242	Cloudflare	CDN adicional
AS16625	Akamai	CDN enterprise de alto desempenho
AS14061	DigitalOcean	Servidores cloud secundários
AS26347	Newfold Digital	Hosting secundário

O uso de 7 ASNs distintos revela que a Rapid7 é uma empresa *cloud-native* com infraestrutura altamente distribuída e redundante. A AWS é claramente a base principal, com Cloudflare e Akamai a fornecerem camadas de CDN e proteção adicionais. Esta arquitetura é típica de empresas *enterprise* que necessitam de alta disponibilidade global.

7.3.3 Descobertas de Maior Relevância

Bloco de *IP's* dos *scanners* InsightVM:

Foi encontrado um bloco sequencial de dezenas de *IP's* no formato 208.118.225.x e 208.118.227.x. Estes são os endereços IP utilizados pelos *scanners* de vulnerabilidades do InsightVM quando realizam *scanning* de clientes. Esta informação tem valor prático real: qualquer organização pode criar regras de *firewall* para explicitamente permitir ou bloquear *scanning* proveniente *de*stes endereços.

```
208.118.225.2 ate 208.118.225.62 (scanners InsightVM)
208.118.227.1 ate 208.118.227.126 (scanners InsightVM)
```

Cliente real hospedado na infraestrutura da Rapid7:

```
1776bank.hosted.rapid7.com
```

Um banco, 1776bank, está hospedado diretamente na infraestrutura da Rapid7, o que revela um cliente real da empresa através de *OSINT* passivo.

***Endpoints* de API expostos:**

```
us.api.insight.rapid7.com
us3.api.endpoint.ingress.rapid7.com
```

Endpoints da plataforma Insight completamente mapeados publicamente, incluindo *URL's* de queries de *log search* com parâmetros reais visíveis nos resultados.

***URL's* de *threat intelligence* com CVEs e grupos de ameaça:**

- Análise de CVE-2026-1731 - RCE critico no BeyondTrust Remote Support
- Relatório sobre grupo Crimson Collective - novo grupo de ameaça a operar na cloud
- Análise de risco de supply chain no Notepad (operação Chrysalis)
- Investigação sobre compromisso de WordPress em operação global de *malware*

7.3.4 Interpretação e Significado

A investigação a Rapid7 demonstra um ponto fundamental: mesmo uma empresa de cibersegurança de topo expõe publicamente, através de *OSINT* passivo, a totalidade da sua infraestrutura cloud distribuída por 7 fornecedores, os *IP's* dos seus *scanners* de vulnerabilidades, clientes reais hospedados na sua plataforma, *endpoints* de API com parâmetros de queries reais e relatórios detalhados sobre ameaças ativas.

Isto não significa que a Rapid7 tem má segurança, significa que qualquer organização suficientemente grande inevitavelmente expõe informação de infraestrutura publicamente. A questão não e eliminar completamente a exposição, mas sim conhecê-la, compreendê-la e gerir o risco de forma informada.

8. Análise Comparativa dos Resultados

A análise dos três domínios investigados permite traçar comparações interessantes sobre os diferentes padrões de exposição de informação pública.

Critério	nasa.gov	lagplace.xyz	rapid7.com
Emails encontrados	1	0	0
Volume de subdomínios	Milhares	14	Centenas
IP's encontrados	7	5 (+1 IPv6)	300+
ASN's identificados	Não reportado	2	7
IP real exposto	N/A	Sim (via subdomínios)	N/A
Serviços de gestão expostos	Sim (admin, accounts)	Sim (phpMyAdmin, FileBrowser)	Não identificado
Ambientes dev/test expostos	Sim	Sim (teste.lagplace.xyz)	Não identificado
Fornecedores cloud	AWS	Cloudflare + Vodafone PT	AWS + Cloudflare + Akamai + DigitalOcean

O padrão mais claro que emerge desta análise é que a dimensão da organização está diretamente correlacionada com o volume de informação exposta mas, não necessariamente com o tipo de exposição. Tanto a NASA como o lagplace.xyz expõe ambientes de desenvolvimento e interfaces de gestão publicamente acessíveis, independentemente da diferença de escala entre as duas organizações.

9. Conclusões

Este projeto demonstrou na prática o poder do reconhecimento passivo baseado em fontes abertas. Utilizando exclusivamente informação publicamente disponível e ferramentas *open-source*, foi possível mapear em detalhe a infraestrutura pública de três organizações distintas, identificar potenciais vetores de ataque em dois deles e revelar informação sensível sobre o servidor pessoal de um colega de curso.

1. A superfície de exposição é sempre maior do que se pensa. Em todos os casos investigados, a quantidade de informação publicamente disponível excedeu o esperado. Subdomínios esquecidos, ambientes de teste e interfaces de gestão raramente surgem na consciência dos administradores mas, surgem nos resultados do theHarvester.

2. Usar Cloudflare não garante anonimato do IP real. O caso do lagplace.xyz demonstra que proteger o domínio raiz com Cloudflare não é suficiente se subdomínios individuais apontarem diretamente para o servidor. O *IP* real da Vodafone Portugal foi exposto apesar da proteção Cloudflare ativa.

3. Nenhuma organização é imune. A Rapid7, uma empresa cujo negocio e a cibersegurança, expõe publicamente centenas de *IP's* e toda a sua infraestrutura cloud. Isto não é falha, é a realidade operacional de qualquer organização de grande dimensão.

4. O *OSINT* defensivo é tão valioso quanto o ofensivo. Conhecer o que está exposto é o primeiro passo para gerir o risco. Este projeto teve tanto valor como exercício defensivo - identificando exposições no lagplace.xyz - quanto como demonstração de técnicas ofensivas éticas.

5. As ferramentas importam, mas a interpretação é o que diferencia. O *theHarvester* produziu listas de milhares de entradas. O valor real não está nos dados em bruto mas na capacidade de identificar o que é relevante, compreender o que significa e comunicá-lo de forma clara.

10. Disclaimer Ético e Legal

Declaração de Responsabilidade

Toda a informação recolhida neste projeto foi obtida exclusivamente através de fontes publicas e abertas, sem qualquer acesso não autorizado a sistemas, redes ou dados privados. Os domínios investigados foram selecionados por serem de organizações públicas de grande dimensão (nasa.gov, rapid7.com) ou com autorização explícita do proprietário (lagplace.xyz — domínio de Guilherme Cruz, colega do curso de Redes e Sistemas Informáticos), tendo todos sido investigados com total conformidade ética e legal.

Este projeto tem fins exclusivamente educativos e de desenvolvimento de competências em cibersegurança. Nenhuma vulnerabilidade foi explorada, nenhum sistema foi comprometido e nenhuma informação privada foi acedida.

- João Pedro Ferraz da Costa Marques Pereira | FerrazWeb | Março 2026

11. Bibliografia

As referências listadas nesta secção correspondem a ferramentas, repositórios e recursos consultados durante a preparação e execução desta investigação OSINT. Importa salientar que o conhecimento aplicado neste projeto não provém exclusivamente destas fontes: parte significativa resulta de interesse pessoal acumulado na área de redes, cibersegurança e tecnologia, desenvolvido de forma contínua através de conteúdo técnico online, redes sociais especializadas (LinkedIn, Reddit, Discord) e comunidades da área. As referências formais servem de complemento e validação a esse conhecimento prévio.

DNSDumpster. (2026).*DNSDumpster — DNS recon & research, find & lookup DNS records*. https://dnsdumpster.com

HackerTarget. (2026).*HackerTarget — Online network tools and security research*. https://hackertarget.com

Hammond, J. [John Hammond]. (2024).*OSINT techniques — Finding information online* [Vídeo]. YouTube. https://www.youtube.com/@_JohnHammond

Martorella, C. (2024).*theHarvester* (Versão 4.10.1) [Software]. Edge-Security Research. https://github.com/laramies/theHarvester

NetworkChuck [NetworkChuck]. (2023).*how I legally hacked a real company's website (OSINT)* [Vídeo]. YouTube. https://www.youtube.com/@NetworkChuck

OSINT Framework. (2026).*OSINT Framework — Focused on gathering information from free tools or resources*. https://osintframework.com

Proton AG. (2026).*ProtonVPN — Secure and free VPN service*. https://protonvpn.com

Rapid7. (2026).*Rapid7 blog — Threat intelligence and security research*. https://www.rapid7.com/blog/

RIPE NCC. (2026).*RIPE NCC — IP address and ASN information for Europe, Middle East and Central Asia*. https://www.ripe.net

Sectigo. (2026).*crt.sh — Certificate transparency search*. https://crt.sh

Shodan LLC. (2026).*Shodan — The search engine for the internet of everything*. https://www.shodan.io

URLScan.io. (2026).*URLScan.io — Scan and analyse websites*. https://urlscan.io

VirusTotal. (2026).*VirusTotal — Analyze suspicious files, domains, IPs and URLs*. Google LLC. https://www.virustotal.com

Nota: O conhecimento demonstrado neste projeto resulta de uma combinação entre as referências formais acima listadas e um interesse pessoal continuado na área de redes e cibersegurança, alimentado por comunidades online, redes sociais especializadas e consumo regular de conteúdo técnico. Não todo o conhecimento aplicado tem uma referência formal associada, parte dele foi absorvido organicamente ao longo do tempo.