“L'OSINT ne sert à rien”

👾 Thibaud “Mhack” PERRARD 👾

🎯 Pentester chez Thucy

🛰️ Renseignement d’Origine Cyber

🔬 Chercheur de vulnérabilités 📱 & 📟

🌐 Fondateur Offensive-Intelligence.com

🪖 Officier de réserve au COMCYBER

Alors, c'est quoi l'OSINT ?

CE QUE C'EST

Tout est déjà là, dehors, en accès libre.
Le talent, ce n'est pas de hacker, c'est de regarder au bon endroit, et de relier ce que personne n'a relié.

Une photo, un favicon, un email oublié dans un commit : pris séparément, du bruit. Recoupés, une identité.

CE QUE CE N'EST PAS

Ce n'est pas du piratage. On ne force aucune porte, on n'entre dans aucun système. On lit ce qui est public.

Ce n'est pas de la magie non plus. Aucun outil ne « trouve le coupable ». Les outils collectent, l'humain font le lien.

« L'OSINT ne sert à rien »

A

Des alternatives peu connues, souvent gratuites, parfois redoutables.

GeoSpy

GÉOLOCALISATION D'IMAGE PAR IA

CE QUE ÇA FAIT

Estime où une photo a été prise à partir des seuls pixels — architecture, végétation, revêtements, mobilier urbain — sans aucune métadonnée EXIF. Le modèle Superbolt vise la précision au niveau de la rue.

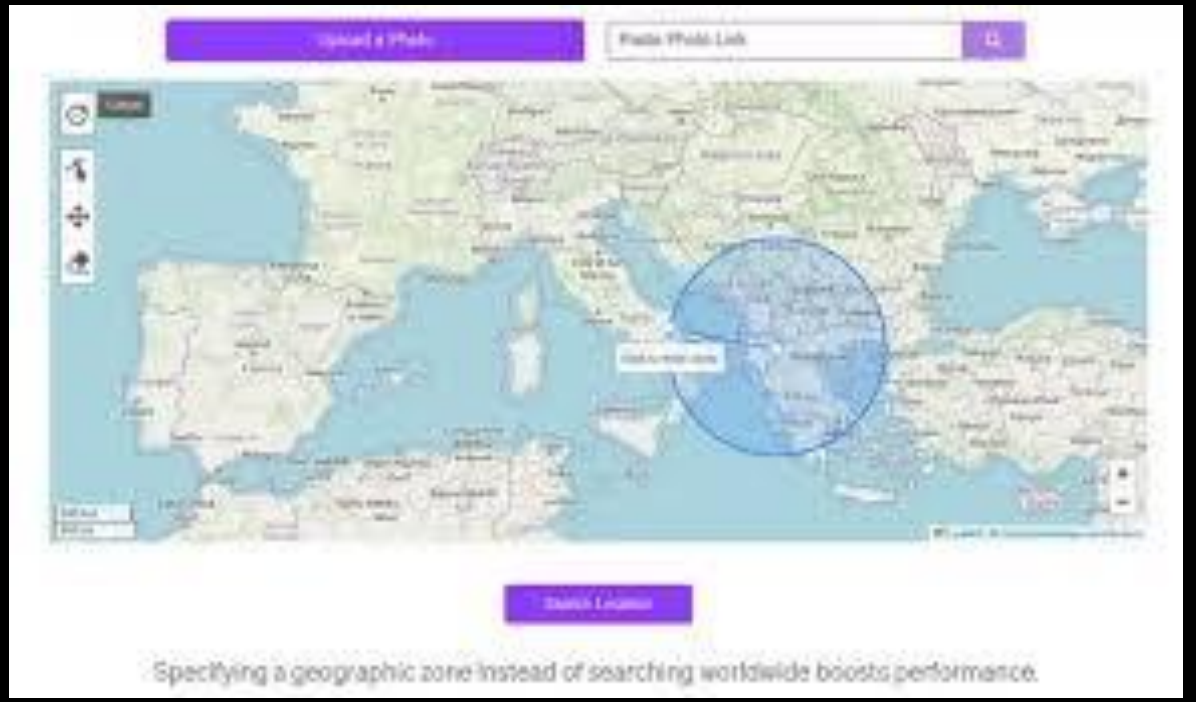
geospy.net

Picarta

GÉOLOCALISATION D'IMAGE — SECONDE OPINION

CE QUE ÇA FAIT

Même principe que GeoSpy, modèle différent : prédit le lieu de prise de vue d'une photo. Croiser deux moteurs qui se trompent rarement de la même façon fiabilise une géolocalisation incertaine.



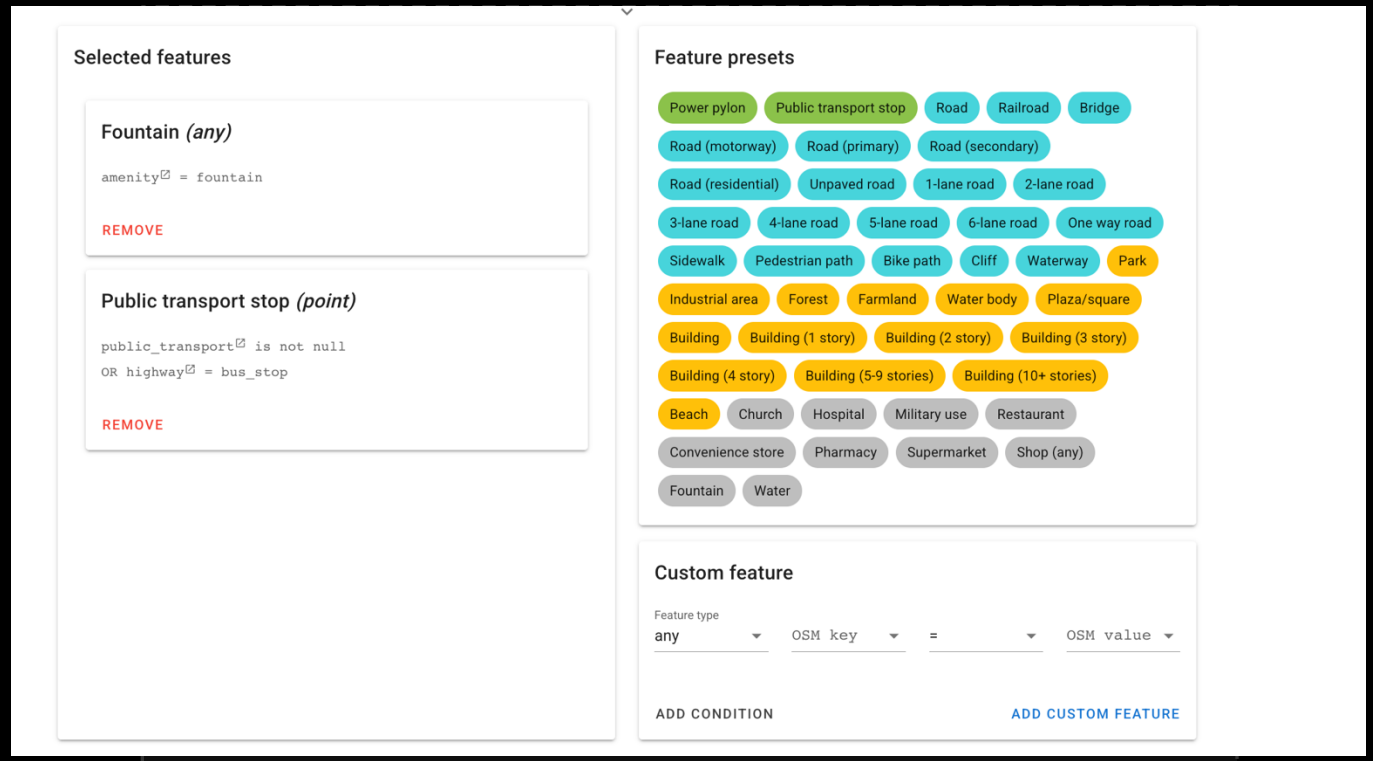
picarta.ai

Bellingcat OSM Search

RECHERCHE D'OBJETS CARTOGRAPHIQUES

CE QUE ÇA FAIT

Surcouche d'OpenStreetMap signée Bellingcat : chercher des combinaisons d'éléments (pont + stade + rivière) sur une zone, sans écrire de requête. Le point d'entrée idéal d'une géolocalisation par recoupement d'objets visuels.



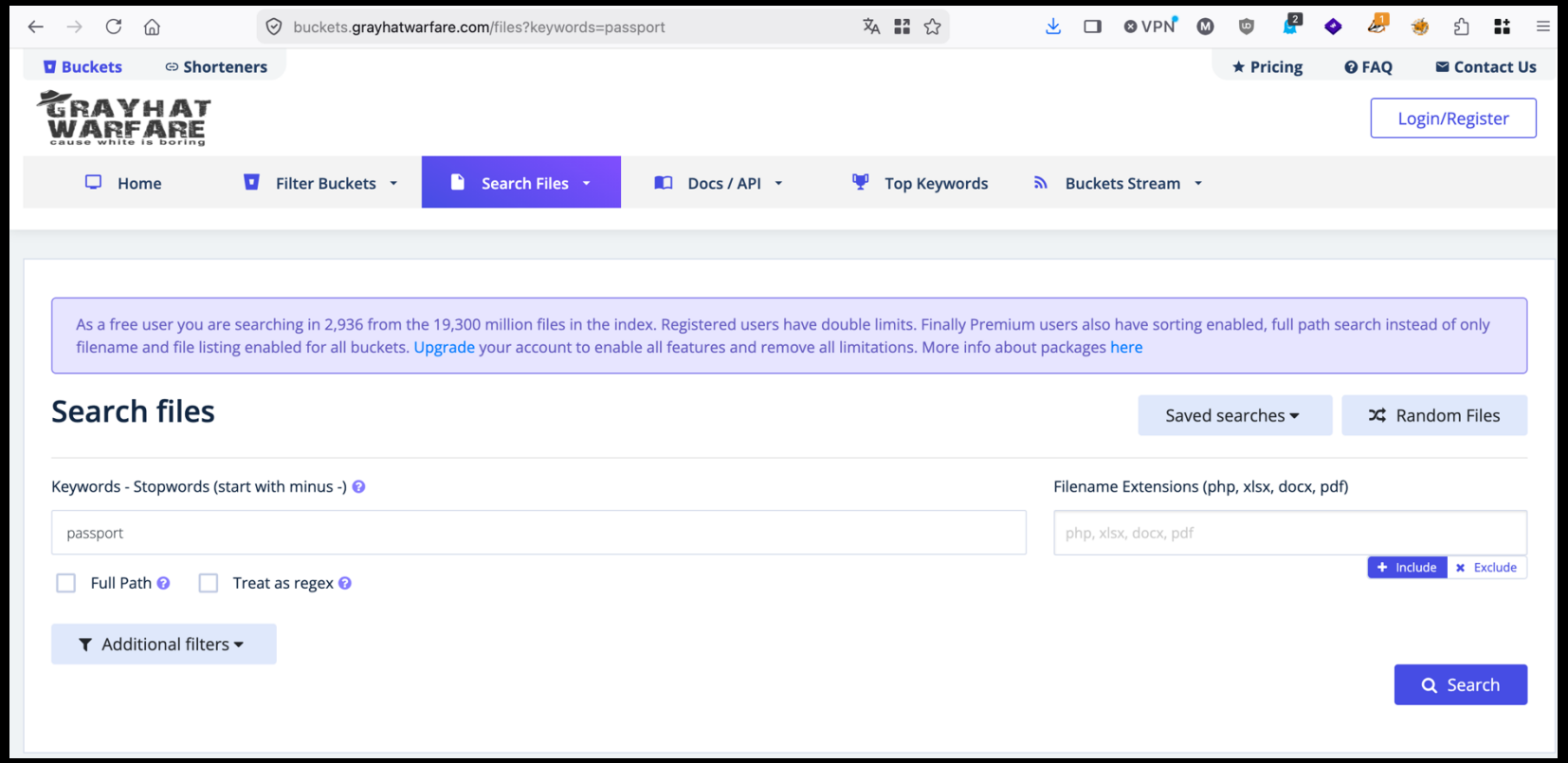
osm-search.bellingcat.com

GrayHatWarfare & scanners de buckets

LE STOCKAGE CLOUD LAISSÉ OUVERT

CE QUE ÇA FAIT

GrayHatWarfare indexe des centaines de millions de fichiers laissés publics sur les buckets cloud (S3, Azure, GCP) : recherche par mot-clé, extension, nom d'organisation. Et quand l'index ne suffit pas, des scanners (cloud_enum, S3Scanner) testent les buckets dérivés d'un nom — société-backup, société-prod, société-dev — que le naming prévisible trahit.



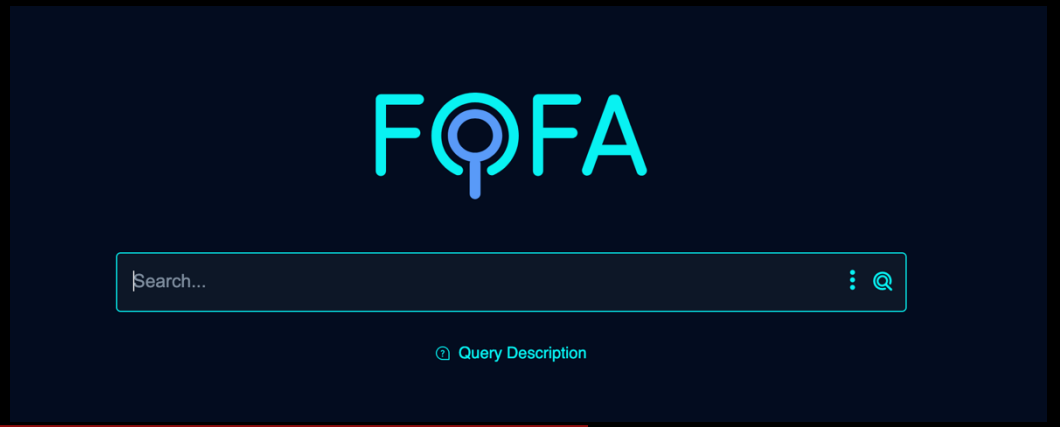
grayhatwarfare.com · github.com/initstring/cloud_enum

Les alternatives à Shodan

QUAND UN SEUL MOTEUR DE SCAN NE SUFFIT PAS

CE QUE ÇA FAIT

Shodan n'est pas seul, et croiser plusieurs moteurs change tout. FOFA excelle au pivot par favicon, certificat ou bannière. Netlas ajoute la recherche par regex dans le contenu HTML. ZoomEye couvre des hôtes que les autres manquent, notamment en Asie. Et Validin remonte l'historique DNS — l'IP d'origine d'un domaine avant son passage derrière un CDN.



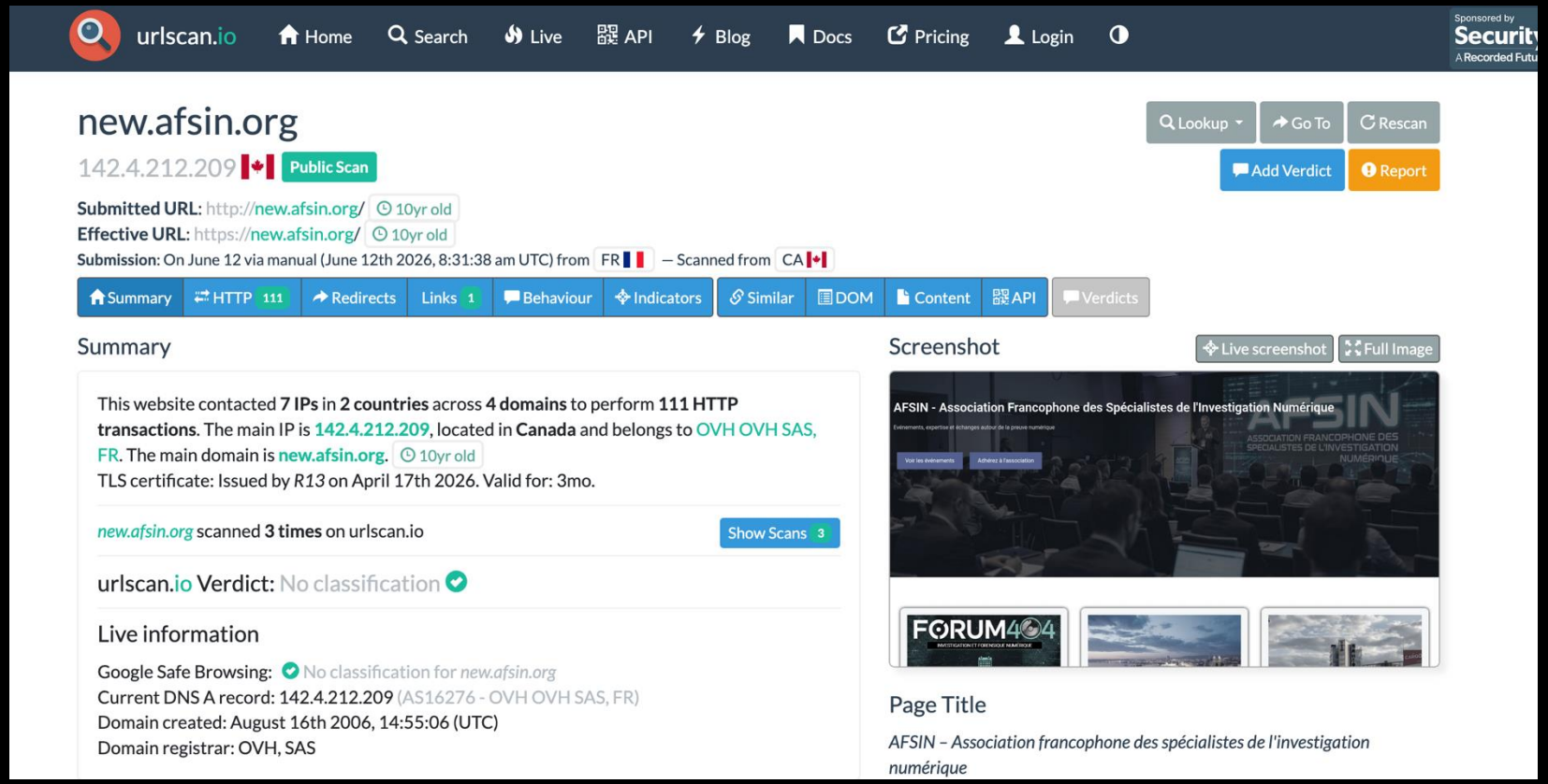
fofa.info · netlas.io · zoomeye.org · validin.com

urlscan.io

SCANNER D'URL ET BAC À SABLE WEB

CE QUE ÇA FAIT

Visite une URL et capture tout : DNS, requêtes, ressources chargées, capture d'écran, technologies, redirections. Et surtout, une base publique de millions de scans passés, interrogeable — où l'on retrouve les sites frères.



urlscan.io

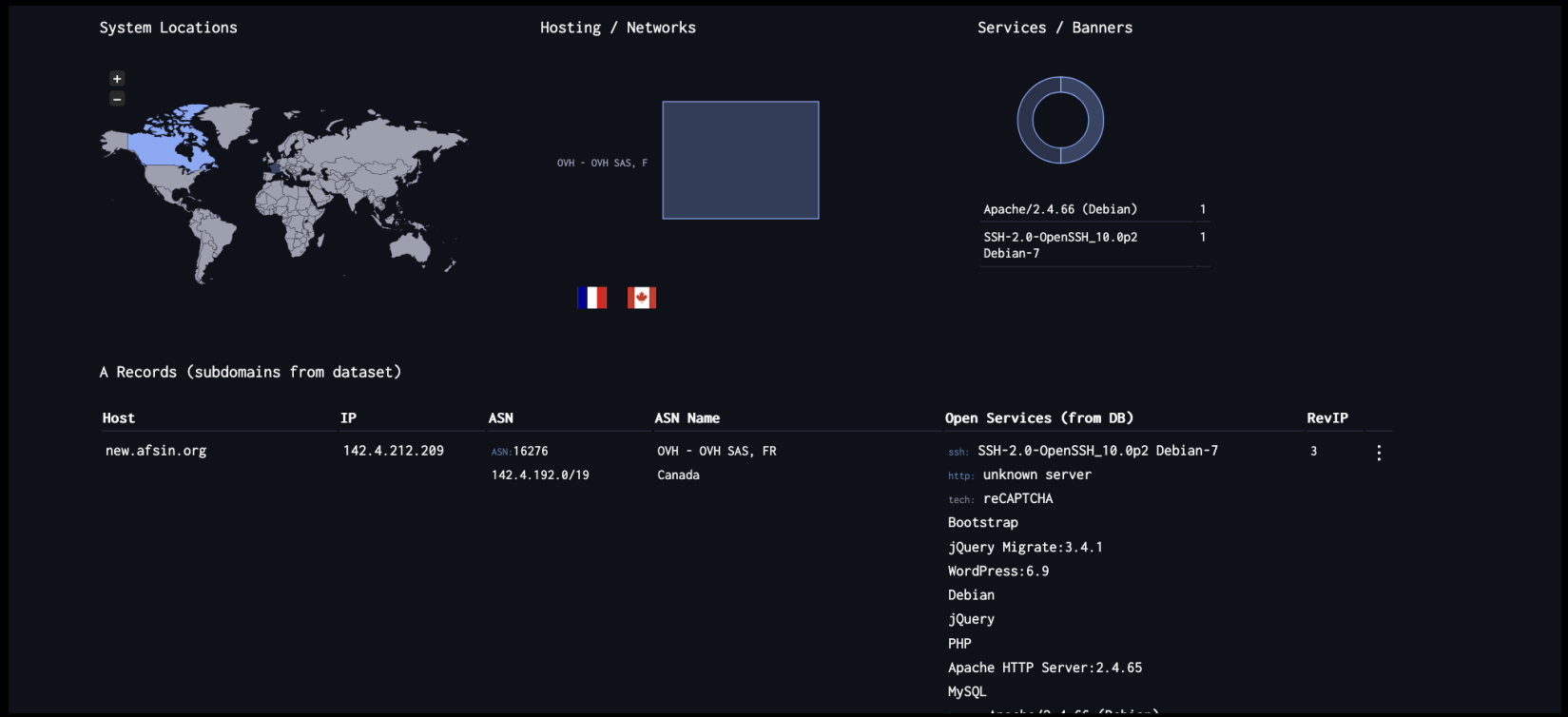
DNSDumpster

CARTOGRAPHIE DNS RAPIDE ET GRATUITE

CE QUE ÇA FAIT

Cartographie d'un domaine en un clic : sous-domaines, enregistrements, hôtes liés, le tout sur un graphe exportable.

Reconnaissance DNS gratuite et instantanée, sans installation.



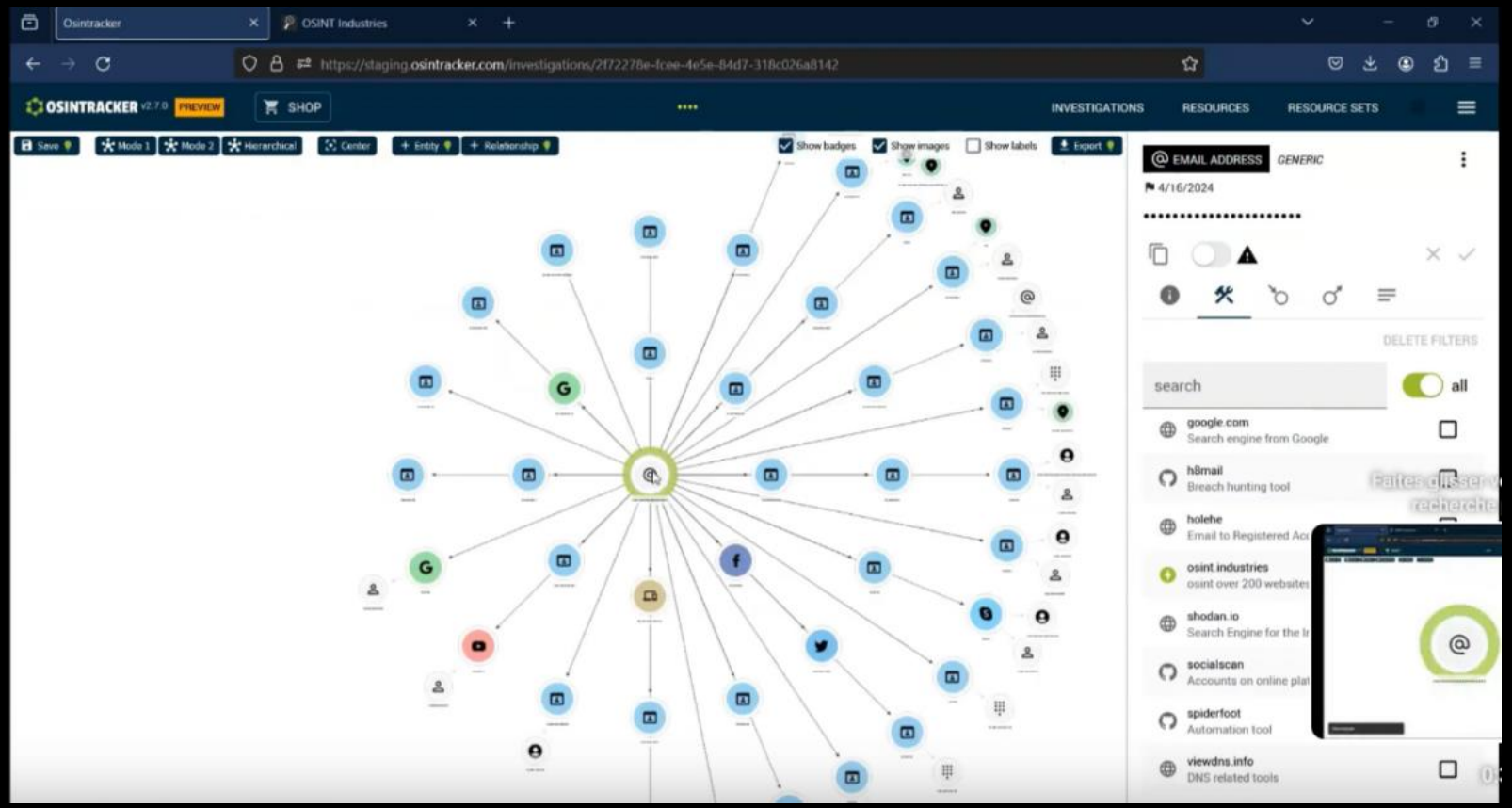
dnsdumpster.com

OSINTracker

CARTOGRAPHIE D'ENQUÊTE TOUT-EN-UN

CE QUE ÇA FAIT

Centralise une investigation : entités, liens, captures, géolocalisations et timeline sur un même graphe. Pensé pour documenter au fil de l'eau et exporter un rapport propre, là où Maltego demande du montage manuel.



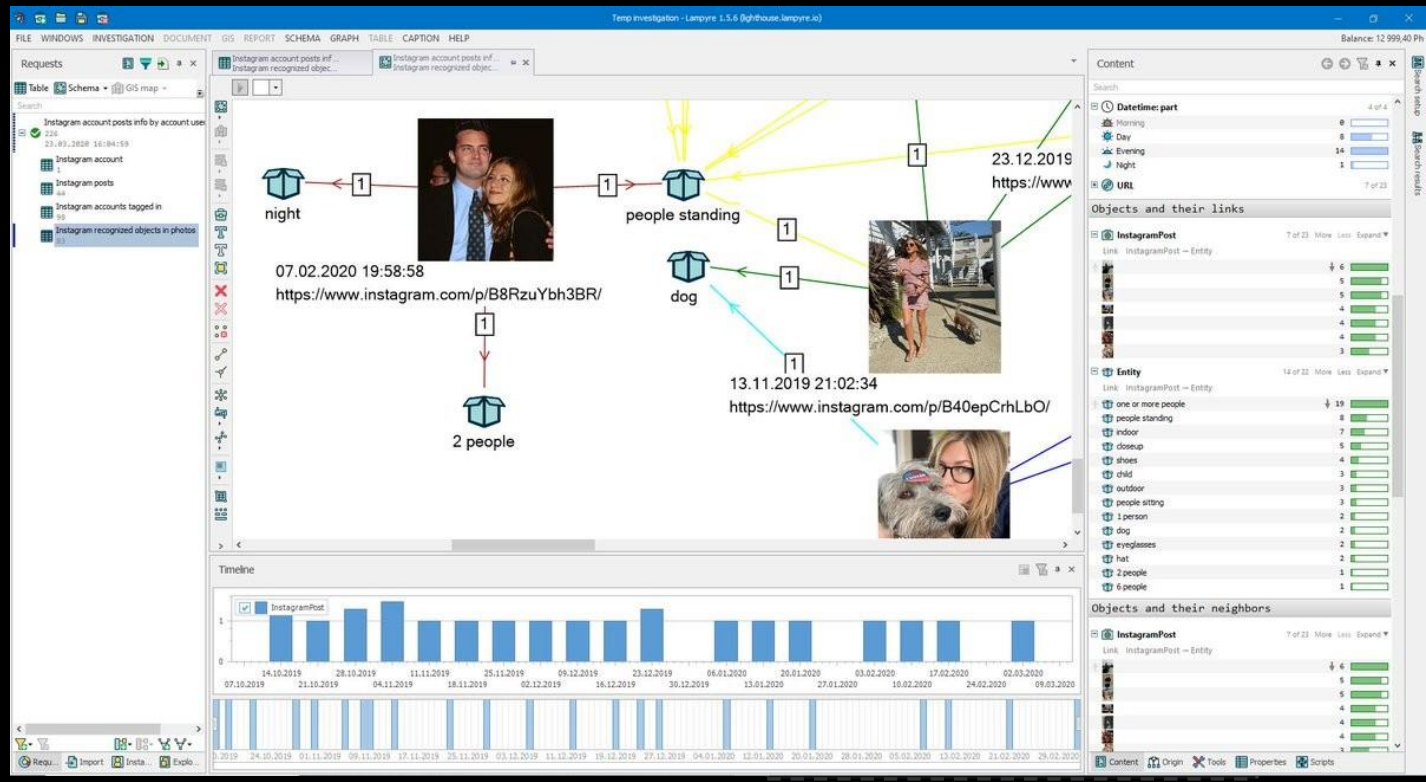
osintracker.com

Lampyre

ANALYSE RELATIONNELLE ET REQUÊTES À LA DEMANDE

CE QUE ÇA FAIT

Outil d'analyse de données et de graphes relationnels avec ses propres requêtes (téléphone, email, domaine, crypto) qui alimentent directement le graphe. Plus accessible que Maltego, paiement à la requête.



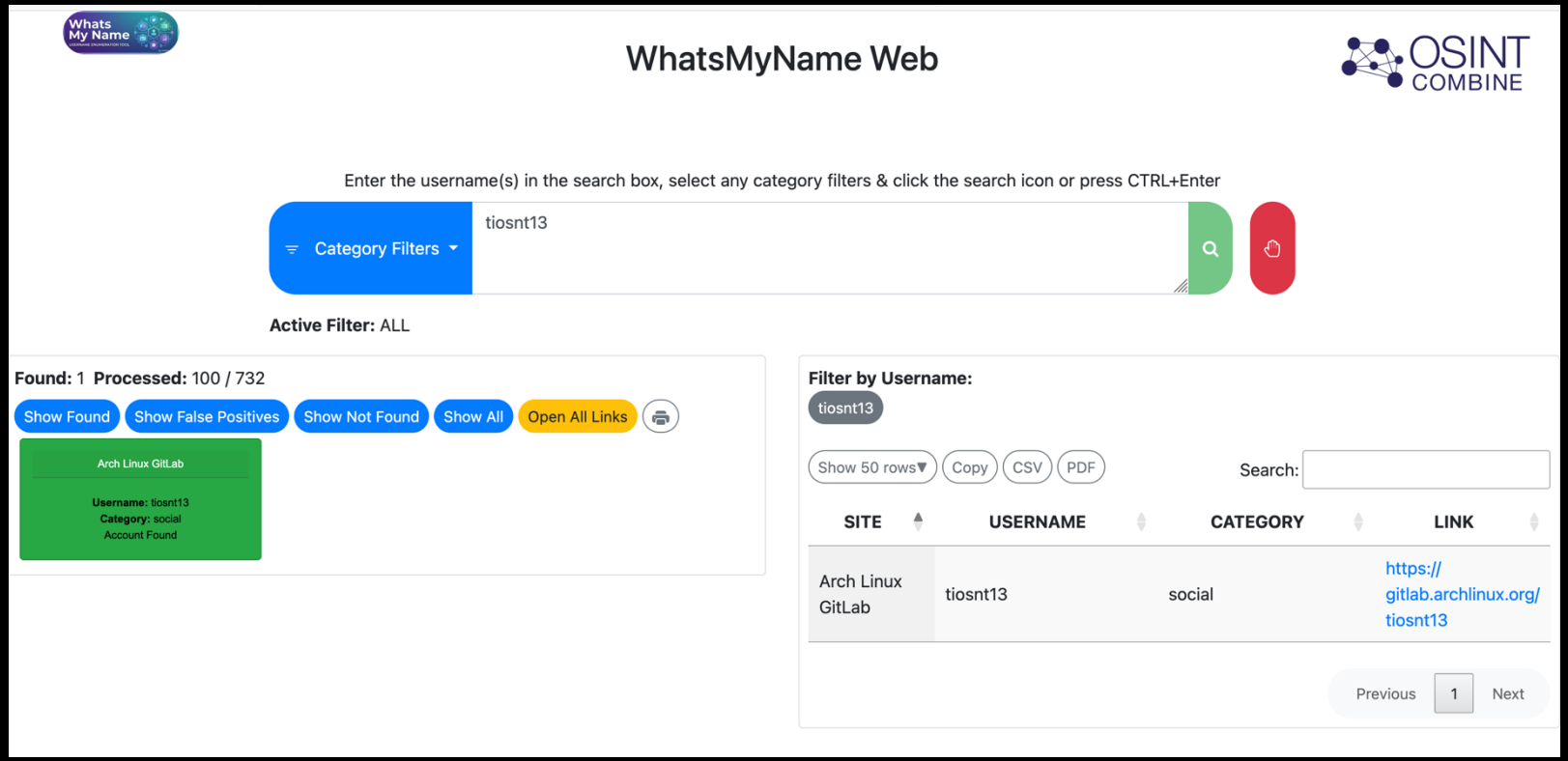
lampyre.io

WhatsMyName

ÉNUMÉRATION DE PSEUDO SUR 600+ SITES

CE QUE ÇA FAIT

Teste la présence d'un pseudonyme sur des centaines de plateformes en quelques secondes. La liste est maintenue par la communauté et limite les faux positifs. Le point de départ de toute enquête sur identité.



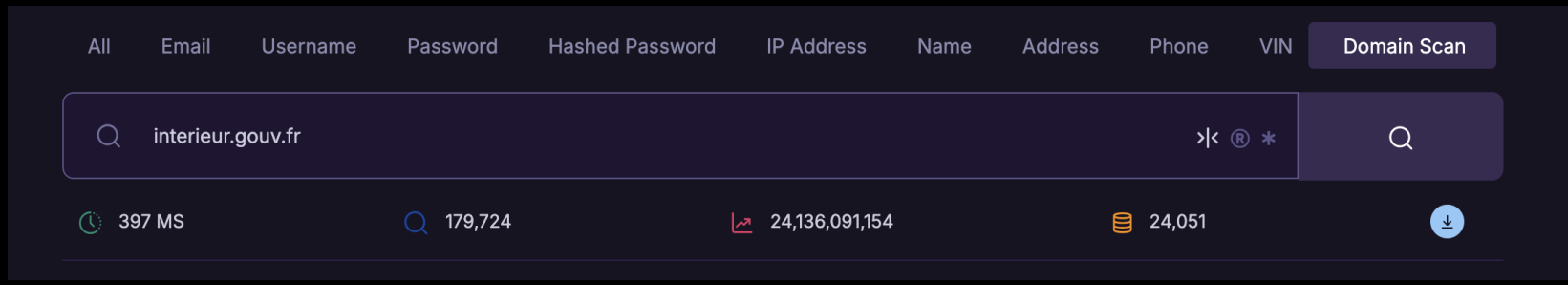
whatsmyname.app

DeHashed

RECHERCHE DANS LES FUITES DE DONNÉES

CE QUE ÇA FAIT

Moteur de recherche dans des milliards d'enregistrements issus de fuites : emails, identifiants, mots de passe, adresses, téléphones. Recherche croisée — un email mène à un mot de passe, qui mène à d'autres comptes.



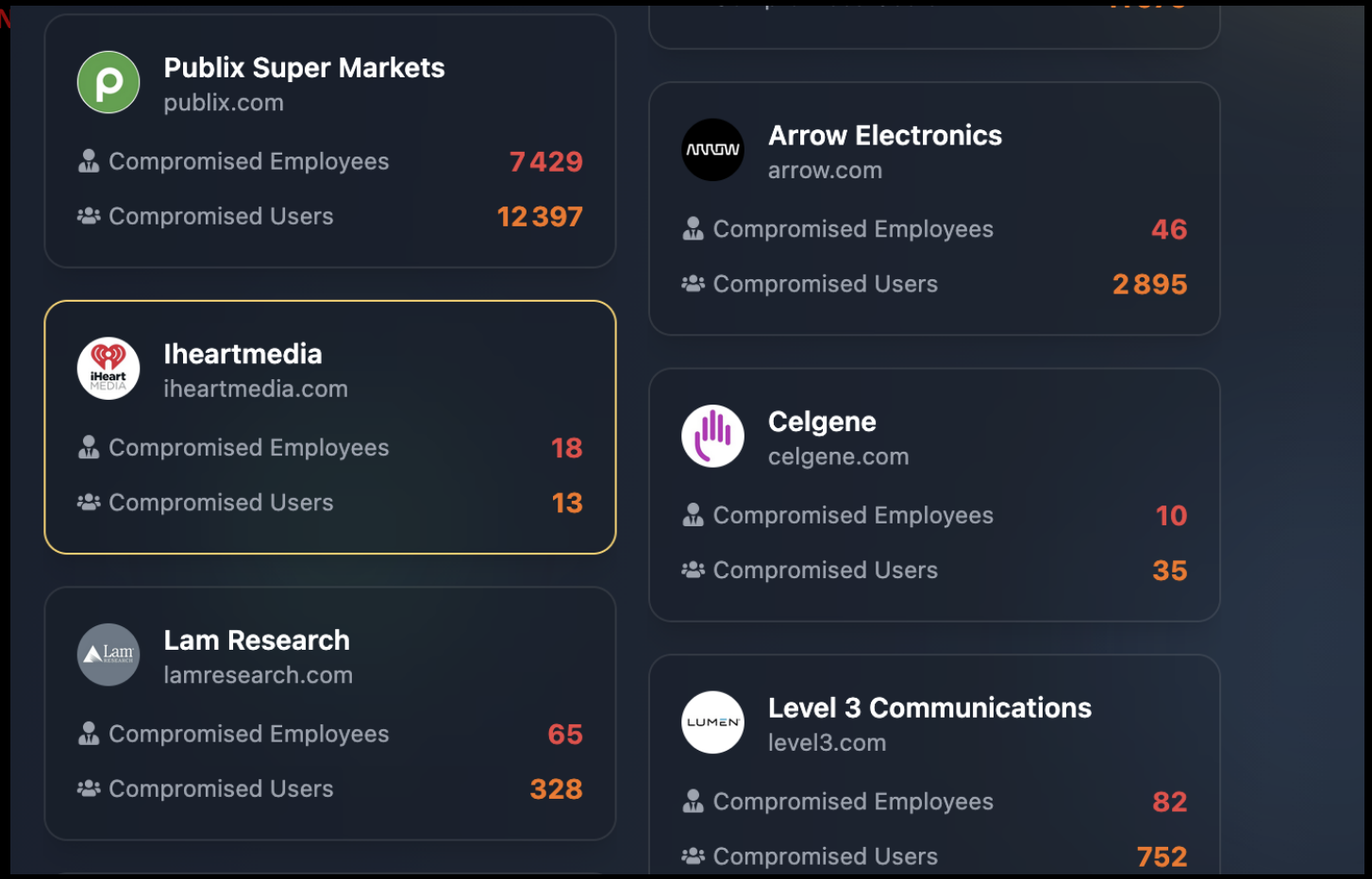
dehashed.com

Hudson Rock — Cavalier

RENSEIGN

CE QUE ÇA FAIT

Base mondiale des machines infectées par des infostealers. Sur un email, un domaine ou un username, révèle si la personne ou l'organisation a été compromise : identifiants volés, machines infectées, dates. API gratuite.



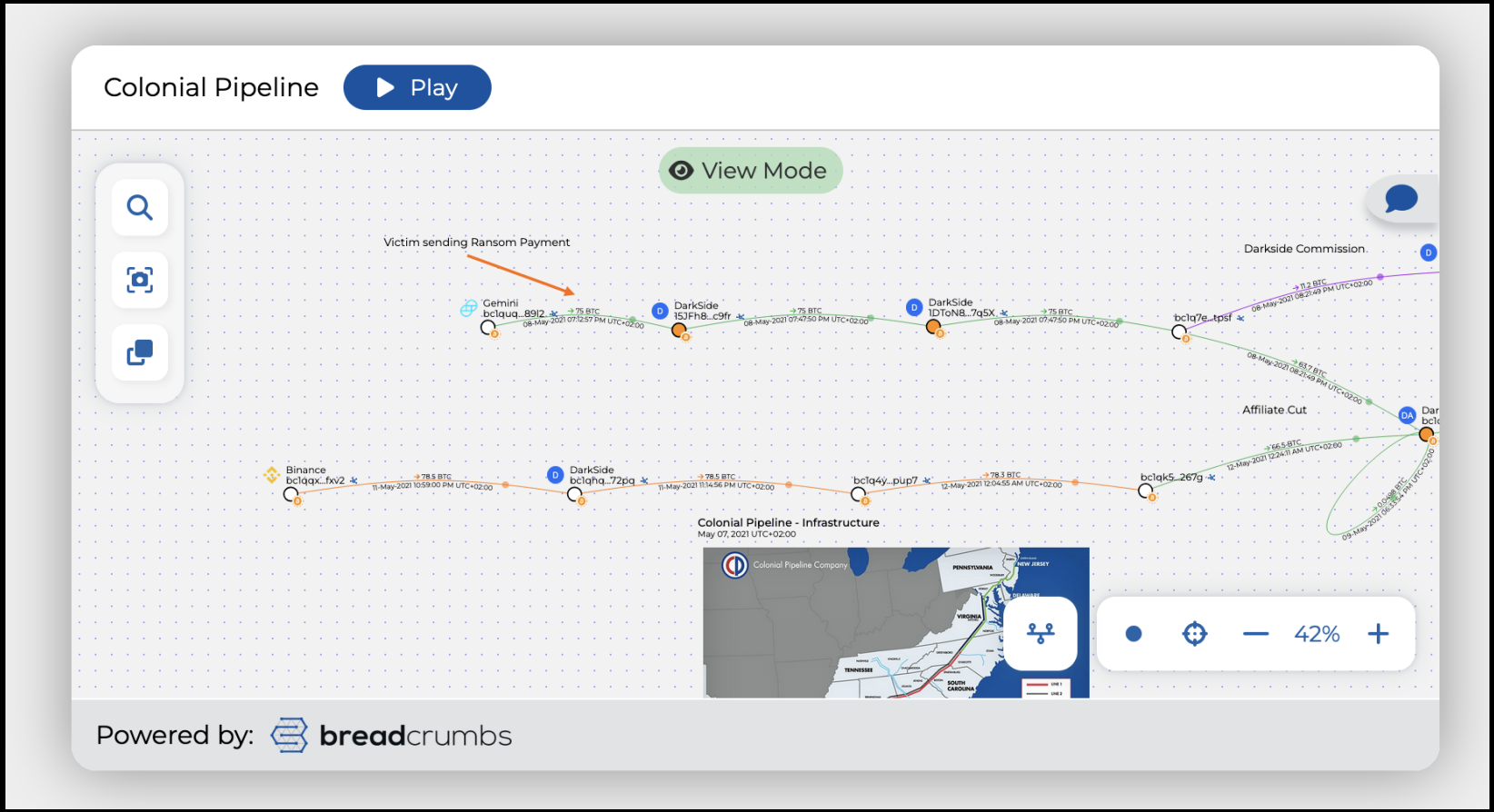
hudsonrock.com

Breadcrumbs

TRAÇAGE DE FLUX CRYPTO VISUEL

CE QUE ÇA FAIT

Outil de suivi des transactions blockchain orienté investigation : visualise les chemins entre wallets, étiquette les acteurs connus, exporte le graphe de flux. Une alternative accessible aux suites professionnelles.



breadcrumbs.app

B

Réflexes côté site web et images

Pas des outils — des gestes. Ce que tout site et toute image trahissent.

Le dépôt .git exposé

RECONSTRUIRE TOUT LE CODE SOURCE

LE RÉFLEXE

Quand un déploiement laisse le dossier .git accessible, on peut reconstituer l'intégralité du code source — et son historique. Les commits révèlent souvent l'email du développeur, des secrets supprimés, la chronologie du projet.

```
$ curl -s https://cible.tld/.git/HEAD
ref: refs/heads/main
$ git-dumper https://cible.tld/.git ./loot
$ cd loot && git log --format='%ae'
dev.scammer@gmail.com
```

/.git/ · git-dumper

Les archives oubliées

LE ZIP DE TOUT LE SITE, À LA RACINE

LE RÉFLEXE

Les développeurs laissent souvent une sauvegarde à la racine : backup.zip, site.tar.gz, www.zip, old/. Téléchargeable d'un clic, elle contient parfois tout le code, les configs et les identifiants. Ton histoire de scammeurs démasqués par un zip entier.

```
# deviner les noms classiques
curl -sI cible.tld/backup.zip
curl -sI cible.tld/site.zip
curl -sI cible.tld/www.tar.gz
HTTP/2 200  ·  content-length: 48 MB
```

backup.zip · site.zip · old/ · www.tar.gz

Les fichiers de configuration

CLÉS ET IDENTIFIANTS EN CLAIR

LE RÉFLEXE

Un .env, un config.php.bak, un wp-config.php~ laissé accessible expose les identifiants de base de données, les clés d'API, les secrets de l'application. L'éditeur a sauvegardé, le serveur l'a servi.

```
curl -s cible.tld/.env
DB_PASSWORD=S3cr3t...
STRIPE_KEY=sk_live_...
MAIL_USER=admin@cible.tld
# on constate. on ne s'en sert pas.
```

.env · config.php.bak · .conf~

Les source maps

DÉ-MINIFIER UNE APPLICATION WEB

LE RÉFLEXE

Les fichiers .js.map, livrés par erreur en production, permettent de reconstituer le code source original — lisible, commenté — d'une application 'minifiée'. On retrouve la logique, les routes d'API, parfois des secrets et des noms internes.

```
# le bundle référence sa source map
//# sourceMappingURL=index.js.map
$ curl -s cible.tld/assets/index.js.map \
   | npx source-map-explorer
→ code original, routes /api/admin, TODO nominatifs
```

.js.map · sourceMappingURL

Les erreurs et stack traces

LE SERVEUR SE CONFIE QUAND IL PLANTE

LE RÉFLEXE

Une erreur PHP ou applicative mal gérée affiche le chemin absolu sur le serveur, la structure des dossiers, la version des composants, parfois une requête SQL ou un identifiant. Provoquer une erreur, c'est faire parler le serveur.

```
Fatal error: in /home/scammer123/www/inc/db.php
Stack trace:
  #0 /home/scammer123/www/...
→ user système, arborescence, version PHP
```

chemins · versions · user système

Les métadonnées et la provenance C2PA

L'IMAGE AVOUE SA FABRICATION

LE RÉFLEXE

EXIF donne l'appareil, la date, parfois le GPS. Le manifeste C2PA, lui, prouve la génération par IA et embarque un certificat de signature — exploitable comme empreinte pour relier tous les visuels d'une même fabrique.

```
$ exiftool logo.png
Claim Generator : OpenAI Media Service
Digital Source  : trainedAlgorithmicMedia
Signature       : SSL.com C2PA (cert)
→ visuel généré par IA, daté, corrélable
```

exiftool · manifeste C2PA

Le favicon comme empreinte

UN HASH QUI DÉPLIE UN RÉSEAU

LE RÉFLEXE

Le favicon d'un site a un hash. Ce hash, requêté sur Shodan ou FOFA, retrouve tous les hôtes servant la même icône — y compris les IP d'origine derrière un CDN. Une icône partagée trahit une fabrique commune.

```
$ python3 -c 'import mmh3,requests,base64;
  print(mmh3.hash(base64.encodebytes(
  requests.get(URL+"/favicon.ico").content)))'
-1234567890
shodan: http.favicon.hash:-1234567890
```

mmh3 · http.favicon.hash

Les identifiants de tracking

RELIER LES SITES D'UN MÊME PROPRIÉTAIRE

LE RÉFLEXE

Un même code Google Analytics, AdSense ou Facebook Pixel sur deux sites trahit un propriétaire commun. Ces identifiants, en clair dans le code source, sont des empreintes de corrélation redoutables.

```
# dans le HTML
gtag("config", "G-AB12CD34")
# qui d'autre utilise ce même ID ?
publicwww.com  "G-AB12CD34"
→ 6 domaines liés au même compte
```

UA- / G- / pixel → publicwww

Le code source et les commentaires

CE QUE LES DÉVELOPPEURS OUBLIENT

LE RÉFLEXE

Le HTML, le JS et les commentaires laissés en clair révèlent chemins de développement, noms internes, environnements de test, TODO nominatifs, anciennes URL. Lire la source brute reste un réflexe payant.

```
<!-- TODO: demander à karim de fix le paiement -->
<!-- old: staging.cible-interne.tld -->
<script src="/js/admin-debug.js">
→ prénom, sous-domaine interne, script admin
```

view-source · Ctrl+U

IA appliquée à l'OSINT

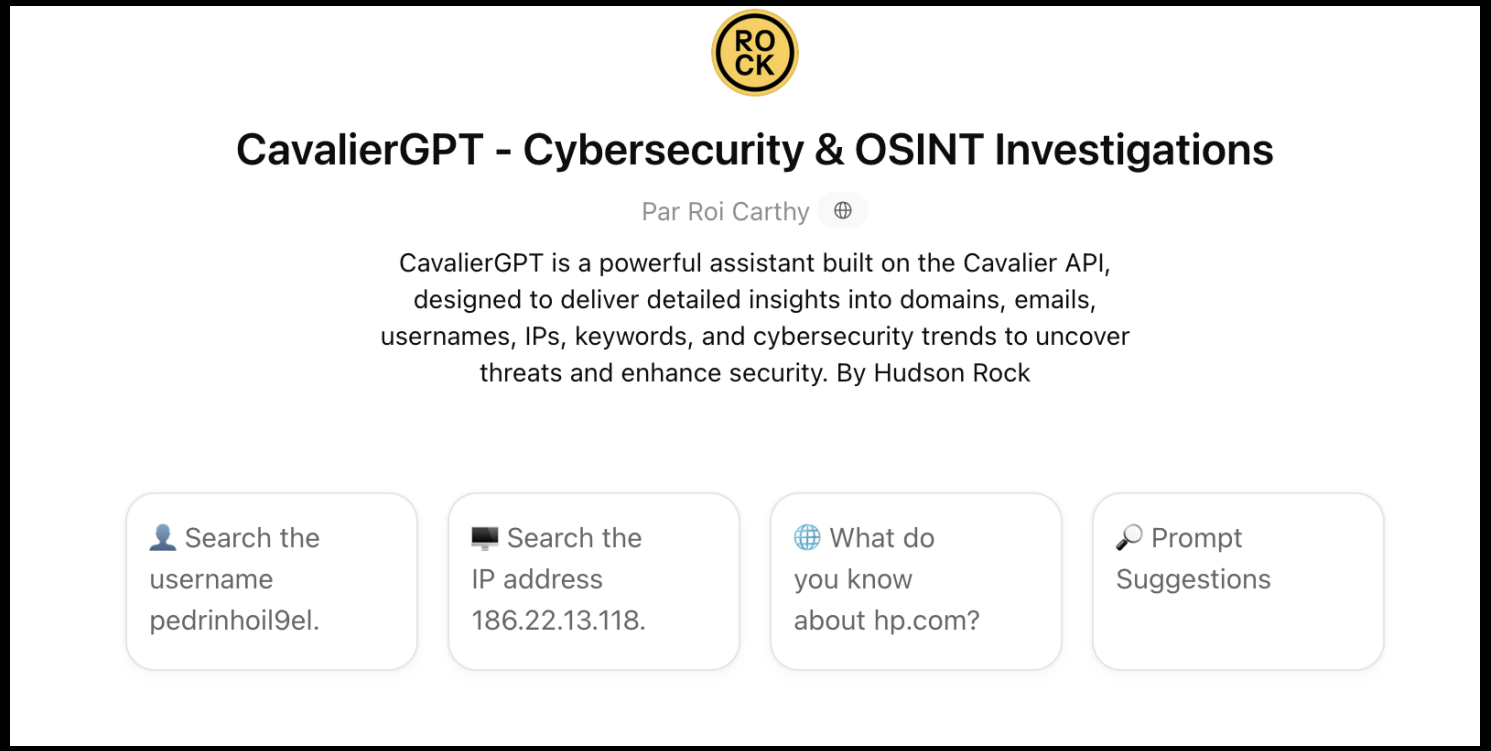
La nouvelle couche : accélérer, transcrire, corréler, géolocaliser.

Cavalier GPT

RENSEIGNEMENT INFOSTEALER EN LANGAGE NATUREL

CE QUE ÇA FAIT

L'interface conversationnelle de la base Hudson Rock : on interroge en langage naturel les compromissions par infostealer d'un email, domaine ou username, et l'IA synthétise machines infectées, identifiants et chronologie.



Disponible dans le GPT Store · hudsonrock.com

Transcription et analyse média par IA

FAIRE PARLER L'AUDIO

CE QUE ÇA FAIT

Les modèles multimodaux transcrivent un audio, traduisent, résument une vidéo, décrivent une scène.
Des heures d'enregistrements ou de captations deviennent un texte cherchable en minutes.

`Whisper (open source) · API multimodales`

D

Créer vos propres outils

Quand aucun outil ne fait le job, on l'écrit. Quelques dizaines de lignes suffisent.

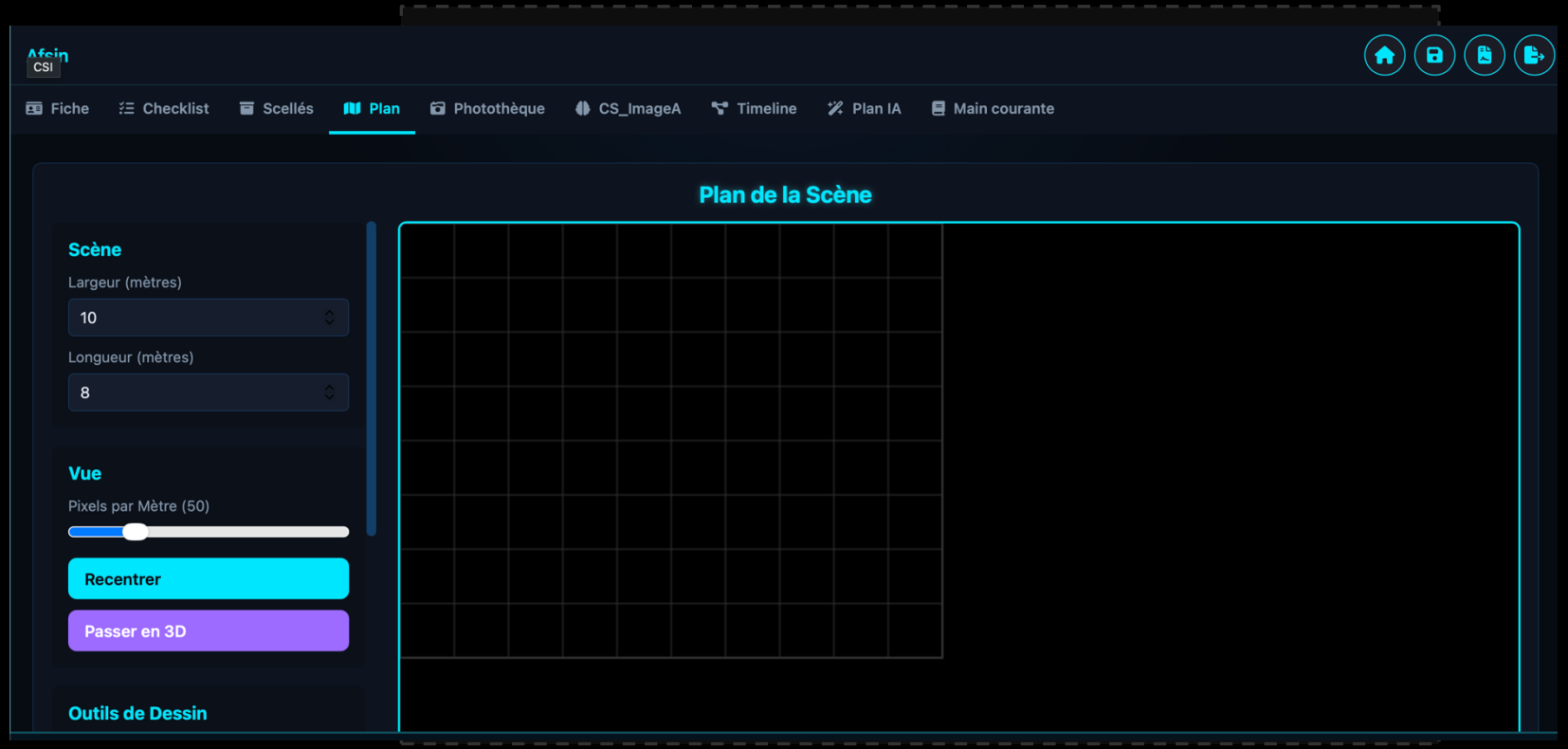
Le principe

POURQUOI CODER SES PROPRES OUTILS

LE PRINCIPE

Les outils publics sont génériques et parfois bridés. Un script maison fait exactement ce dont l'enquête a besoin, automatise le répétitif, et reste sous votre contrôle — aucune donnée ne part chez un tiers.

Trois briques suffisent : récupérer, analyser, enrichir. Assemblées, elles couvrent la majorité des besoins d'automatisation.



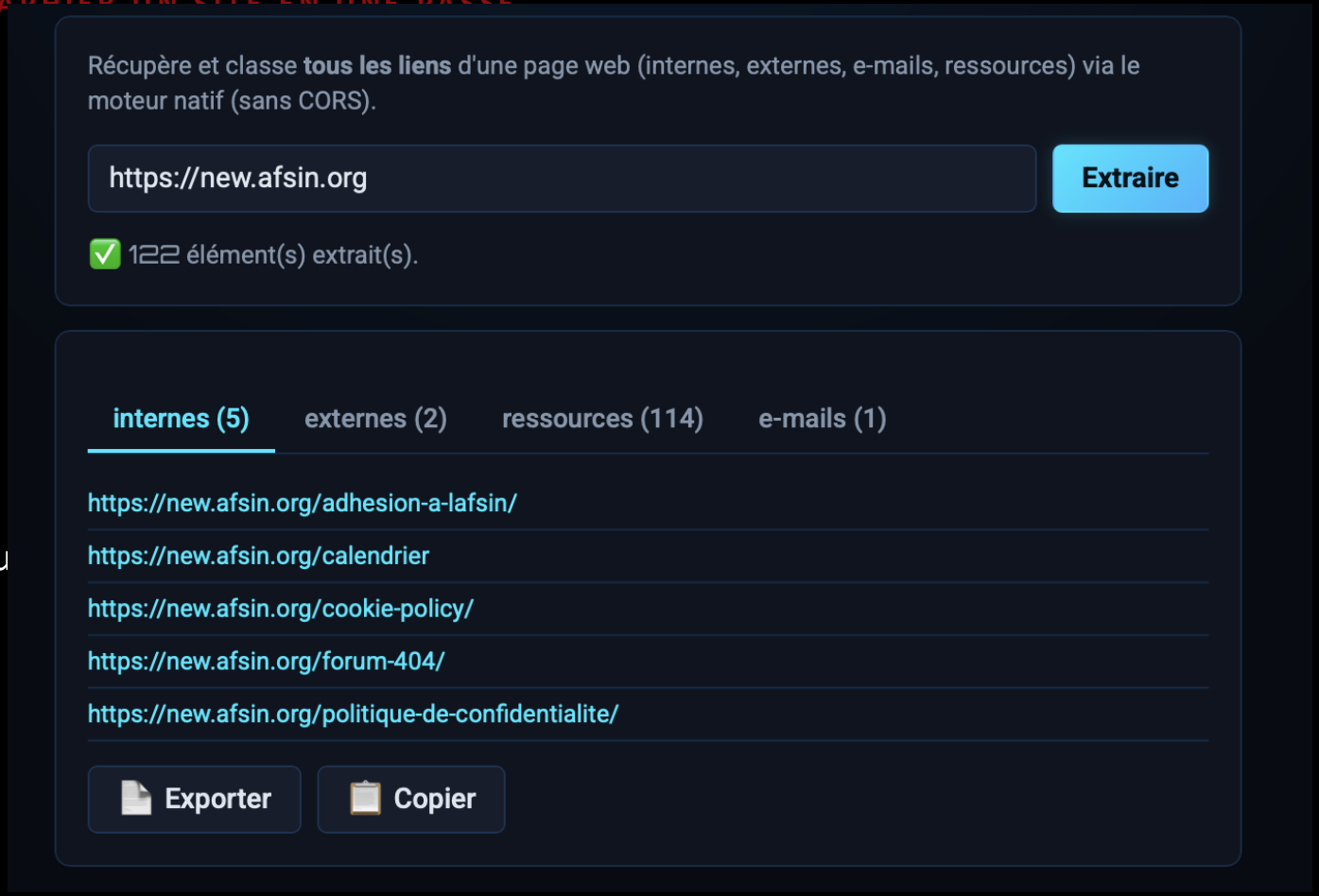
`requests · BeautifulSoup · exiftool · API IA`

Extracteur de liens

CARTOGRAPHIER UN SITE EN UNE PASSE

LE PRINCIPE

Un crawler parcourt un site et extrait toutes les URL : pages internes, liens externes, adresses mailto, numéros tel. En quelques secondes, on a la structure complète et les points de contact d'une cible — souvent un compte de paiement ou des emails cachés en pied de pages secondaires.



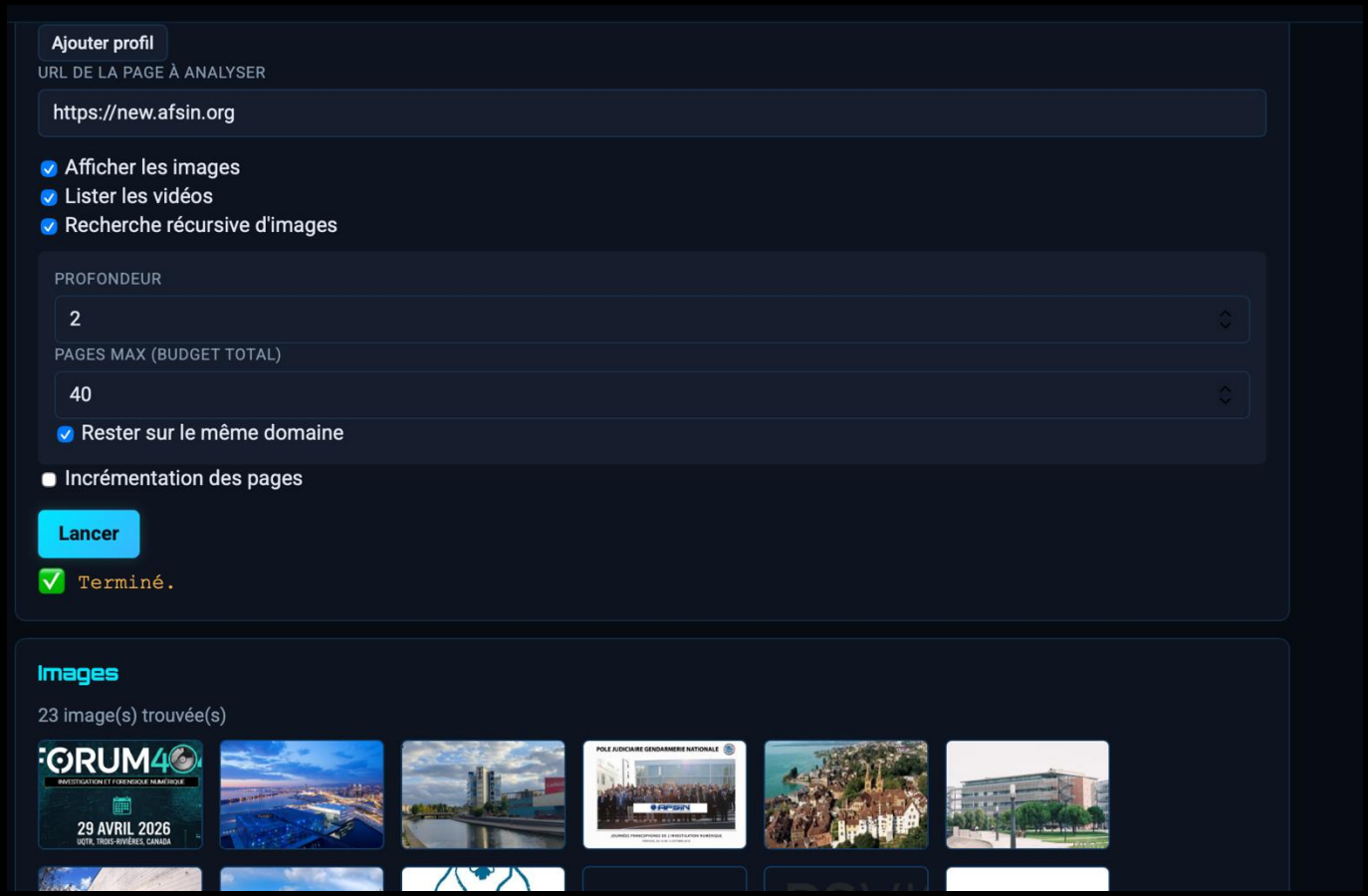
`requests · BeautifulSoup · urllib.parse`

Extracteur automatique d'images

RÉCUPÉRER TOUS LES VISUELS D'UNE CIBLE

LE PRINCIPE

Un script collecte automatiquement toutes les images d'un site ou d'un profil : balises img, arrière-plans CSS, ressources liées. On constitue en une passe le corpus visuel complet — là où le clic-droit manuel prendrait une heure — prêt pour l'analyse de métadonnées en masse.



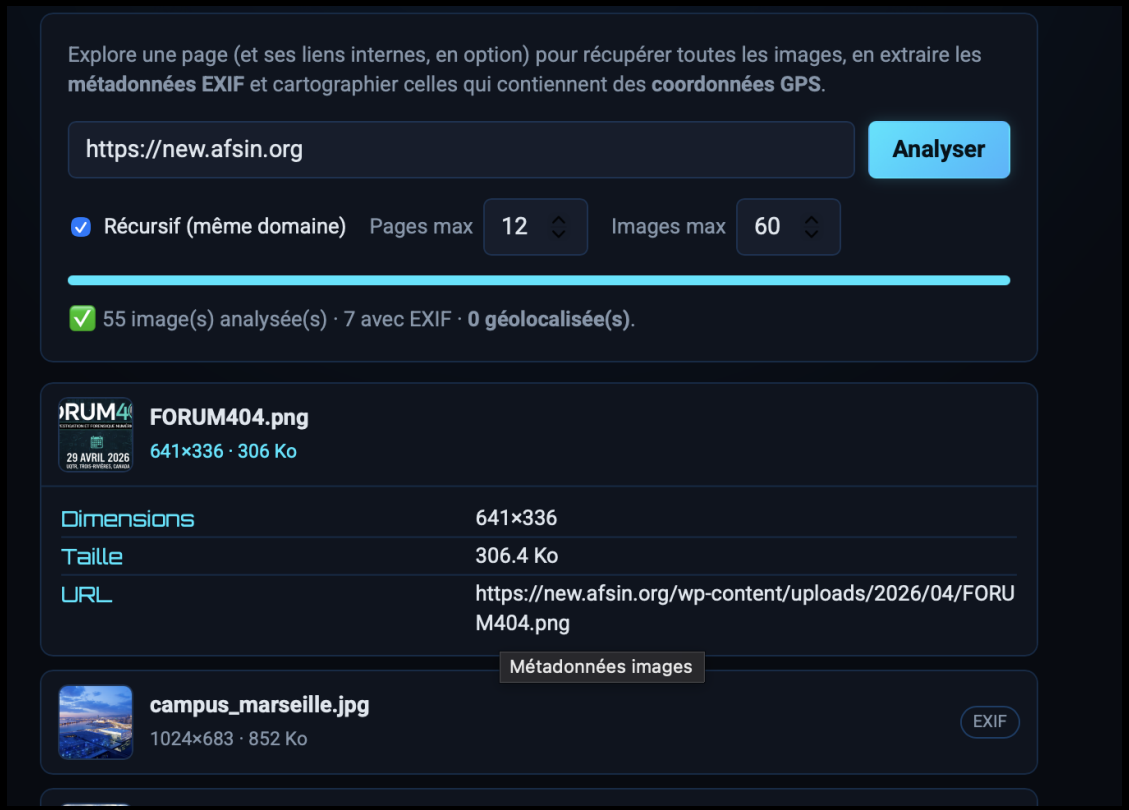
`requests · BeautifulSoup · re`

Analyse de métadonnées en masse

TOUTES LES IMAGES D'UN SITE, PASSÉES AU CRIBLE

LE PRINCIPE

Projet maison : on enchaîne l'extraction d'images d'un site et l'analyse automatique de leurs métadonnées — EXIF, GPS, logiciel créateur, provenance C2PA. Un site entier audité d'un coup, pour repérer fuites et fabrications.



requests · exiftool · traitement par lot

Appel et transcription par IA

FAIRE PARLER LES RÉPONDEURS

LE PRINCIPE

Projet maison : on fournit une liste de numéros, un service de téléphonie les appelle, enregistre l'annonce du répondeur, puis une IA transcrit l'audio et détermine si l'annonce révèle des informations sur le titulaire.

Twilio · Whisper · API IA

Croque-escrocs

🔥 LIQUIDATION TOTALE — Stocks limités ! Jusqu'à -70% sur toute la gamme · Livraison OFFERTE dès 49€

⚡VoltKraft

AccueilBatteriesChargeursPowerbanksProContact

🛒 Panier (0)

✓ Paiement 100% sécurisé✓ Garantie 5 ans✓ Expédition sous 24h✓ +12 000 clients satisfaits

L'énergie premium au prix le plus bas du marché.

Batteries lithium haute performance, chargeurs rapides et powerbanks certifiés. Importés directement de nos usines — sans intermédiaire.

-70% AUJOURD'HUI

L'offre expire dans : 02:14:09

Nos meilleures ventes

Sélection du moment — quantités très limitées

-65%

PowerCell X9 — 20 000 mAh

★★★★★ (842)

29,90€ 89,90€

⚠ Plus que 3 en stock

-70%

Chargeur UltraFast 120W

★★★★★ (1 203)

24,90€ 82,90€

⚠ Plus que 5 en stock

-60%

Powerbank Solar Pro

★★★★☆ (597)

34,90€ 87,90€

⚠ Plus que 2 en stock

-68%

Batterie Garage 48V

★★★★★ (311)

59,90€ 189,90€

⚠ Bientôt épuisé

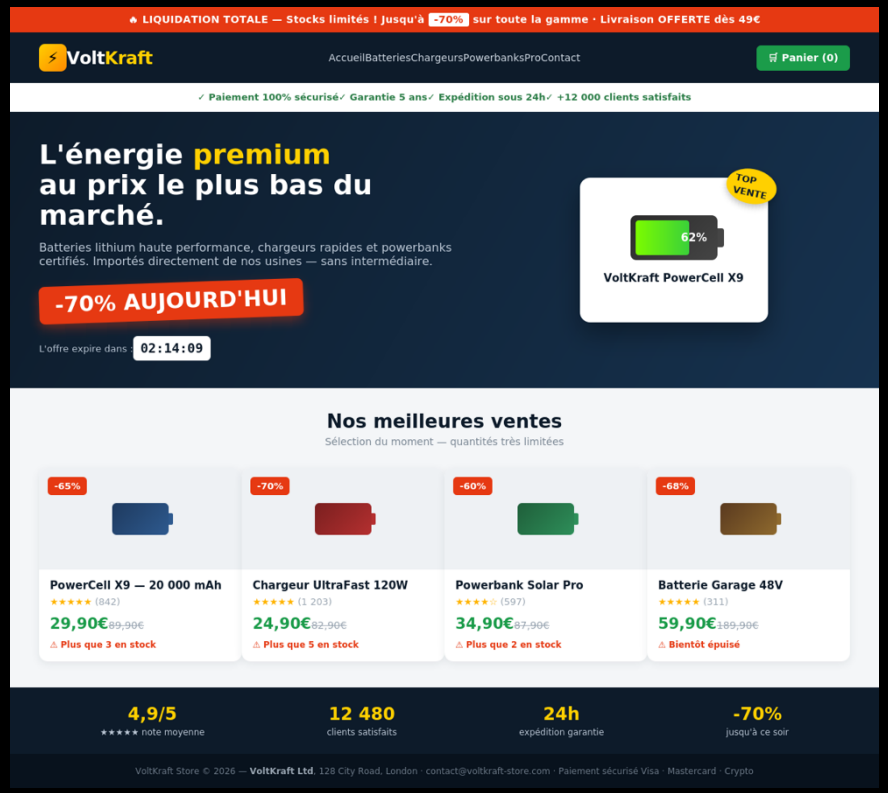
4,9/5	12 480	24h	-70%
★★★★★ note moyenne	clients satisfaits	expédition garantie	jusqu'à ce soir

VoltKraft Store © 2026 — **VoltKraft Ltd**, 128 City Road, London · contact@voltkraft-store.com · Paiement sécurisé Visa · Mastercard · Crypto

VoltKraft

Une fausse boutique de batteries « haut de gamme » à prix cassés. Encaisse, ne livre jamais.

Paiement carte et crypto. Beaucoup de victimes. Juste une URL.



ACTE 1

Le terrain

On commence sans rien toucher. La réputation confirme l'arnaque — mais une trouvaille change l'échelle : la structure de la page a déjà été vue ailleurs.

Le site n'est peut-être pas seul.

OUTILS MOBILISÉS

- `Scamdoc`
- `signal-arnaques`
- `urlscan.io`

→ l'historique urlscan montre la même page sous deux autres domaines. On creuse l'infra.

ACTE 2

L'infrastructure

Domaine de cinq semaines, derrière Cloudflare : l'IP réelle est masquée. Mais l'historique DNS garde la mémoire des premières heures, avant l'activation du CDN.

Dans cette fenêtre, l'IP d'origine était exposée.

OUTILS MOBILISÉS

- whois / dig
- Validin
- DNSDumpster
- crt.sh

→ Validin livre l'IP d'origine brute. crt.sh sort un sous-domaine old. oublié, encore en ligne.

ACTE 3

Le site se déshabille

Le sous-domaine old. sert un listing ouvert : un backup.zip de 60 Mo, le kit complet. À la racine, un .git exposé. git-dumper reconstruit le dépôt — et les commits lâchent l'email du développeur.

OUTILS MOBILISÉS

- `.git → git-dumper`
- `backup.zip`
- `fichier .env`
- `favicon hash`
- `ID Analytics G-…`

→ email dev récupéré, ID de tracking et favicon notés comme empreintes de corrélation.

ACTE 4

Un site devient un réseau

Les empreintes uniques croisées sur plusieurs moteurs convergent. L'IP d'origine, le favicon et l'ID de tracking pointent les mêmes cibles.

La fabrique apparaît : six boutiques jumelles, un seul opérateur.

OUTILS MOBILISÉS

- FOFA
- ZoomEye
- Netlas
- PublicWWW

→ intersection des trois méthodes = une grappe de six sites. On passe d'un scam à un sériel.

ACTE 5

Les images parlent

On aspire les visuels des six sites et on les passe au crible. La plupart sont propres — mais quelques-uns ont échappé au nettoyage.

Un EXIF GPS oublié, un logo signé par l'IA, un arrière-plan géolocalisable.

OUTILS MOBILISÉS

- `Extracteur d'images (maison)`
- `Analyse métadonnées (maison)`
- `exiftool / C2PA`
- `GeoSpy · Picarta`

→ GPS sur trois photos produit, logo prouvé généré par IA, entrepôt géolocalisé en Europe de l'Est.

ACTE 6

On remonte à l'humain

L'email du .git devient le fil. Une plateforme d'agrégation déballe les comptes liés, un pseudo récurrent, une timeline. On déroule pseudo et fuites — puis vient le coup de grâce.

OUTILS MOBILISÉS

- OSINT Industries
- WhatsMyName · Maigret
- DeHashed
- Hudson Rock / Cavalier

→ l'opérateur a lui-même une machine infectée par infostealer. L'arroseur arrosé.

ACTE 7

Le téléphone

Le numéro sorti des fuites colle au profil géographique. On le qualifie, puis on le fait parler : un appel automatisé capte l'annonce du répondeur, qu'une IA transcrit et analyse.

OUTILS MOBILISÉS

- `PhoneInfoga`
- `Twilio (maison)`
- `Whisper + IA`

→ l'annonce du répondeur contient un prénom et un nom. Ils recoupent le GitHub de l'acte 6.

ACTE 8

L'argent

Une partie des victimes a payé en crypto. Le wallet n'est pas une impasse :
Les sauts successifs se suivent jusqu'à un point où l'anonymat se brise.

OUTILS MOBILISÉS

- Arkham
- Breadcrumbs

→ les fonds convergent vers un dépôt sur un exchange identifié — soumis à obligation d'identification.

ACTE 9

Le faisceau

LA CHAÎNE, D'UN BOUT À L'AUTRE

URL → réputation → IP d'origine → grappe → images (GPS/C2PA) → email → identité → téléphone → argent → **faisceau**

CE QU'ON A ÉTABLI

Un opérateur unique derrière six boutiques jumelles — même template, même backend, identité visuelle générée par IA. Relié à un email et un pseudo exposés par ses propres négligences (.git, fuites, infostealer), un téléphone confirmé à l'oral, un wallet traçable jusqu'à un exchange identifiable.

Aucune intrusion. On signale : abuse hébergeur, processeurs de paiement, exchange, autorités. L'OSINT documente et transmet — il ne franchit jamais la porte.

Merci