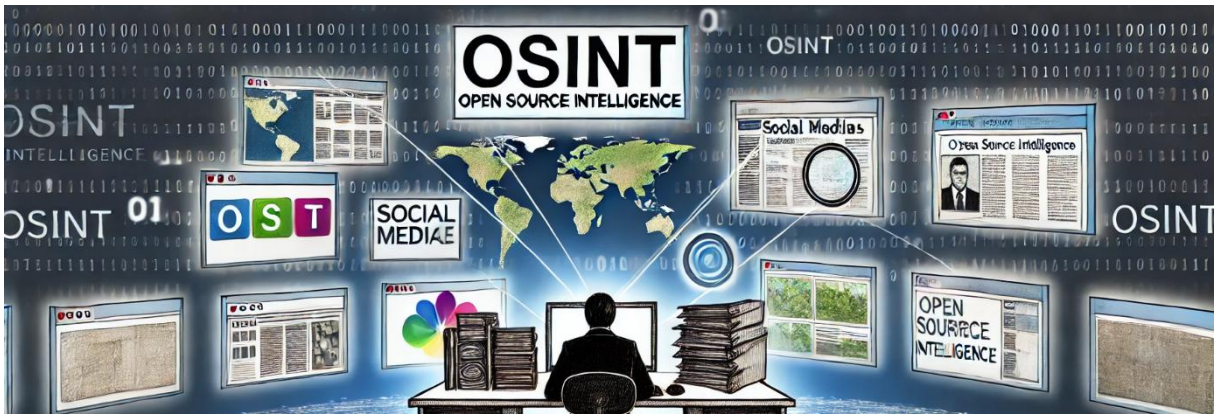


Introduction à l'OSINT

Comprendre la **collecte et l'analyse d'informations publiques**.

OSINT (Open-Source INTelligence)

Max Grégoire (contact@mgregoire.be)

À Anne-Marie

Un grand merci à toi.

CERTIFICATE OF COMPLETION

This certifies that

MAX GRÉGOIRE

has completed Open-Source Intelligence (OSINT) Fundamentals

Date: 01/06/2025 | 9 CEU Hours

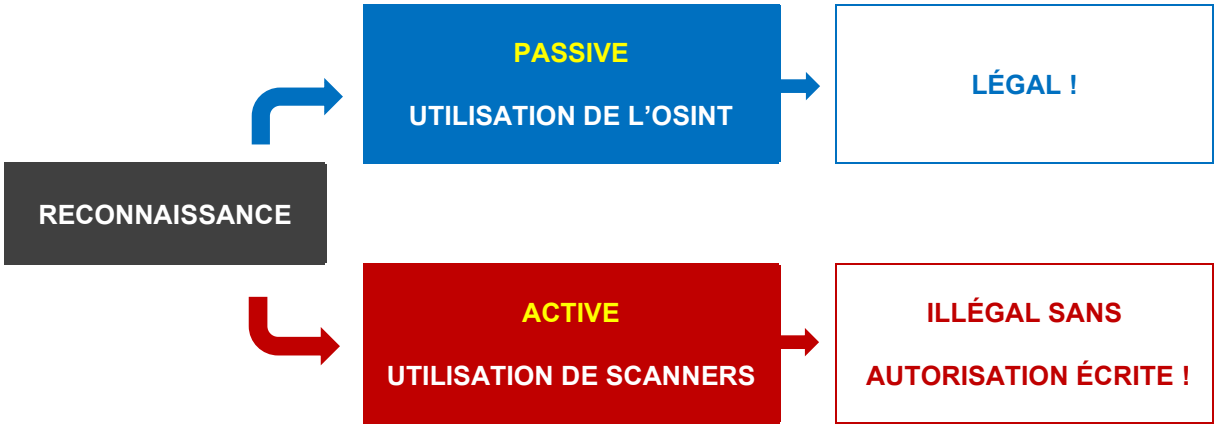
Certificate ID: cert_s9ddcx6c

HEATH MAVERICK ADAMS
Founder & CEO

TCM SECURITY

Table des matières

OSINT



OSINT (EN) / ROSO (FR)

La reconnaissance en bref

LA RECONNAISSANCE (*information gathering*, collecte d'informations)

Reconnaissance passive

Pas d'interaction avec la cible

OSINT
Renseignement d'origine sources ouvertes

WHOIS
Enregistrements de domaines

Google Dorking
Opérateurs avancés de Google

Enregistrements DNS
A, AAAA, CNAME, MX, TXT, …

Machines connectées
Moteur de recherche Shodan

Data leak / Data Breach
Have I Been Pwned ?, DeHashed

Wayback Machine
Versions d'un site web dans le passé

Moins précis mais légal

Reconnaissance active

Interaction avec la cible

SCAN de ports
Nmap, ZenMap

SCAN de vulnérabilités
OpenVAS, Nessus, ZAP

Banner Grabbing
Analyse de bannière

Brute force de sous-domaines
Avec outil local ou en ligne

Brute force de répertoires
Dirbuster

Plus précis mais illégal sans autorisation écrite

Le Google Dorking utilisé à des fins malveillantes sera plutôt appelé Google Hacking !

COLLECTE D'INFORMATIONS

COLLECTE PASSIVE

- ➔ Pas d'interaction avec la cible
- ➔ C'est notamment l'OSINT
- ➔ Les informations collectées se situent dans des sources accessibles publiquement
- ➔ Les informations collectées peuvent être incomplètes ou obsolètes
- ➔ Cette collecte est impossible à détecter pour la cible

OUTILS :

- ✓ Google Dorking
- ✓ Shodan
- ✓ Censys
- ✓ WHOIS
- ✓ TheHarvester
- ✓ Maltego
- ✓ Recon-ng

COLLECTE SEMI-PASSIVE

- ➔ On interagit avec la cible, en imitant un trafic réseau normal
- ➔ Via le téléchargement de documents
- ➔ Via des requêtes DNS
- ➔ Via l'analyse des métadonnées
- ➔ Cette collecte est difficile à détecter pour la cible

OUTILS :

- ✓ Exiftool (métadonnées)
- ✓ Metagoofil (métadonnées)
- ✓ FOCA (métadonnées)
- ✓ Wireshark
- ✓ TCPDump
- ✓ DNS Dumpster
- ✓ Centralops.net

COLLECTE ACTIVE

- ➔ On interagit avec la cible de manière voyante
- ➔ Cette collecte est facile à détecter pour la cible

OUTILS :

- ✓ Nmap (scan de ports, de services)
- ✓ Zenmap (scan de ports, de services)

OSINT : définition et exemples

L'OSINT (Open-Source Intelligence), ou en français ROSO (renseignement d'origine sources ouvertes) est la collecte et l'analyse d'informations publiquement accessibles pour produire du renseignement.

L'OSINT est très utile pour :

- ➔ **Les journalistes (qui cherchent à acquérir ou à vérifier des informations)**
- ➔ **Forces de l'ordre / enquêteurs (personne disparue, criminel recherché, …)**
- ➔ **Entreprises (pour l'évaluation de la concurrence, pour évaluer les antécédents d'un postulant, …)**
- ➔ **Les experts en cybersécurité.**
- ➔ **Le renseignement militaire.**
- ➔ **Les hackers éthiques (toujours dans un but légal)**
- ➔ **Les cybercriminels (dans un but illégal)**
- ➔ **Vous et moi, dans un but privé : par exemple, pour vérifier d'une identité en ligne sur un site de rencontre (pour ne pas être victime d'une arnaque sentimentale), pour rechercher un proche dont on a perdu la trace, …**

On peut distinguer :

- ✓ **Website OSINT**
- ✓ **SOCMINT**
- ✓ **People OSINT**
- ✓ **Username OSINT**
- ✓ **Email OSINT**
- ✓ **Phone Number OSINT**
- ✓ **Gmail OSINT**
- ✓ **Image OSINT**
- ✓ **Video OSINT**
- ✓ **Maps OSINT**
- ✓ **IoT OSINT**
- ✓ **Dark Web OSINT (concerne les forums et marketplaces du darknet)**
- ✓ **Deep Web OSINT (concerne les pages non indexées par Google et Bing)**
- ✓ **Etc.**

Nous verrons quelques exemples dans les pages qui suivent…

Website OSINT

WHOIS	Affiche les informations d'enregistrement d'un nom de domaine. **https://whois.domaintools.com/, https://who.is/, …**
crt.sh	Certificats SSL/TLS émis pour un domaine et sous-domaines **https://crt.sh/**
SSL Server Test	Analyse la configuration d'un serveur web SSL. **https://www.ssllabs.com/ssltest/**
Securityheaders	Analyse les en-têtes de réponse HTTP. **https://securityheaders.com/**
Builtwith	Affiche les technologies utilisées par un site. **https://builtwith.com/**
Subdomain finder	Trouve les sous-domaines d'un domaine donné. **https://osint.sh/subdomain/, …**
Wayback machine	À quoi ressemblait le site dans le passé. **http://web.archive.org/**

Trouver des sous-domaines d'Amazon avec Google :

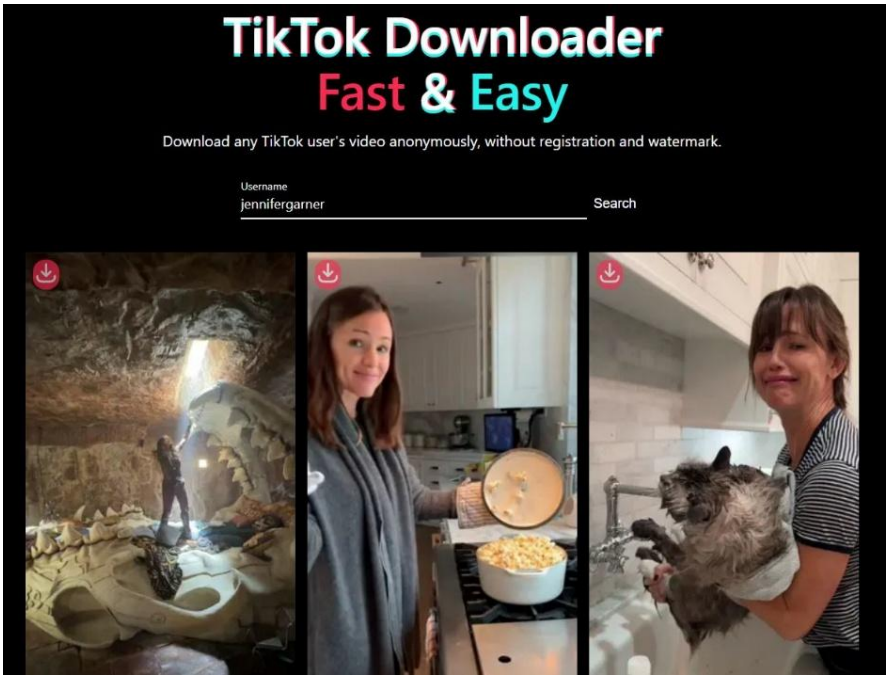
site:amazon.com -inurl:https://amazon.com

Trouver des sous-domaines de Google avec Yandex :

rhost:com.google.*

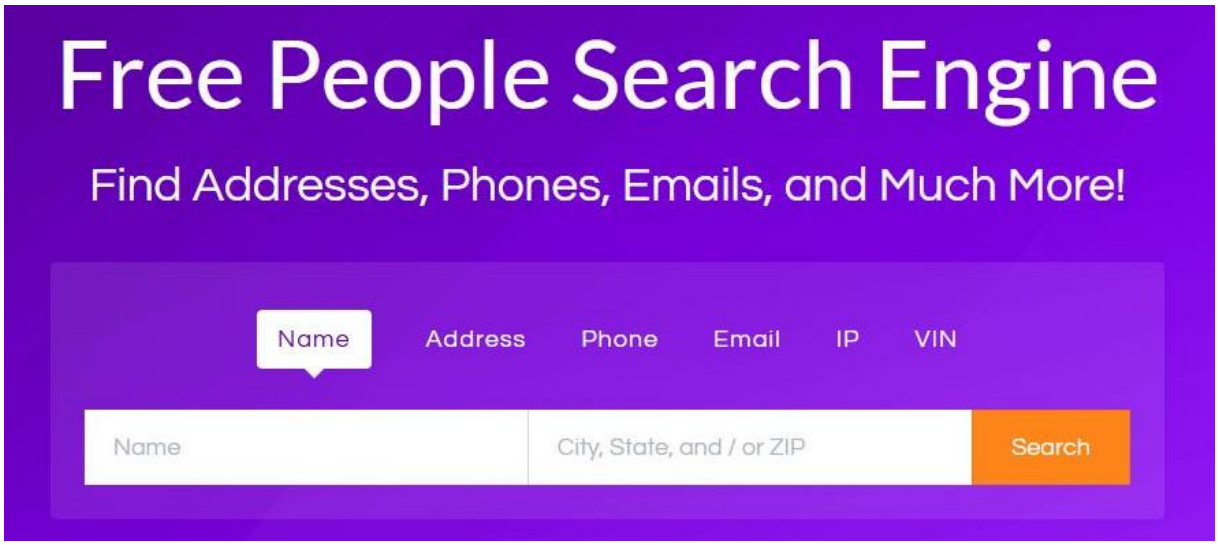
SOCMINT (Social Media Intelligence)

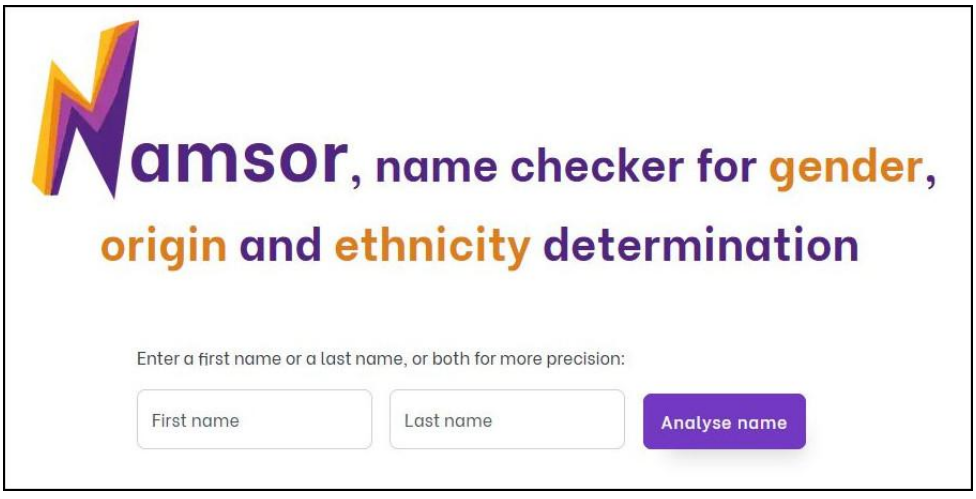
Facebook	Trouver l'User ID d'un compte Facebook : **https://lookup-id.com/, https://findidfb.com/** ou code source **: chercher « id »**	**Barrack Obama : 100044322825129**
Instagram	Trouver l'User ID d'un compte Instagram : **https://fameswap.com/tool-instagram-user-id** ou code source **: chercher « profile_id »**	**@barackobama : 10206720**
X (Twitter)	Trouver l'ID d'un compte X : **https://twiteridfinder.com/** ou inspecter l'élément sur l'username (**@...**) puis rechercher **profile_banners.**	**@BarackObama : 813286**
LinkedIn	Afficher le code source : rechercher « **member:** ». S'il y en a deux : l'un est le vôtre (celui du visiteur) et l'autre est celui du compte visité.	**Barrack Obama : 11932467**
TikTok	Télécharger les vidéos d'un compte TikTok : **https://freetik.co/, https://ssstik.io/, https://snaptik.kim/**	
Pour trouver les ID sur différents médias sociaux	**https://commentpicker.com/**	



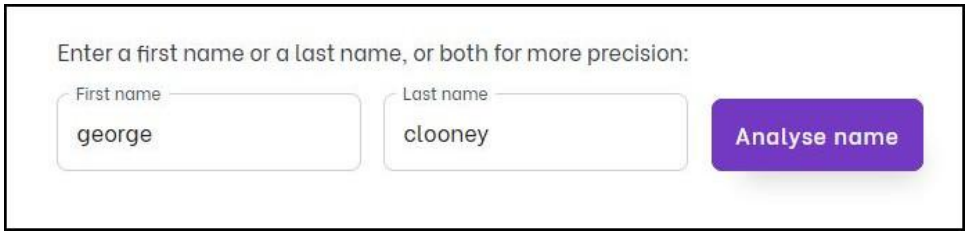
People OSINT

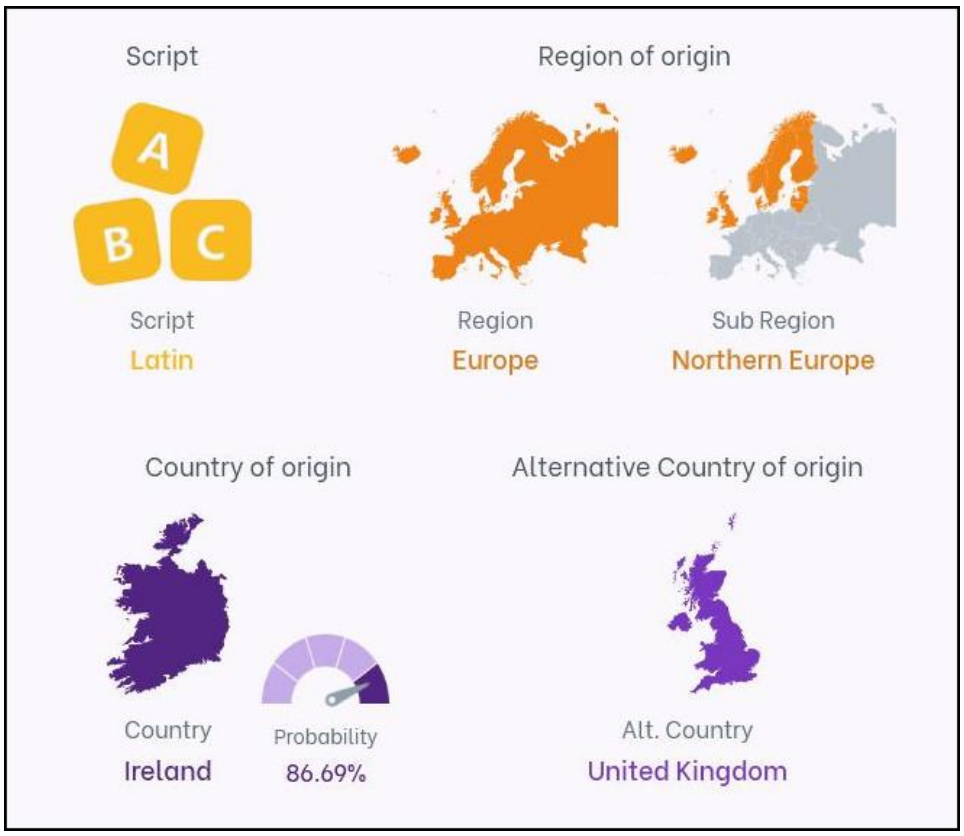
https://www.idcrawl.com/	Informations sur une personne (USA et hors USA)
https://www.truepeoplesearch.com/	Informations sur une personne vivant aux USA. Il faut utiliser un serveur VPN situé aux USA pour pouvoir visiter le site !
https://webmii.com/	Informations sur une personne.
https://thatsthem.com/	Informations sur une personne.
https://namsor.app/	Quelle est l'origine et l'ethnicité d'un nom.





Recherchons avec Namsor l'origine du nom de George Clooney :





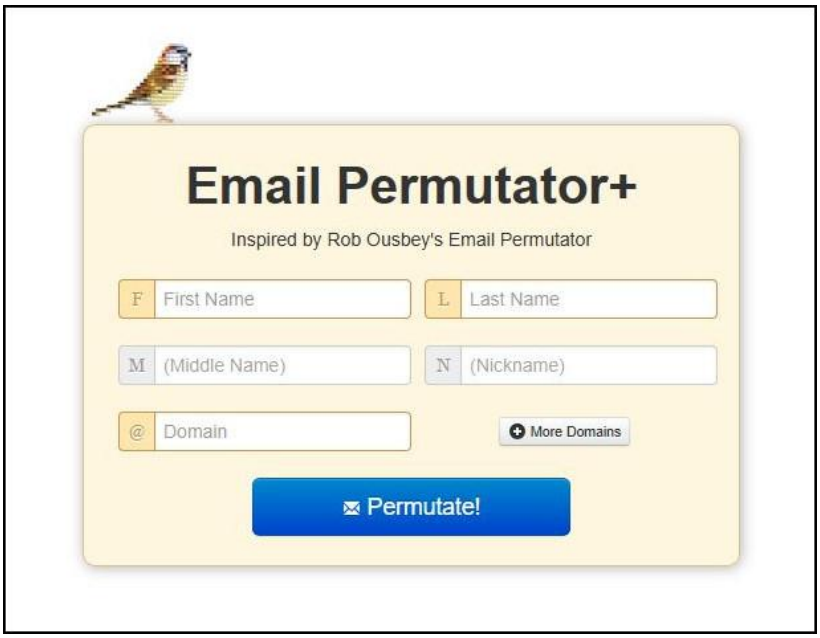
L'acteur américain est bien d'origine irlandaise.

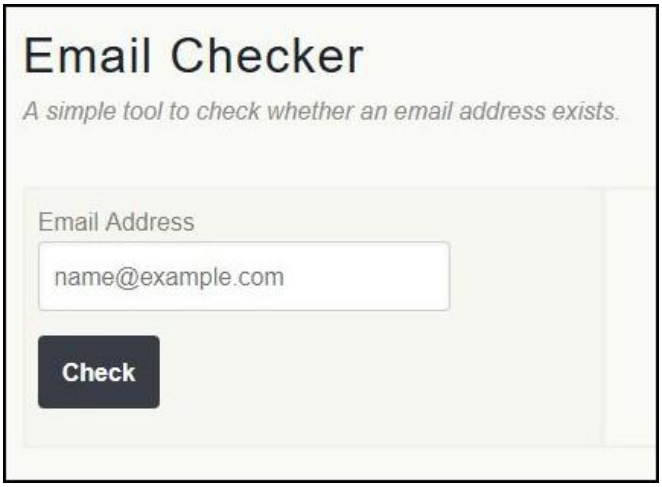
Username OSINT

https://whatsmyname.app/	Username OSINT
https://www.idcrawl.com/username-search	

Email OSINT

http://metricsparrow.com/toolkit/email-permutator/	Génère une liste d'email possible à partir du nom, du prénom et d'un domaine.
https://email-checker.net/	Un email existe-t-il ?
https://hunter.io/	Trouver et vérifier des emails
https://mxtoolbox.com/EmailHeaders.aspx	Analyse de l'en-tête d'un email.
https://www.whatismyip.com/email-header-analyzer/	
https://www.iptrackeronline.com/email-header-analysis/	
https://emailrep.io/	Réputation d'un email.

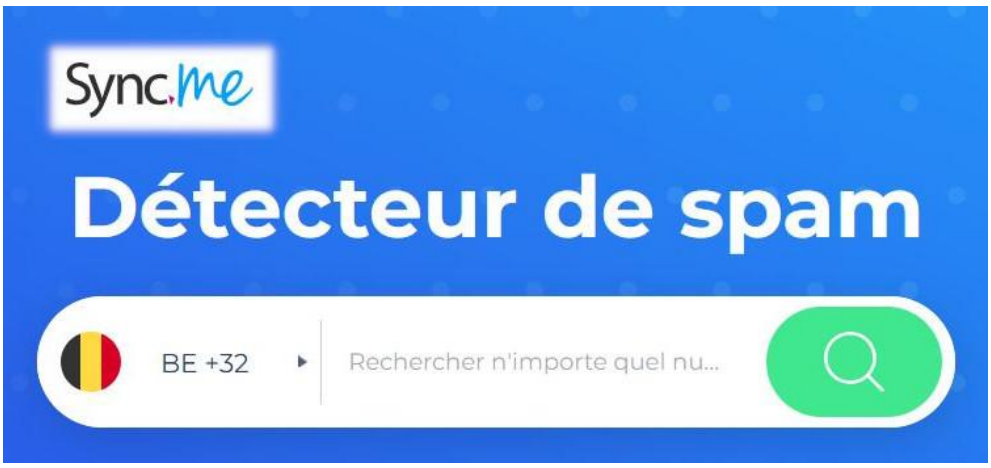




Phone Number OSINT

https://www.ipqualityscore.com/free-phone-number-lookup	Analyse d'un numéro de téléphone international
https://www.comfi.com/abook/reverse	
https://www.numberingplans.com/index.php?page=analysis&sub=phonenr	
https://www.searchyellowdirectory.com/	Pages blanches américaines
https://www.truecaller.com/	Identification de l'appelant et blocage des spams. Fonctionne en ligne ou comme application pour Android et iOS.
https://sync.me/fr/	Blocage des spams. Fonctionne en ligne ou comme application pour Android et iOS.

Des applications pour smartphones comme Truecaller et Sync.me sont très pratiques mais ont un inconvénient majeur : elles transfèrent toute votre liste de contacts sur leur serveur. Pour éviter cette violation de votre vie privée, je recommande de les utiliser en ligne avec un compte Google secondaire ayant une liste de contacts limitée.



Pour rechercher un numéro de téléphone parisien avec Google, la requête sera : "0123456789" OR "01 23 45 67 89" OR "01.23.45.67.89" OR "01-23-45-67-89"

Pour rechercher le numéro de téléphone en France de Jean Dupond, les requêtes pourraient être : "jean dupond" France "téléphone" OR "contact" OR "numéro"
ou "jean dupond" "+33" OR "33" OR "0033"

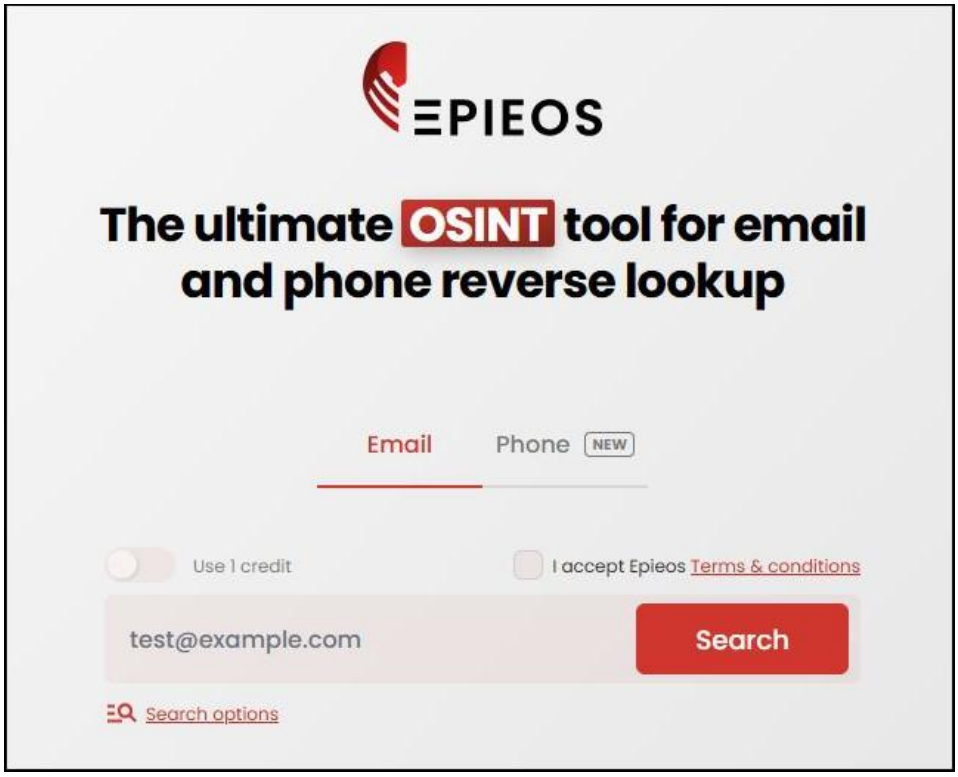
Pour rechercher le numéro de téléphone londonien de John Smith avec Google, la requête sera : "john smith" "+44 20" OR "+4420"

ASTUCE

Trouver des comptes liés à un email ou numéro de téléphone sur de nombreux sites

Les applications ci-dessous seront utiles :

- ➔ **Epieos :** **https://epieos.com/** **Version gratuite et payante**
- ➔ **Castrickclues :** **https://castrickclues.com/** **Version gratuite et payante**
- ➔ **OSINT Industries :** **https://app.osint.industries/** **Version payante**



Gmail OSINT

https://gmail-osint.activetk.jp/	Obtenir le GAIA ID d'un compte Google

Chaque compte Google est associé à un identifiant : le GAIA ID. L'outil GHUNT permet d'obtenir diverses informations sur un compte Google (dont le GAIA ID). Le site ci-dessus est une version en ligne de GHUNT.

Email : john@gmail.com

Gaia ID : 100050327862319454221

Ghunt sur Github :

https://github.com/mxrch/GHunt

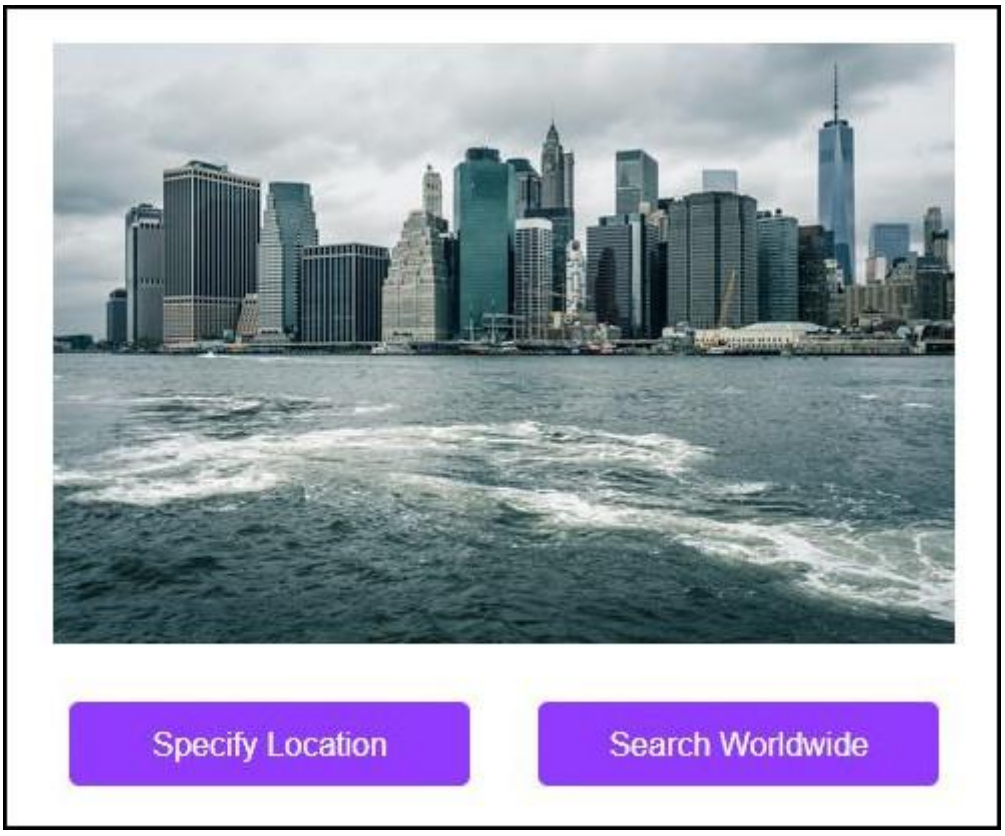
Image OSINT

https://images.google.com/	Moteurs de recherche inversée d'images
https://tineye.com/	
https://yandex.com/images/	
https://www.bing.com/visualsearch	
https://picarta.ai/	Localise une photo avec l'IA
https://facecheck.id/fr	Moteurs de recherche de visages (sites payants)
https://pimeyes.com/en	
https://www.metadata2go.com/	Récupération des métadonnées d'une image
https://onlineexifviewer.com/	

Beaucoup de réseaux sociaux aujourd'hui effacent les métadonnées des images uploadées par leurs utilisateurs. Ces données sont dangereuses puisqu'elles contiennent des informations sensibles comme la date de création et même parfois des coordonnées GPS…

Je vais tester ci-dessous picarta.ai en lui soumettant une photo prise à New York :

Voici l'image (obtenue sur Pixabay) :



Voici le résultat obtenu avec Picarta :

I think the image was taken in one of these locations:

1. Oakland, United States. GPS location around: 40.979897, -74.267580 Confidence: 93.72%
2. New York City, United States. GPS location around: 40.699722, -73.997190 Confidence: 93.61%
3. New York City, United States. GPS location around: 40.699413, -74.030320 Confidence: 93.39%
4. New York City, United States. GPS location around: 40.702488, -73.997050 Confidence: 93.24%
5. New York City, United States. GPS location around: 40.702698, -73.997120 Confidence: 92.73%
6. New York City, United States. GPS location around: 40.703415, -73.995240 Confidence: 92.53%
7. New York City, United States. GPS location around: 40.703370, -73.994540 Confidence: 92.33%
8. New York City, United States. GPS location around: 40.703010, -73.998024 Confidence: 92.29%
9. New York City, United States. GPS location around: 40.702854, -73.994500 Confidence: 91.94%
10. New York City, United States. GPS location around: 40.701717, -73.997820 Confidence: 91.91%

À vous de tester et de juger ce service !

Deux sites de recherche de visages (sites payants) :

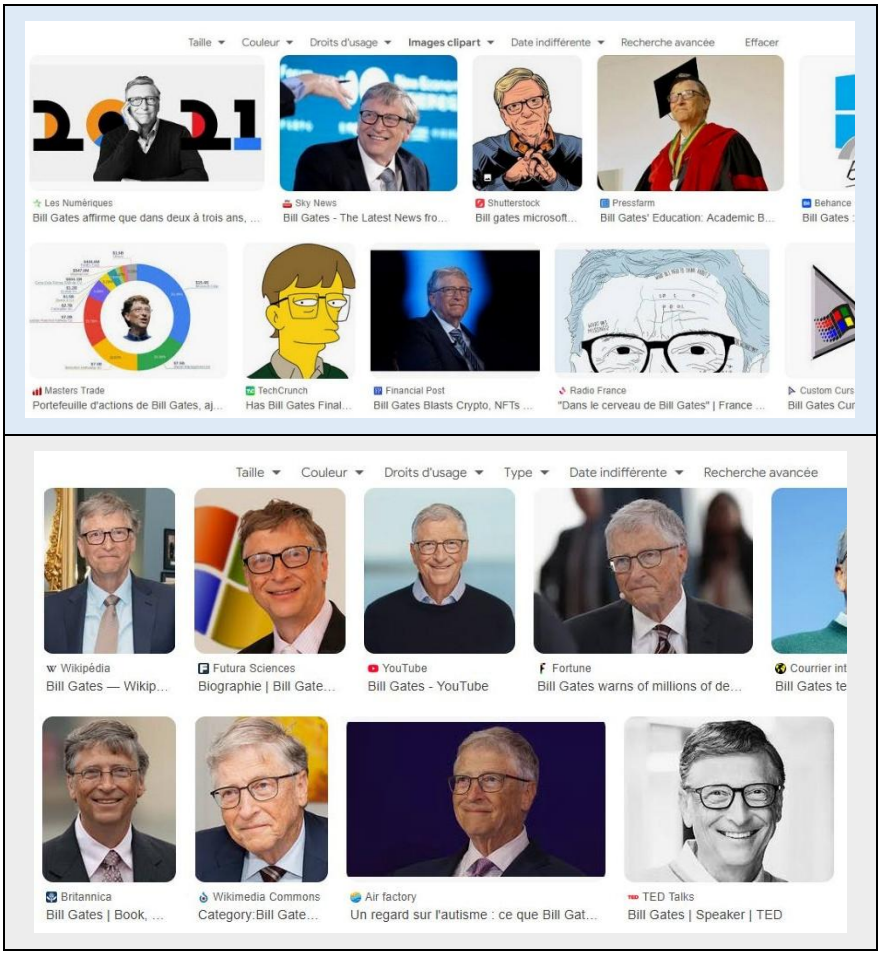
Lorsqu'on fait une recherche d'image concernant une personne avec Google, on peut, si l'on souhaite afficher seulement les visages, cliquer sur Outils/Images clipart :

Dans l'URL, on remplace alors manuellement le mot « clipart » par le mot « face » :

➔ **https://www.google.be/search ?q=bill+gates&sca_esv=…c&udm=2&source=lnt &tbs=itp:clipart&sa=X&ved=…&dpr=1**

➔ **https://www.google.be/search ?q=bill+gates&sca_esv=…c&udm=2&source=lnt &tbs=itp:face&sa=X&ved=…&dpr=1**

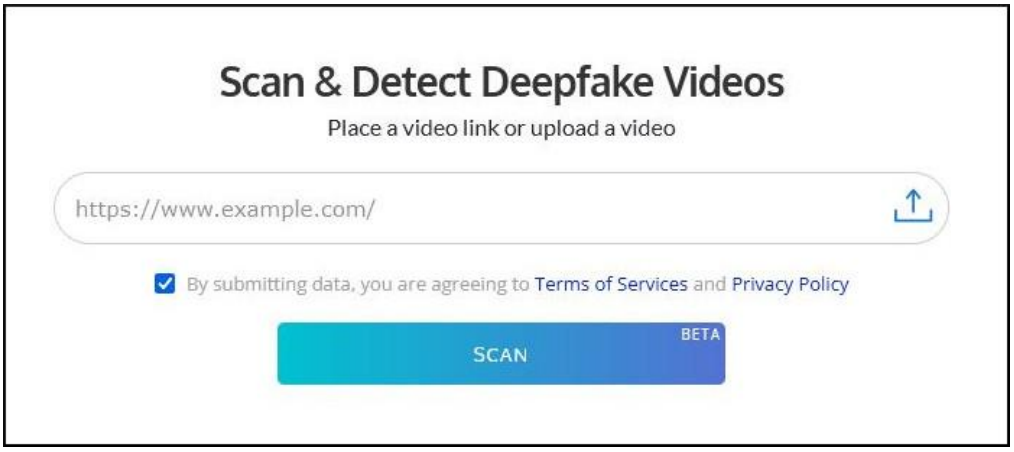
Seuls les visages s'afficheront alors :



Video OSINT

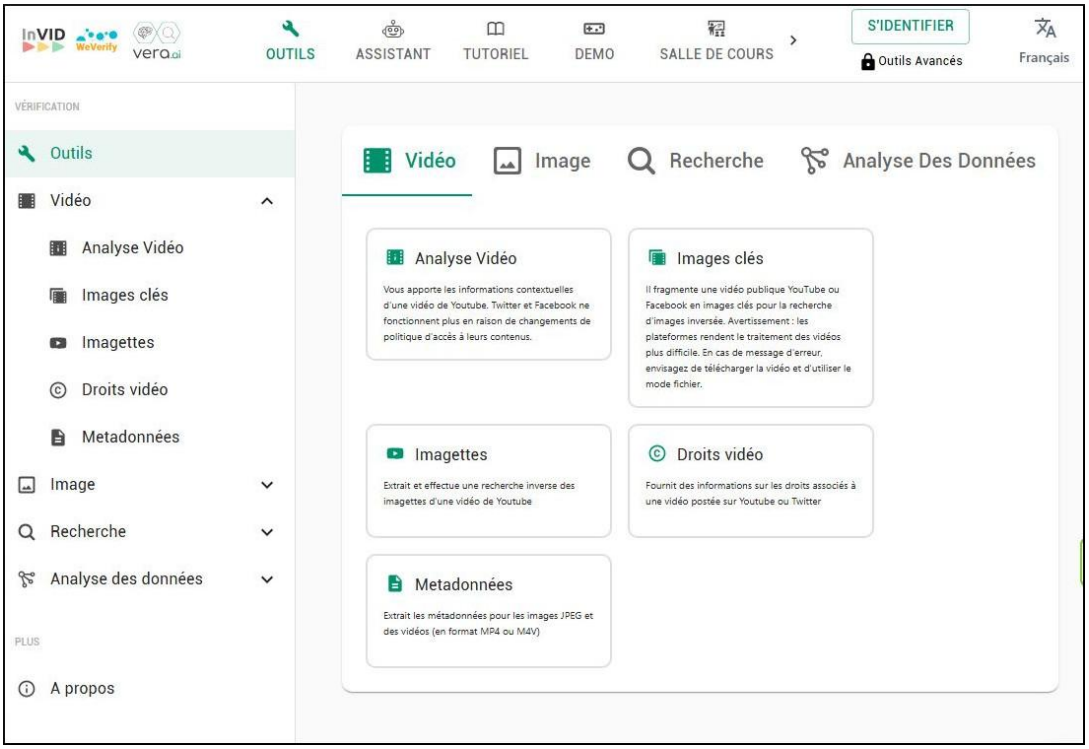
Détection des deepfakes

Deepware Scanner (https://scanner.deepware.ai/) permet de détecter des deepfakes :



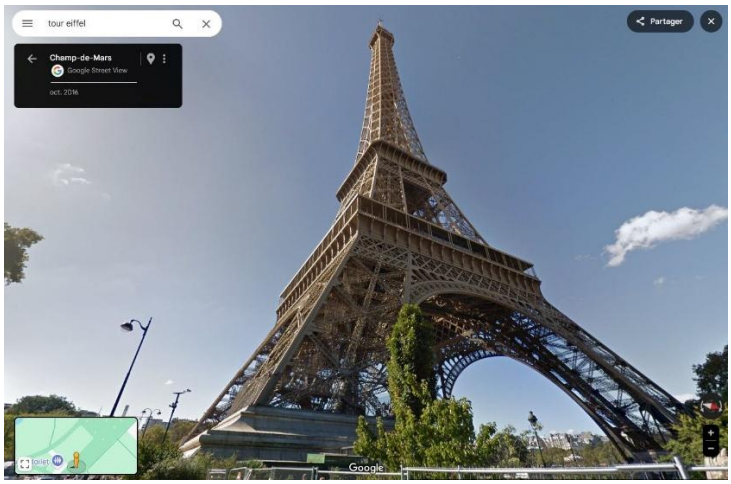
Vérification des vidéos Youtube avec InVID/WeVerify

Cette extension pour Chrome permet aux journalistes de faire du fact-checking et par exemple d'analyser les vidéos présentes sur Youtube :



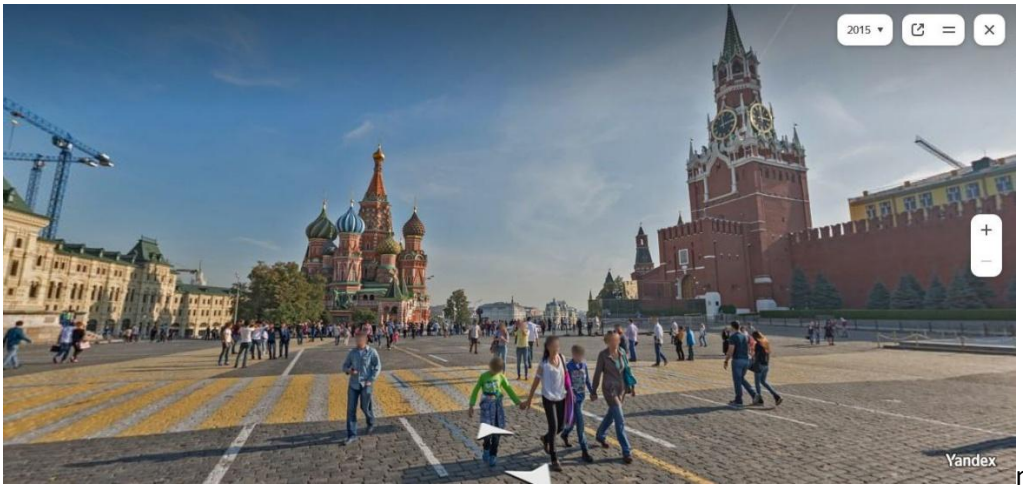
Maps OSINT

https://www.google.com/maps/	**Carte du monde de Google. Sa fonctionnalité *Street View* est remarquable.**
https://www.bing.com/maps	**Carte du monde de Bing. Sa fonctionnalité *Vue Aérienne* (Bird's Eye) est moins bonne que la vue *Satellite* de Google.**
https://yandex.com/maps	**Carte du monde de Yandex. Très bonne couverture de la Russie et des pays russophones.**
https://www.openstreetmap.org/	**Carte du monde sous licence libre.**
https://osm-search.bellingcat.com/	**Site utilisé par les chercheurs et journalistes pour effectuer des géolocalisations. Il faut se connecter avec un compte Google ou avec un email.**



Tour Eiffel à Paris avec Google Street View.

Place Rouge à Moscou avec Yandex.

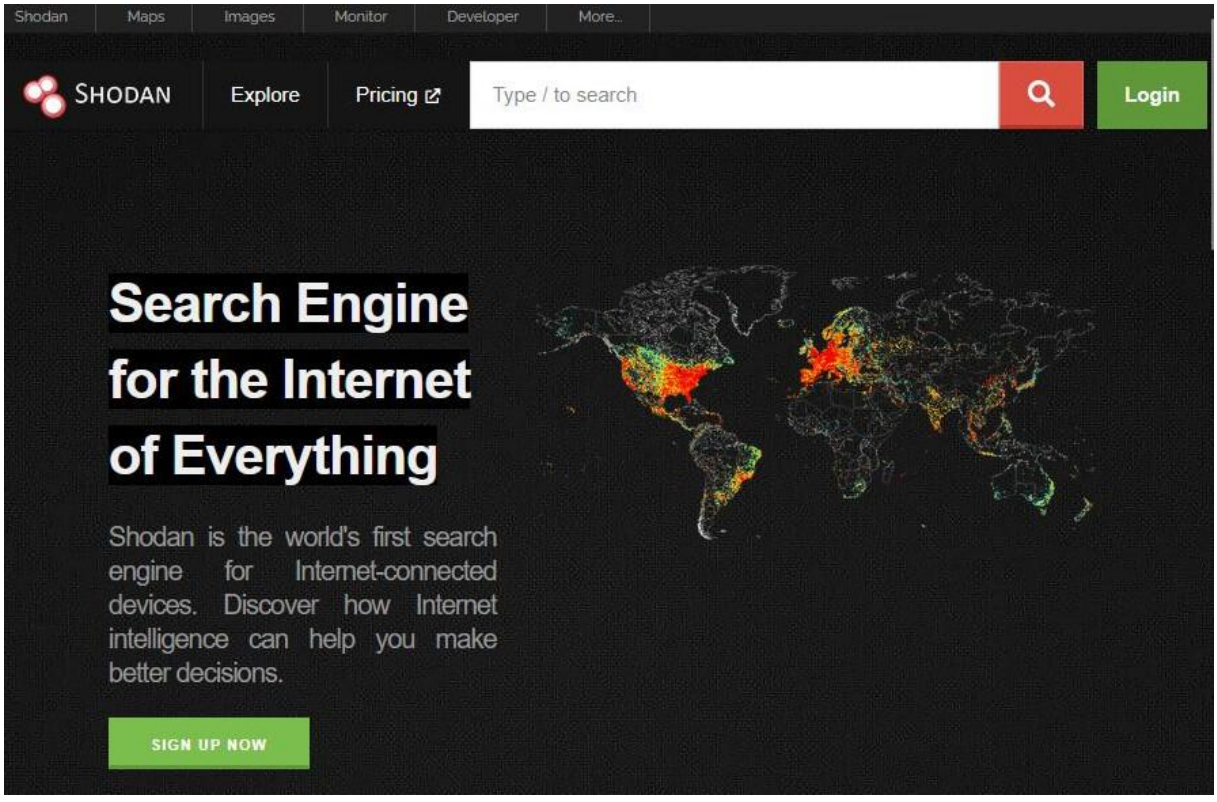


m

IoT OSINT

Le site Shodan (https://www.shodan.io/) est un moteur de recherche spécialisé qui permet de trouver des appareils connectés à Internet (ordinateurs, IoT, …). Il indexe donc des appareils connectés plutôt que des pages HTML ou PHP (comme le font Google ou Bing).

Si une caméra connectée à Internet n'a pas de mot de passe ou a gardé son mot de passe par défaut, Shodan permet d'y avoir accès via son adresse IP et son port et vous pourrez voir en direct les images enregistrées par cette caméra. C'est assez incroyable ! Shodan peut bien sûr être détourné par des cybercriminels pour accéder à vos appareils connectés non protégés. C'est une raison majeure pour laquelle il vous faut vérifier la configuration de vos ordinateurs et IoT et ne jamais laisser leur mot de passe par défaut en vous disant que personne n'y aura accès. Il est également très utile de vous familiariser avec Shodan afin d'en savoir plus.



Dark Web OSINT

Le navigateur Tor, que l'on peut installer sur Windows (même s'il est conseillé de l'utiliser plutôt sur une machine Linux) permet de visiter le Dark Web, dont les URL se termine par l'extension .onion. Les services cachés du Dark Web ne sont donc pas accessibles avec les navigateurs classiques comme Chrome ou Firefox.

Lorsqu'on se connecte à un serveur avec Tor, le trafic passe successivement par trois nœuds : un nœud d'entrée, un nœud intermédiaire et un nœud de sortie. Cela rend la connexion théoriquement intraçable (attention, ce n'est pas toujours vrai en pratique).

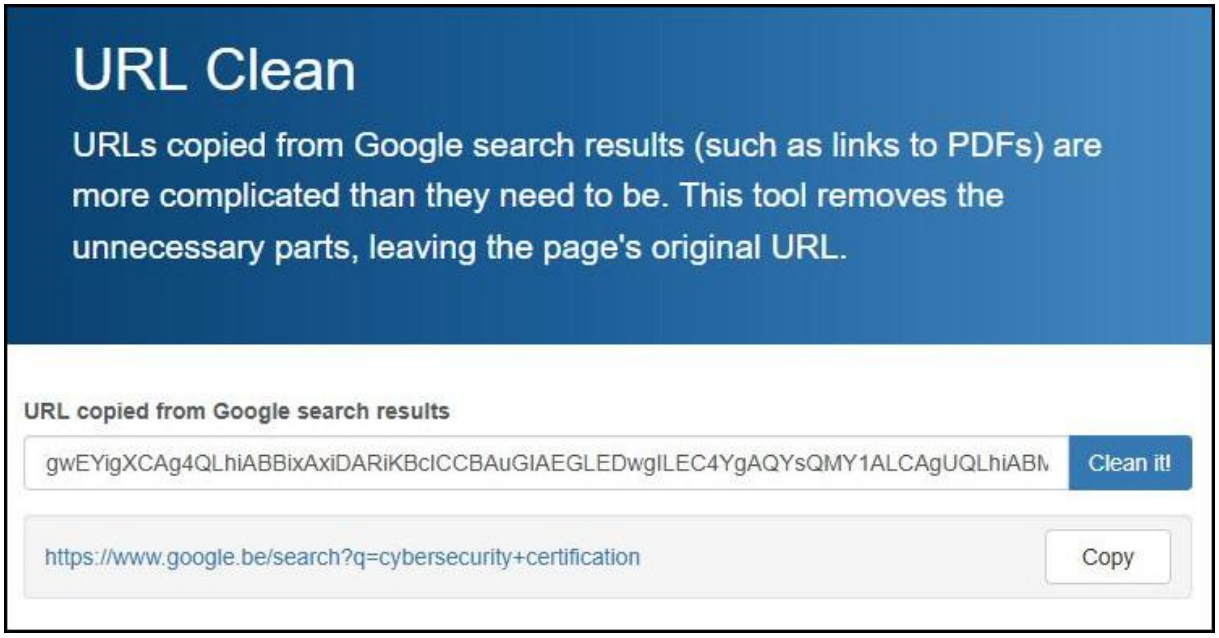
Le service ExoneraTor permet de savoir si une adresse IP était un relais Tor à une date précise. Ce service est très utilisé par la police dans le cadre de leurs enquêtes.

OSINT et moteurs de recherche

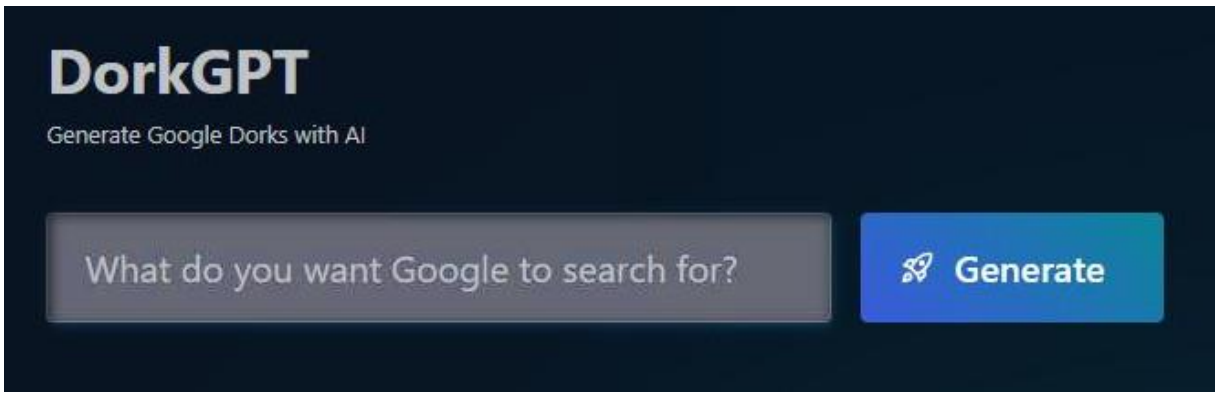
Les **six principaux** moteurs de recherche dans le monde

Google	**Moteur de recherche qui viole la vie privée en collectant massivement des données.**
Bing	**Moteur de recherche de Microsoft (collecte de données moins massive).**
DuckDuckGo	**Moteur de recherche qui respecte votre vie privée.**
Qwant	**Moteur de recherche récent (2013) qui respecte votre vie privée. Surtout utilisé en France.**
Yandex	**Moteur de recherche russe qui collecte des données en Russie. Très utilisé dans les pays russophones.**
Baidu	**Moteur de recherche chinois qui collecte des données et est censuré par le gouvernement chinois. Peu utilisé hors de Chine.**

Simplifier une URL à rallonge issue d'une recherche Google est facile avec URL Clean (https://urlclean.com/) :



Pour générer des Google Dorks (utilisation des opérateurs avancés de recherche de Google) avec l'IA, vous pouvez utiliser l'outil suivant (https://www.dorkgpt.com/) :



REMARQUES :

1. **Bing possède des opérateurs qui n'existent pas chez Google, par exemple :**
   - ➔ **loc:us**
   - ➔ **language:us**

   Le paramètre us peut être remplacé par fr, be, ca, es, it, …

2. **Si on désire filtrer par pays avec Google, il faudra utiliser un artifice :**
   - ➔ **site:us**

3. **Pour filtrer avec la langue, Google possède l'opérateur lang :**
   - ➔ **lang:us**

4. **Autres opérateurs spécifiques de Bing :**
   - ➔ **linkfromdomain:site.net (utilisé seul, il affiche les liens trouvés sur le site cible)**
   - ➔ **ip:192.145.37.11 (affiche les sites hébergés à une adresse IP)**
   - ➔ **contains:pdf (ex : « site:cnn.com contains:pdf » : affiche les pages HTML sur cnn.com qui contiennent un lien vers un fichier PDF) ≠ filetype !**

5. **Deux opérateurs de Yandex :**
   - ➔ **date:2024 (ou encore date:202401, date:20240101, date:2022..2024)**
   - ➔ **rhost:net.mon-site.* (recherche les sous-domaines de *mon-site.net*)**

6. **Les opérateurs de DuckDuckGo sont similaires aux opérateurs de Google.**

OSINT : se créer un compte intraçable (sock puppet account)

Toutes les recherches seront effectuées à l'aide de ce compte.

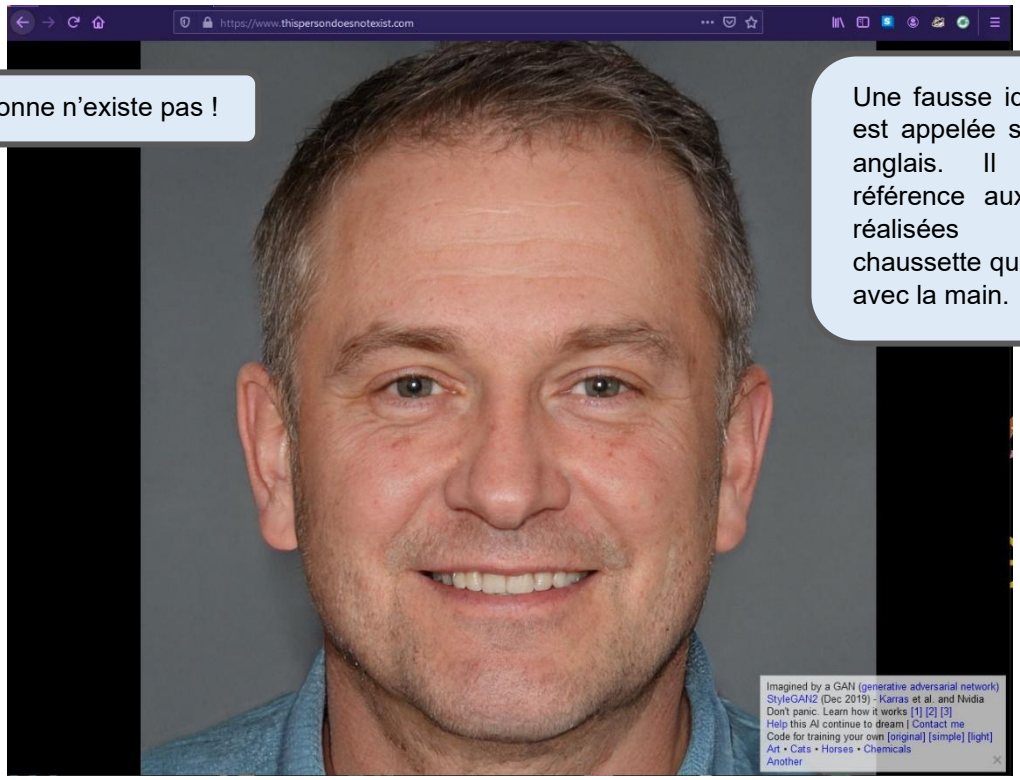
Procédure à suivre

1	**Utiliser une machine virtuelle (je conseille la machine virtuelle *Trace Labs,* dédiée à l'OSINT), pour les raisons évidentes :** • **Cela cache l'OS de votre machine hôte aux sites visités.** • **Cela vous protège des malwares.** • **Cela permet un effacement complet après investigation.**
2	**Utiliser https://www.fakenamegenerator.com/ pour se créer une identité fictive. Ne sélectionner que les informations intéressantes (nom et prénom, date et lieu de naissance, *user agent*, …). Se servir de l'extension de navigateur UserAgent Switcher pour définir son *user agent* (agent utilisateur).**
3	**Choisir une photo de visage en concordance avec le nom et l'âge de l'identité fictive choisie sur un site qui génère, grâce à l'IA, des visages uniques qui n'existent pas réellement (ainsi une recherche inversée d'image ne donnera aucun résultat permettant de vous démasquer). Deux sites permettent cette prouesse :** • **https://thispersondoesnotexist.com/** • **https://thispersonnotexist.org/**
4	**Créer un mail avec Fastmail (qui ne demande pas d'email de récupération)**
5	**Obtenir un numéro de téléphone jetable (temporaire) qui permet d'activer ses comptes et de préserver son anonymat. Un smartphone Android ou iOS est nécessaire. Ex : https://numeroesim.com/**
6	**Utiliser le Wi-Fi gratuit d'une bibliothèque ou d'un café.**
7	**Créer des comptes Facebook, Instagram et X (Twitter).**
8	**Alimenter ses comptes et s'y connecter régulièrement.**

Se créer un visage avec **thispersondoesnotexist.com**

Voici deux visages très réalistes créés par cette intelligence artificielle :

Le site thispersondoesnotexist.com **permet, grâce à une intelligence artificielle, de générer une photographie d'une personne qui n'existe pas. Cela est utile au testeur d'intrusion qui désire se forger une fausse identité en ligne (sock puppet).**



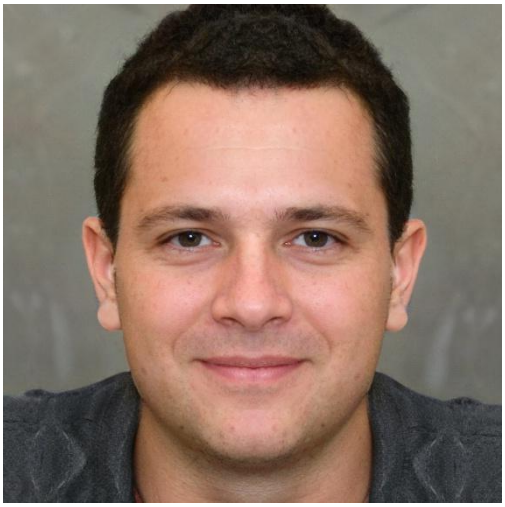
Cette personne n'existe pas !

Une fausse identité en ligne est appelée sock puppet en anglais. Il s'agit d'une référence aux marionnettes réalisées avec une chaussette que l'on manipule avec la main.

Une alternative existe : **thispersonnotexist.org**

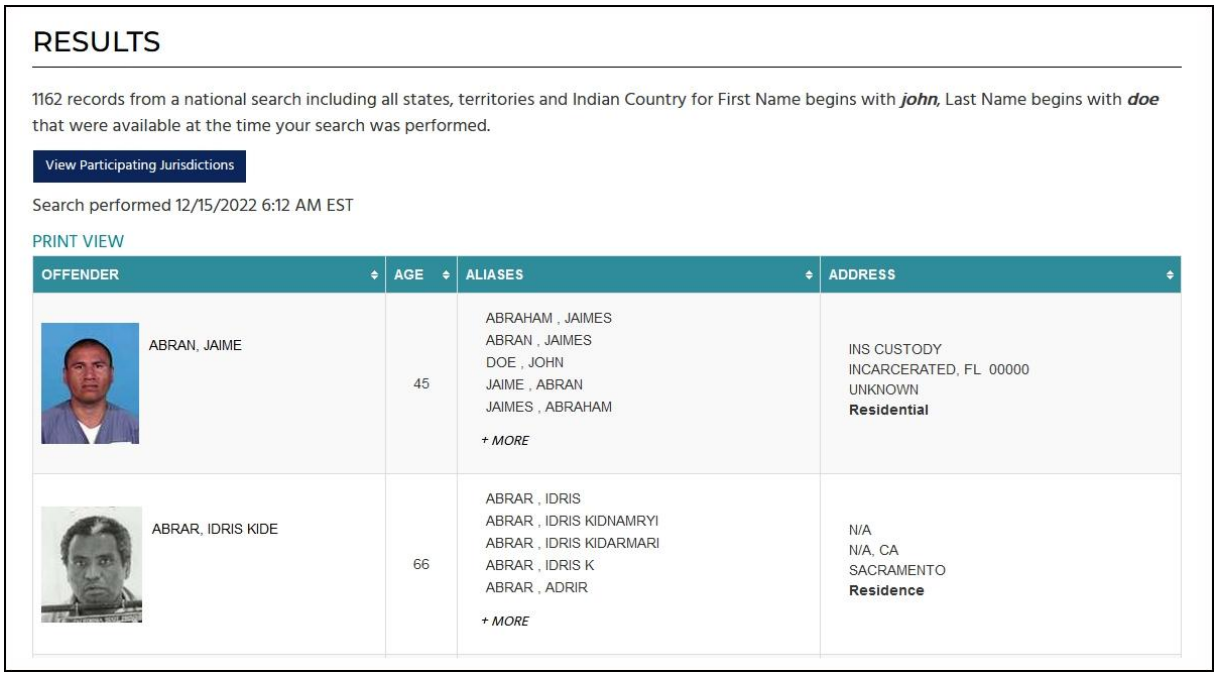
Sur ce second site, il est possible de créer une image d'un visage ou même d'un corps entier d'une personne n'existant pas réellement grâce à l'intelligence artificielle.

Exemples :



USA - OSINT : le registre des délinquants sexuels

Totalement impensable chez nous en Europe, il est possible de consulter sur **https://www.nsopw.gov/** le registre des délinquants sexuels américains (avec photos et adresses) :



RESULTS

1162 records from a national search including all states, territories and Indian Country for First Name begins with ***john,*** Last Name begins with ***doe*** that were available at the time your search was performed.

View Participating Jurisdictions

Search performed 12/15/2022 6:12 AM EST

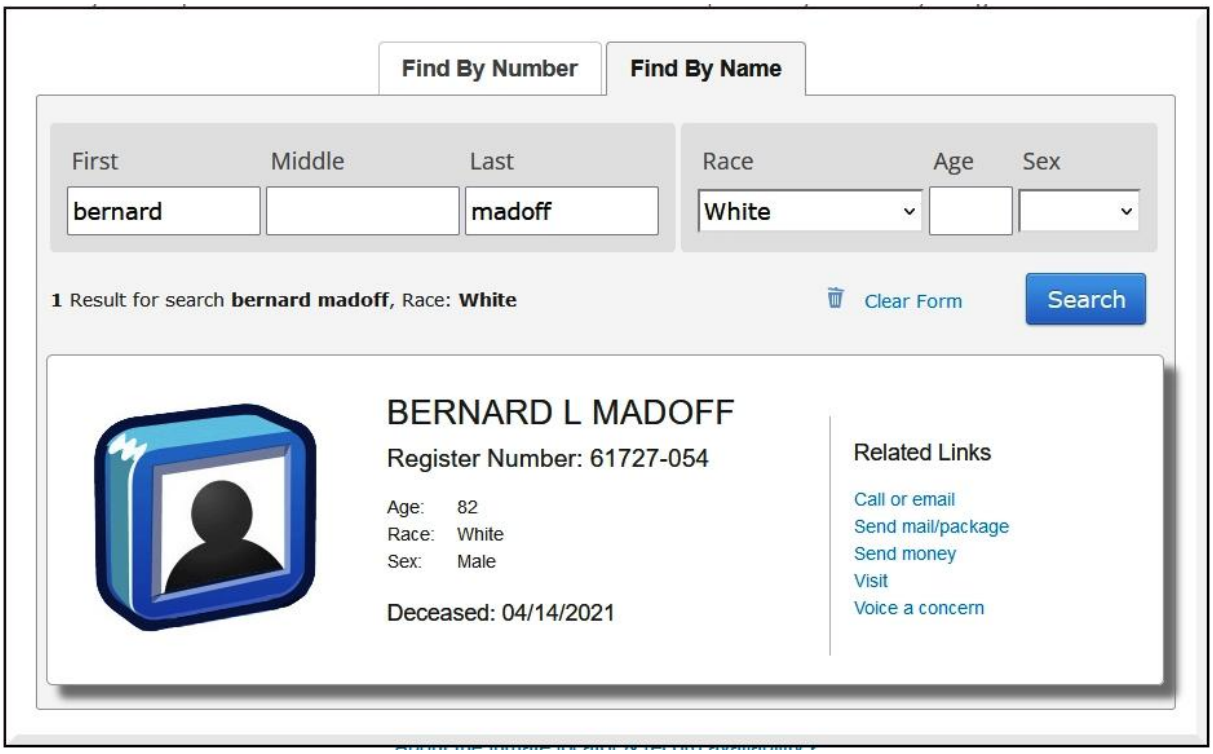
PRINT VIEW

OFFENDER	AGE	ALIASES	ADDRESS
ABRAN, JAIME	45	ABRAHAM , JAIMES ABRAN , JAIMES DOE , JOHN JAIME , ABRAN JAIMES , ABRAHAM *+ MORE*	INS CUSTODY INCARCERATED, FL 00000 UNKNOWN **Residential**
ABRAR, IDRIS KIDE	66	ABRAR , IDRIS ABRAR , IDRIS KIDNAMRYI ABRAR , IDRIS KIDARMARI ABRAR , IDRIS K ABRAR , ADRIR *+ MORE*	N/A N/A, CA SACRAMENTO **Residence**

USA - OSINT : localiser un détenu

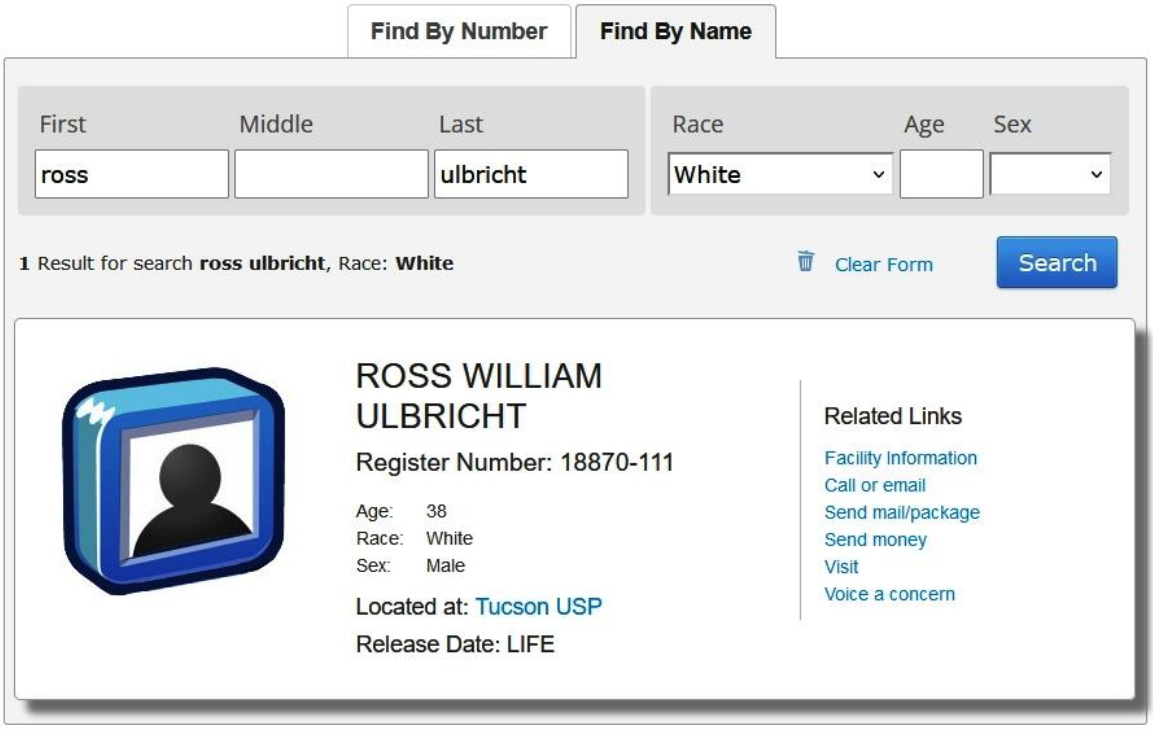
Toujours aux USA, il est possible de localiser le lieu de détention d'un criminel ou délinquant sur le site **https://www.bop.gov/inmateloc/** :

Recherchons le lieu de détention de **Bernard Madoff**, l'auteur de la plus grosse escroquerie de tous les temps, ayant été condamné à 150 ans de prison :

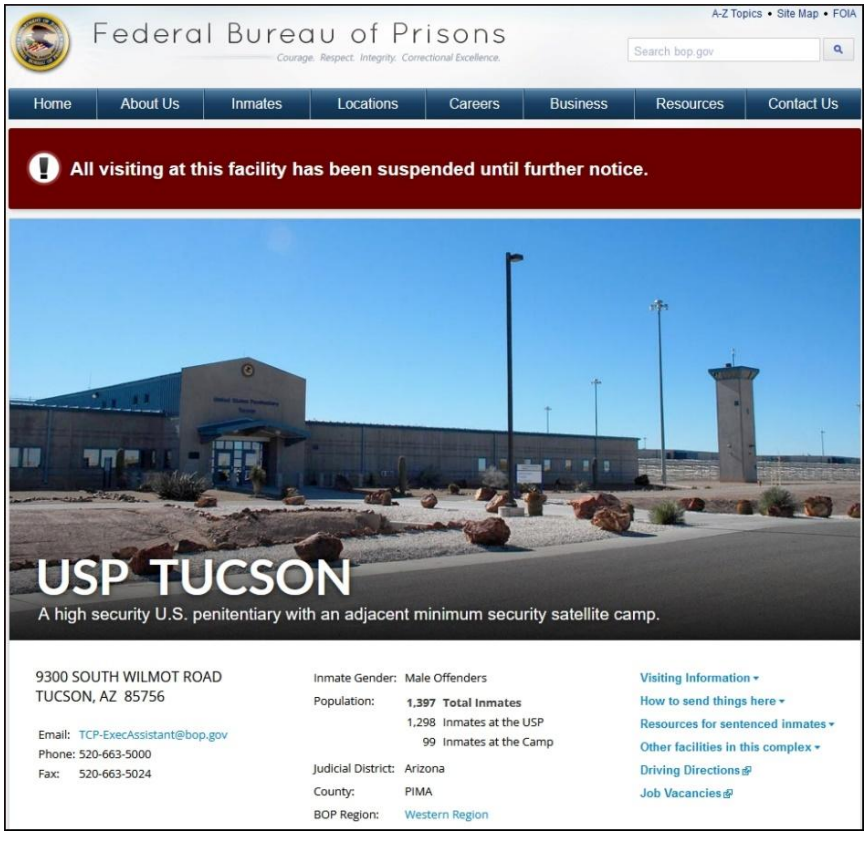


Ce lieu n'apparaît plus puisque Madoff est décédé en détention en 2021.

Lorsqu'on recherchait en 2023 le lieu de détention de **Ross Ulbricht**, le sulfureux créateur du site Silk Road sur le dark web, nous obtenions ceci :

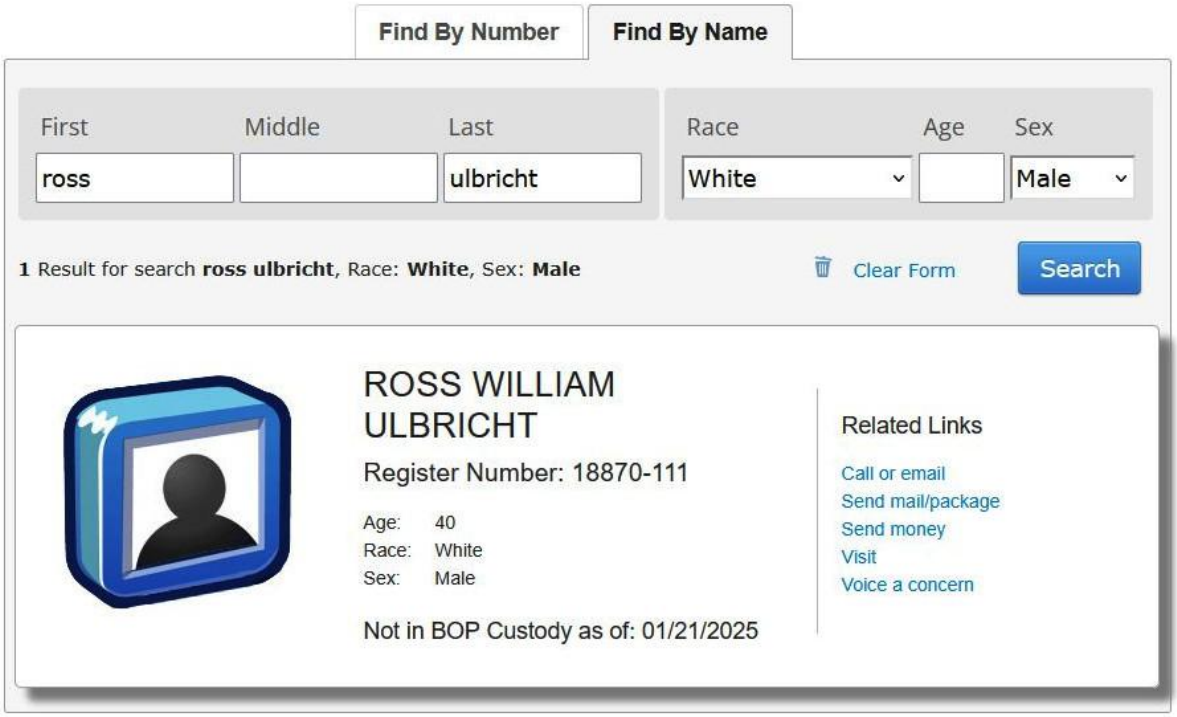


On apprenait que sa peine était la prison à perpétuité et qu'il était détenu à **Tucson USP** :



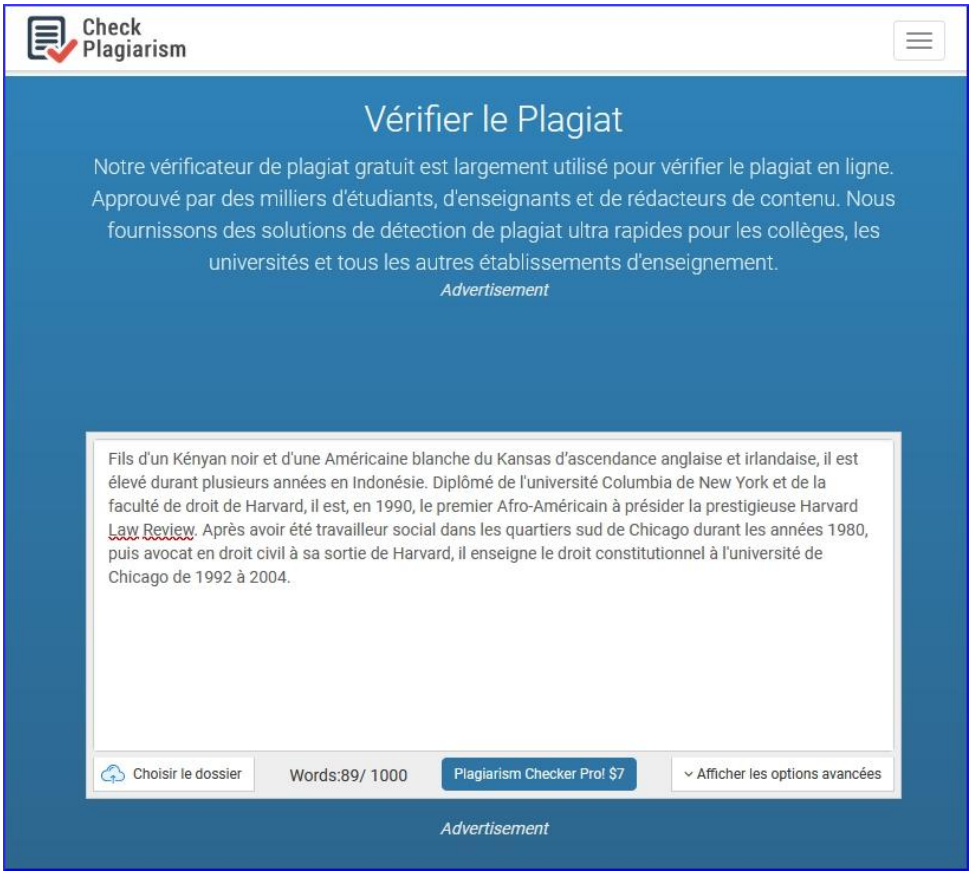
La situation changea le 21 janvier 2025 lorsque le président américain Donald Trump, à la suite d'une promesse faite durant sa campagne électorale, libéra Ross Ulbricht pour remercier le mouvement libertarien qui l'avait fermement soutenu. Ce mouvement prône un marché libre, sans interférence de l'État, et jugeait la peine de Ross Ulbricht disproportionnée.

Voici la fiche de Ross Ulbricht mise à jour après sa grâce présidentielle de janvier 2025 :



OSINT : vérifier le plagiat

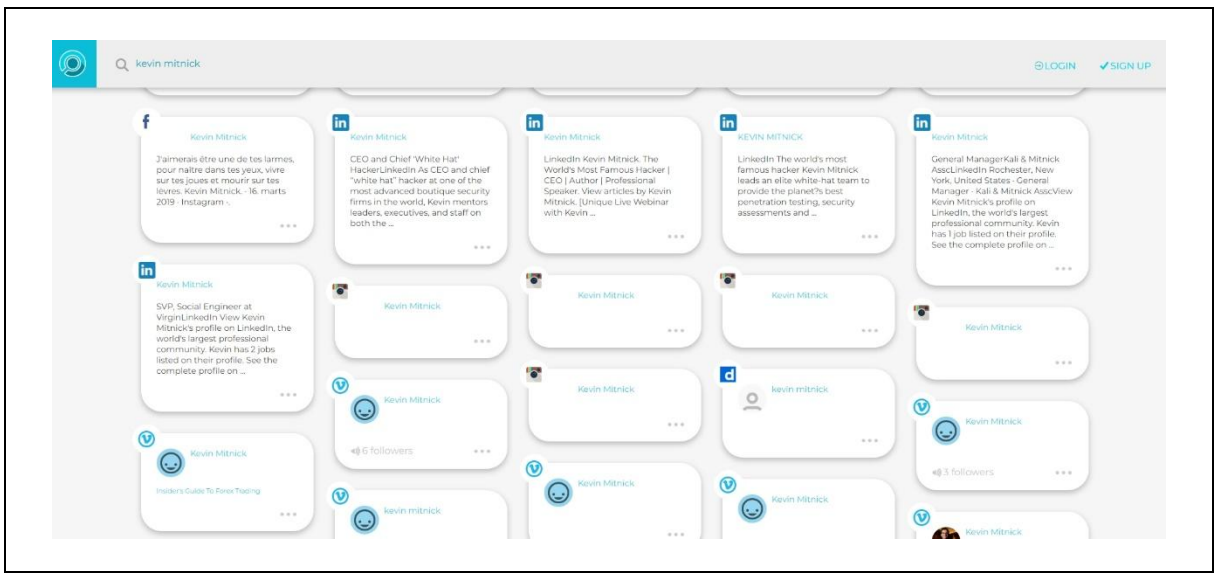
J'ai copié quelques phrases de l'article de Wikipedia consacré à Barack Obama et j'ai soumis le tout au site **https://www.check-plagiarism.com/fr/** :



Le résultat est bien celui attendu, le plagiat est détecté :

OSINT : un moteur de recherche sur les médias sociaux

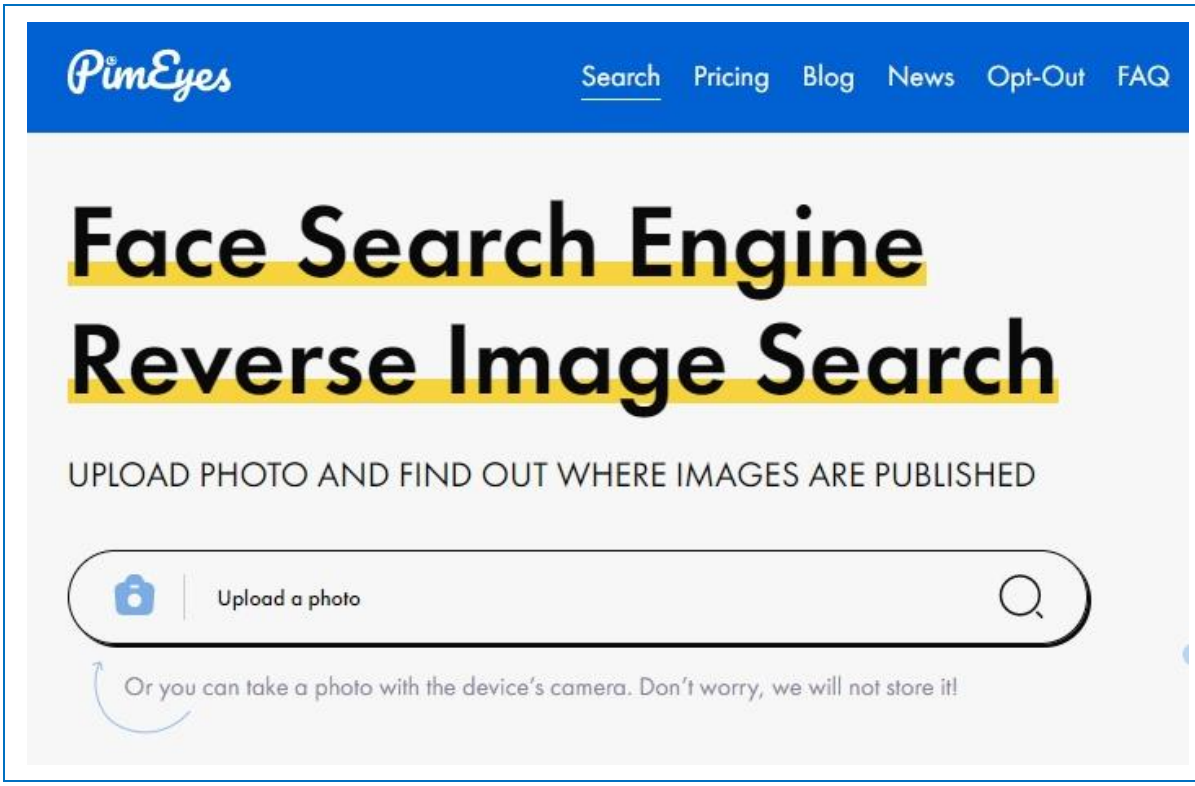
Recherchons le célèbre **Kevin Mitnick** sur les principaux médias sociaux :



Notre recherche donne des résultats sur Facebook, LinkedIn, Instagram, Dailymotion, Vimeo, …

OSINT : recherche inversée d'image

Mis à part la recherche inversée sur Google, il est encore possible de réaliser ce type de recherche sur deux sites majeurs **: TinEye** et **PimEyes.**



PimEyes, qui fait aussi de la reconnaissance faciale (voir page suivante) est payant et coûte, selon l'option choisie de 350 € à 3.500 € par an !

OSINT : La reconnaissance faciale avec PimEyes et Clearview AI

Il existe deux moteurs de recherche spécialisés dans la reconnaissance faciale : PimEyes (https://pimeyes.com) et Clearview AI (https://www.clearview.ai).

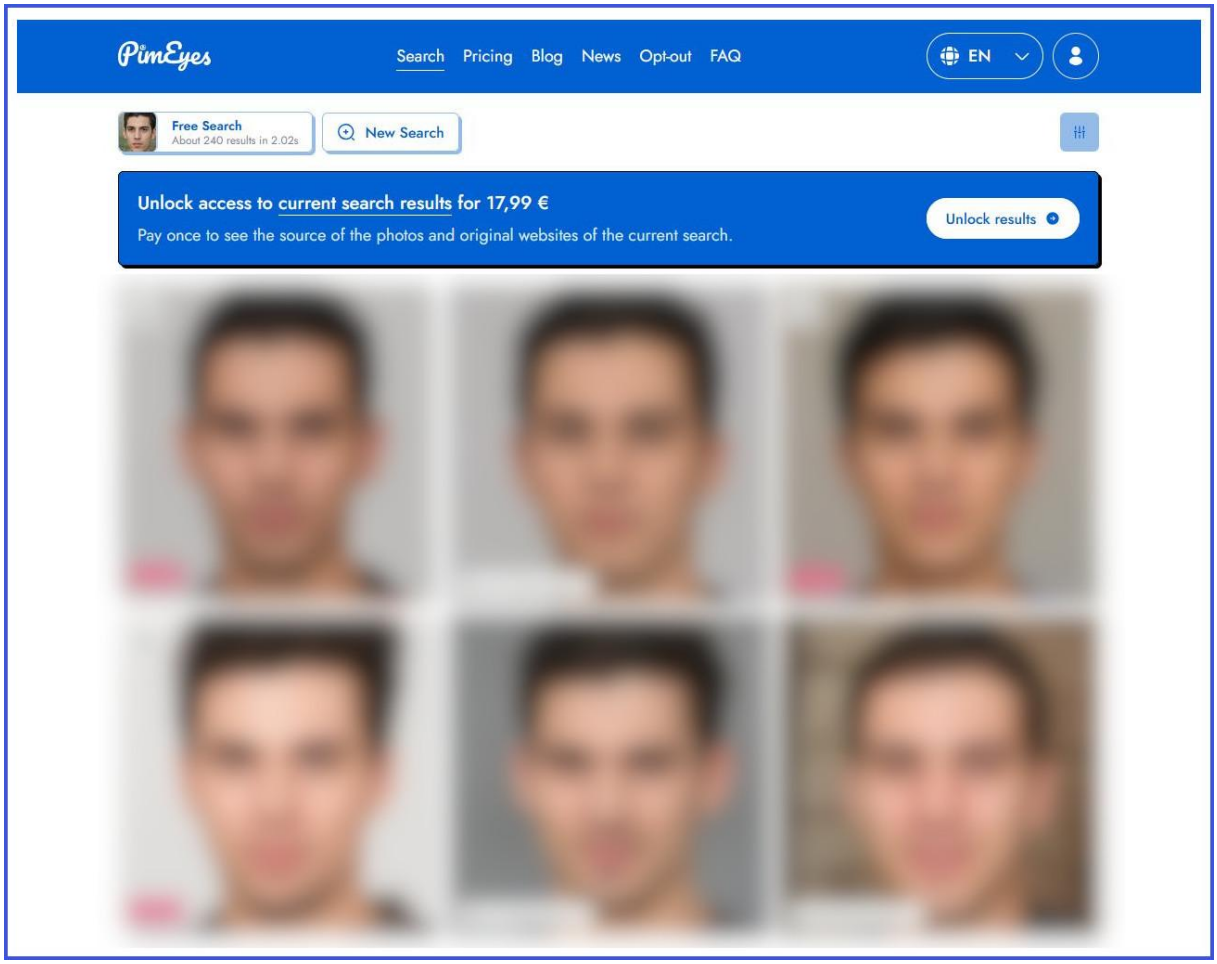
Voici un tableau comparatif concernant ces deux puissants outils :

PimEyes	Clearview AI
Destiné au grand public en général	**Destiné aux forces de l'ordre et agences gouvernementales**
➔ Il suffit d'uploader une photo de vous pour obtenir les images correspondantes disponibles en ligne. Cela vous permet de retrouver les photos que vous laissez sur Internet, parfois sans le savoir, ou sans vous en souvenir. Cela permet encore de démasquer les utilisation illicites de vos photos (usurpation d'identité ou autres). Cela permet enfin de mettre en évidence des utilisations non autorisées de vos photos, par un ex-conjoint par exemple. ➔ Les photos de la base de données sont collectées parmi les sites accessibles publiquement. La base de données est de taille réduite. ➔ La version gratuite est limitée tandis que la version payante peut coûter 35 à 350 € par mois selon la formule choisie (prix au début 2025).	➔ Clearview AI n'est donc pas accessible au grand public ➔ Les photos et selfies de la base de données sont issus des sites accessibles publiquement, mais aussi des différents réseaux sociaux (Facebook, X et Youtube) et des forums. La base de données est de grande taille : elle contient plus de 30 milliards d'images ! ➔ Le but est principalement, pour les forces de l'ordre, de retrouver un criminel présumé ou au contraire un témoin non identifié dans une affaire criminelle. ➔ Clearview AI est plus controversé que PimEyes, notamment en ce qui concerne la taille gigantesque de la base de données et la surveillance de masse que cela implique par les gouvernements… ➔ L'Europe considère par exemple que Clearview AI viole le RGPD.

Pour l'exemple, j'ai soumis à PimEyes la photo suivante :

Cette photo, fabriquée par un algorithme, représente une personne qui n'existe pas réellement (https://thispersondoesnotexist.com/).

Voici le résultat obtenu avec PimEyes (j'ai flouté les photos par respect de la vie privée des personnes concernées). Les photos sont cependant vraiment ressemblantes :



OSINT : détecter un hypertrucage vidéo

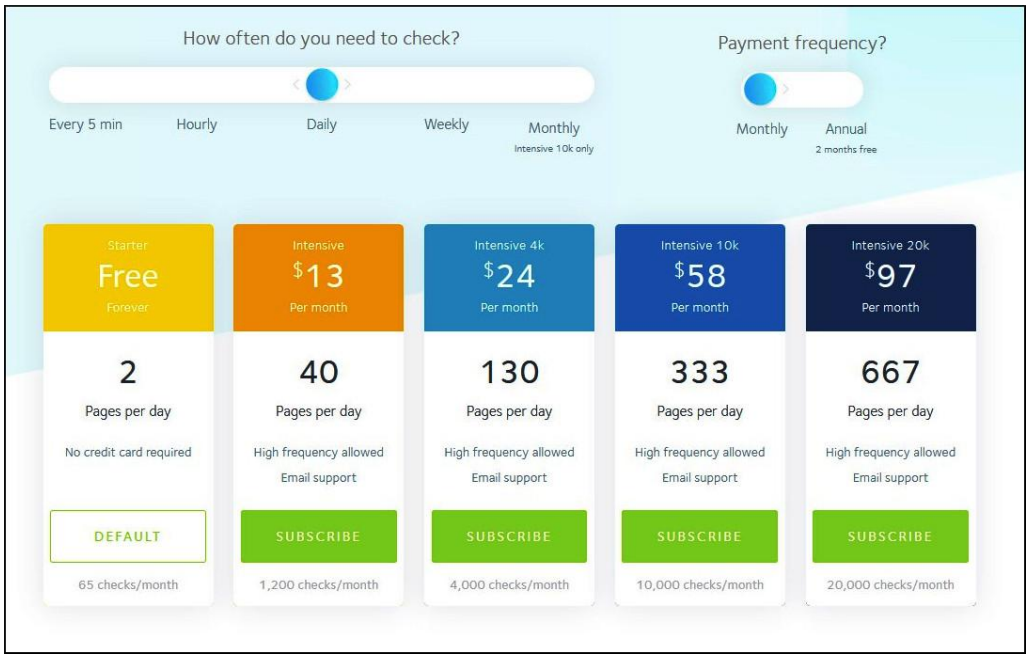
Un hypertrucage (en anglais : deepfake) est une technique qui permet de remplacer un visage dans une vidéo par un autre visage, ou de truquer un fichier audio afin de faire dire des choses imaginaires à un personnage public. On peut ainsi créer des fake news, fausses informations que vous trouverez de plus en plus souvent durant vos recherches d'informations sur le Net.

Le site deepware (https://deepware.ai/) affirme pouvoir détecter les hypertrucages vidéos, choisissons un hypertrucage sur Youtube et soumettons-le à deepware :

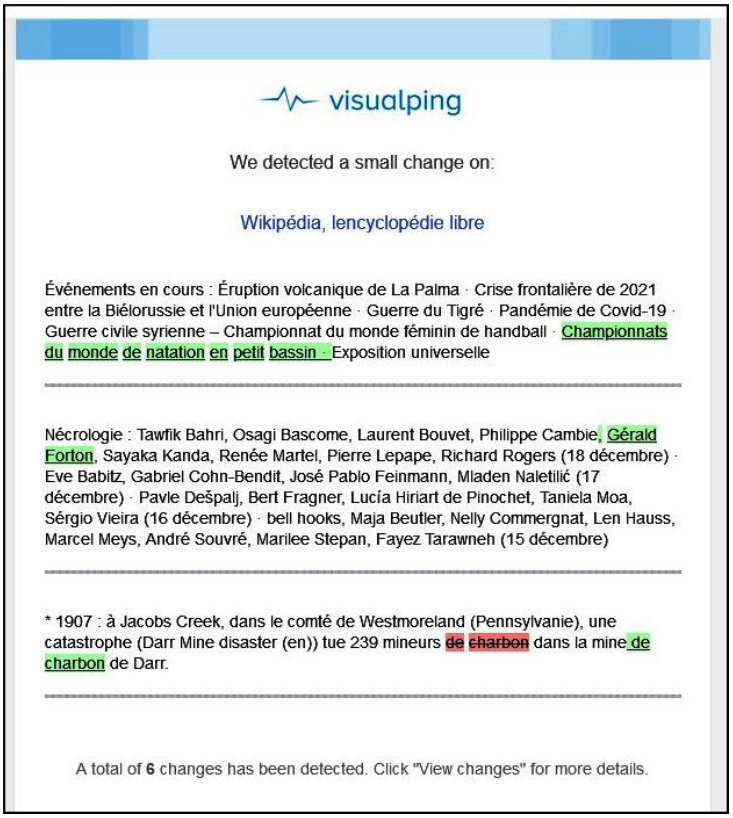
L'hypertrucage est clairement détecté (on est dans la zone rouge) :

OSINT : suivre les modifications d'une page web avec visualping.io

Vous désirez être alerté par courriel lorsqu'une page web est modifiée ? Le site canadien visualping (https://visualping.io) vous rend ce service pour quelques dollars par mois (le service est gratuit en version limitée).



Pour l'exemple, je désire surveiller la page d'accueil de wikipedia :



OSINT : arrestation de John Mc Afee grâce aux données EXIF

En 2012, John Mc Afee, recherché par les autorités judiciaires, a pu être arrêté par la police au Guatemala, après avoir été localisé grâce aux données EXIF d'une photo publiée en ligne.

La photo incriminée :

John McAfee, le créateur du célèbre logiciel antivirus, a été retrouvé mort le 23 juin 2021, à l'âge de 75 ans, dans une prison espagnole où il attendait une probable extradition vers les USA. Il s'agirait d'un suicide.

Que trouve-t-on dans les données EXIF de cette photo (programme **exiftool.exe) :**

```
GPS Altitude                    : 7.1 m Above Sea Level
GPS Latitude                    : 15 deg 39' 29.40" N
GPS Longitude                   : 88 deg 59' 31.80" W
```

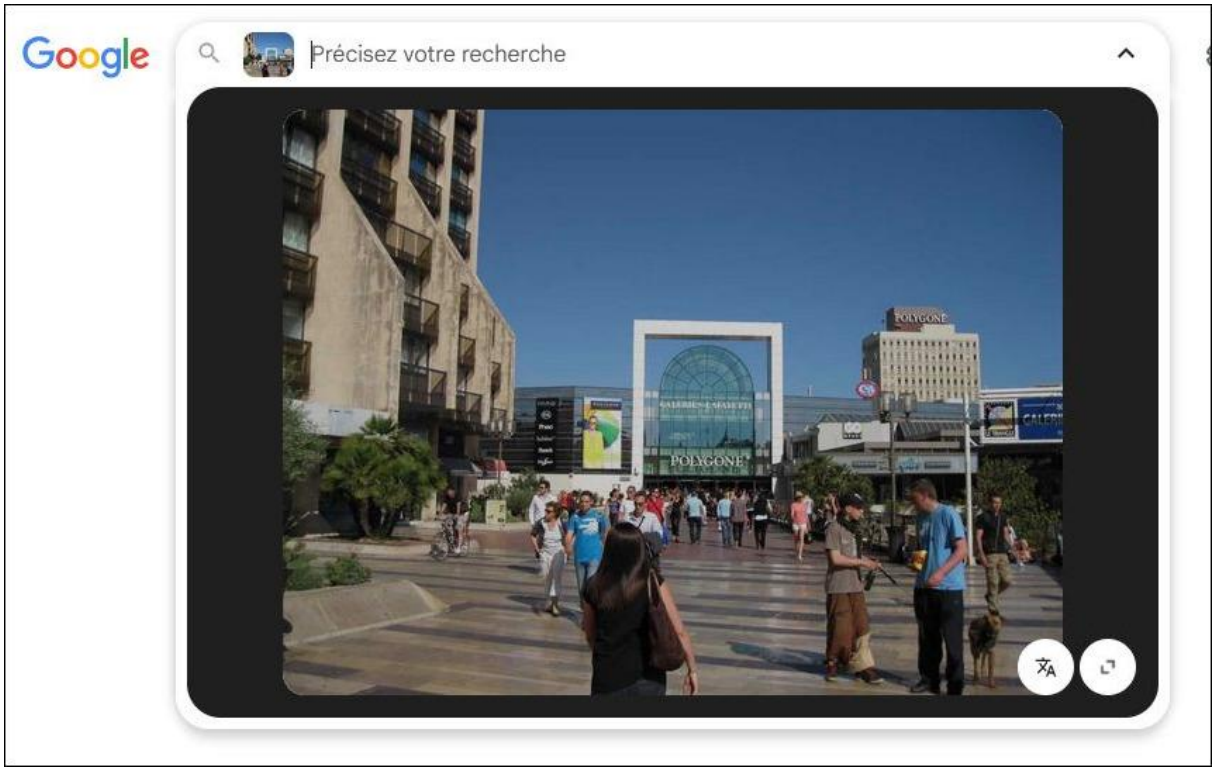
Google Maps localise ces coordonnées au Guatemala (ville de Rio Dulce) :

De nombreux criminels ont pu être localisés et arrêtés de façon similaire. Vous devez savoir que les données EXIF vous concerne vous aussi. Elles peuvent notamment permettre à des gens mal intentionnés de localiser votre maison dont la photo a été publiée sur Internet. Vous pouvez imaginer les conséquences désastreuses pour votre tranquillité.

OSINT : localiser une photographie avec Google

Voici une photo que j'ai prise à Montpellier (dont j'ai enlevé des coordonnées GPS) :

Soumettons-la à Google Image :



Zoomons sur le bâtiment principal :

En quelques secondes, Google trouve la localisation exacte de ce bâtiment :

OSINT : enlever le background d'une image

Le site https://www.remove.bg permet d'effectuer automatiquement cette tâche. Voici deux exemples avec des images trouvées sur Pixabay :

OSINT : le site crt.sh

Dans le cadre du Website OSINT, le site crt.sh permet d'afficher les certificats SSL/TLS pour un nom de domaine.

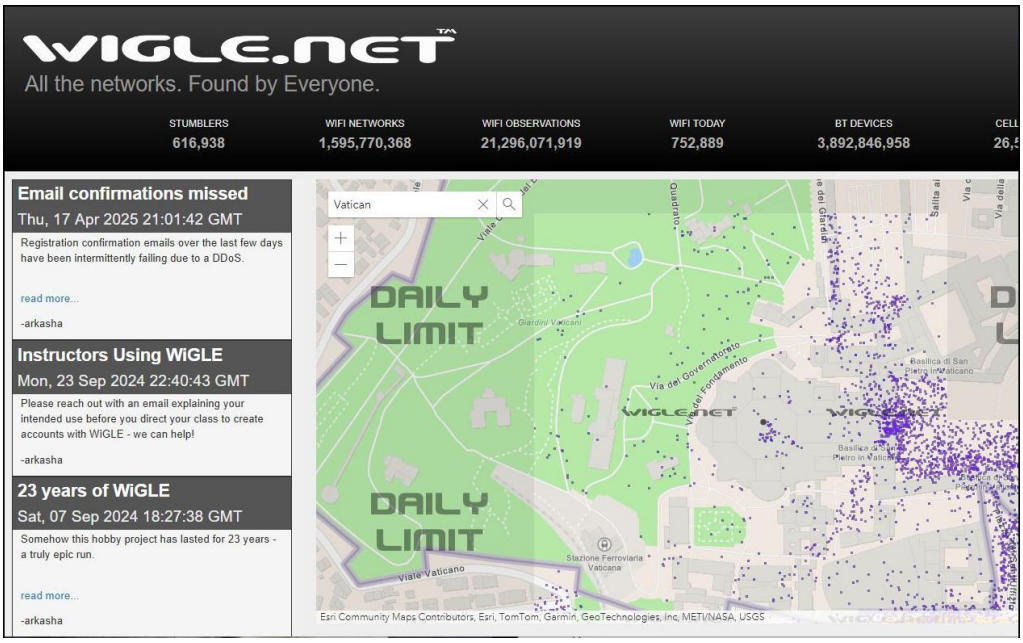
Cela permet :

1. **La surveillance des certificats pour un domaine (sont-ils valides ou expirés)**
2. **La surveillance par un administrateur de site des certificats émis pour son domaine (un certificat a-t-il été émis sans autorisation)**
3. **La découverte, dans le cadre de l'OSINT, des sous-domaines.**

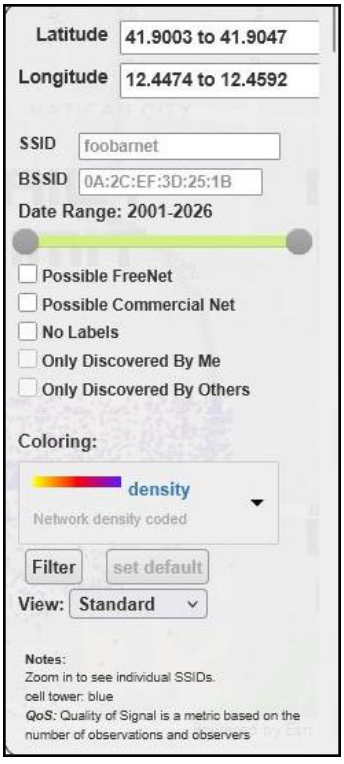
Si vous tapez %.google.com dans la barre de recherche de crt.sh, s'afficheront tous les certificats émis pour les sous-domaines de Google, comme mail.google.com, accounts.google.com, adwords.google.com, …

OSINT : la recherche de réseaux Wi-Fi avec wigle.net

WiGLE.net (Wireless Geographic Logging Engine) est un site web et une base de données qui sert principalement à **cartographier les réseaux sans fil** (Wi-Fi, GSM, Bluetooth, etc.) à travers le monde. On peut y rechercher des réseaux Wi-Fi publics ou ouverts dans la ville de son choix.

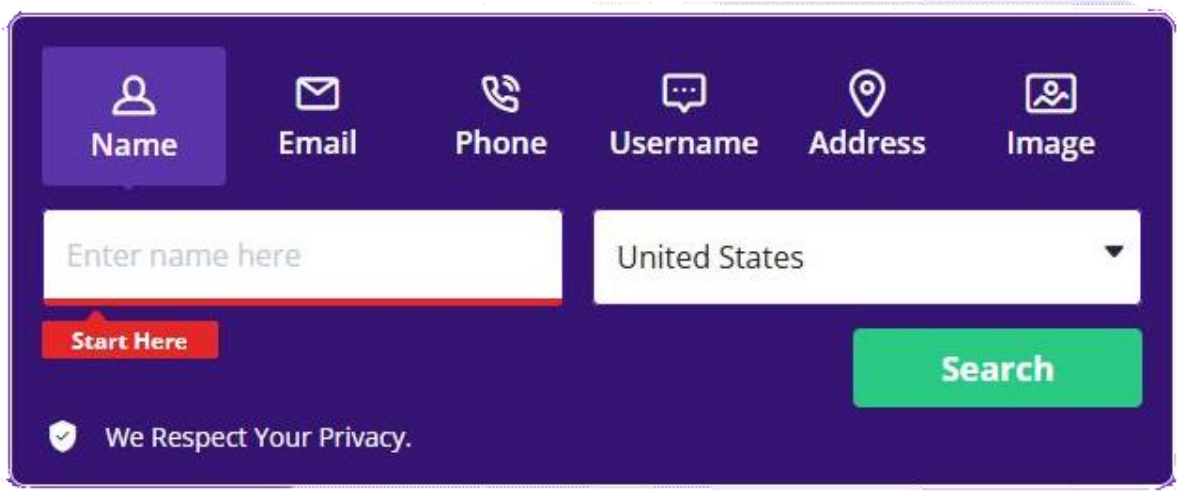


Vous pouvez indiquer un BSSID (adresse MAC d'une carte Wi-Fi) et wigle vous donnera le SSID correspondant. En Europe, fonctionne surtout pour les grandes villes.

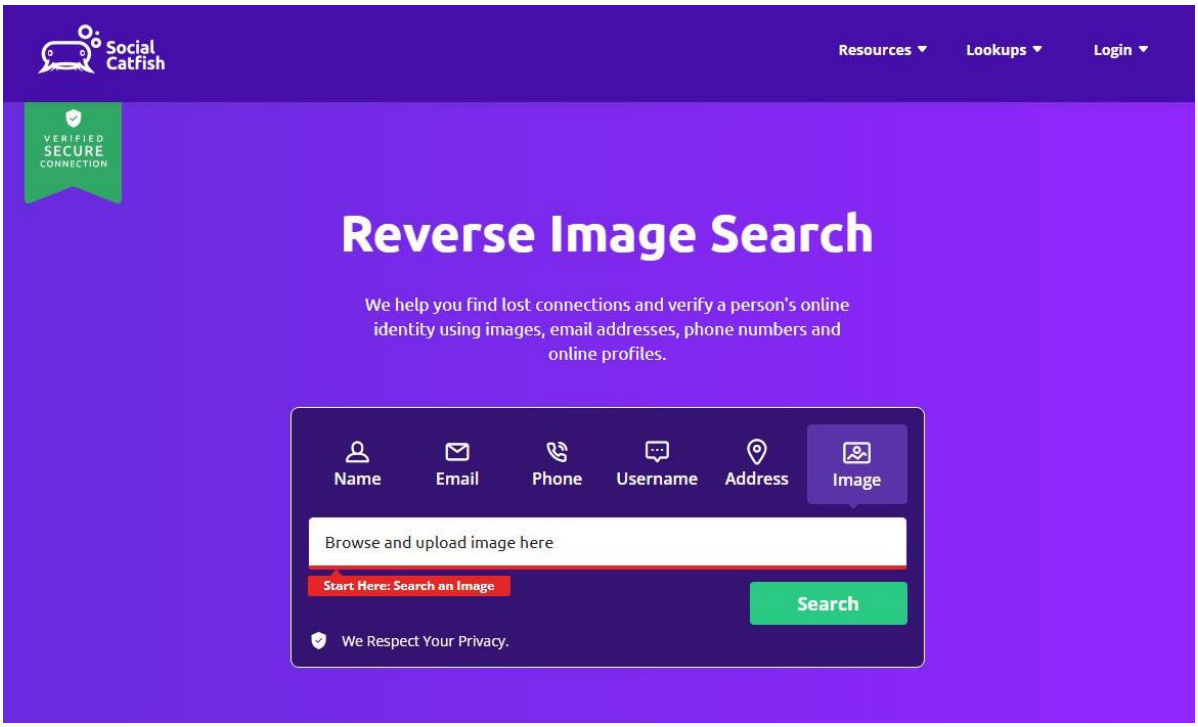


OSINT : la vérification d'identité avec CATFISH

Social Catfish (https://socialcatfish.com/) est un site web américain permettant la vérification d'identité à partir d'un nom, email, numéro de téléphone, username ou encore une adresse et même une image :

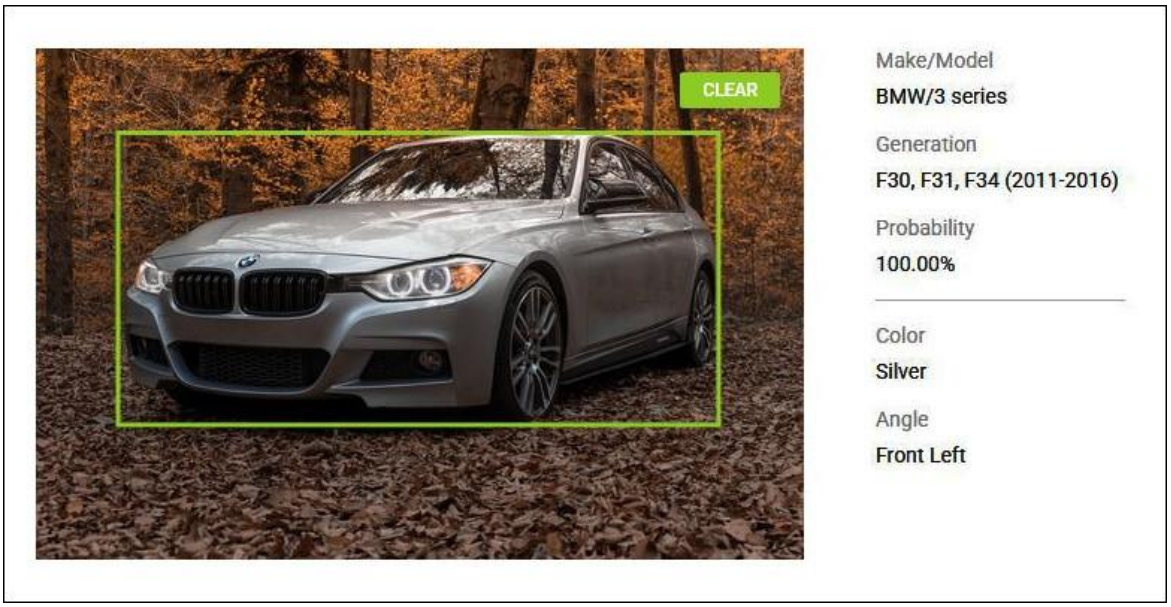


Ce site est principalement utilisé pour enquêter sur des personnes rencontrées sur Internet en cas de soupçon d'arnaques sentimentales ou de fraudes numériques.
Le fait de tromper quelqu'un sur son identité lors d'une arnaque sentimentale ou financière est appelé catfishing.



OSINT : les véhicules

Le site https://carnet.ai/ permet d'obtenir le modèle d'une voiture d'après une simple photo. Exemple 1 (j'utilise une photo de Pixabay) :



Exemple 2 (encore une photo trouvée sur Pixabay) :

Data leak vs data breach

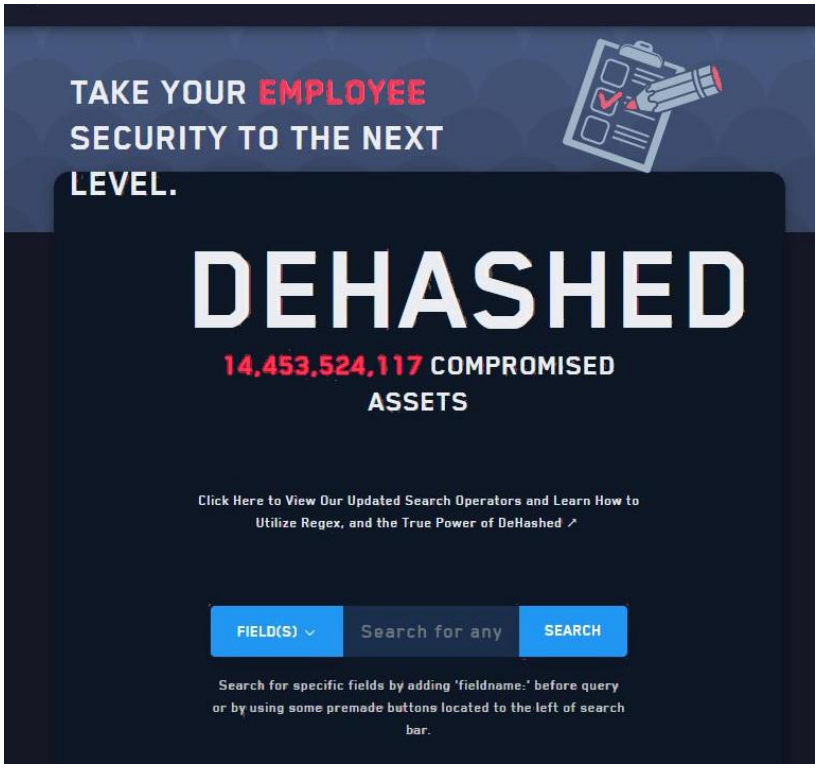
Il est possible, dans le cadre de l'OSINT, de trouver des informations sensibles dans des fuites de données ou à la suite d'une violation de données.

Il est important de distinguer les *data leaks* des *data breaches* :

- **Data leak = fuite de données : exposition involontaire de données sensibles causée par une erreur humaine, par une négligence**
- **Data breach = violation de données : accès intentionnel non autorisé à des données sensibles grâce à une exploitation de vulnérabilité (piratage)**

On peut trouver de telles informations dans des sites spécialisés :

- ➔ **HIBP (Have I Been Pwned ? = https://haveibeenpwned.com/) : site qui vous permet de savoir si votre adresse email a été compromise.**
- ➔ **DeHashed (https://dehashed.com/) : moteur de recherche payant dédié aux data leaks et data breaches (il peut même parfois afficher les mots de passe qui ont fuités pour un email donné)**
- ➔ **IntelligenceX (https://intelx.io/) : autre moteur de recherche dédié aux data leaks et data breaches avec versions gratuite et payante.**



Les pages « Index of »

Il est intéressant, dans le cadre de l'OSINT, de rechercher avec Google les pages « index of ».

La requête Google est : intitle:"index of" site:mon-site.com

Ces pages apparaissent lorsqu'un serveur web est mal configuré et qu'un répertoire ne contient pas de fichier index.html ou index.php.

Le contenu du répertoire est alors visible dans le navigateur.

Il est alors possible d'avoir accès à des fichiers normalement non accessibles aux internautes :

- ➔ **Documents contenant des informations sensibles (PDF, DOCX, XLSX, TXT)**
- ➔ **Des archives (ZIP, 7Z, RAR)**
- ➔ **Des configurations (config.php)**
- ➔ **Des logs (log.txt)**
- ➔ **Des fichiers audio ou vidéo privés (intitle:"index of" mp3 OR wav OR mp4)**

Exemple de page « index of » :

Notes :

Notes :

CE FASCICULE EST EXTRAIT DE :

COMPRENDRE LA CYBERSÉCURITÉ !

Introduction à la cybersécurité

Une série de volumes pour comprendre les **techniques des cybercriminels.**

Recon / OSINT / Scan

m@x (contact@mgregoire.be)