Fuente: Bing Chat AI

INTRODUCCIÓN AL OSINT

DEPARTAMENTO DE INFORMÁTICA. UNIVERSIDAD DE OVIEDO

Técnicas de Investigación con Fuentes Abiertas | 2023 – 2024 (V1.5 *"A-21 'Poseidón' Lite"*)

José Manuel Redondo López

Contenido

¿Qué es esto y para que me sirve? 6

Introducción 6

El delito del doxing 7

¿Qué clase de información se recoge? 8

To "engage" o "to not engage" he ahí la cuestión 10

Técnicas pasivas 10

Técnicas activas 10

El "ciclo de inteligencia" 11

Anonimato "personal": Crear una identidad online falsa 12

Uso de emails temporales para registrarse en servicios 12

Creación de identidades falsas 27

¿Y si para crear la identidad falsa necesito algo más que datos personales? 36

Anonimato de máquinas 38

Remote Browsing 38

Sandboxing de programas 38

VPNs 39

Conexiones vía ToR 40

Búsquedas desde otra localización 44

Preparando el estudio OSINT. *¿Qué tipo de información buscar?* 45

OSINT "Manual": El arte de buscar e interpretar los datos 51

OSINT sobre empleados o personas 51

Cosas acerca de personas en general: Nombres / Imágenes / Vídeos / Fotos... 52

Direcciones de email / Nombres de dominio 62

Redes sociales 65

Redes de búsqueda de personas 76

Aplicaciones de citas 77

Números de teléfono 78

Registros públicos 82

Teléfonos y tarjetas de crédito (u otros documentos de los que aparezcan fotografías en Internet) 87

Uso de medios de transporte varios 88

Geolocalización 89

Foros / Blogs / IRC... 91

Anuncios 92

OSINT sobre máquinas y código 94

Emulación de terminales móviles 94
Direcciones IP / dominios 95
Análisis offline de archivos en la red 96
Qué pasa si te dejas tus dispositivos desbloqueados 98

OSINT "Automático". OSINT Ex Machina 100

OSINT sobre personas 100

OSINT sobre máquinas: Metadatos masivos con FOCA 102

Introducción 102
Instalación 103
Uso de FOCA 103

Minería de datos para OSINT: Maltego 109

Instalación 109
Primer arranque 111
Búsqueda automatizada de información 117
Detalle de una entidad de *Maltego* 124
Creando gráficos manualmente 126
Buscando en redes sociales 130
Otras entidades de Maltego 131
El *Centro de Transformaciones* 132
Exportar la información 137
¡Quiero saber más! 138

OSINT "Automático": Motores de búsqueda y *Threat Intelligence* 141

Introducción 141

APIs para OSINT: ¡Programas que usen OSINT! 143

Árboles interactivos de herramientas OSINT clasificadas por tarea a realizar 145

Malfrat's OSINT (https://map.malfrats.industries/) 145
OSINT Framework (https://osintframework.com/) 145

Motores de búsqueda para máquinas y redes 147

Binary Edge (https://www.binaryedge.io/) 148
Censys (https://search.censys.io/) 148
DNS: Herramientas específicas para servidores y dominios 149
FOFA (https://en.fofa.info/) 151
FullHunt (https://fullhunt.io/) 152
GreyNoise (https://viz.greynoise.io/) 153
IVRE (https://ivre.rocks/) 154
Netlas (https://app.netlas.io/responses/) 155
OnHype (https://www.onyphe.io/) 156
PulseDive (https://pulsedive.com/) 157

Shodan (https://www.shodan.io/) 158
Wigle (https://wigle.net/) 160
ZoomEye (https://www.zoomeye.org/) 161
Motores de búsqueda para sitios web ***162***
Advance Online Utilities (https://centralops.net/co/) 162
Cisco Reputation Check (https://www.talosintelligence.com/) 162
Cloudflare Radar URL Scanner (https://radar.cloudflare.com/scan) 163
CyberGordon (https://cybergordon.com/) 163
IBM Reputation Check (https://exchange.xforce.ibmcloud.com/) 164
IP Reputation Check (https://www.abuseipdb.com/) 164
PublicWWW (https://publicwww.com/) 165
URL Category Finder (https://sitereview.bluecoat.com/#/) 165
URLScan (https://urlscan.io) 166
Website Profiler Tool (https://builtwith.com/) 167
Motores de búsqueda de vulnerabilidades ***168***
CyberChef (https://gchq.github.io/CyberChef/) 168
Packet Storm Security (https://packetstormsecurity.com/) 168
Vulners (https://vulners.com/) 169
Motores de búsqueda para filtraciones de datos generales ***171***
CRT.sh (https://ui.ctsearch.entrust.com/ui/ctsearchui) 171
FilePursuit (https://filepursuit.com/) 171
Gray Hat Warfare (https://buckets.grayhatwarfare.com/) 172
IntelligenceX (https://intelx.io/tools) 173
LeakIX (https://leakix.net/) 174
TinEye Reverse Image search (https://tineye.com) 175
Motores de búsqueda específicos para filtraciones relativas a cuentas de usuario ***177***
Hunter.io (https://hunter.io/) 177
LeakCheck (https://leakcheck.io/) y Have I been pwnd? (https://haveibeenpwned.com/) 177
Motores de búsqueda para código fuente o ejecutable ***179***
Grep.App (https://grep.app/) 179
Joe SandBox (https://www.joesandbox.com/) 179
SearchCode (https://searchcode.com/) 180
Virustotal (https://www.virustotal.com/gui/home/upload) y VT Intelligence (https://www.virustotal.com/gui/intelligence-overview) 182
Cuando el fin justifica los medios **184**
Phishing y Spear phishing para la obtención de información ***184***
SET 185
Gophish 196
Evilgnx2 197

Investigaciones en la Dark web ***197***

😇 Anexo: Un amigo/a quiere saber más de estos temas (bueno, vale, el amigo/a soy yo): Sitios donde aprender más sobre OSINT **199**

Guías y artículos ***199***

Repositorios de documentación, guías y herramientas ***199***

To defeat an enemy, you must know them

Not only their battle tactics

But the history, philosophy, art… […]

So, we can always be one step ahead of them

Almirante Thrawn

¿Qué es esto y para que me sirve?

Introducción

OSINT significa ***Open Source Intelligence*** (inteligencia sobre fuentes de información abierta, es decir, sacar información de "cosas" que uno se encuentra en Internet) y es **uno de los aspectos clave** para comprender la ciberseguridad que rige *Internet* hoy en día. Algunos alumnos lo llamaron el año pasado "ser un/una cibermaruja", y yo creo que es una definición que capta la idea base 😊.

Hacer investigaciones con este tipo de información abierta implica **hacerse muchas preguntas sobre individuos o entidades a estudiar (en adelante, objetivos) y dar una respuesta más o menos precisa a ellas**. Para ello se deben hacer una serie de actividades que, a ojo de una persona que crea que la seguridad tiene un componente puramente técnico, pueden ser sorprendentes o "no serias". **Pero nada más lejos de la realidad**.

En un mundo cada vez más y más interconectado, los **delitos basados en el mal uso de información expuesta** (normalmente de forma indebida) no van a ir a menos, **sino todo lo contrario**. Estos delitos ocurrirán sobre personas, sobre empresas y sobre los productos software que ejecutan. Nadie se libra, y la única "vacuna" es estar lo más formado y alerta posible, aunque nada es infalible. Trabajar con esta clase de información es **el medio favorito para recopilar información que desencadene ataques de *phishing* y / o *spear phishing* muy efectivos**, que han hecho caer gigantes. Para muestra la brecha de Febrero de 2023 de *Reddit*. Y tantas, tantas otras que han pasado y pasarán.

¿Por qué? Porque hemos concentrado nuestros esfuerzos en defender máquinas e infraestructuras, lo cual está muy bien **y es muy necesario**. Pero quizá no le hemos dedicado el mismo esfuerzo a defender a las personas. Y esas personas operan esas máquinas e infraestructuras. Y tienen un perfil de conocimientos técnicos extremadamente variable. **Y son el eslabón más débil**. El error humano es el responsable del 82% de las brechas de datos (https://www.grcelearning.com/blog/human-error-is-responsible-for-85-of-data-breaches). Por tanto, en esta parte de la asignatura vamos a dedicar un apartado a las personas, ya que en las otras se lo estamos dedicando a las máquinas (¡y al código!). Creemos que así saldréis mucho mejor formados para el futuro.

Y esta es una disciplina de la seguridad que es muy útil y atractiva, con muchas ramificaciones que tocan gran cantidad de aspectos técnicos y también sociales, usables desde el lado "defensivo" y el "ofensivo". Esto lo hace un aspecto de la seguridad muy atractivo para mucha gente, y es mi intención usar este documento como "*core*" para **desarrollar una asignatura completa acerca de este tema (el A-21 "Poseidón", pero no Lite** 😊**)**, con una organización más extensa y con más conceptos, organizados para tener una idea de esta disciplina más amplia y transversal. Pero por el momento, ¡disfruta de este "*crash course*" en OSINT! 😊

El delito del *doxing*

No debe confundirse las investigaciones basadas en OSINT con el ***doxing*** (también llamado *'doxxing'*), una abreviatura de "*Drop Documents*". Este es **un tipo de ataque cibernético** (es decir, **un delito**) en el que el atacante **roba** la información personal de la víctima y la filtra a Internet.

En cierto sentido ambas operaciones están relacionadas. En ambos casos se intenta obtener toda la información posible sobre un objetivo, pero en el caso del OSINT se hace con fuentes abiertas, o dicho de otra forma, información de libre acceso. Esto implica que se hace **de manera legal**, salvo si luego usamos esa información para extorsionar o causar alguna clase de daño a la víctima. En el caso del *doxing* se hace de fuentes restringidas o no públicas a las que se tiene acceso, lo que lo hace **una actividad ilegal**. *¿Motivo?* acceder a información privada de una persona **es un delito**, salvo que cuentes con un mandato legal o la ley te ampare a hacerlo por ser un caso muy específico. **Divulgar información privada luego es siempre un delito**.

En esta asignatura NO queremos delincuentes, sino gente que sepa cómo actúan para poder defenderse contra ellos. Es muy distinto, aunque a algunas personas les parezca que no. Sin saber las armas con las que cuenta el enemigo, difícilmente podremos a) establecer contramedidas que funcionen y b) detectar puntos vulnerables a las armas del enemigo. Dentro de nuestras posibilidades, en este documento no nos hemos dejado nada en este sentido. Lo que ves es crudamente lo que hay, **y te aseguro que hay cosas mucho peores** que no he incluido simplemente por falta de tiempo o porque son refinamientos de cosas que verás aquí. Esto te servirá para darte cuenta de lo mal que estamos, y actuar en consecuencia en tu entorno o futuro trabajo.

Por lo general, el objetivo del *doxing* es **acosar o dañar** a la víctima de alguna manera. Las variantes dependen del contexto y la víctima, y no es posible enumerarlas aquí. Pero te puedes hacer una idea a medida que vayas leyendo los ejemplos que te pongo y dejes volar tu imaginación.

Desafortunadamente, **el *doxing* es más común de lo que piensas**. Todo tipo de personas son potenciales víctimas, no importa si eres o no famoso. Hay gente "mala" por ahí fuera que puede querer hacerte daño a ti o a los tuyos por alguna razón (envidia, superioridad moral supuesta, no considerarte "digno", por pertenecer a un colectivo determinado (edad, empresa, ideología, ...), hay miles de razones...). **Esta parte de la informática enlaza directamente con temas de siquiatría o sicología**. O quizá no haya ninguna, sino solo por "diversión" (esto es directamente un caso de estudio para siquiatría).

Para evitar el *doxing*, es muy importante **tener mucho cuidado con el tipo de información que pones en Internet**. A veces puede haber un escenario de *doxing* "controlado", donde una persona haciendo una auditoría ofensiva de seguridad tiene la autorización para "robar" determinada clase de información que pueda usar luego para continuar con una cadena de ataque como la comentada del MITRE ATT&CK (**Tema 7**), **aunque jamás exponerla fuera, usarla para dañar a otros o publicarla**. Somos profesionales, **NO delincuentes**. Aquí se trata de saber qué se podría robar simulando ser un adversario, no ser un "actor de método" (seguro que entiendes lo que digo, *¿verdad?*).

¿Qué clase de información se recoge?

En cualquier caso, ya sea de fuentes públicas o no públicas, se trata de **saber cosas del objetivo** que luego puedan usarse para **ataques** (*phishing* que implique suplantación o convencer al objetivo de algo, averiguar credenciales, suplantar a alguien...) u obtener **cualquier clase de ventaja** táctica o estratégica sobre dicho objetivo. Esto incluye aspectos como los siguientes, aunque luego veremos más:

- Nombre real / apellido
- Características físicas (peso, altura...), aficiones, hobbies, gustos personales....
- Direcciones de correo electrónico
- Números de teléfono (con lo que tendremos acceso a la víctima vía mensajería instantánea)
- Perfiles en redes sociales
- Dirección o ubicación
- Empleos (actuales y pasados)
- Información de cuentas bancarias, tarjetas de crédito...
- Opiniones sobre determinados asuntos, ideologías, fe religiosa...

Los "**mandamientos**" de la recopilación de inteligencia de fuentes abiertas u OSINT son cuatro:

1. ***Open-Source***
2. **Público**
3. **Gratuito**
4. **Legal**

Es decir:

- Debe ser inteligencia **objetiva**, libre de sesgos. Por ejemplo, si es la opinión de alguien en una red social de algo, debe ser **contrastada** con una o varias fuentes fiables para evitar caer en bulos o ***fake news***.
- Los datos deben estar **disponibles en fuentes públicas**, como las disponibles en Internet, aunque el término no se limita estrictamente a Internet, sino que incluye todas las fuentes disponibles públicamente (físicas o digitales) donde el auditor pueda obtener la información. Debe poder obtenerse **de forma gratuita**, salvo casos muy concretos justificados. No importa si se encuentra dentro de periódicos, *blogs*, páginas web, tweets, redes sociales, imágenes, podcasts, videos...siempre y cuando sea público, gratuito y legal.

Con la información correcta podemos obtener una gran ventaja en nuestras operaciones de reconocimiento (fase ***TA0043 Reconnaissance*** del ***MITRE ATT&CK***, **Temas 5 y 7**) y acelerar cualquier investigación de la empresa / personas. Aunque se suele asociar el OSINT a la guerra cibernética, ataques cibernéticos, la ciberseguridad, etc. (y ciertamente es parte de ello), y por ello parecen acciones realmente complejas, las operaciones para buscar entre este tipo de información suelen ser muy poco complicadas una vez conocidas las técnicas, las herramientas y qué buscar a partir de cada información encontrada. Para muestra de lo que hay que hacer tenemos estas operaciones:

- Hacer preguntas de muchos tipos en cualquier **motor de búsqueda**. Si, esto incluye los motores -chat modernos basados en IA como ***ChatGPT*** o ***Bing Search***. Puedes preguntarle datos de una persona y seguramente sea capaz de darte una respuesta veraz y coherente en bastantes casos.
- Investigar **foros públicos, blogs o redes sociales** (identificándolos previamente) sobre el comportamiento de una determinada persona en ellos, para ver qué cosas cuenta, revela, "se le

escapan" o admite conocer creyendo que lo hace anónimamente. La gente suele revelar muchas cosas en estas situaciones.

- Investigar **documentos públicos** acerca de cosas que le hayan pasado al objetivo (concesiones de becas, premios, proyectos, multas, impagos...).
- Investigar si tiene alguna **posesión o propiedad** a su nombre (coches, casas en venta…)
- Cualquier **información cotidiana** que creamos que nos pueda servir para nuestros fines
    - Nombres completos de los **empleados**, los **roles de trabajo** y el **software** que utilizan
    - **Datos antiguos** almacenados en caché de *Google* / *Wayback Machine*: a menudo revela información interesante.
    - **Números de teléfono**, **direcciones de correo**, **redes sociales** o resultados de *Google*.
    - **Fotografías y videos** en sitios sociales comunes para compartir fotos
    - *Google Maps* y similares para recuperar imágenes de la **ubicación geográfica** de los usuarios.

Fuentes de datos no nos van a faltar (¡mira esta imagen!), ahora hay que saber cómo buscarlas y qué buscar

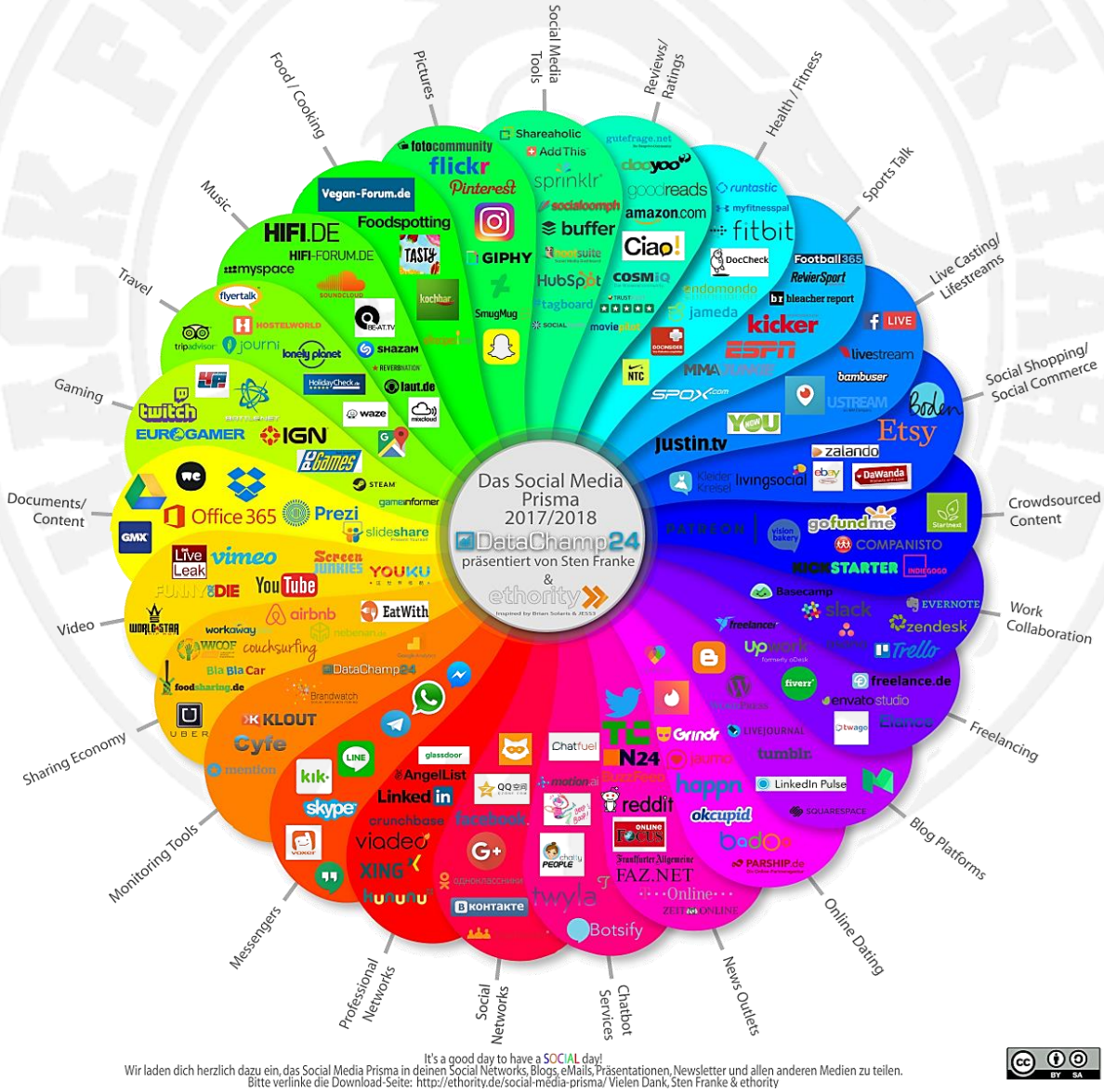


Figura 1. Orígenes de información según el tipo de documento o dato buscado

To "*engage*" o "*to not engage*" he ahí la cuestión

Para hacer estudios OSINT hay **dos tipos de técnicas**, que dependen del grado de interacción que se haga con el objetivo.

Técnicas pasivas

La principal característica de estas técnicas es que **el investigado NUNCA puede saber que está siendo investigado**, es decir **JAMAS debes contactar con él/ella** (**o cualquier miembro de la organización/empresa**), bajo ningún concepto. **Ni siquiera bajo una identidad falsa**. Esto se hace cuando no debes despertar sospechas de ninguna clase en nadie investigado, y así el "rastro" de tus investigaciones quedará "sumergido" en logs enormes de muchos sistemas ajenos que nadie mirará casi con total probabilidad, porque en el fondo tú no has hecho nada ilegal ni de lo que se te pueda acusar. **La información estaba ahí de manera pública y tú solo has dado con ella**. Si no hay uso criminal de la información ni das razón alguna para que alguien rastree tus pasos, todo lo que hagas se perderá, como lágrimas en la lluvia 😉

Además de minimizar el riesgo de detección, este tipo de OSINT usa mayoritariamente **servicios de terceros para hacer su labor** (por ejemplo, resultados de motores de búsqueda como *Google* o *Shodan* como los que se vieron en el **Seminario 1** y en el **Laboratorio 2**). Dicho de otra forma, el contacto con el objetivo no lo haces tú, sino que **lo hacen otros por ti**, que procesa la información y te la da clasificada y ordenada de alguna forma para que tú la consumas como más te interese. Como ese tercero se dedica a eso, es raro (aunque posible) que el objetivo lo bloquee. También depende del tercero, claro. Bloquear los agentes de investigación ("arañas web") de máquinas de *Shodan* es más frecuente que bloquear los de *Google*, *Bing* o similares, salvo que uno decida no ser indexado y pertenecer a la ***Deep web*** (https://es.wikipedia.org/wiki/Internet_profunda).

No obstante, esa capacidad de pasar desapercibido tiene un precio: **mayor nº de falsos positivos** y **"ruido" en los resultados** (resultados que parece que son correctos, pero al mirarlos resultan ser de otra entidad o persona con nombre o características parecidas).

Técnicas activas

Este tipo de técnicas implican el **contacto directo con el objetivo**. El ejemplo más típico es un escaneo de puertos con `nmap`, como los que se describen en los **laboratorios 1, 2 y 8-9**. Salvo que tomes **medidas de anonimato** a nivel de persona o de máquina que hace la operación (que veremos en la siguiente sección), **van a saber que eres tú**. En cualquier caso te van a detectar, y por tanto, **te pueden bloquear** si consideran que lo que haces es anómalo o de carácter ofensivo. Todo se basa en una carrera entre las capacidades de detección del objetivo y tu habilidad para "ocultarte". A cambio de perder el anonimato y de aumentar la probabilidad de bloqueo, obtienes una **mayor fiabilidad de resultados** en general y normalmente **más detalles del objetivo**.

Por tanto, debido al riesgo de descubrir al auditor que estas técnicas conllevan, y salvo que la operación que estés haciendo te autorice por contrato a usar este tipo de técnicas directamente sobre el objetivo, **debe contemplarse incluir acciones para aumentar el anonimato**, tanto a nivel de persona como de máquina auditora, para aumentar el realismo de las acciones. Obviamente estas técnicas de anonimato también pueden combinarse con técnicas pasivas para un nivel extra de precaución que roza seguramente lo excesivo, pero evita ser descubierto por descuido (por ejemplo, mirando una historia de *Instagram*). Veamos qué técnicas son.

NOTA: En esta asignatura queda completamente prohibido el uso de técnicas que impliquen el "*engagement*" de cualquier clase con personas, salvo que sea una simulación y la otra persona sea un compañero que sepa exactamente que lo estás haciendo.

El "ciclo de inteligencia"

Los trabajos donde se recoge información de un objetivo **no consisten solo en recabar información y nada más**, sino que dicha información debe procesarse en un ciclo: una vez obtengamos información sobre un objetivo es posible que podamos usarla para obtener de más objetivos asociados, o bien para profundizar en lo que hemos encontrado, por lo que es un proceso que fácilmente se retroalimenta. Además, el proceso se puede estructurar en cinco fases.

- **Preparación**: Que es un poco lo que hablaremos en la siguiente sección, preparando el entorno para recoger información e identificando las mejores fuentes de información o herramientas
- **Recolección**: De lo que irá la mayor parte del documento 😊
- **Procesamiento**: La información recogida debe organizarse y ordenarse correctamente para facilitar su análisis
- **Análisis**: Se interpreta la información y se sacan todo tipo de conclusiones de ella.
- **Presentación**: Se hace un informe con la información encontrada y sus conclusiones, que es el objetivo final de toda esta actividad.

La siguiente imagen describe este proceso:

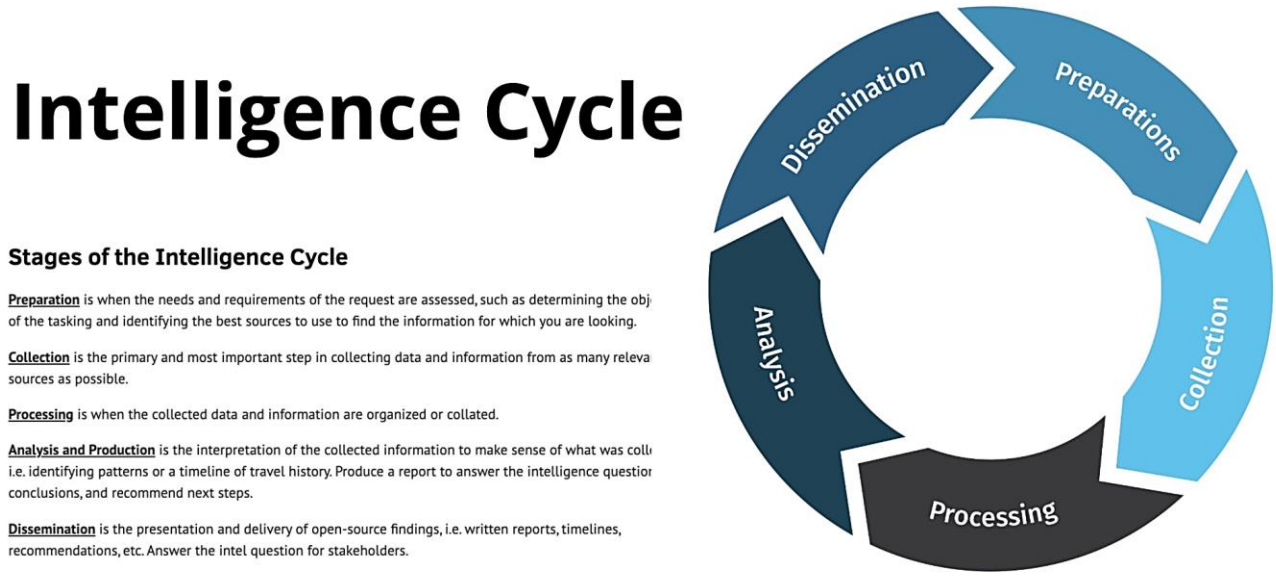


Figura 2. Ciclo de inteligencia. Fuente: https://www.sans.org/blog/what-is-open-source-intelligence/

Anonimato "personal": Crear una identidad online falsa

Uso de emails temporales para registrarse en servicios

Una de las principales razones por las que recibimos *spam* es por usar nuestras cuentas principales para registrarnos en servicios cuya política de privacidad deja mucho que desear, se roba su base de datos de clientes, o directamente está pensada para que aceptemos que nos lleguen todo tipo de comunicaciones por defecto (¡la letra pequeña que no nos leemos!). También al registrarnos en algún servicio online podemos dejar la cuenta visible por descuido.

En operaciones ofensivas como las que describimos aquí, una cuenta de esta clase nos permite registrarnos en algún servicio o foro que use el objetivo para estudiar lo que hace sin revelar ninguna clase de información de nosotros mismos. Obviamente, **no deben usarse para recibir notificaciones periódicas de nada, porque como veremos son desechables**.

Podemos mitigar parte de estos problemas usando **una cuenta de *email* desechable**. También se aplica en escenarios donde tenemos que dejar una persona de contacto para una web, servicio técnico, etc. pero no queremos exponer la identidad de una persona real, creando así una ilusión que "engañe" al interlocutor. Un proceso de registro en una RRSS normalmente no admite este tipo de cuentas, y tampoco convendría usarlas si queremos mantenerlas durante un tiempo (el correo de recuperación desaparecería), por lo que debemos inclinarnos por crear una cuenta nueva en un proveedor más estándar en estos casos, **obviamente usando datos que no nos vinculen con la misma**. En otras palabras, siempre hay que investigar con emails desechables que no puedan asociarse a nuestra identidad, lo que cambia es la facilidad o velocidad con la que se desechan y el "tiempo de vida" planificado que deben tener 😊.

Los *emails* desechables son un servicio que permite **recibir correos electrónicos en una dirección temporal** que se **autodestruye** tras transcurrir un tiempo o cuando voluntariamente decidamos hacerlo. Son de especial aplicación en foros, portales Wi-Fi (hoteles, congresos…), sitios web y blogs que solicitan a los visitantes que se registren antes de que puedan ver el contenido, publicar comentarios o descargar algo. Aunque parezca increíble, si lo hacemos con nuestro email real en muchos casos estamos abriendo nuestra dirección de correo a ser incluida en una lista de distribución de *spam* y por tanto ser "asediado" por correos de este tipo, además de poder ser identificado. Lo mismo ocurre en cualquier dirección de correo pública que dejemos en una web. Con una cuenta desechable, esto no es posible: en cuanto "matemos" la cuenta no quedará rastro de ella.

Aparte de la temporalidad de estas direcciones, hay otra característica de estas que debemos tener en cuenta, y es que, si bien no tienen problemas en recibir correos de otras direcciones, **el enviar correos desde las mismas está muy limitado**, falla mucho o directamente es un servicio de pago. No obstante, para registrarse en diferentes sitios solo es necesario recibir correos.

Email On Deck

Otro problema adicional de estos servicios es que algunos sitios detectan su uso y no permiten usar una de estas direcciones para registrarlos, y por eso es útil conocer varios 😊. Aunque hay varios proveedores de este tipo de correos, vamos a describir 3 para daros material de trabajo suficiente. **El primero de ellos es *Email On Deck*** (https://www.emailondeck.com/es/eod.php) que nos creará inmediatamente una cuenta de correo de nombre aleatorio acabada en un dominio predeterminado y que podemos usar al momento:

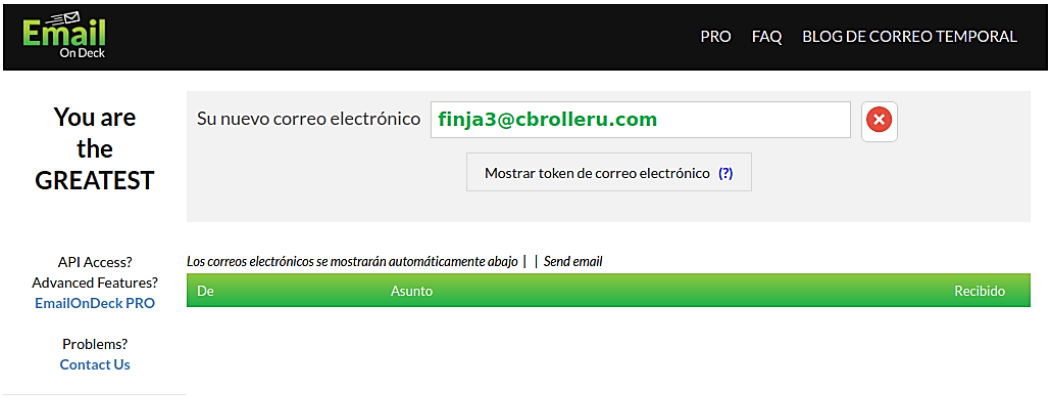


Figura 3. Cuenta de *Email on Deck* recién creada

Como vemos, puede recibir *emails* de *Uniovi* sin problemas:

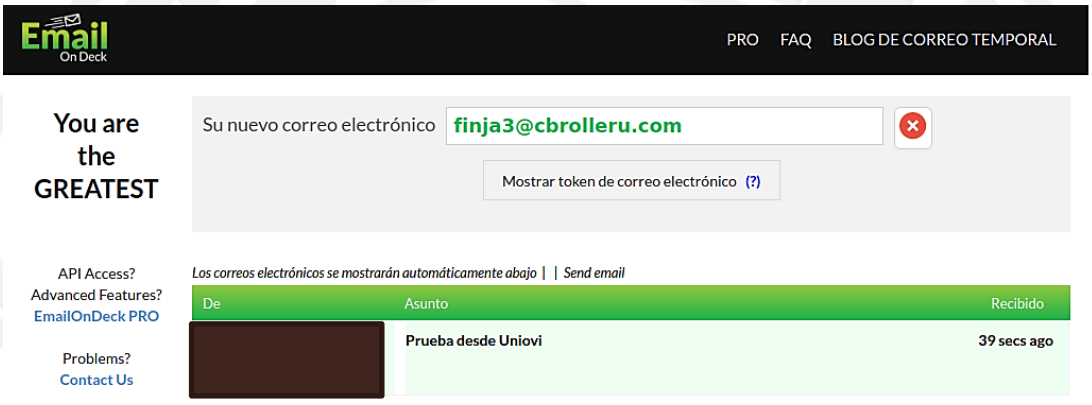


Figura 4. *Email on Deck* recibe correos sin problema

Cuando terminemos de usar esta dirección temporal, podemos pulsar el botón de borrar y ya no quedará rastro de esta:

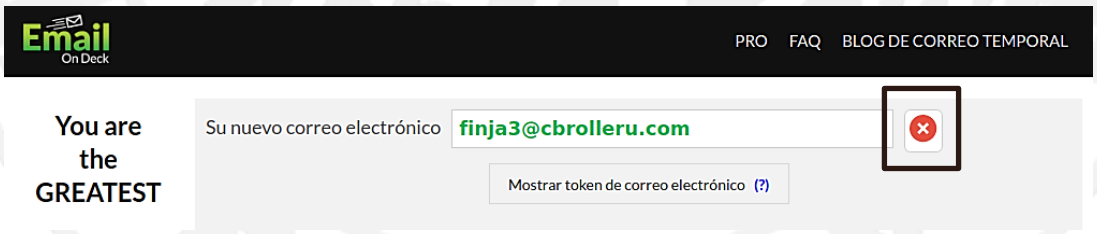


Figura 5. Eliminar todo rastro nuestro de *Email On Deck*

Nuestro siguiente acceso a la web generará una nueva dirección aleatoriamente y no habrá rastro de los correos recibidos en la anterior, **que son irrecuperables**. Enviar *emails* con este servicio requiere una suscripción de pago, pero no lo necesitamos en esta asignatura. No obstante, esto abre la puerta a usarlo en investigaciones "serias" en el futuro, si decidís dedicaros a esto 😉.

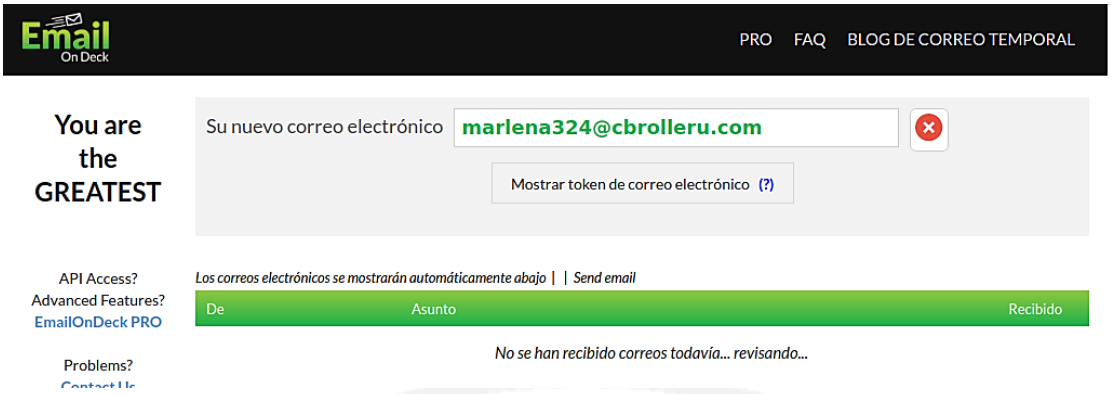


Figura 6. Nueva cuenta de *Email on Deck*

GuerrillaMail

El siguiente servicio que vamos a ver es *Guerrillamail.com* (https://www.guerrillamail.com/es/), de características similares al anterior y que, en principio, permite **enviar también *emails* gratis**, si bien las pruebas de funcionamiento de esta última opción **no han sido satisfactorias** al menos con Uniovi. A priori no puede saberse si fue un problema temporal del servicio, una política de bloqueo/*spam* que tiene Uniovi (con lo que quizá para otros destinos funcione) o por otro motivo. De nuevo, simplemente accediendo a la página ya tenemos un *email* temporal que podemos usar:

Figura 7. Cuenta recién creada con *GuerrillaMail*

Se puede modificar "en caliente" el nombre de la cuenta desde la que enviar un *email* y, como vemos, las funcionalidades para redactar un *email* son las típicas. Como característica añadida, tenemos que se permite enviar adjuntos (hasta 150Mb). Estos *emails* y adjuntos se borrarán al cabo de un día y cabe destacar que **no ofrecen privacidad del todo**, ya que la IP desde la que enviamos el mismo formará parte de las cabeceras del *email* para evitar ser clasificado automáticamente como *spam*. No obstante, si hacemos esto desde una red ToR (sección siguiente) para que también nuestra máquina sea más anónima...(*guiño* *guiño* *codo* *codo*). **NOTA**: En realidad usar la red ToR en investigaciones reales siempre es aconsejable por si acaso 😊.

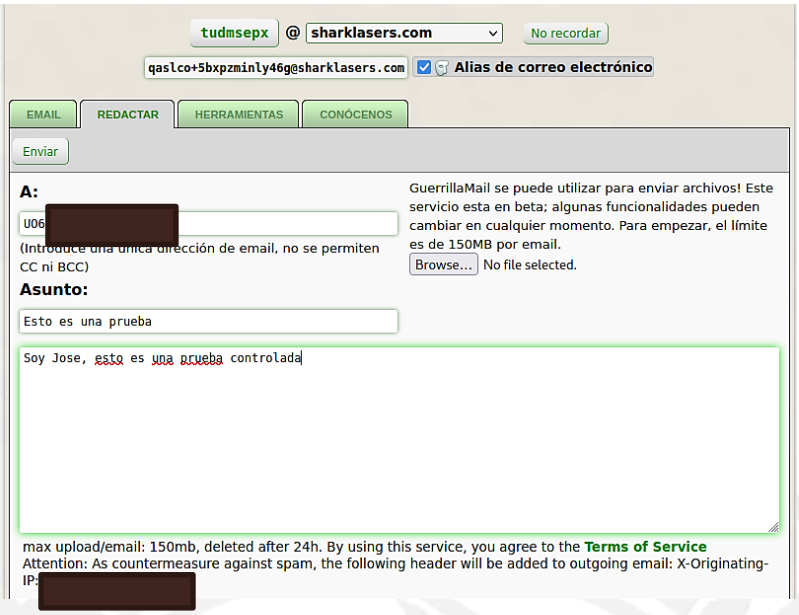


Figura 8. Prueba de envío con *GuerrillaMail*

Otra característica interesante es que podemos cambiar muy fácilmente el dominio desde el que se origina el *email* por si alguno de ellos no fuese admitido por algún servicio:

Figura 9. Dominios disponibles con *GuerrillaMail*

Vemos finalmente como la recepción de *emails* desde *Uniovi* ocurre sin problema. Pulsando en "No recordar" destruiremos la dirección de *email* y los correos recibidos:

Figura 10. Recepción de un email con *GuerrillaMail*

TempMail

Finalmente un tercer servicio de capacidades equivalentes es ***TempMail*** (https://temp-mail.org/es/), de nuevo **solo apto para recibir *emails***.

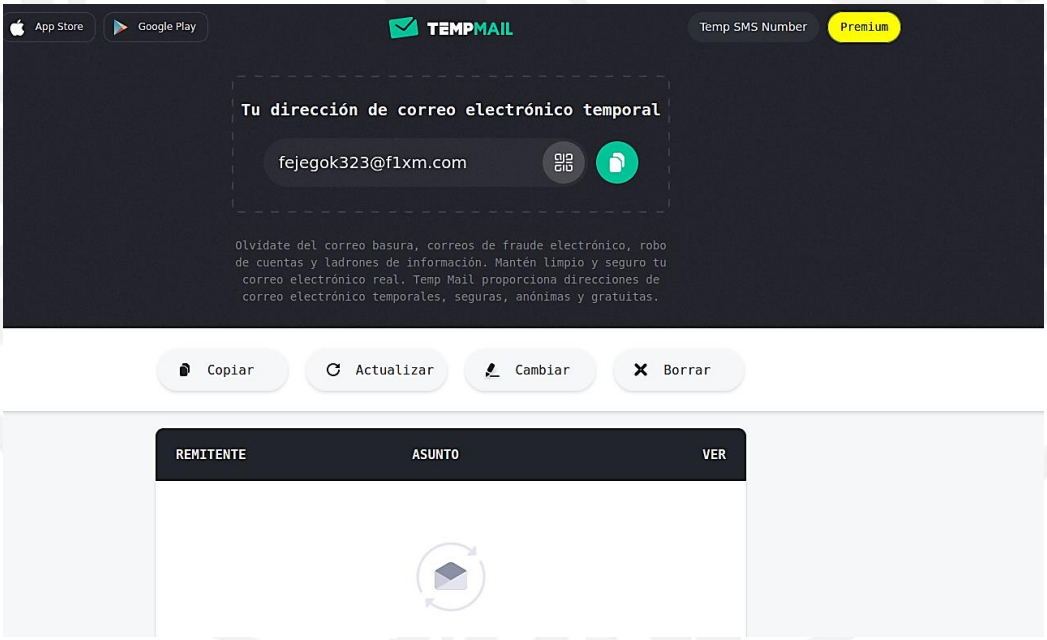


Figura 11. Servicio de correo desechable *TempMail*

Vemos como también recibe correos de *Uniovi* sin problema alguno:

Figura 12. Recepción de correos con *TempMail*

Y nuevamente la opción "*Borrar*" genera inmediatamente una cuenta nueva:

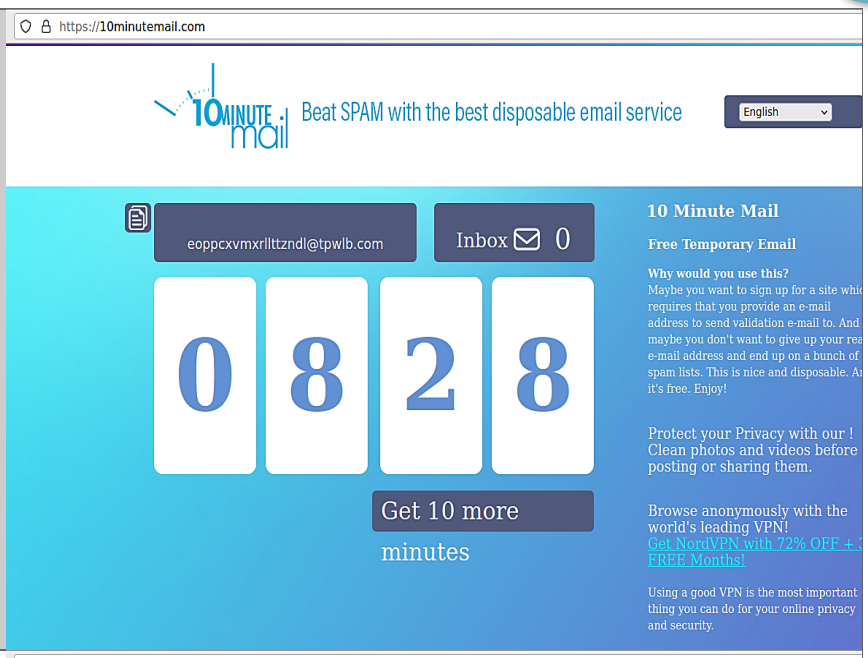
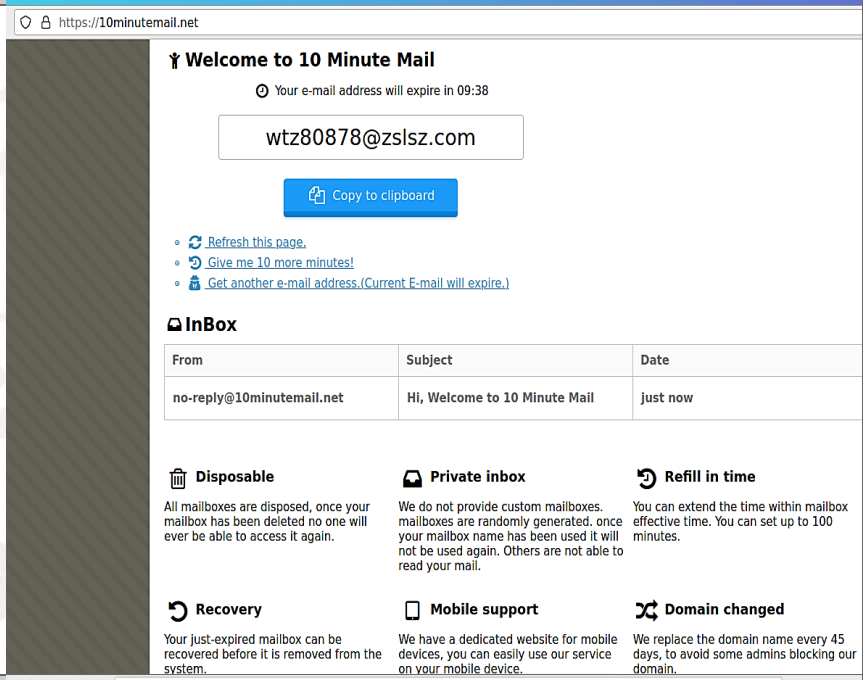
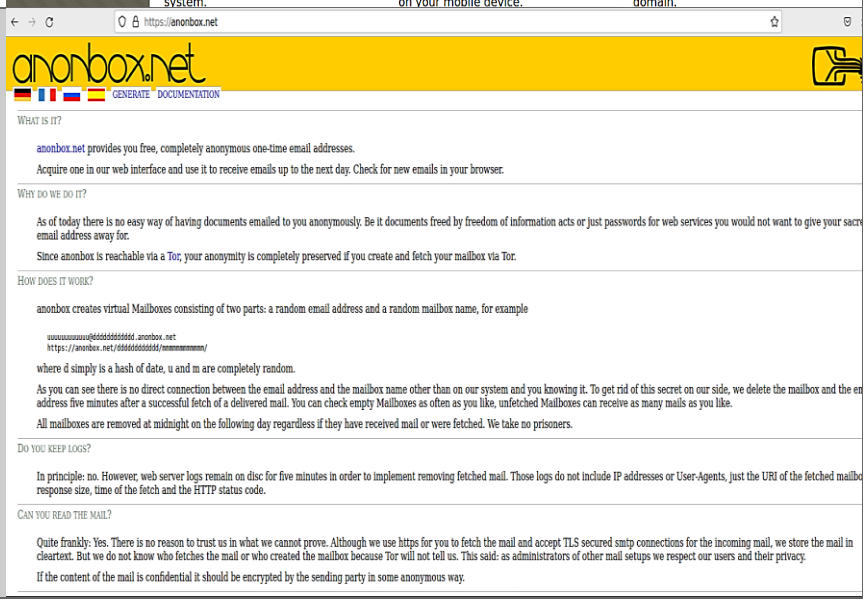
Figura 13. Borrado de correos con *TempMail*

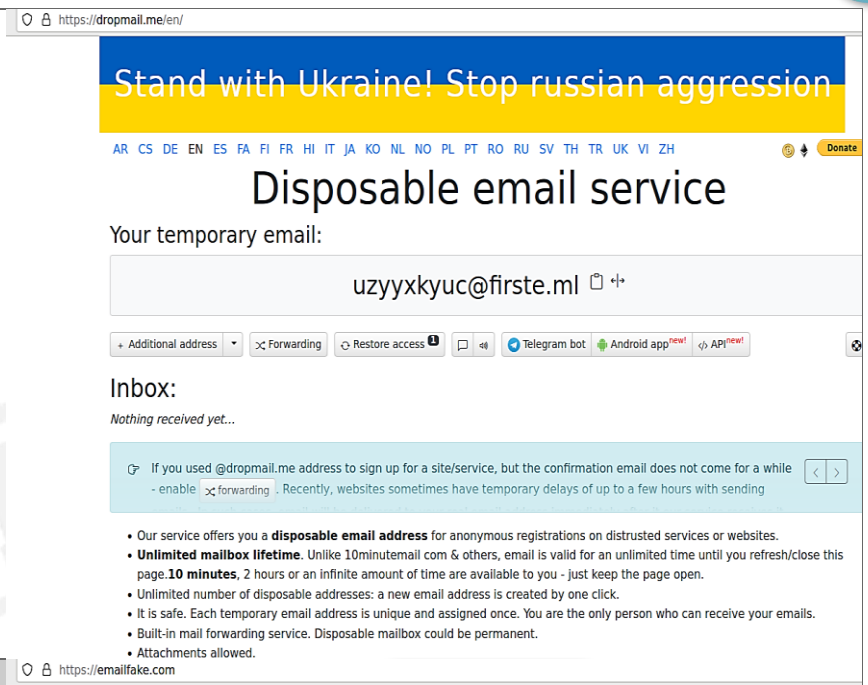
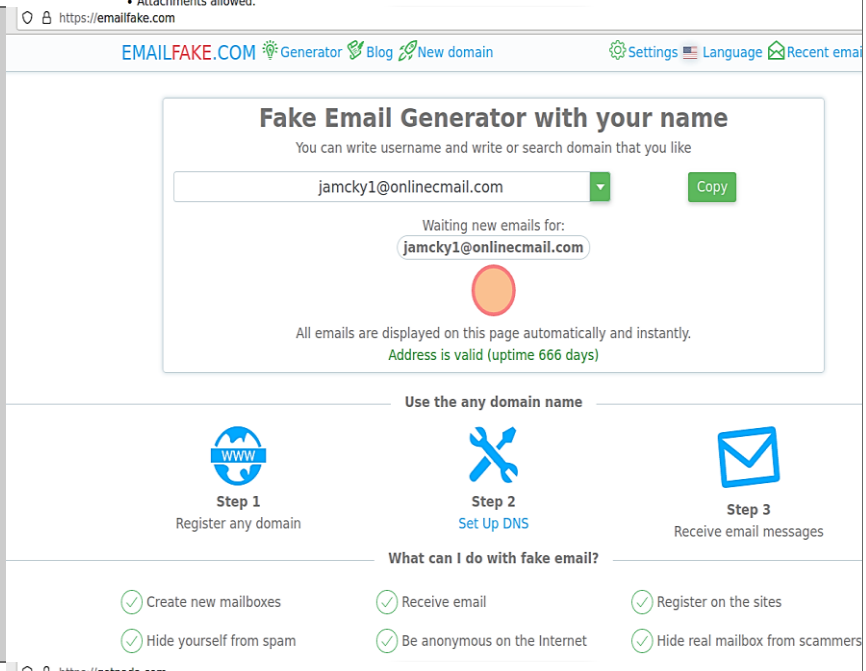
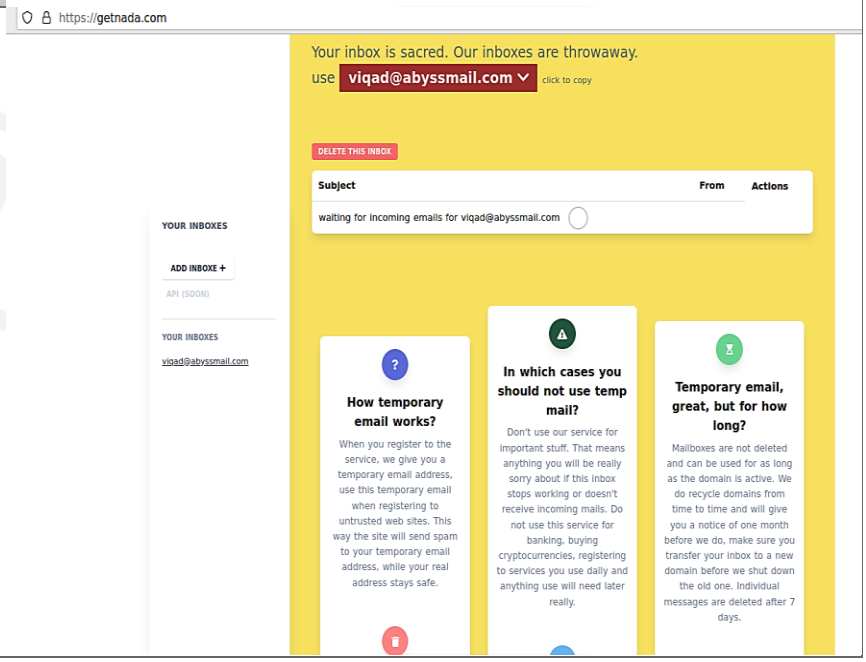
Otros proveedores de email temporales

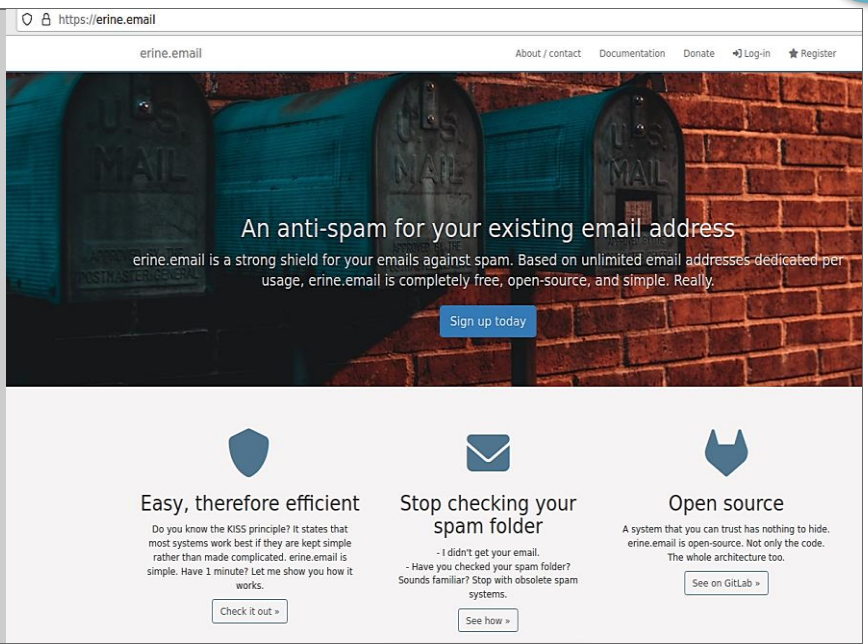
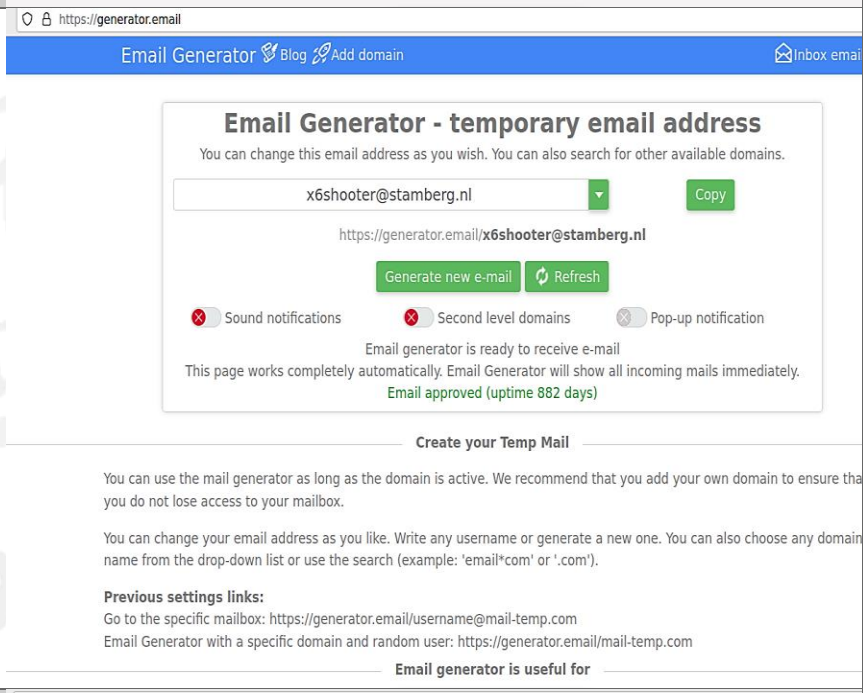
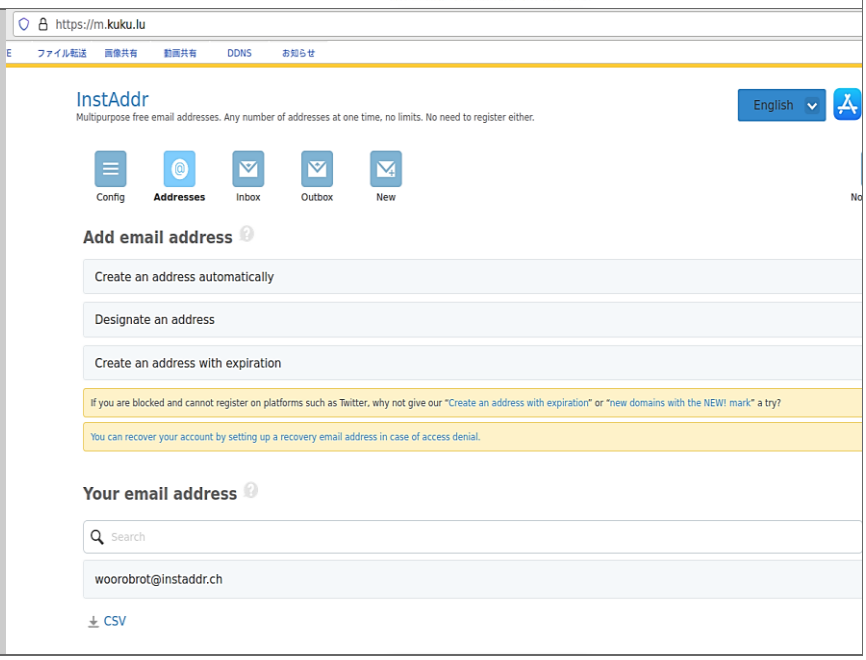
Como decíamos, estos proveedores temporales pueden ser bloqueados por nuestro objetivo o bien pueden desaparecer pasado un tiempo. Por eso es muy importante tener muchos recopilados para usarlos cuando sea necesario. También es posible que algunos de ellos sí admitan enviar correos de forma gratuita, aunque sea con limitaciones. La siguiente tabla lista un gran nº de proveedores equivalentes a los vistos en las subsecciones anteriores. Todos ellos tienen un funcionamiento muy similar a los vistos.

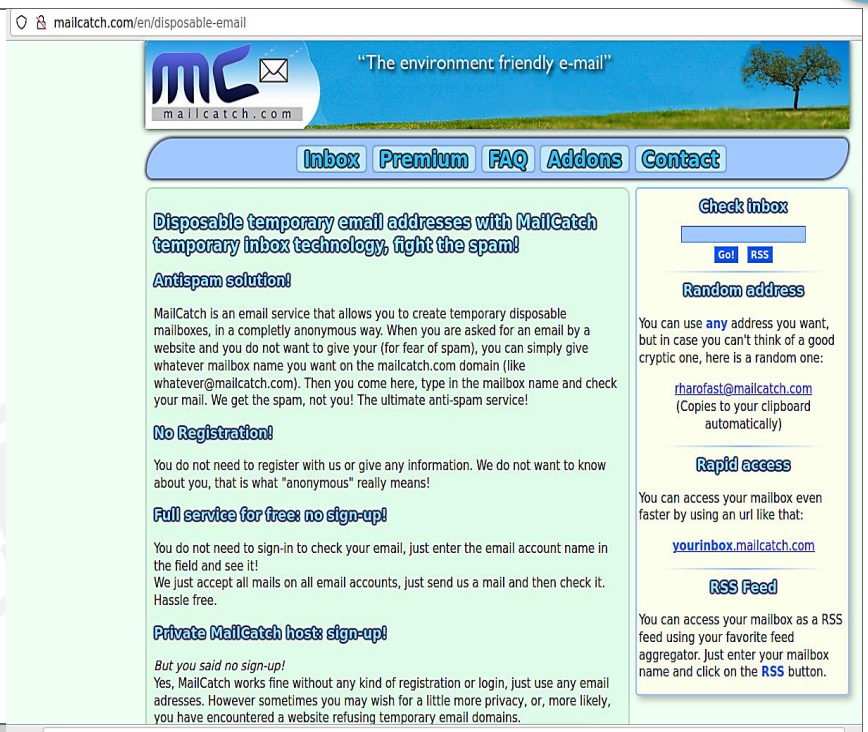
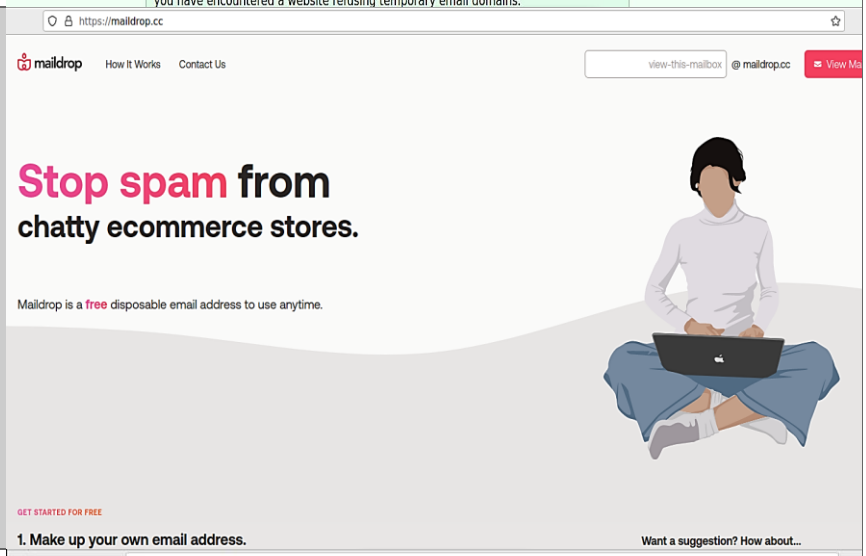
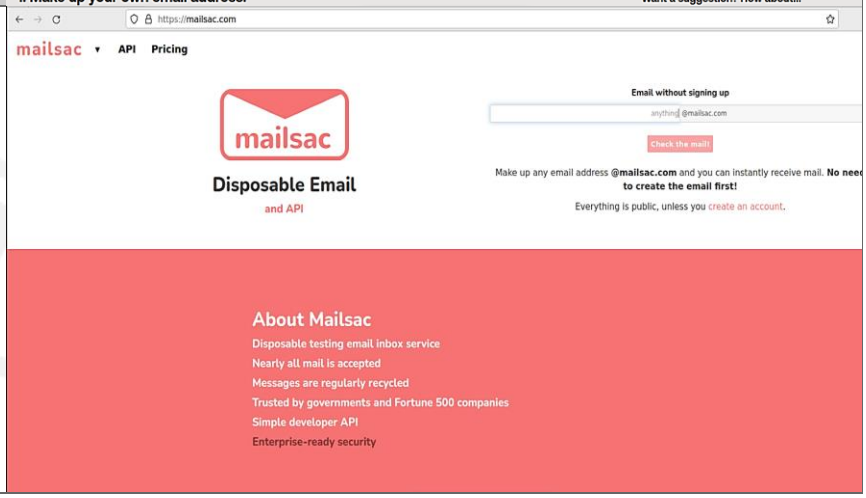
Tabla 1: Proveedores de emails temporales

Proveedor	Imagen del servicio

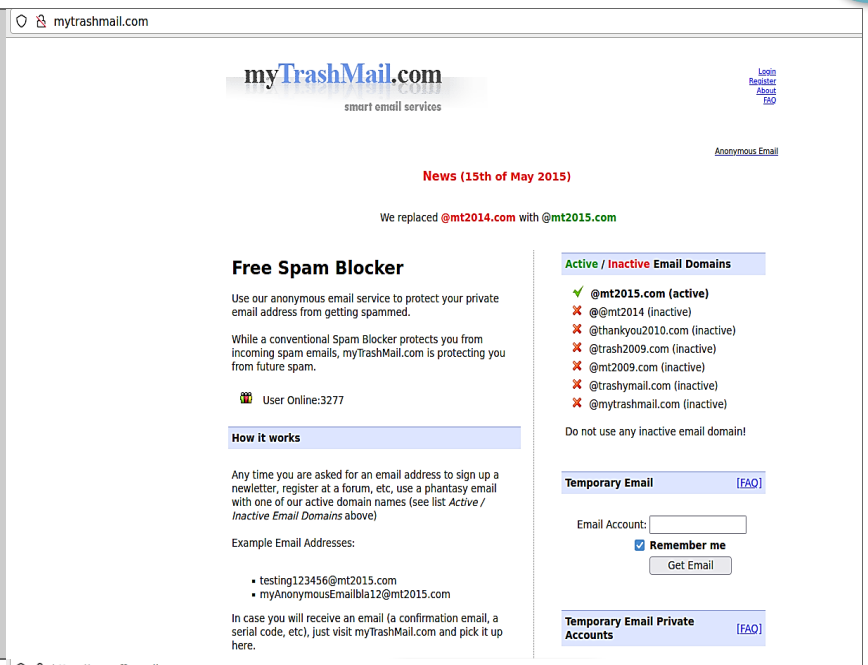
https://10minutemail.com	
https://10minutemail.net	
https://anonbox.net	

https://dropmail.me/en	
https://emailfake.com	
https://en.getairmail.com o https://getnada.com	

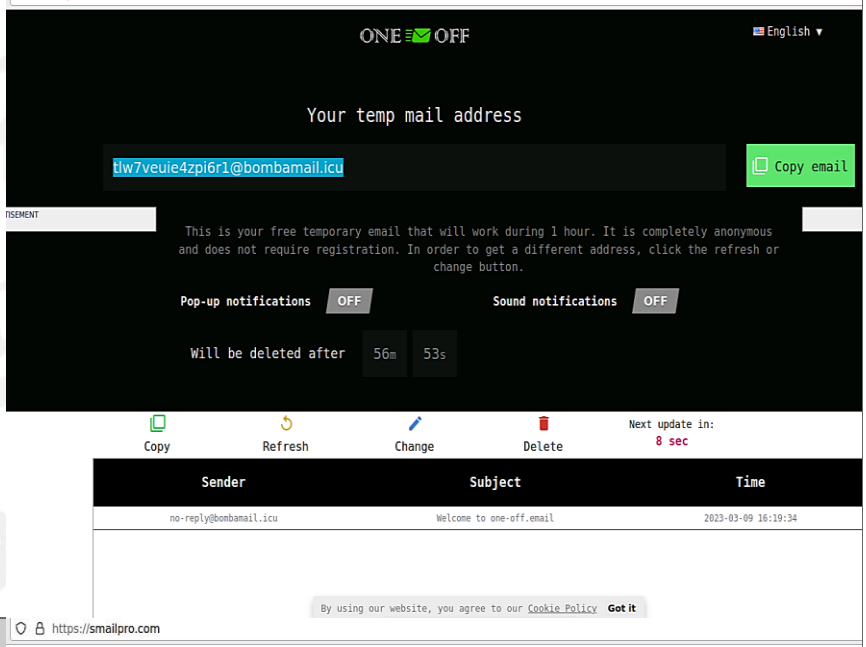
https://erine.email	
https://generator.email	
https://m.kuku.lu	

http://mailcatch.com/en/disposable-email	
https://maildrop.cc	
https://mailsac.com	

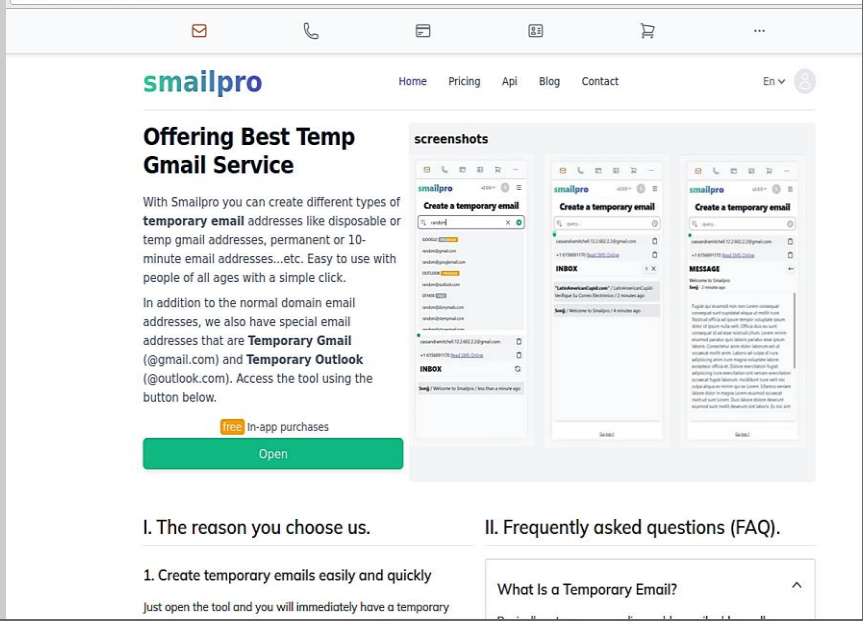
https://mytrashmail.com

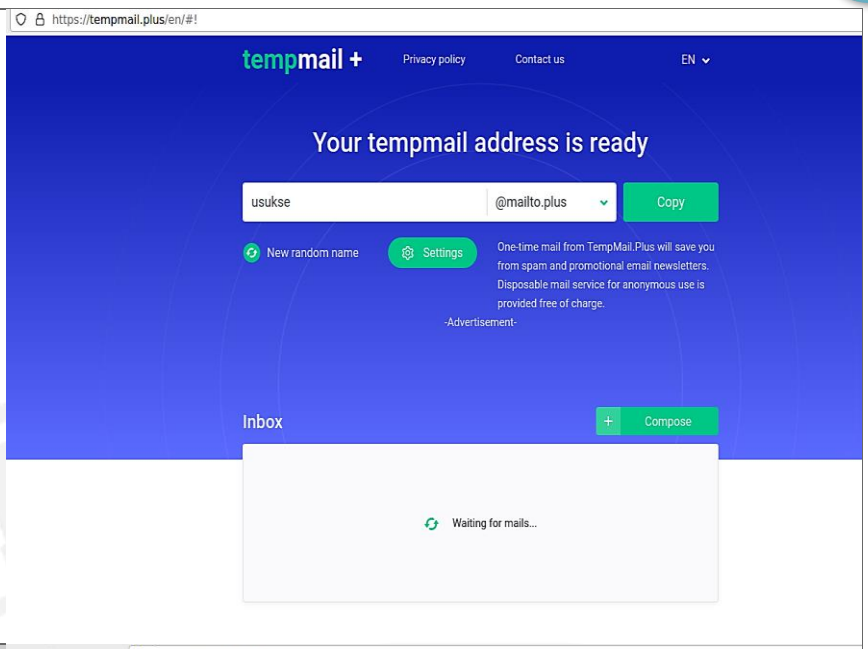
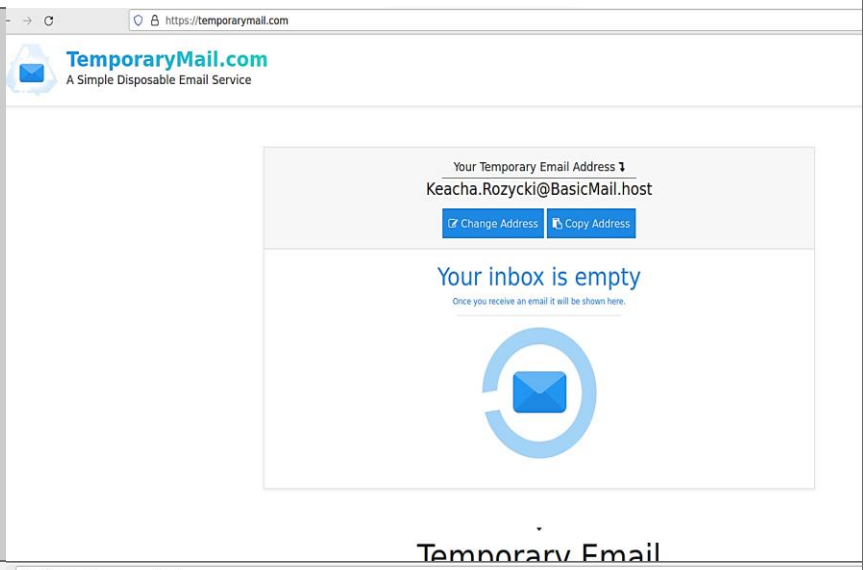
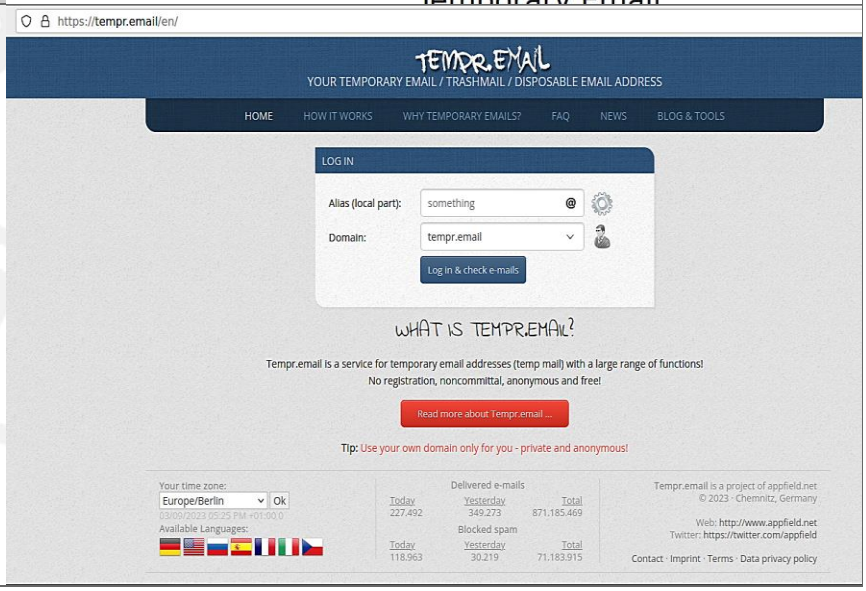


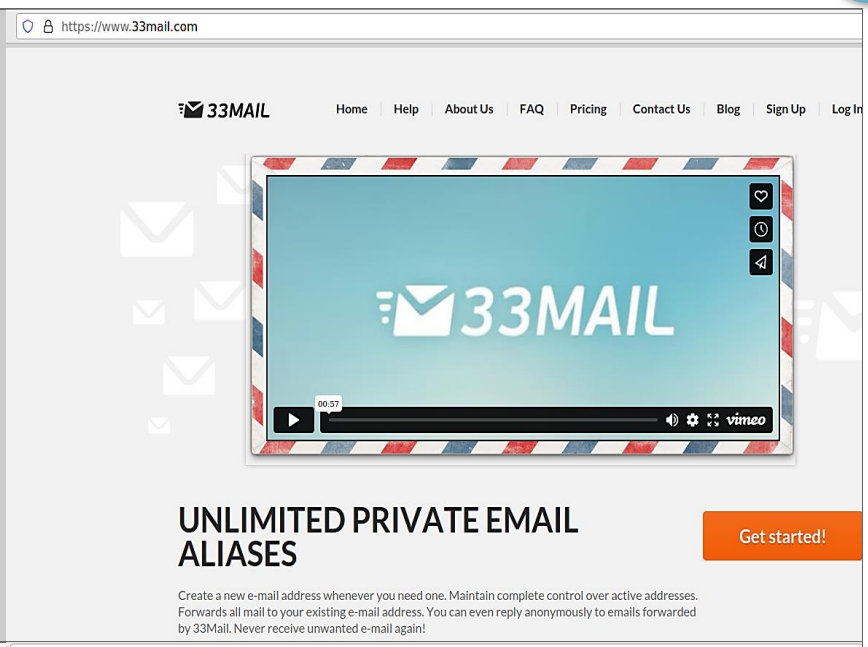
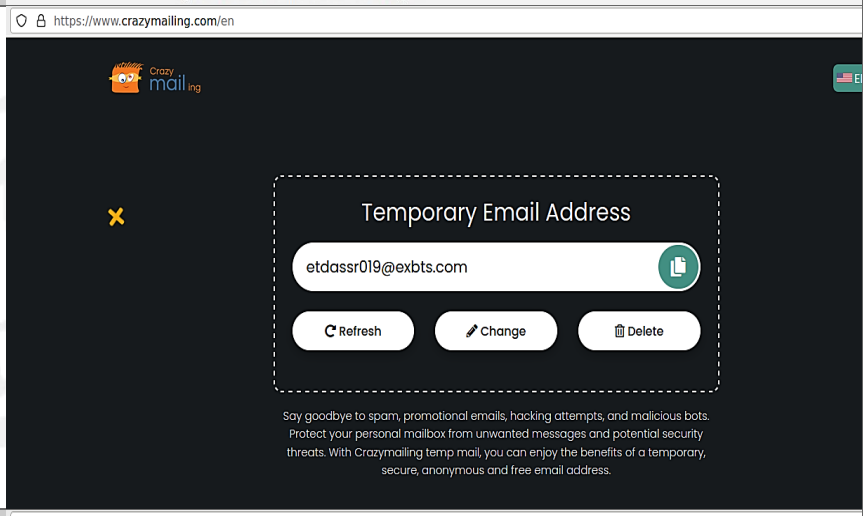
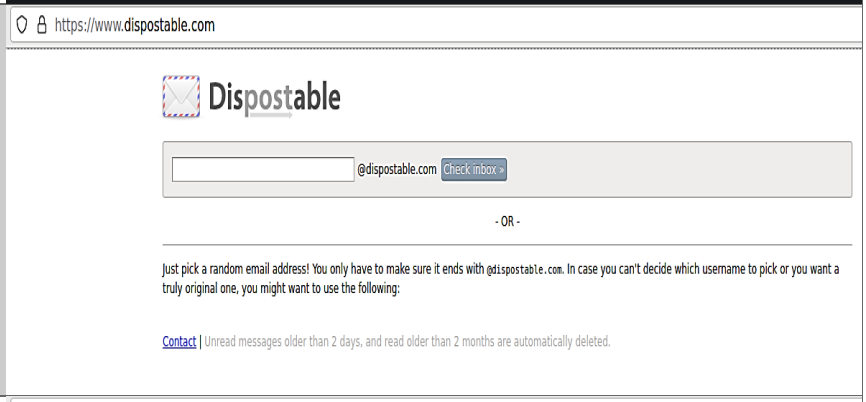
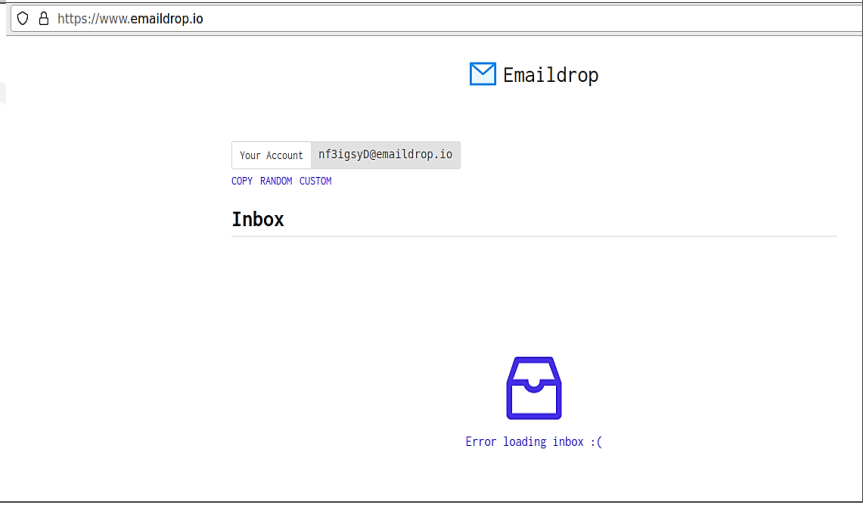
https://one-off.email/ru

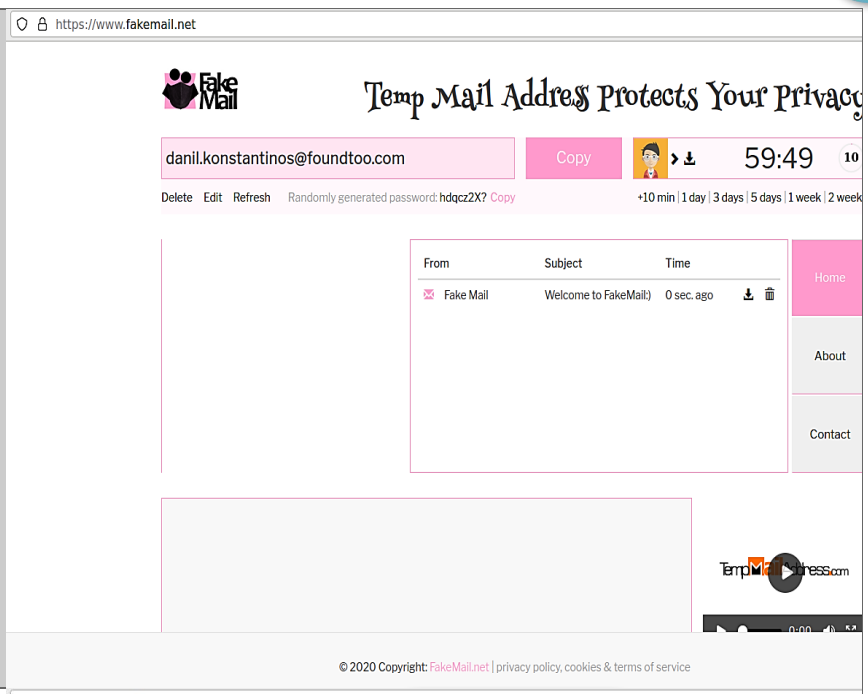
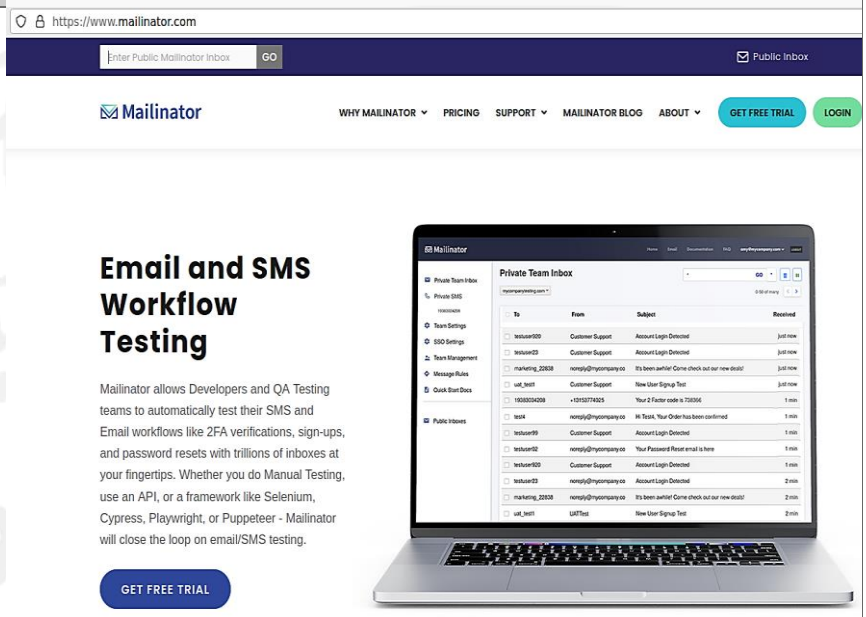
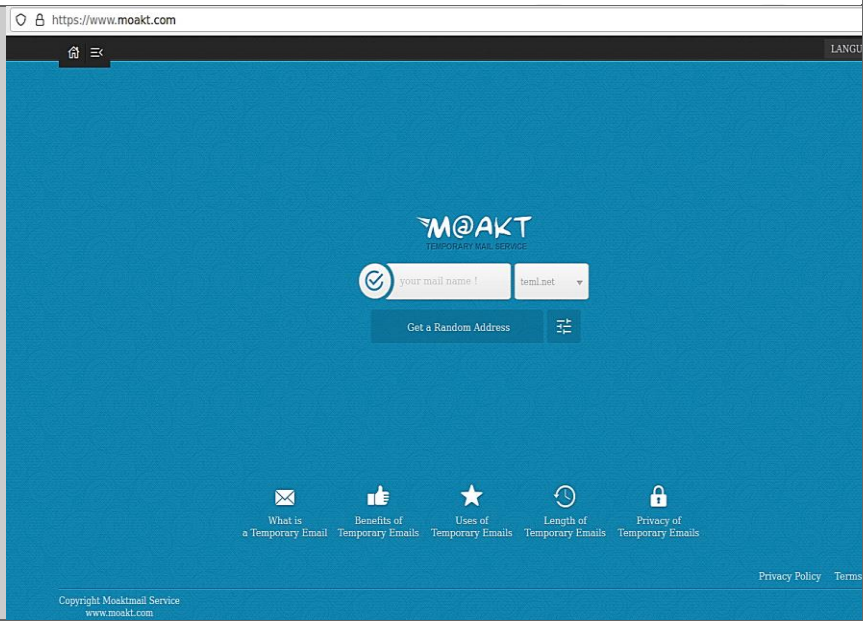


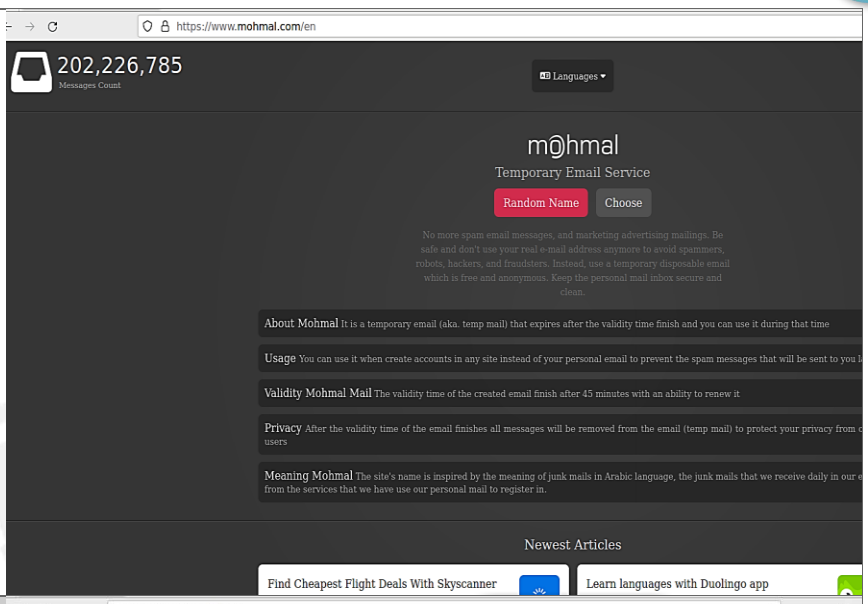
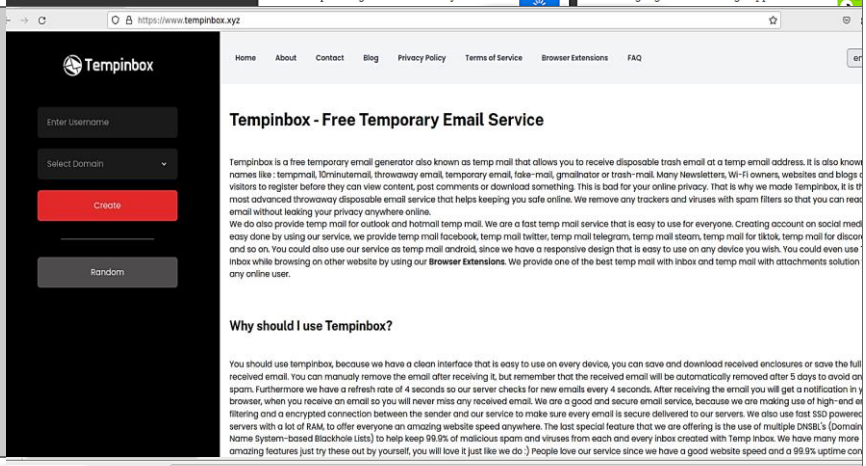
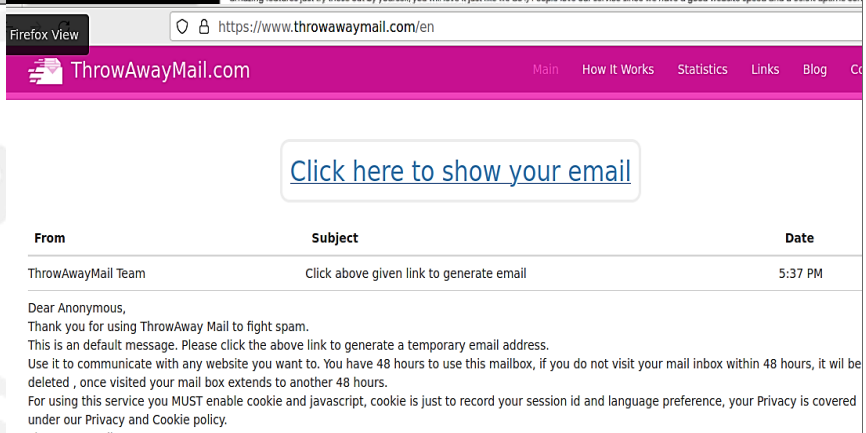
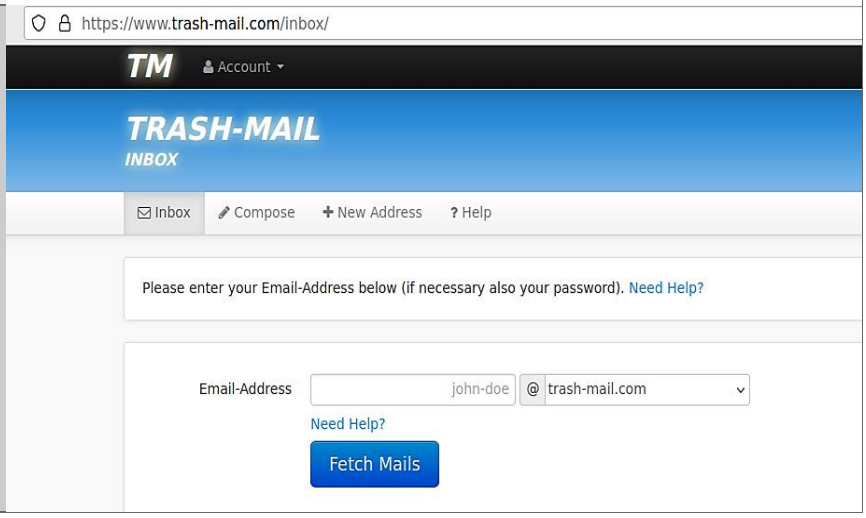
https://smailpro.com

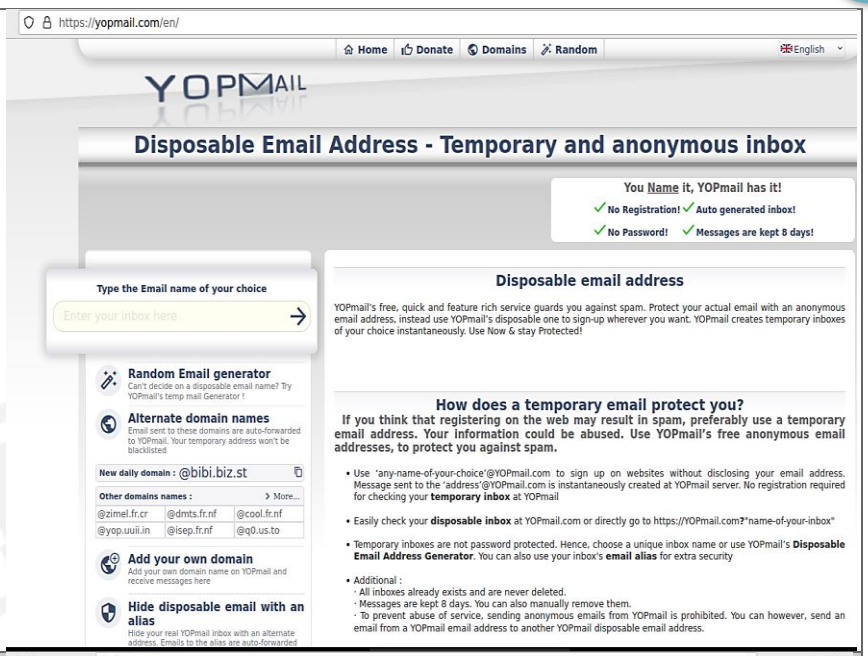
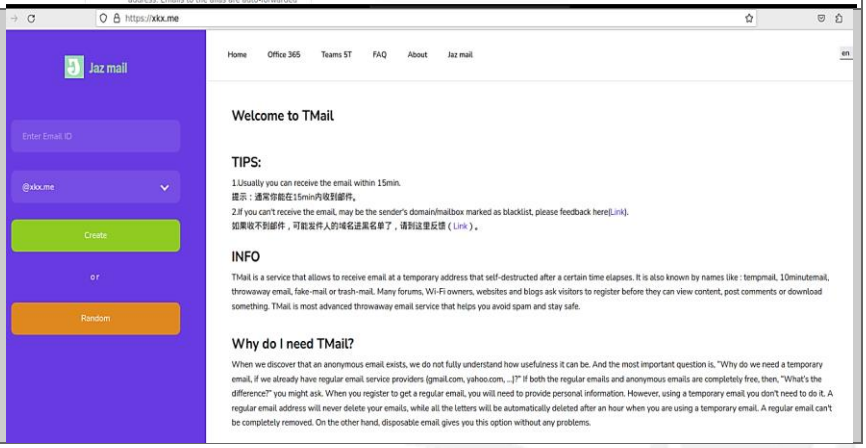


https://tempmail.plus	
https://temporarymail.com	
https://tempr.email/en	

https://www.33mail.com	
https://www.crazymailing.com/en	
https://www.dispostable.com/	
https://www.emaildrop.io	

https://www.fakemail.net	
https://www.mailinator.com	
https://www.moakt.com	

https://www.mohmal.com/en	
https://www.tempinbox.xyz	
https://www.throwawaymail.com/en	
https://www.trash-mail.com/inbox/	

https://www.yopmail.com/en	
https://xkx.me	

Creación de identidades falsas

Si bien en la actividad anterior hemos visto que la creación de *emails* temporales es una posibilidad real para cubrirnos ante robos de identidad, datos personales, proteger nuestras cuentas contra *spam* o contactar con algún investigado de forma que se use una dirección de email que no puedan trazar hasta nosotros, en ocasiones es necesario ir más allá y crearnos una ***online persona* falsa completa**. Esta sección se incluye para hacer pruebas que impliquen **no interactuar** con objetivos investigados, como registrarse en sitios con el objetivo de observar lo que hace o escribe la gente sin dar datos propios. **Recuerda, en esta asignatura queda terminantemente prohibido el "*engagement*" con desconocidos**.

Suplantar a alguien es delito. Ser alguien falso para engañar a otra persona y obtener datos privados es delito, salvo que seas una fuerza del orden y tengas autorización a hacerlo por alguna clase de investigación en marcha. En realidad este apartado se detalla más por el motivo contrario: tienes muchas más opciones de **acabar siendo tú una víctima de delincuentes que hagan esto**. Así que lee esto para estar más preparado, y **saber más fácilmente cuando "algo huele mal con una persona"**, pero contándotelo como si fueras a hacerlo tú, que yo creo que es más fácil de ver y entender. Y ojo, esto no es *Hollywood*. **Esto puede pasarte en un ataque de *spear phishing* de alto nivel o un timo elaborado**. Si algún día tienes un puesto de responsabilidad puedes ser un objetivo más fácilmente de lo que crees. Así que cuidado.

¿Qué quiere decir esto? Que hay servicios que nos dan datos para permitir crearnos una identidad falsa completa, y que podemos usar para registrarnos en otros servicios cuando no queramos exponer nuestra identidad real. **Si bien esto es ilegal o no está permitido por algunas plataformas o servicios** (por ejemplo,

algunas redes sociales prohíben en sus términos y condiciones hacerlo)**, sí es una opción real en otras muchas** (o sea hace de todas formas, hasta que lo descubran y eliminen). En resumen: **se puede falsificar cualquier dato personal de una forma muy creíble**.

Una de las primeras cosas que se falsifica es la **foto de perfil**. **Robar la identidad de alguien real es un delito**. Colocar un dibujo o avatar puede también ser **una infracción** (si usamos uno de pago, pero no hemos pagado por él) o **no estar permitido** por el servicio (que requiere la foto realista de una persona). Algunos servicios para crear identidades falsas ya nos dan una (lo veremos luego), pero puede notarse o averiguarse fácilmente (Ej.: buscando en *Google*) que es una foto "de *stock*". Para estos casos hay un servicio llamado https://thispersondoesnotexist.com/ que nos genera usando algoritmos de IA una foto de la cara de una persona que no existe, pero con un realismo increíble (incluye distintos gestos, fondos, edades, sexos, razas, etc.). Cada vez que recarguemos la página se nos generará una foto nueva aleatoria que no se repetirá, **por lo que si alguna nos gusta debemos guardarla**. Todas las fotos de la figura siguiente son completamente falsas, y podemos comprobar como muchas de ellas son muy realistas:

Figura 14. No, ninguna de estas personas *exist* ;)

Obviamente, hay que escoger una que encaje con la información que demos de nuestra identidad falsa (sexo, raza, características físicas...). Otra ventaja de usar este tipo de servicios es que buscar la foto de perfil con la **búsqueda inversa de imágenes de *Google*** rara vez dará algún resultado, ya que, como decíamos, **estas fotos se generan de manera aleatoria**.

No obstante, si bien esto nos cubre el apartado de la imagen de perfil, **aún necesitamos el resto de los datos** para crear una "persona falsa". **En una auditoría real**, es importante tener estos datos prefijados para darlos de manera consistente en distintos perfiles que creemos, y que nuestra "tapadera" no sea descubierta por una inconsistencia. Esto no es trivial. Si vamos a investigar a alguien es posible que esa persona tenga conocimientos técnicos y tengamos que ser muy profesionales para que no nos descubran:

- El investigado **puede tratar de corroborar nuestra historia** contactando con terceros supuestamente asociados a nosotros (que nosotros mismos habremos creado, porque cuando se crea una persona se crea también su entorno si uno se pone serio o se requiere un realismo alto).
    - Crear un conjunto de personas falsas que formen una "piña" o grupo asociado no es descabellado. Queremos controlar las interacciones que puedan delatarnos, pero también puede usarse para que una de esas personas falsas "sonsaque" información sobre el investigado si nos montamos una "película" lo suficientemente compleja, y de esta forma sacar más información aún. Este documento no se llama "*hacking* humano" (sólo) por hacer un chiste 😊.

  - Piensa en ti mismo: *¿en qué confías más? ¿En un perfil con amigos o en otro sin ellos?* La respuesta es obvia y la tienes dentro. Ponte en la piel del investigado y actúa en consecuencia. Es una persona, como tú.
  - Si vas a crear alguien falso que forma parte de un entorno o contexto conocido, prepárate para preguntas de tipo *"¿Conoces a X?"*, *"¿Has estado en Y?"* o *"¿Qué te parece lo que paso en T (tiempo pasado)?"*. En una investigación real hay que "hacer los deberes" haciendo investigaciones OSINT extendidas al entorno objetivo (en esta asignatura no, esto solo te lo digo para que veas cómo va la cosa en contextos más serios donde la víctima puedas ser tú).
- **Siempre hay que asumir que el investigado es alguien muy hábil e inteligente**, y puede tratar de buscar activamente "grietas" en nuestro personaje. Por ejemplo, preguntándonos por un mismo dato personal en varios momentos de tiempo y buscando incoherencias. Para evitar incoherencias hay que diseñar bien la "obra de teatro", sus personajes y su profesión. Obviamente, esto minimiza la opción de ser incoherente por accidente también. También os digo que es algo que requiere práctica.
- Hazte a la idea de que estás **jugando una partida de rol o estás en una obra de teatro**. Tu público o tus compañeros deben sumergirse en el "mundo de fantasía" que has creado para que puedas obtener los datos que necesitas. Si no, fracasarás y todo el esfuerzo no servirá para nada.
- **Las identidades falsas se "queman"**. No reutilices entidades falsas. El mundo es muy pequeño, más de lo que imaginas, y es muy triste fracasar porque alguien ha conocido previamente tu identidad falsa por una tercera persona. Los servicios que te doy generan todo aleatorio. Tranquilo, no se van a acabar 😊.
- Puedes mezclar la información aleatoria con partes de perfiles reales para darle más consistencia si quieres, pero **nunca te bases mucho en alguien** o perderá sentido la acción y podrías incurrir en un delito de suplantación. Tienes que seguir siendo falso. Creíble 100%, pero falso.
- También es fácil dejarse llevar por el ego en estos casos. **Eres un profesional y te han contratado para sacar información de alguien para cumplir con tu trabajo**. Usar estos poderes para otra cosa hace que no seas un profesional y te lleva fácilmente a **ser un delincuente**. No lo hagas. Tu labor es descubrir información para prevenir desgracias, **no convertirte en la desgracia**. En esta asignatura **NO queremos delincuentes**.

Un servicio que nos permite hacer esto es ***Fake Person Generator***: https://www.fakepersongenerator.com/Index/generate. Con él podemos generar una persona de cualquier género, edad y estado (solo permite generar personas residentes en EEUU, si bien obviamente luego podemos modificar cualquier dato a mano para "personalizar" nuestra identidad). Este servicio nos da ya una imagen de perfil que podemos cambiar por una de las que vimos antes si queremos.

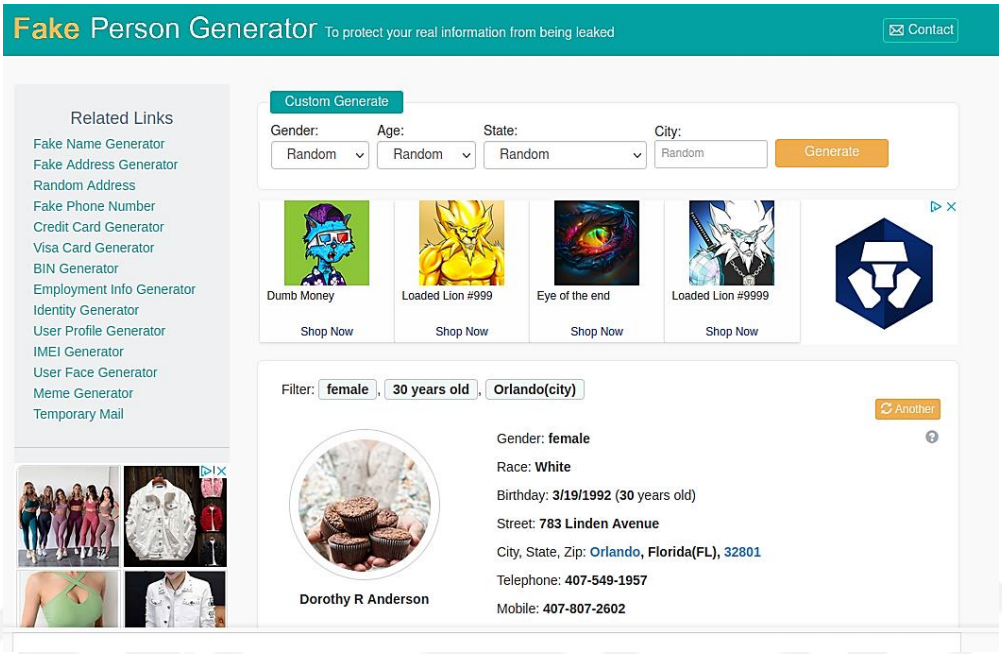


Figura 15. Perfil creado por *Fake person generator* (I)

Vemos como se genera aleatoriamente una serie de datos básicos de la persona (**todos falsos y aleatorios**, obviamente, pero que cumplen con el formato de un dato real).

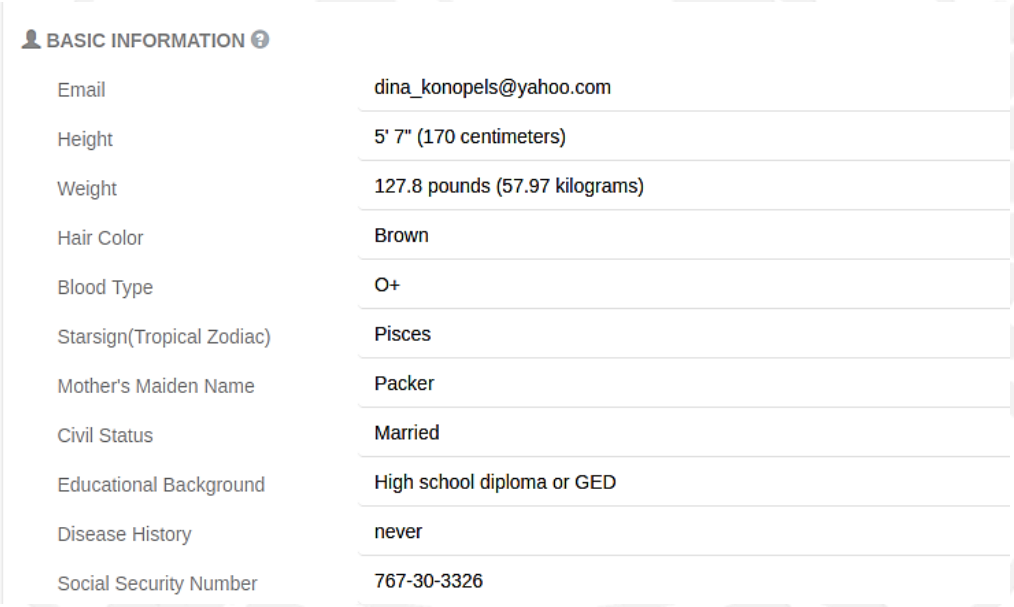
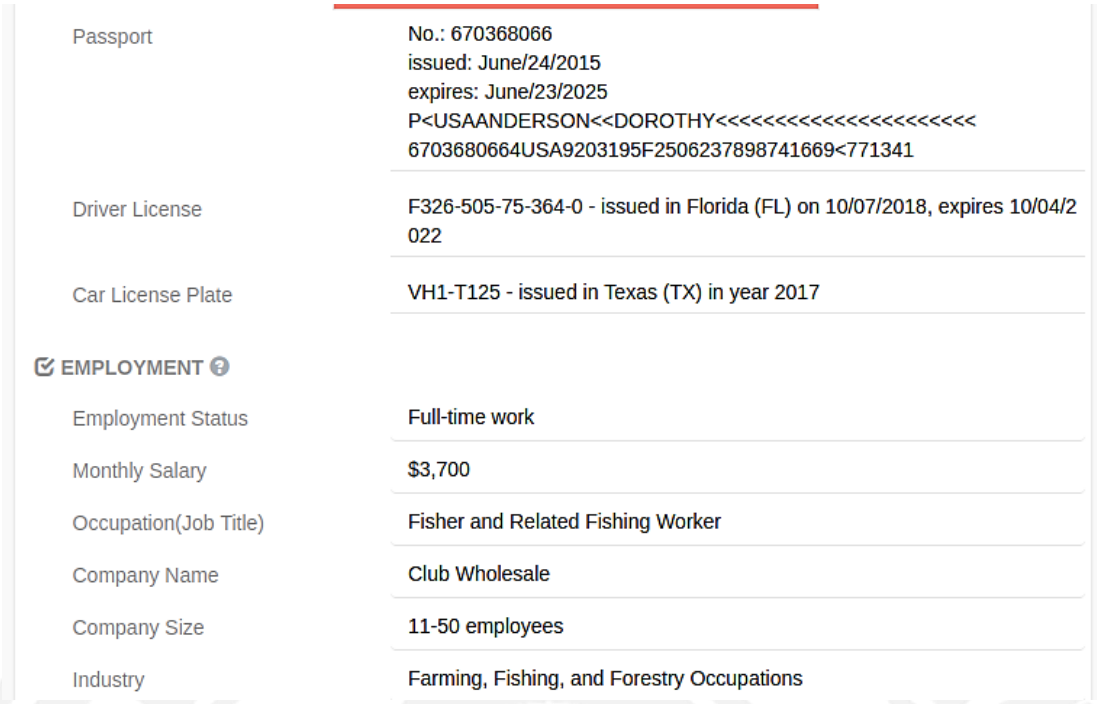


Figura 16. Perfil creado por *Fake person generator* (II)

La página no se queda ahí, si no que nos genera datos de pasaporte, carnet de conducir y matrícula de coche falsos, así como empleo actual y finanzas. Date cuenta de que, como decíamos antes, esta información puede pedirla algún interlocutor para comprobar que no está delante de un *bot*, por ejemplo (y no lo está, solo eres una persona...falsa 😊):



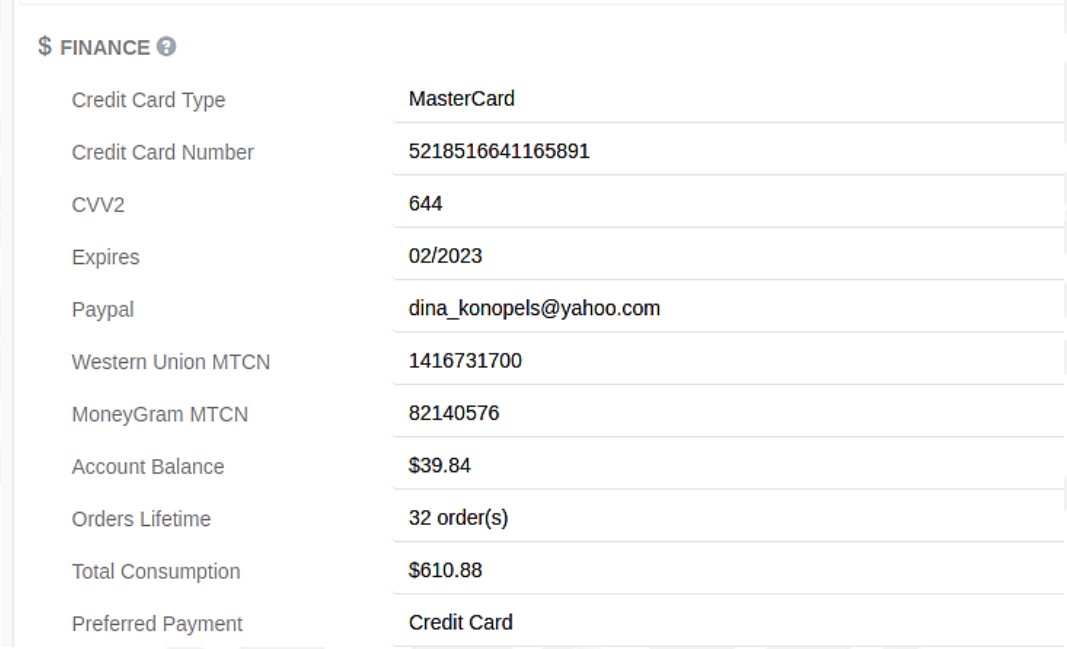
Passport	No.: 670368066 issued: June/24/2015 expires: June/23/2025 P<USAANDERSON<<DOROTHY<<<<<<<<<<<<<<<<<<<<<< 6703680664USA9203195F2506237898741669<771341
Driver License	F326-505-75-364-0 - issued in Florida (FL) on 10/07/2018, expires 10/04/2022
Car License Plate	VH1-T125 - issued in Texas (TX) in year 2017

EMPLOYMENT

Employment Status	Full-time work
Monthly Salary	$3,700
Occupation(Job Title)	Fisher and Related Fishing Worker
Company Name	Club Wholesale
Company Size	11-50 employees
Industry	Farming, Fishing, and Forestry Occupations

Figura 17. Perfil creado por *Fake person generator* (III)

Puedes ver también que se proporcionan datos financieros completos (falsos) para usar en "tapaderas" más especializadas donde tengas que mantener una historia coherente con datos de este tipo.

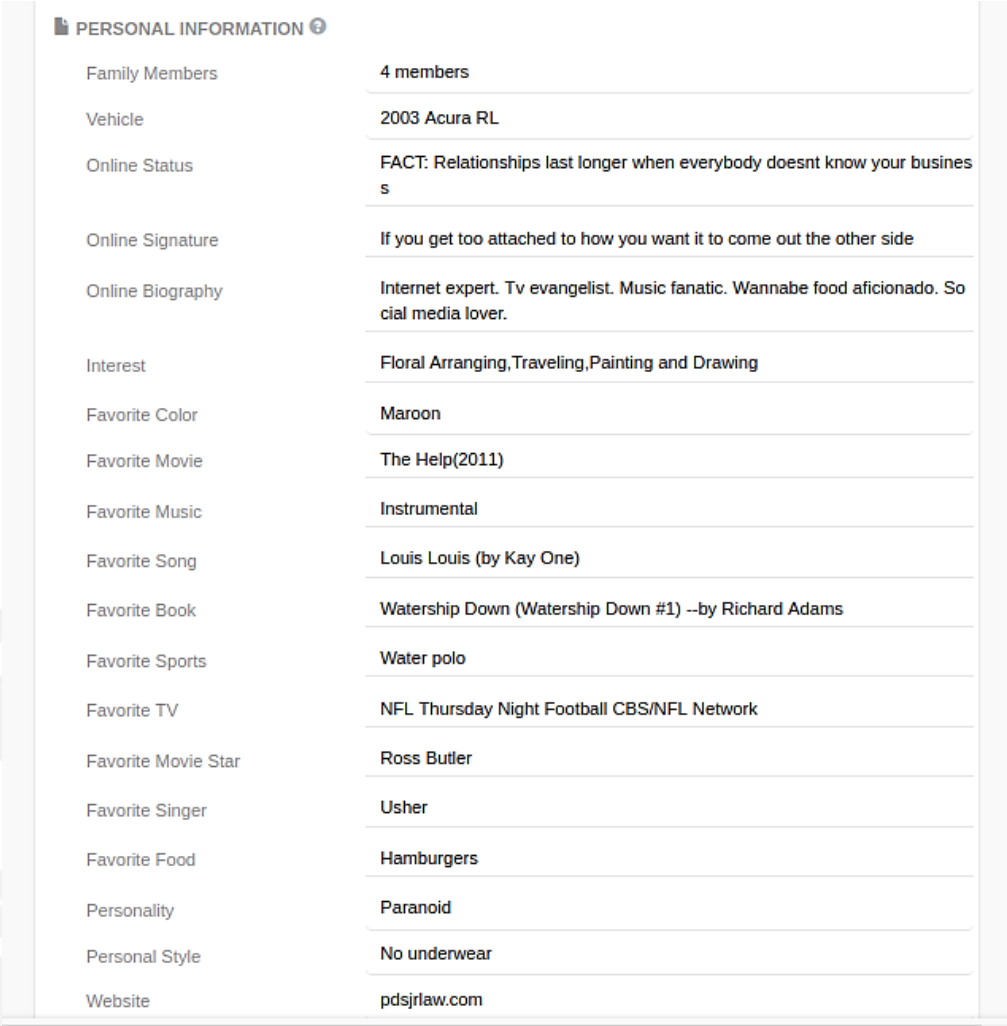


$ FINANCE

Credit Card Type	MasterCard
Credit Card Number	5218516641165891
CVV2	644
Expires	02/2023
Paypal	dina_konopels@yahoo.com
Western Union MTCN	1416731700
MoneyGram MTCN	82140576
Account Balance	$39.84
Orders Lifetime	32 order(s)
Total Consumption	$610.88
Preferred Payment	Credit Card

Figura 18. Perfil creado por *Fake person generator* (IV)

Adicionalmente, se incluyen datos personales, familiares, hobbies, gustos, etc. utilizables en determinados entornos donde esto sea de aplicación. La "tapadera" que podemos crear es así mucho más completa, ya que realmente contempla casi cualquier aspecto típico que se puede comentar o preguntar en una red.

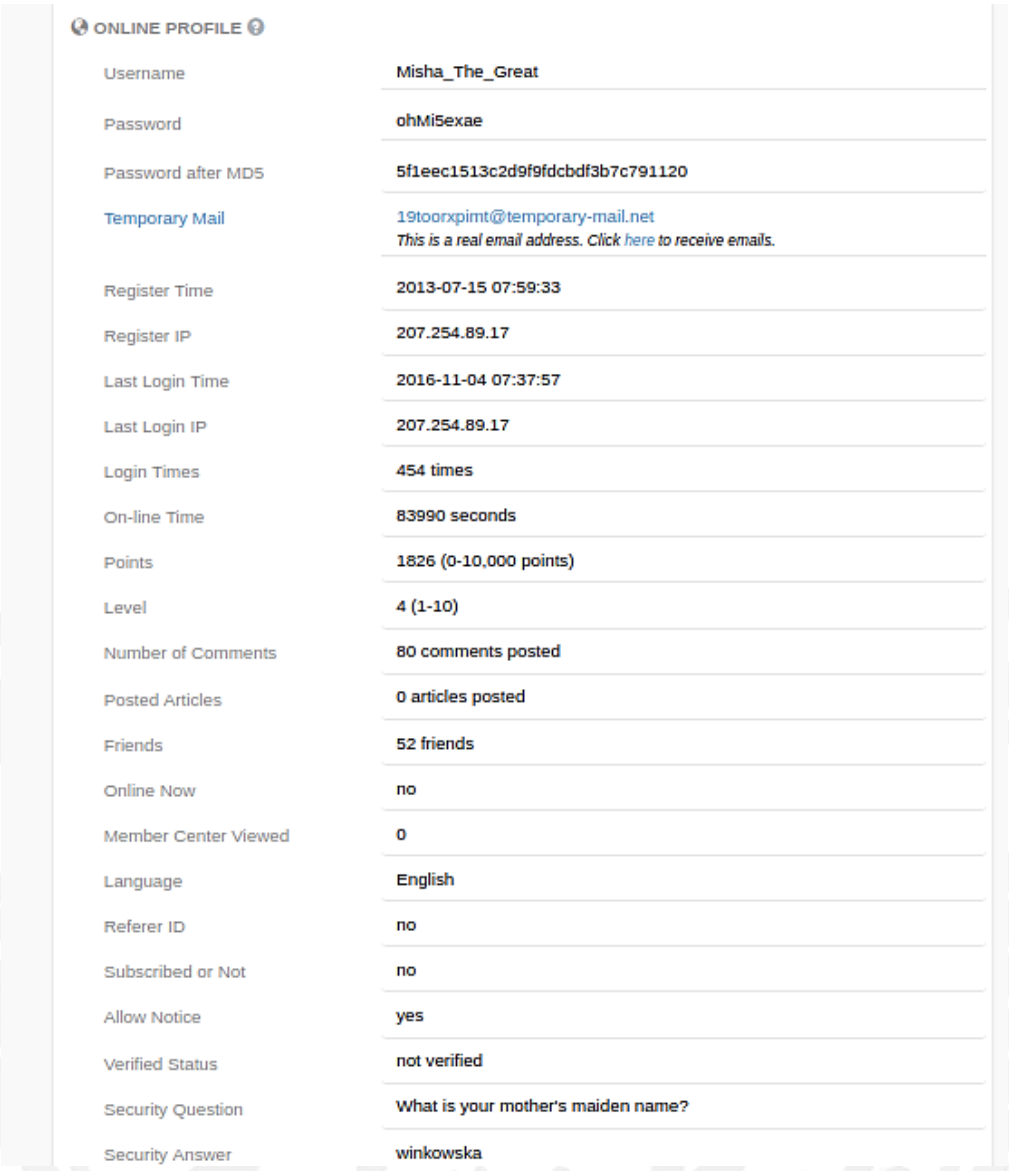


PERSONAL INFORMATION

Family Members	4 members
Vehicle	2003 Acura RL
Online Status	FACT: Relationships last longer when everybody doesnt know your business
Online Signature	If you get too attached to how you want it to come out the other side
Online Biography	Internet expert. Tv evangelist. Music fanatic. Wannabe food aficionado. Social media lover.
Interest	Floral Arranging,Traveling,Painting and Drawing
Favorite Color	Maroon
Favorite Movie	The Help(2011)
Favorite Music	Instrumental
Favorite Song	Louis Louis (by Kay One)
Favorite Book	Watership Down (Watership Down #1) --by Richard Adams
Favorite Sports	Water polo
Favorite TV	NFL Thursday Night Football CBS/NFL Network
Favorite Movie Star	Ross Butler
Favorite Singer	Usher
Favorite Food	Hamburgers
Personality	Paranoid
Personal Style	No underwear
Website	pdsjrlaw.com

Figura 19. Perfil creado por Fake person generator (V)

Finalmente, también se generan datos de carácter mucho más técnico que pueden usarse para registrarse en servicios (incluye una cuenta de correo temporal alternativa a las vistas), dar de alta al usuario de forma artificial en algún servicio en pruebas, y un largo etcétera. El servicio por tanto apuesta por generar de manera aleatoria gran cantidad de información diferente y **que seamos nosotros los que decidamos cómo usarla en el contexto de aplicación**.



ONLINE PROFILE

Username	Misha_The_Great
Password	ohMi5exae
Password after MD5	5f1eec1513c2d9f9fdcbdf3b7c791120
Temporary Mail	19toorxpimt@temporary-mail.net *This is a real email address. Click here to receive emails.*
Register Time	2013-07-15 07:59:33
Register IP	207.254.89.17
Last Login Time	2016-11-04 07:37:57
Last Login IP	207.254.89.17
Login Times	454 times
On-line Time	83990 seconds
Points	1826 (0-10,000 points)
Level	4 (1-10)
Number of Comments	80 comments posted
Posted Articles	0 articles posted
Friends	52 friends
Online Now	no
Member Center Viewed	0
Language	English
Referer ID	no
Subscribed or Not	no
Allow Notice	yes
Verified Status	not verified
Security Question	What is your mother's maiden name?
Security Answer	winkowska

Figura 20. Perfil creado por *Fake person generator* (VI)

Este servicio, aunque extremadamente potente, tiene el problema de que solo genera identidades de EEUU. Otra opción es ***Fake Name Generator*** (https://es.fakenamegenerator.com) que, si bien no es tan potente, y bastantes de los nombres que genera (al menos en Español) son un tanto "discutibles", sí que nos permite tener una identidad más adecuada a ciudadanos de distintos países (entre ellos **España**), cuyos datos se pueden completar con los obtenidos en el servicio anterior.

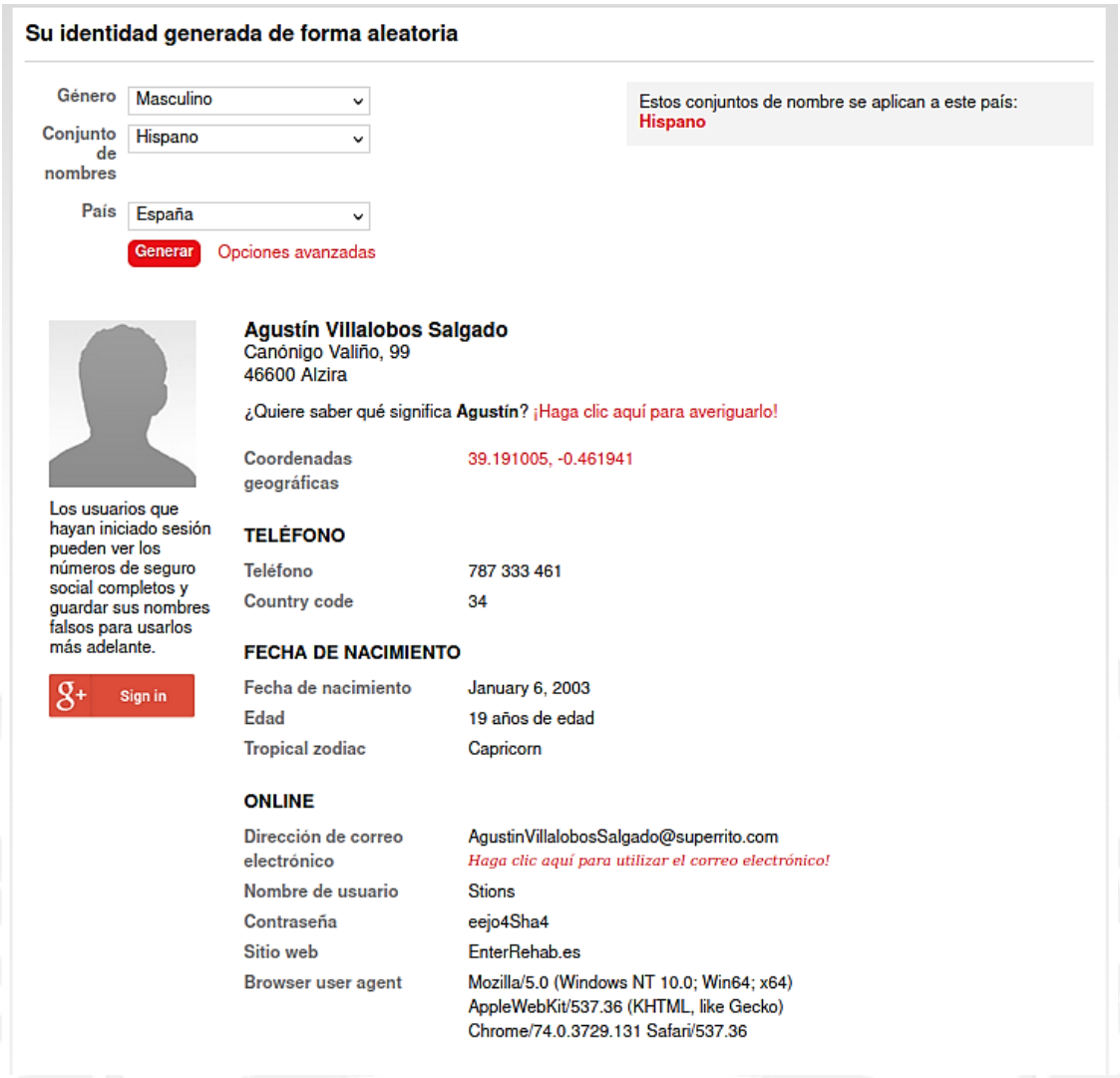


Figura 21. Perfil creado por *Fake name generator* (I)

No obstante, como vemos los datos falsos que genera **son muchos menos** que los del servicio anterior (aunque incluye una cuenta de correo temporal real):

FINANCE

MasterCard	5134 4656 8731 2465
Vence	6/2027
CVC2	777

EMPLOYMENT

Empresa	Gart Sports
Ocupación	Aircraft pilot

PHYSICAL CHARACTERISTICS

Altura	5' 7" (170 centímetros)
Peso	164.8 libras (74.9 kilos
Tipo sanguíneo	B+

TRACKING NUMBERS

Número de seguimiento de UPS	1Z 4V6 59E 41 1184 740 8
Western Union MTCN	4671732598
MoneyGram MTCN	86202379

OTHER

Color favorito	Azul
Vehículo	2006 Lexus RX
GUID	88341456-78bc-4b52-ae61-f0bca5c8d255
Código QR	Haga clic aquí para ver el código QR de esta identidad

Figura 22. Perfil creado por *Fake name generator* (II)

En definitiva, usando uno de ellos, una **combinación de ambos**, o bien, como decíamos, una "historia" que nos creemos en base a una combinación de datos falsos que nosotros aportemos más los que aporten estos servicios, junto con una cuenta de correo (temporal o no) no asociada a nuestra persona, **podemos ser en Internet prácticamente quien nosotros queramos** y, si mantenemos la coherencia en nuestras historias y no revelamos información que nos delate, resultar ser 100% creíble a ojos de cualquier usuario estándar. Interactuar con alguien para tareas de OSINT activo por tanto puede ser viable desde un perfil completamente anónimo, preservando nuestra identidad.

Un aviso final, aún a riesgo de ser un pesado: usar bien este tipo de técnicas para "meterse en la piel" de otra persona requiere **interiorizar** que eres esa otra persona para evitar "fugas". Esto muchas veces se ve favorecido si **tienes cierta experiencia en la actuación y el teatro**, o si has hecho ejercicios o pruebas en este estilo previamente que te den la experiencia suficiente para ello. **No uses este conocimiento para atacar o acosar a otra persona**: tu anonimato se puede revelar con una orden judicial que dé en el punto preciso: recuerda que los *logs* de comunicaciones se guardan, y hay herramientas y personas que pueden identificarte bajo mandato judicial. Tu solo lo pones más difícil, **pero no imposible.**

Y NO ERES UN DELINCUENTE. Así que no te comportes como uno. Pero ahora sabes qué tipo de recursos puede tener un delincuente de este tipo que tenga preparación, así que usa este conocimiento para no caer en su trampa. Algún día puedes tener un puesto de responsabilidad y manejar dinero, y ese día puedes fácilmente **ser el objetivo de uno de ellos**. De nuevo: cuidado.

¿Y si para crear la identidad falsa necesito algo más que datos personales?

Si necesitas aportar a tu identidad falsa **algo más que datos personales**, como por ejemplo objetos, casas, mascotas, etc. el uso de IAs generativas esta ya tan extendido que puedes generar cualquier cosa con ellas y que, al ser una "cosa" aleatoria, no podrán encontrar en Internet. Tampoco tienes que instalarte nada en local si no quieres (que ya hay IAs locales de todo tipo), **puedes usar esta web** que ofrece gran cantidad de posibilidades acerca de las cosas que genera: https://thisxdoesnotexist.com/

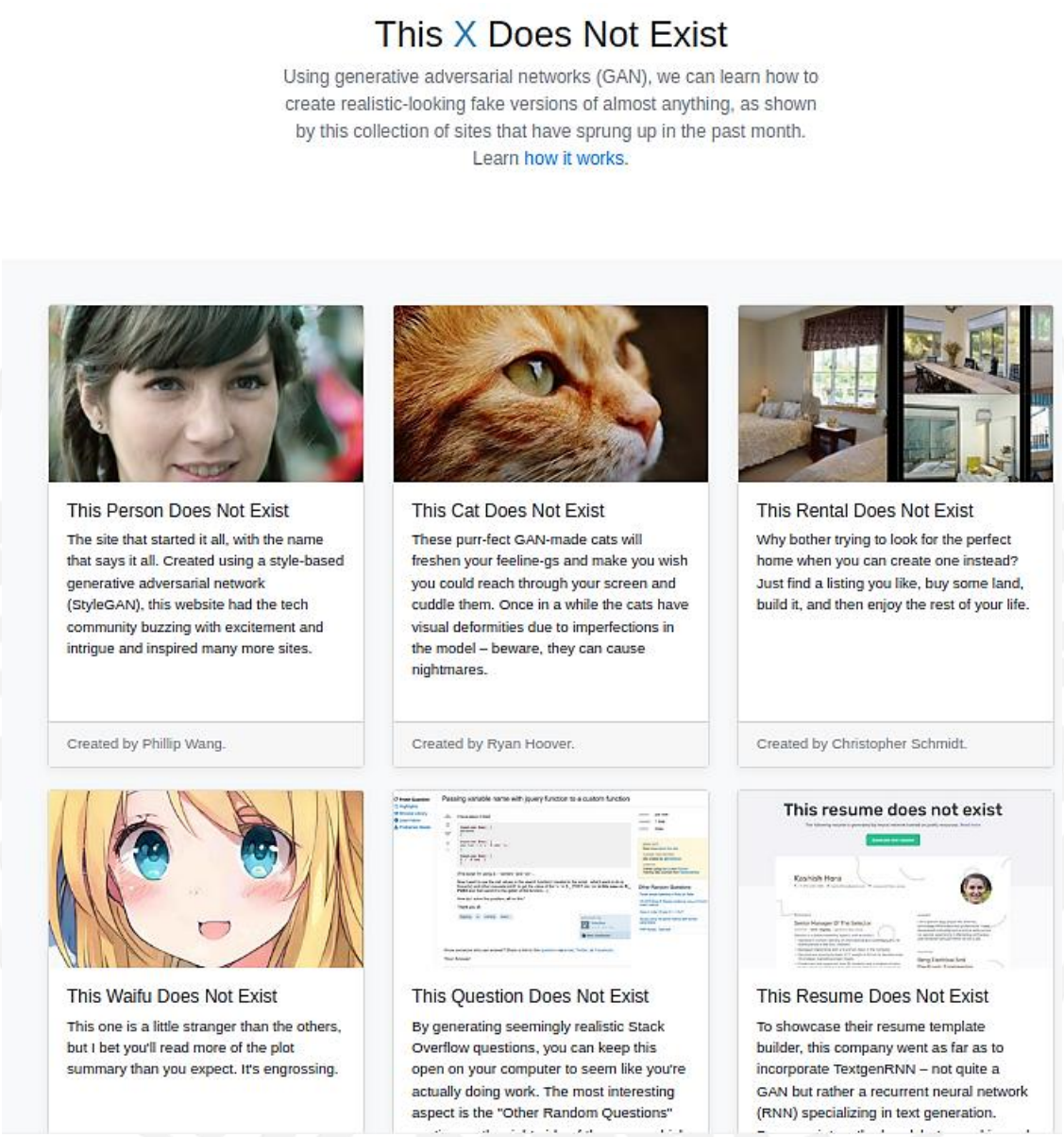


Figura 23. La cantidad de cosas generables con IA hoy en día es abrumadora. Y con esta web ni siquiera tienes que instalar nada para eso...

Otra forma de generar imágenes adaptadas a distintos propósitos es ***Canva***:
https://www.canva.com/apps/

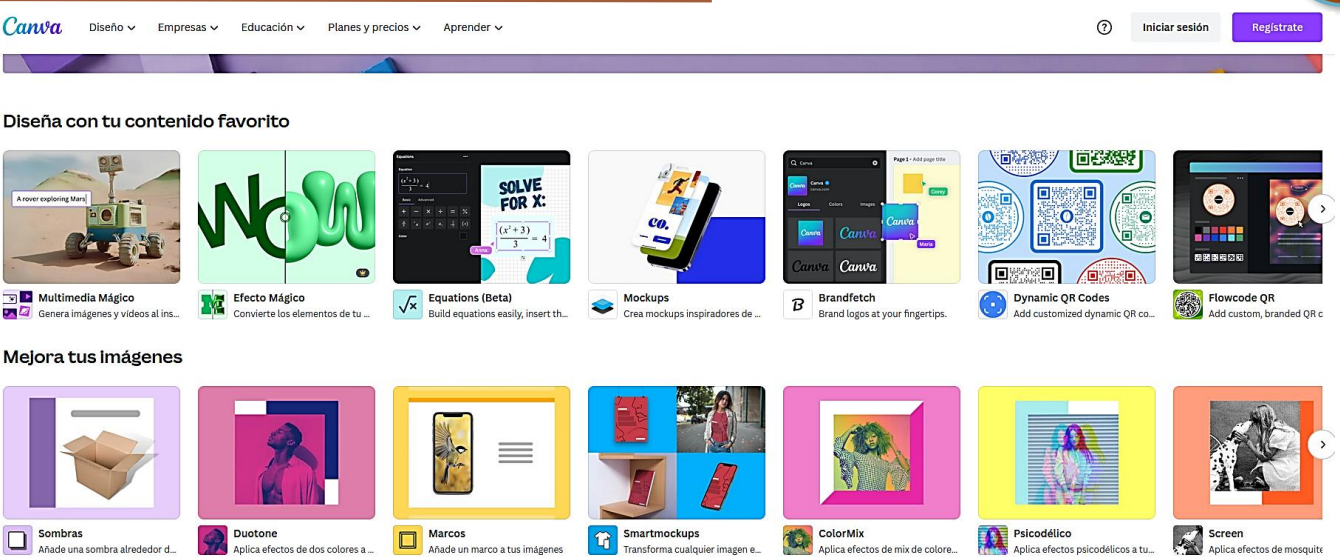


Figura 24. *Canva* es otra herramienta de generación de imágenes con IA que puedes usar para cosas muy concretas que necesites en investigaciones

Anonimato de máquinas

Remote Browsing

Esta es una técnica que te **permite navegar por páginas webs con un navegador "de otro"**, normalmente el de un servicio que contratas que te ofrece esta funcionalidad. Obviamente mantienes el anonimato porque será ese otro el que "de la cara" por ti, y no solo ante las páginas que navegues, sino también ante el posible malware que recibas. Por tanto, combinar un servicio de esta clase con una máquina virtual proporciona un **nivel de aislamiento enorme** ante cualquier amenaza que nos llegue por navegar por una web no fiable.

Estos servicios son altamente especializados y usados en entornos de mucha seguridad, por lo que **suelen ser de pago.** No obstante, el servicio ***BrowserLing*** (https://www.browserling.com/) nos permite navegar (e incluso interaccionar con las páginas web) gratis durante tres minutos en una máquina remota, con SO / navegador configurable, para ver posibles webs peligrosas o en las que non queramos ser registrados. Simplemente **recuerda no enviar información personal en estas sesiones**, puesto que podría ser capturada y usada por el proveedor del servicio si ha sido vulnerado o como parte de sus políticas de usuario.

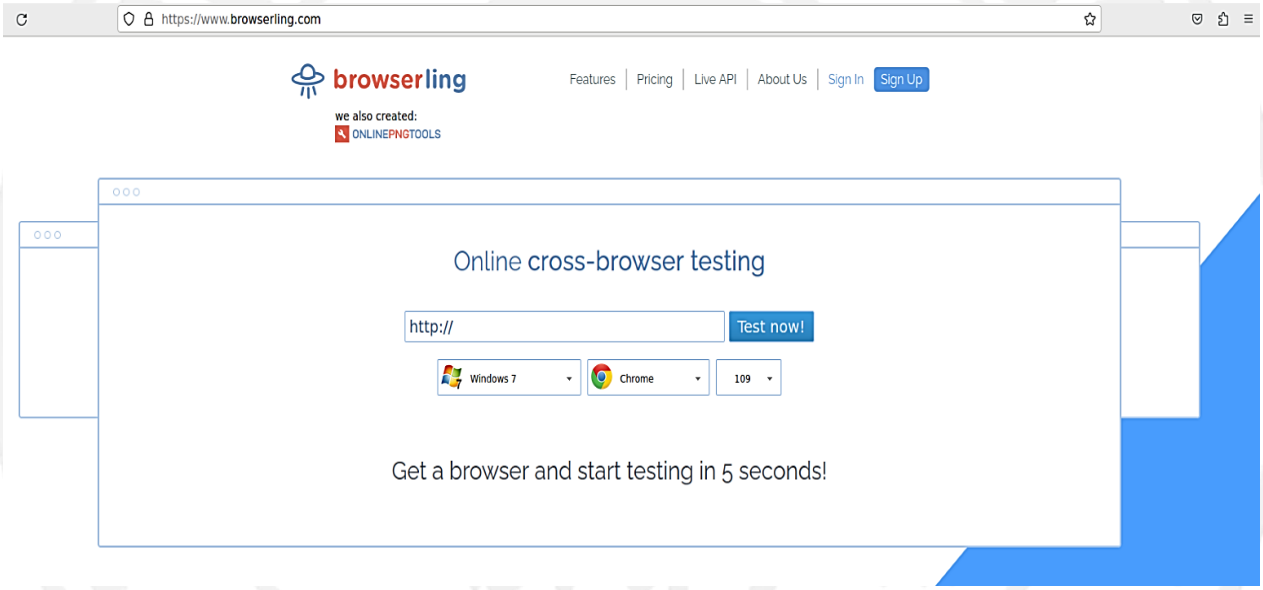


Figura 25. Navegador remoto *BrowserLing*

Sandboxing de programas

La ejecución de ciertos programas puede también exfiltrar datos nuestros que "recolecten" como parte de sus políticas de usuario o, mucho peor, **por estar infectados con malware de diferentes clases**. Si bien ejecutar un programa sospechoso en una máquina virtual proporciona un nivel de aislamiento elevado, este **no es infalible**, puesto que ya se conocen casos de malware que reconoce que está en una MV y usa técnicas que se aprovechan de vulnerabilidades del software de virtualización para "escapar" de la misma. En estos casos, para ejecutables muy sospechosos podemos usar servicios remotos de ***sandboxing***: Ejecutar programas en un servidor remoto de la empresa que nos provee el servicio, y obtener un informe de lo que hacen dichos programas.

Estas webs tienen principalmente usos para informática forense avanzada y análisis de malware, y suelen ser de pago. No obstante, existen algunas que tienen un "*tier*" gratuito que nos deja hacer un nº limitado de escaneos, suficiente para usos esporádicos. Una de ellas es ***Joe's Sandbox*** (visto en los seminarios), pero otro es ***Any.run*** (https://any.run/).

Por tanto, un servicio de esta clase puede determinar si alguno de los programas que tenemos que usar hace algo peligroso debido a que analizan su comportamiento y nos hacen un informe de este, que además nos puede decir si "recolectan" información privada que no queremos revelar.

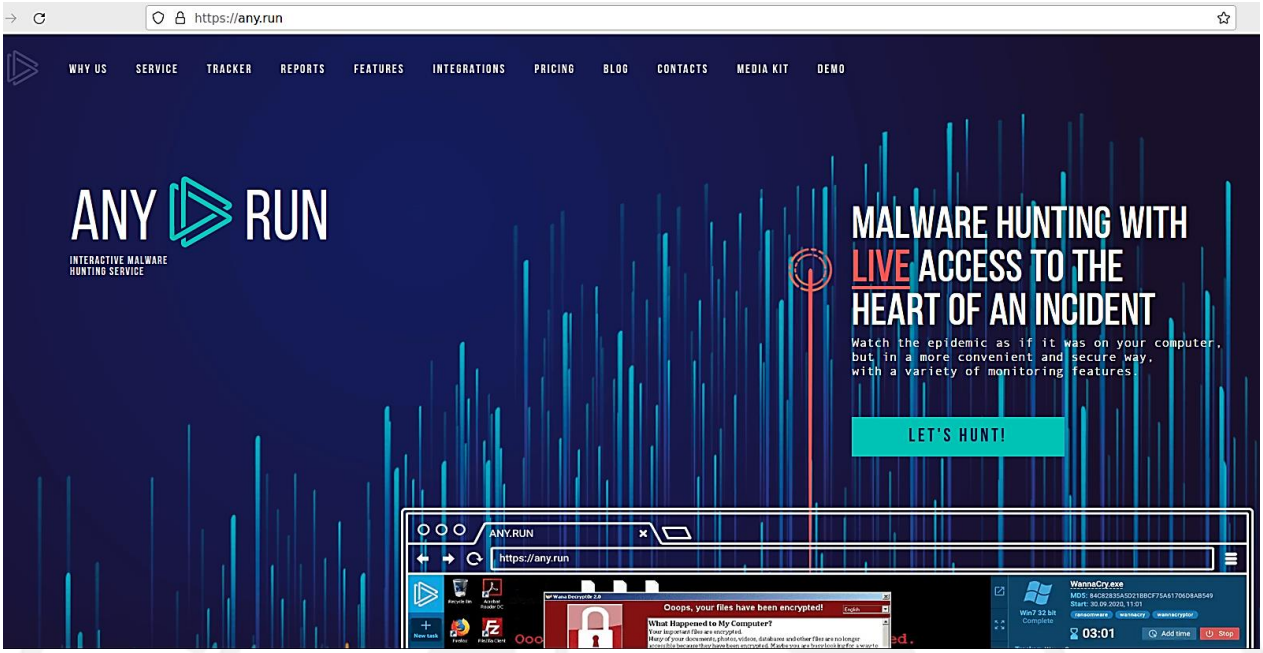


Figura 26. Servicio de *sandboxing* de programas Any.run

VPNs

Es posible lograr anonimato gracias a un **servicio comercial de VPN** que te haga de "pasarela" entre tu máquina y una máquina remota en un país que elijas (según los que te ofrezca el servicio que has contratado), de manera que **sea esa máquina remota del país elegido la que "de la cara" por ti en Internet**. En otras palabras, a ojos de quien recibe tus peticiones, si el servicio que contratas está bien configurado, tú serás la máquina remota que te ofrece el servicio VPN. Además, al ser una VPN el tráfico entre tu máquina y ese servidor será privado y nadie más podrá verlo.

Hay muchos servicios VPN gratuitos y comerciales, y algunos de los gratuitos no están exentos de polémica porque hay sospechas de que venden sus datos de navegación o la información no es tan privada como debería. De los gratuitos, uno que tiene muy buena fama es ***ProtonVPN*** (https://protonvpn.com/es/), que si bien la versión gratuita limita el nº de servidores disponibles, es suficiente para lograr la privacidad que hemos mencionado (se ve en el **Laboratorio 7**). Se da la circunstancia de que es un servicio que acompaña a otros interesantes como ***ProtonDrive*** (almacenamiento cifrado en nube, gratuito hasta 1Gb) y ***ProtonMail*** (*webmail* que garantiza la privacidad y el secreto de las comunicaciones entre clientes de *ProtonMail* y, si ambos interlocutores conocen una clave simétrica común, también con otros clientes de correo).

Figura 27. ProtonVPN

Conexiones vía ToR

La famosa red *ToR* (**Tema de Teoría 2**) es una red gratuita **que te permite ser (al menos teóricamente) anónimo en *Internet***. En lugar de exponer tu IP a los servidores a los que te conectas, en realidad te conectas a una red de servidores a nivel mundial intercomunicados entre ellos. Estos servidores enrutan cada mensaje entre ellos de acuerdo **con un patrón de comunicación aleatorio que cambia cada X tiempo**, por lo que hace que trazar la IP de origen no sea posible en condiciones normales (salvo que se guarde un *log* detallado de conexiones, algo que "en teoría" no se hace...salvo que haya muchos nodos "controlados" por entidades / gobiernos que lo hagan...). Además, elimina cualquier información que tu dispositivo proporcione y que permita identificarlo.

Por otro lado, el tráfico entre los nodos ToR que forman tu "circuito de transmisión" **está cifrado** entre ellos, por lo que **ningún nodo intermedio, solo el de salida**, podrá ver los datos reales que estás enviando (Ver: https://es.wikipedia.org/wiki/Tor_(red_de_anonimato)). El nodo de salida envía la información al destino tal cual tú la has remitido, por lo que sí podrá ver su contenido. Por eso es aconsejable que además uses cifrado a nivel de aplicación, como TLS, para mantener el la información segura del todo.

En otras palabras, **estas engañando al destino de tu comunicación** acerca de quién eres y desde dónde te conectas, y además todos los componentes del mecanismo que usas para ello no pueden, en teoría, saber lo que estás enviando. El siguiente diagrama describe esto con más detalle (**Fuente**: https://www.f5.com/labs/articles/threat-intelligence/snooping-on-tor-from-your-load-balancer)

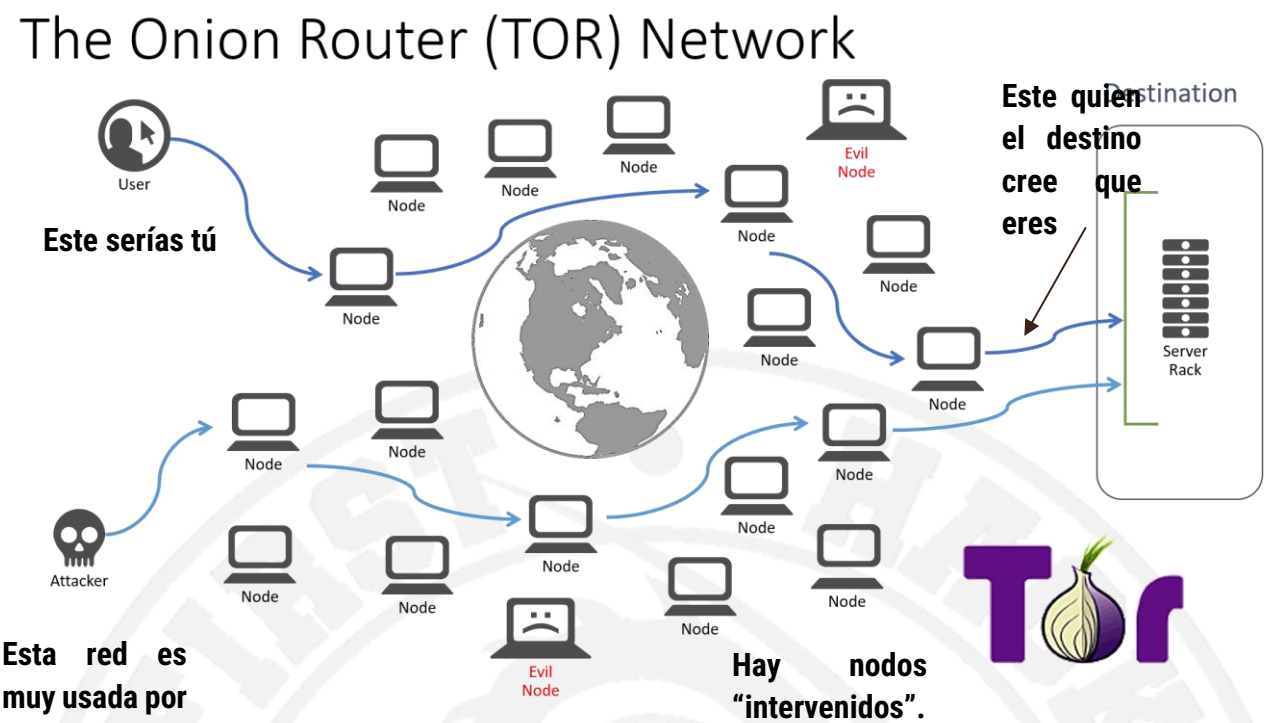


Figura 28. Esquema general de lo·que es la red ToR

Como se ve en la imagen, el anonimato que proporciona la red *ToR* se puede usar tanto **legítimamente** (gente amenazada / investigadores con técnicas activas que deban mantener su anonimato en *Internet*...) como por delincuentes. Además, como decíamos su anonimato puede estar también potencialmente comprometido por la presencia de nodos intervenidos por autoridades, o troyanizados para revelar los datos de conexiones que pasan por la misma.

En el **Laboratorio 7** ya vimos una forma de conectarse a la red ToR, a través del navegador ***TOR Browser***. El problema es que esto solo anonimiza nuestra navegación, y no lo hace con el resto de las herramientas que se puedan conectar a Internet. Ahora vamos a arreglar esto, puesto que **el propósito de esta sección es múltiple**

1. Aprender a instalar el *software* que te permite acceder a la red *ToR*
2. Poder usar *ToR* con cualquier comando que requiera acceso a Internet y/o con un *browser*
3. Comprobar si realmente ralentiza las conexiones a Internet hasta el punto de hacerlas no usables
4. Comprobar si realmente hay webs conocidas que bloquean conexiones que provengan de la red ToR.

Para hacer esto sigue este procedimiento:

- Necesitas la versión más moderna de *Ubuntu* estable que haya en este momento (u otro SO soportado de tu elección). Las pruebas se han hecho con éxito con una *Ubuntu 22.04 LTS*. **La última versión de *ToR* no funciona adecuadamente en estos momentos en *Ubuntu 18.04***.
- Instala la red *ToR* de los paquetes oficiales de *Ubuntu*: `sudo apt install tor`
- Por defecto *ToR* se ejecuta en el puerto `9050`, por lo que **hay que comprobar que el *firewall*** (por ejemplo, `ufw`) **no lo está bloqueando**. Por otro lado, se puede comprobar si hay una conexión activa a esa red en ese puerto con el comando `ss` (estoy de acuerdo contigo si estás pensando que el nombre del comando no es muy afortunado 😊).
- Para comprobar que *ToR* funciona, lo primero que debemos saber es que **es una red que se usa a voluntad** (es decir, bajo demanda, sólo cuando activemos su uso). Por tanto, aunque se instale, por

defecto todo comando se conecta a *Internet* como lo hace normalmente y, por ejemplo, la orden `wget -qO - https://api.ipify.org; echo` nos devuelve nuestra IP real actual.

- Si repetimos la misma orden anteponiendo `torsocks` (`torsocks wget -qO - https://api.ipify.org; echo`), ahora deberíamos ver otra IP, correspondiente a **un nodo de salida a Internet** de la red ToR
- Como ejecutar cada comando anteponiendo `torsocks` es incómodo y propenso a errores (vamos, que se nos puede olvidar y "liarla" 😊), podemos **activar la red *ToR* para cualquier comando que use Internet** desde un momento dado en adelante ejecutando `source torsocks on`. Si todo es correcto, el comando anterior debería devolver la IP de la red *ToR*, no la real, aunque no usemos `torsocks`. Obviamente, `source torsocks off` la desactiva.
- Por defecto activar la red *ToR* como se indica en el punto anterior **dura una sesión de terminal**, por lo que **nuevas terminales o reiniciar el PC desactivan su uso**. Se puede activar permanentemente con `echo ". torsocks on" >> ~/.bashrc`. Piénsate un poco las cosas antes de hacer esto: usar la red ToR tiene unas desventajas que veremos luego que te pueden no venir bien si la máquina la usas para otras cosas. Si es una máquina virtual que usas solo para navegar por ToR es otra cosa, claro 😊.
- Finalmente, si cambiamos la configuración de la red de un navegador cualquiera para que use como *SOCKS Host v5* nuestra propia máquina (`localhost`) y el puerto de ToR (`9050`), **toda nuestra navegación se hará por la red *ToR***.

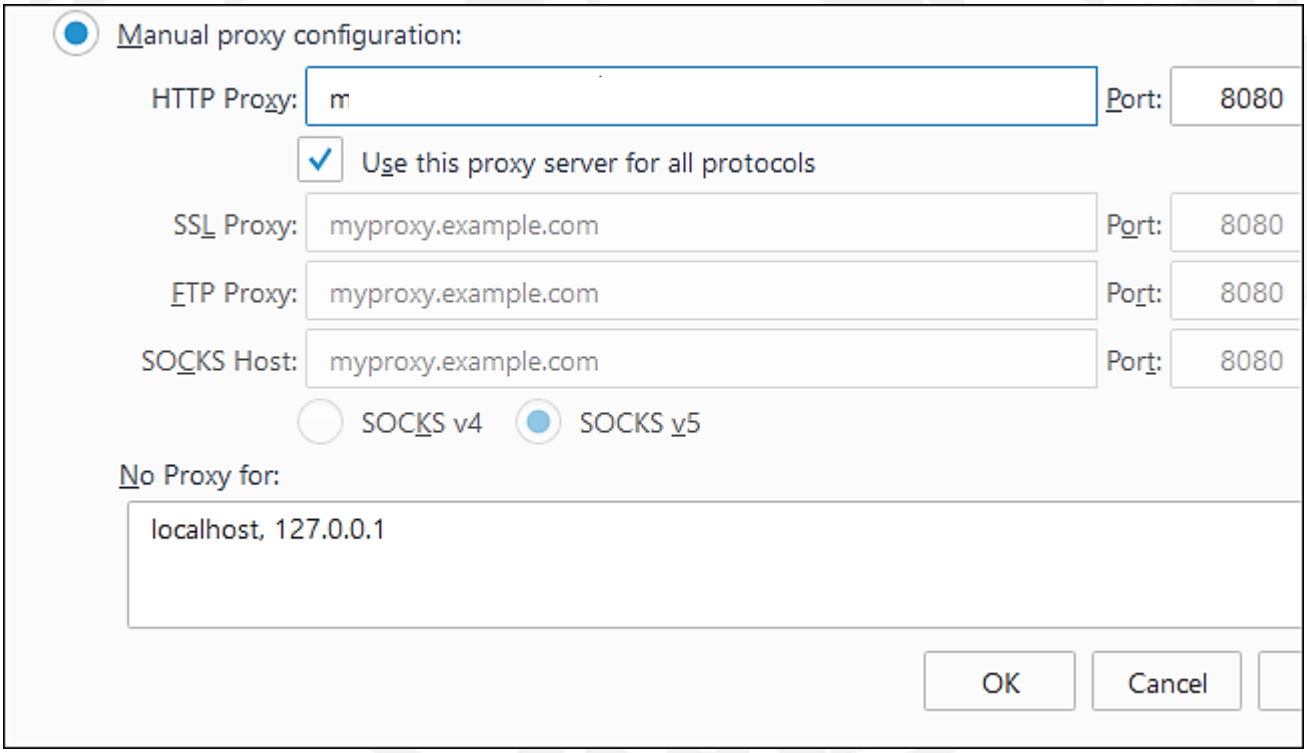


Figura 29. Configuración del proxy en *Firefox*

- Para comprobar que realmente estás conectado a la red *ToR* mientras navegas se puede preguntar a un servicio online como https://ipchicken.com/ para que identifique el origen de tu IP. **Debería mencionar un nodo de la red *ToR***.

¿Y qué pasa entonces si "sales a Internet" por un nodo de salida de la red ToR? Pues que el destino verá una IP que se conecta al mismo siempre, **pero en lugar de la tuya será la del nodo de salida** (el que te toque en cada momento), como se indicó en la imagen anterior. Trazar tu IP real "hacia atrás" es imposible en condiciones teóricas, ya que cada X tiempo las máquinas que hacen "los saltos" que enrutan tu conexión

cambiarán y no se debería guardar *log* de lo que se hizo antes. No obstante, el uso de la red *ToR* tiene una serie de consideraciones que debes contemplar:

- **Otro elemento de tu conexión puede estar revelando datos sobre ti**. Imagina por ejemplo que usas una web que pasa en parámetros `POST` o `GET` algún dato que te pueda identificar como la "IP de origen" o simplemente tu nombre de usuario. Por mucho "salto" que la red ToR dé, y por muy cifrada que esté la información entre nodos, te va a delatar. En realidad, **cualquier información personal que se revele a nivel de aplicación, cabeceras HTTP, etc. puede usarse para identificarnos** aunque usemos *ToR*, ya que esta red no la puede eliminar.

  Si necesitas algo así, debes instalar un **proxy anónimo** como `privoxy` (https://en.wikipedia.org/wiki/Privoxy, para cualquier aplicación que use *SOCKS*) o usar ***ToR Browser*** (https://en.wikipedia.org/wiki/Tor_(network)#Tor_Browser), si vas a navegar. Es muy fácil instalarlo en un *Kali Linux* estándar, por ejemplo: https://www.kali.org/docs/tools/tor/ . Este navegador ya usa *ToR* por defecto cuando lo lanzas, no es necesario configurarlo como uno "estándar". También es un poco absurdo usar un servicio donde tengas que identificarte como tu usuario real a través de una red *ToR*, salvo que lo que tengas que ocultar más que tu identidad sea tu localización. Otra cosas es, por supuesto, que tu identidad sea falsa como las que vimos en la sección anterior. En ese caso esto da igual 😊.
- **La velocidad de conexión va a resentirse**: Esto es lógico si piensas en lo que hace la red *ToR*, que básicamente hace que tus datos "salten" entre nodos. Cada "salto" es un retardo extra, pero es un precio a pagar y para la navegación "normal" no es algo limitante.
- **Algunas páginas no van a permitirte acceder a ellas vía ToR**: Sí, dado que los nodos de salida *ToR* son conocidos, algunas webs (por ejemplo bancos, grandes empresas...) mantienen una lista actualizada de nodos de salida existentes en cada momento y bloquean cualquier acceso que provenga de dichos nodos o bien carezca de determinados valores en las cabeceras HTTP que hayan podido eliminarse, asumiendo que **cualquier conexión que provenga de dichas fuentes es inherentemente maliciosa**.

  Contra esto realmente no podemos hacer más que intentar dar un "salto" más entre el nodo de salida y una máquina "intermedia" en el destino, pero que obviamente tampoco puede ser asociada a nosotros porque si no todo esto carece de sentido. Supongo que ahora te imaginas por qué **robar un PC cualquiera de un hogar es interesante para muchos atacantes**, *¿verdad?* 😊. Obviamente, en esta asignatura no vamos a llegar a este nivel de sofisticación, alquilando con una tarjeta prepago anónima y una identidad falsa (como en la sección anterior) una máquina virtual en un proveedor de nube o algo así para dar ese último "salto" (***TA0042. Resource Development*** del *MITRE ATT&CK*, **Tema 7**).

Figura 30. Esta empresa ha decidido que nada perteneciente a su infraestructura de nube *Office 365* va a ser accesible vía *ToR* o *proxies* que anonimicen la conexión

Búsquedas desde otra localización

Si lo que necesitas es simplemente hacer búsquedas desde otra localización y no te importa mucho la privacidad de lo que buscas porque solo necesitas los resultados, todo esto es muy complejo para esta necesidad. Puedes en su lugar usar ***I Search From*** (https://isearchfrom.com/) para así hacer búsquedas simulando estar en otras localizaciones o usando distintos dispositivos. A veces uno encuentra información solo así (como páginas con una versión para móvil que dan más información que su equivalente para PC).

Figura 31. Si solo te interesa aparentar buscar desde otro sitio o ser un dispositivo diferente, esta herramienta sirve

🕵️ Preparando el estudio OSINT. *¿Qué tipo de información buscar?*

Una vez hemos introducido estas técnicas, queda claro que gracias a ellas podremos acceder a muchos, muchos datos de un gran nº de personas, entidades o empresas. Obviamente, acceder a todos estos datos de primeras puede parecer una tarea muy compleja (porque realmente lo es 😊), y para facilitar este trabajo es necesario **dar un orden a la información que se podría obtener de una persona o entidad**. En otras palabras, organizarte y tener un diagrama que te guíe acerca de **qué cosas te van a interesar en tu escenario de uso**, en lugar de "lanzarse" a lo loco a obtener información que igual no te sirve para nada, o no tiene aplicación en tu caso concreto. Un posible conjunto de información aparece representado en esta imagen:

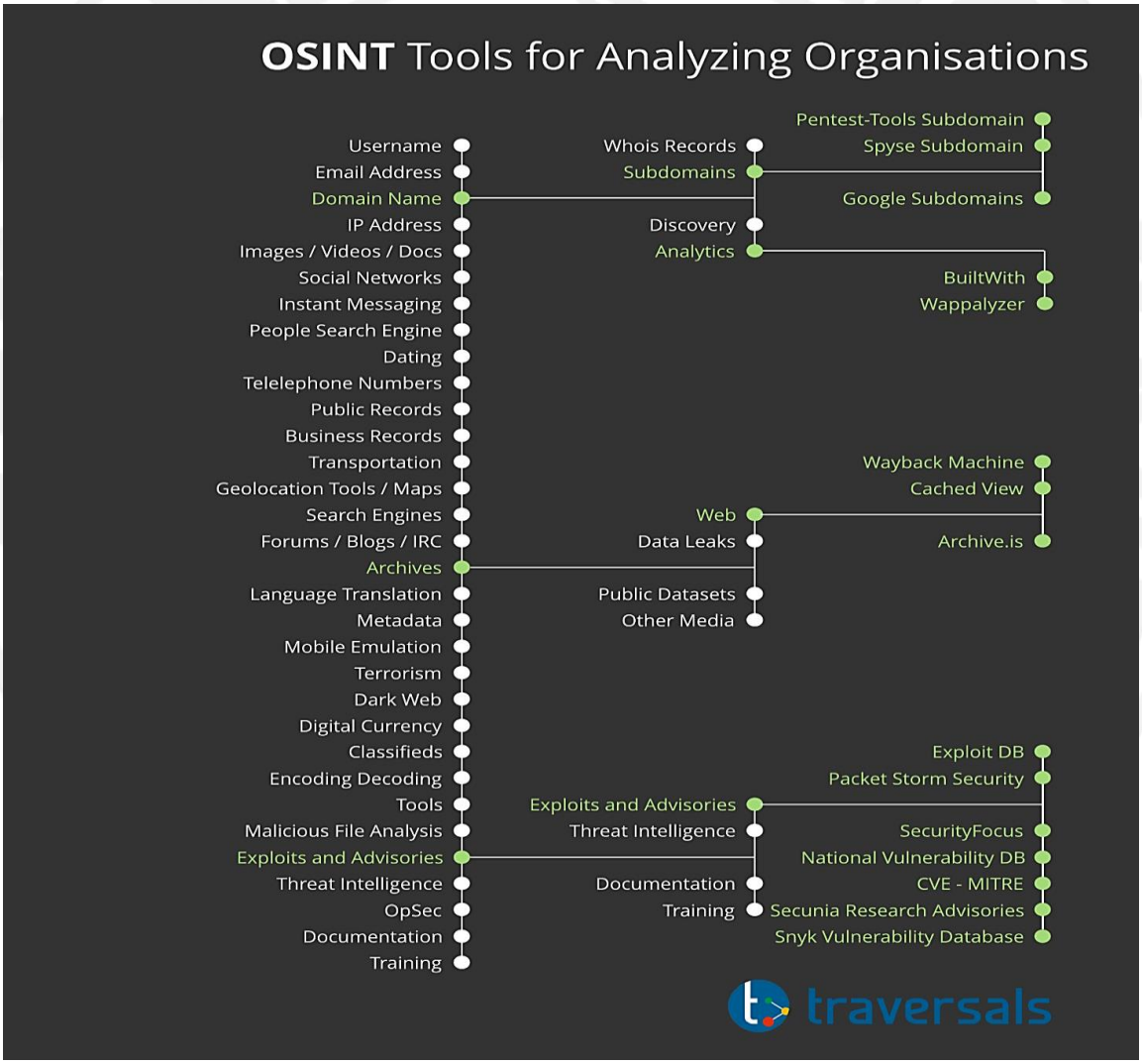


Figura 32. Ejemplo de información personal que se puede localizar vía OSINT. Fuente: https://traversals.com/blog/osint-tools/

Pero decimos posible porque realmente **no hay un orden "canónico"**, sino que pueden usarse diferentes en función del tipo de objetivo o el país en el que reside. Por ejemplo, esto es adecuado para un **ciudadano de los EEUU** ya que menciona datos que solo tienen sentido o están disponibles de forma pública allí (aunque estarás conmigo en que buena parte puede aplicarse a uno Español).

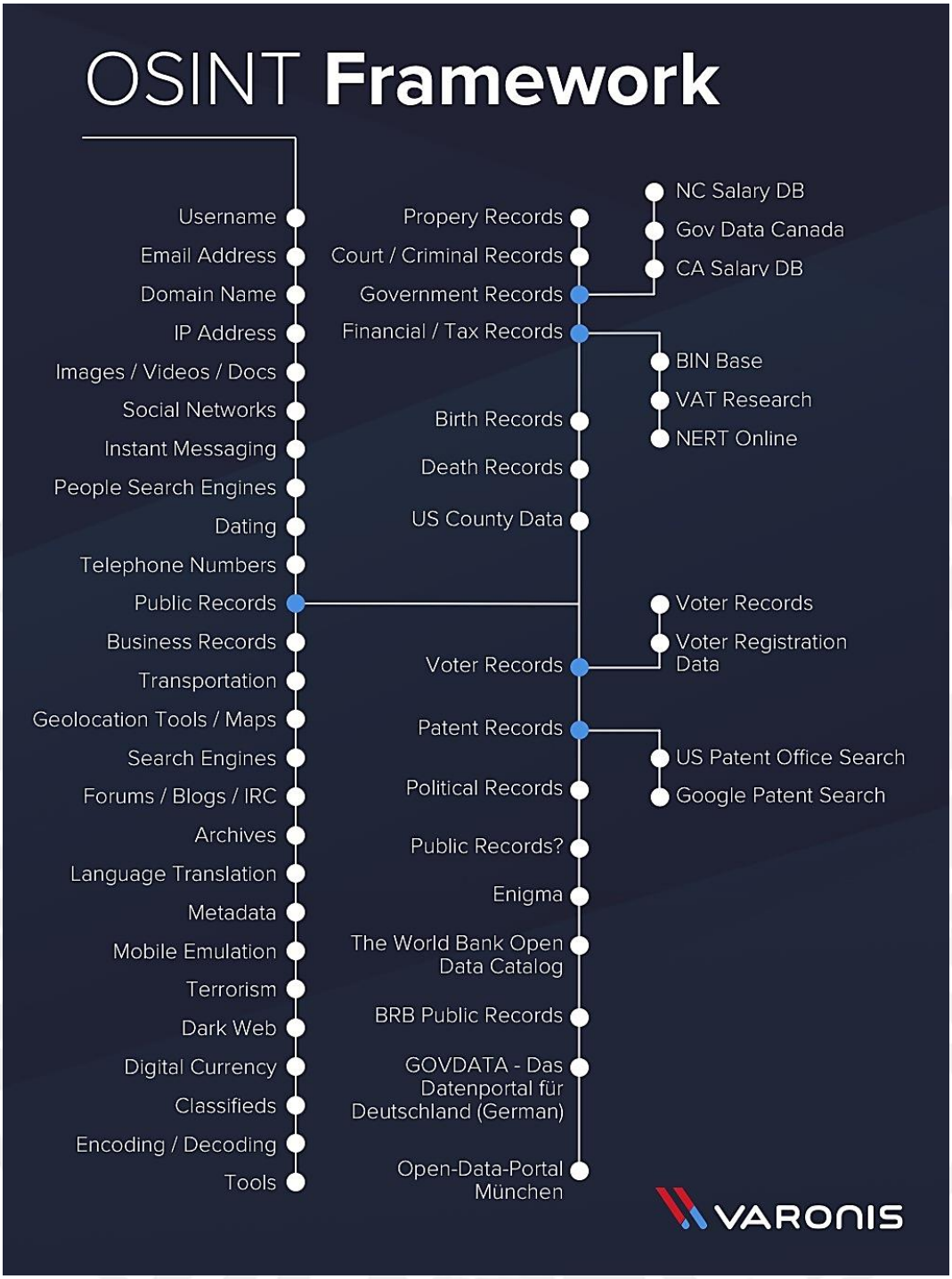


Figura 33. Ejemplo de datos obtenibles de un ciudadano de EEUU: Fuente: https://www.varonis.com/blog/what-is-osint

Si la investigación se orienta a una **empresa u organización** y es más general, quizá este esquema abreviado resulte ser más adecuado.

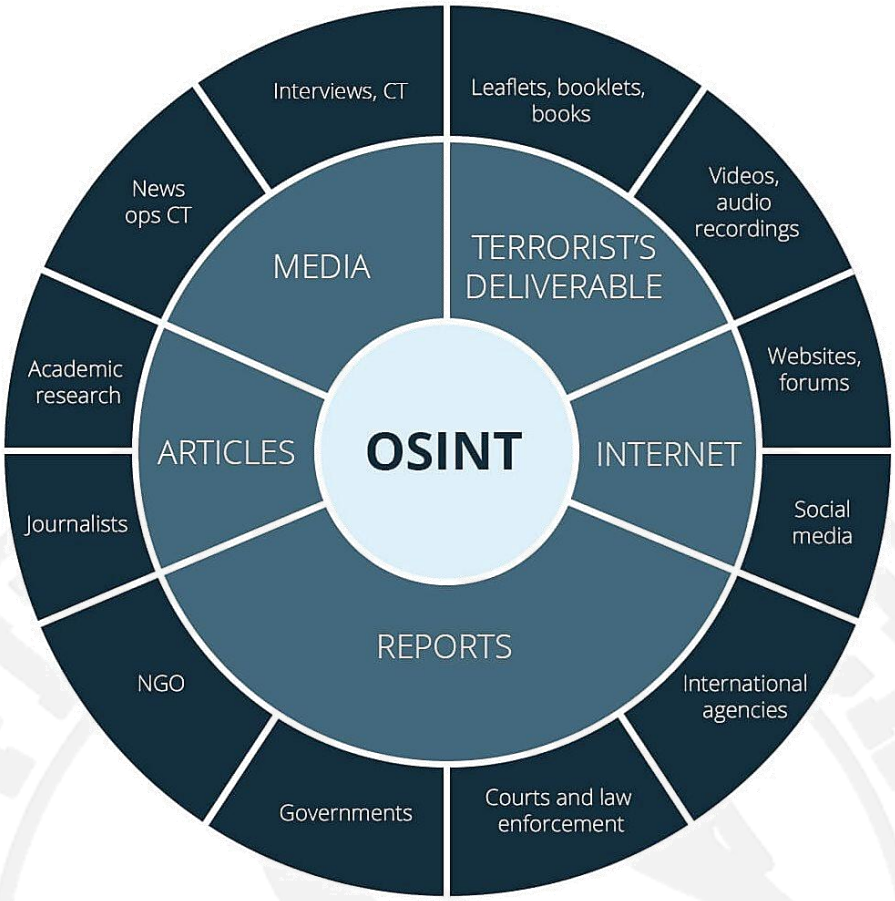


Figura 34. Esquema abreviado de información OSINT a obtener de una empresa u organización. Fuente: https://www.defensa.com/cyberseguridad/osint-como-obtener-inteligencia-explotando-fuentes-informacion

Y ya si queremos un estudio en profundidad de **aspectos técnicos y "humanos" al mismo tiempo** (y el tiempo para hacer la investigación nos lo permite), podemos usar este esquema de ***PortSwigger*** (https://portswigger.net/daily-swig/osint-what-is-open-source-intelligence-and-how-is-it-used). En este último vamos a basar el resto del documento.

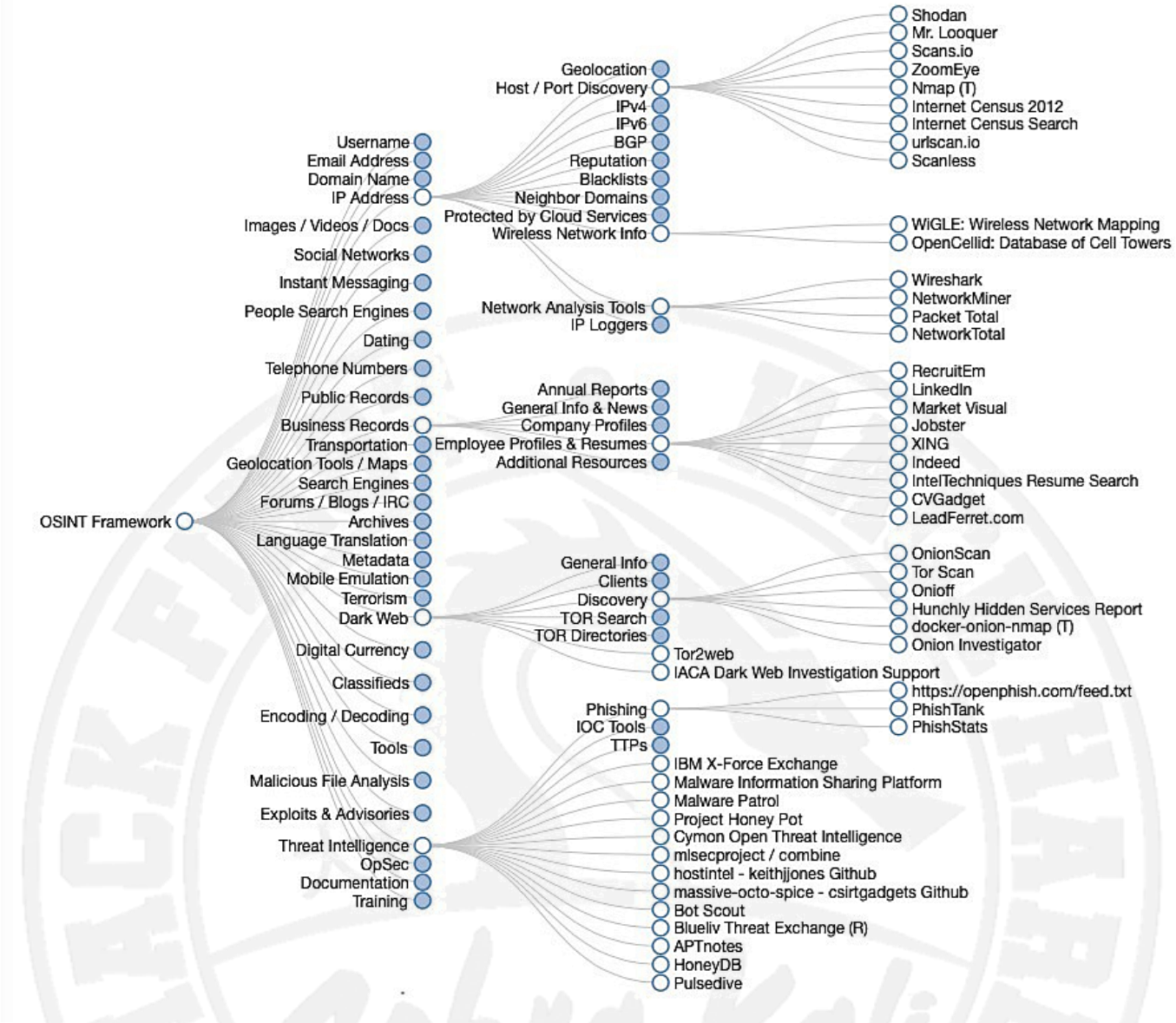


Figura 35. Esquema de información técnica y personal de una investigación OSINT de *PortSwigger*

Por otro lado, quizá tu interés sea averiguar solamente **datos técnicos de un negocio concreto y del personal asociado a él**. En ese caso este esquema puede serte útil para guiar tus investigaciones:

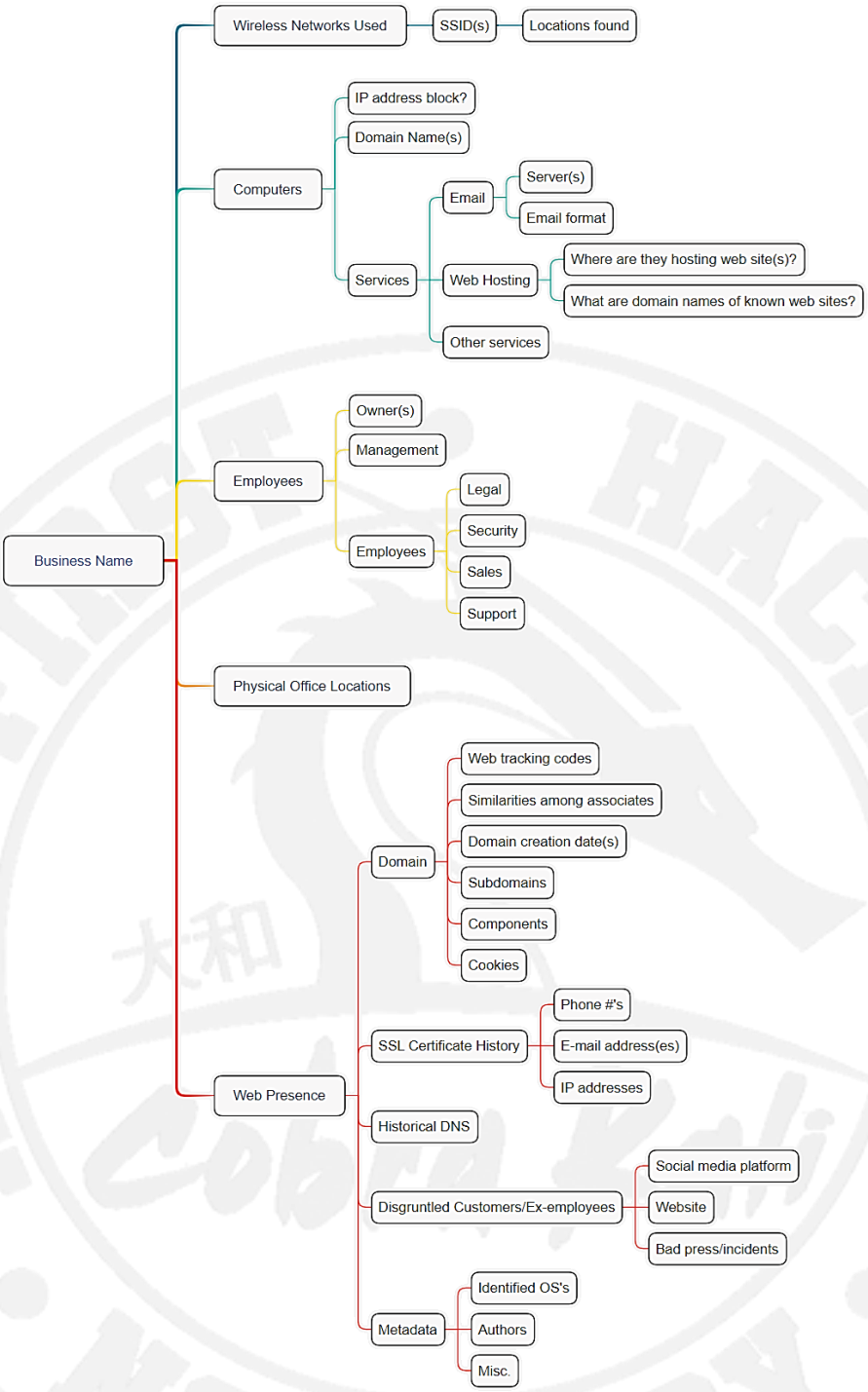


Figura 36. Datos OSINT asociados a un negocio. Fuente: https://es.linkedin.com/in/asier-rohde

Como ves, todo esto se trata de:

- **Obtener un conjunto determinado de datos públicos** (más o menos extenso, según necesidades, y **siempre de forma planificada previamente**) de una persona, entidad, organización y/o sus allegados.
- Hecho esto, **se clasifica** por entidad, persona o tipo de dato
- Una vez se tiene toda la información, **se analiza y se extraen conclusiones de ella**. De dichas conclusiones uno puede sacar de todo: posibles nombres de usuario válidos en un sistema, posibles indicios de claves, aspectos de la vida personal y laboral no obvios a priori...
- Muchas veces esto se usa para labores de *phishing* de distintas clases o para facilitar mecanismos de acceso inicial (***TA0001. Initial Access*** del *MITRE ATT&CK*, como vemos en el **Tema 7** de teoría y

en los **laboratorios 11 y 12**), al ser básicamente una técnica de reconocimiento más. Aunque, como puedes ver, en función de cuanto la "estires" puede englobar a otras muchas....

¿Y cómo saco toda esta información? Pues lo cierto es que **no hay un método automatizado que valga en todos los casos**: depende de la víctima, de su contexto y su entorno. Lo que si es cierto es que en la inmensa mayoría de ocasiones se hace con una **combinación de búsquedas manuales y automatizadas**. Es más, normalmente las búsquedas manuales "alimentan" a las automatizadas, aunque realmente ambas deben actuar como complemento (una búsqueda manual puede servir para confirmar que una automatizada ha encontrado algo real y no un falso positivo).

A continuación vamos a ver técnicas manuales para sacar información y su posible aplicación práctica, relacionándolas cuando sea necesario con el resto de la asignatura. Hecho esto, veremos una herramienta de referencia para recopilar esa información y automatizar búsquedas: ***Maltego***. Finalmente, veremos técnicas para sacar esta información "por cualquier medio", puesto que a veces la persona que nos contrata nos da posibilidades adicionales para extraer información, como hacer **campañas de *phishing* simuladas** contra un objetivo para comprobar su "ciber-concienciación" (no estoy de broma, muchas empresas las hacen, mismamente Uniovi hizo una hace pocos años). Con todas estas técnicas, tu habilidad de "cibermarujeo" se verá incrementada a más de 9000 puntos 😉.

OSINT "Manual": El arte de buscar e interpretar los datos

OSINT sobre empleados o personas

La cantidad de información que se puede encontrar sobre una persona en la red es muchas veces **abrumadora**, y por ello si no sabemos por dónde empezar podemos sentirnos perdidos. En esta sección vamos a hablar de muchas clases de información diferentes, pero un **esquema general** lo tienes en este diagrama, para que te hagas una idea de cómo de expuestos podemos estar en Internet.

Figura 37. No somos conscientes de lo que podemos exponer en Internet...hasta que lo buscamos

Si esto te parece demasiado, **otro esquema más concreto** puede ser este:

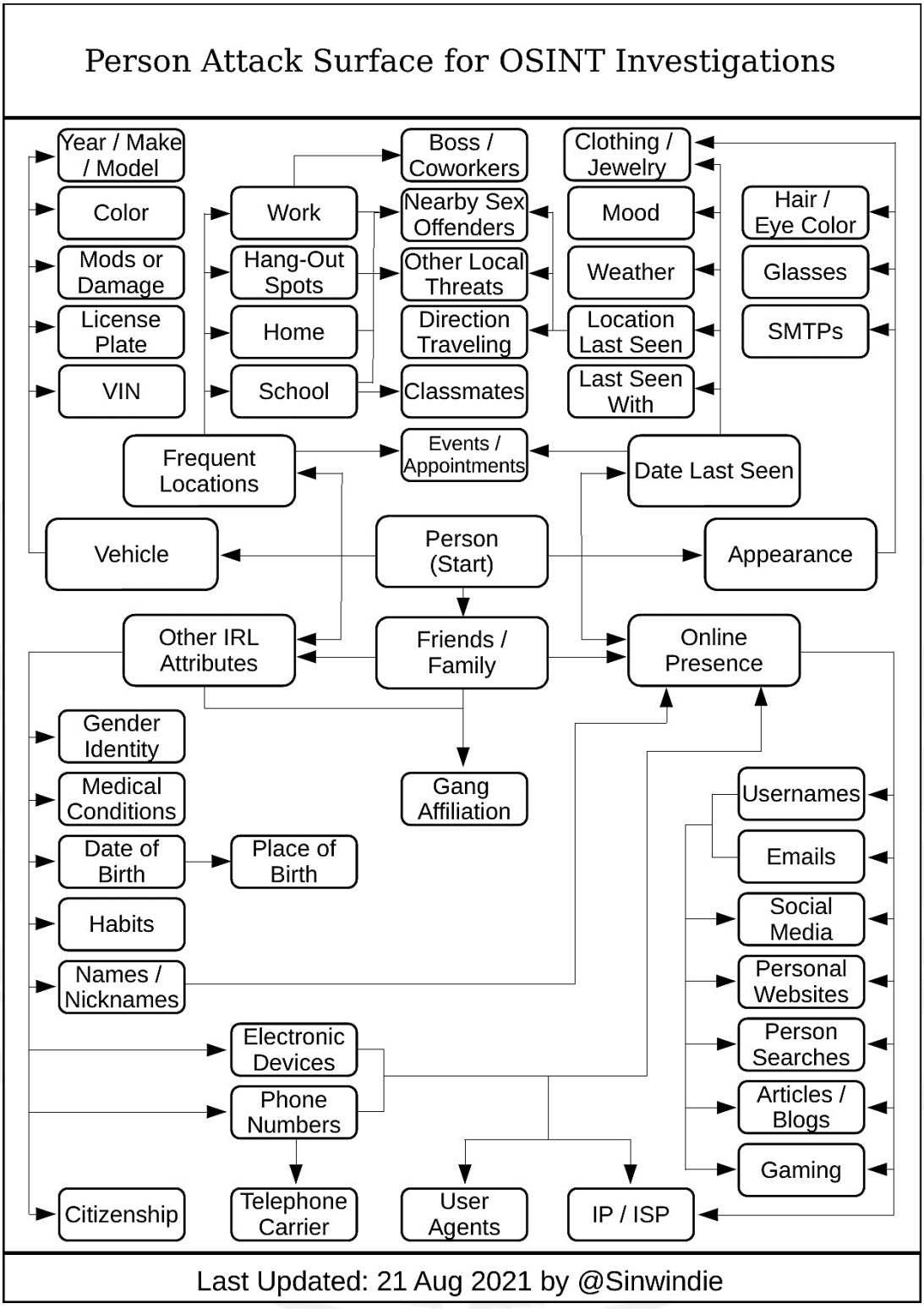
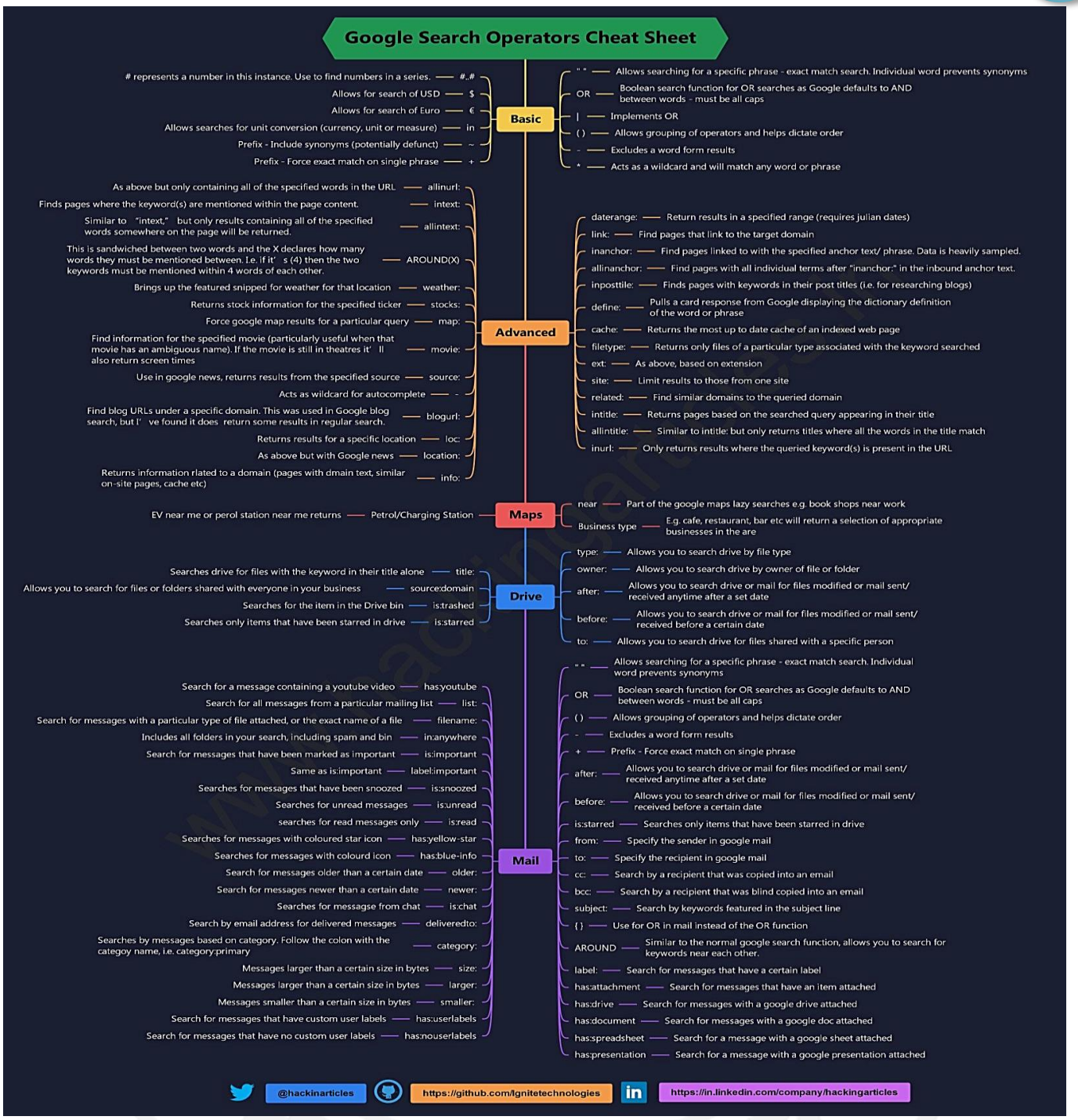


Figura 38. Información más abreviada de lo que puede averiguarse de una persona por Internet

Cosas acerca de personas en general: Nombres / Imágenes / Vídeos / Fotos...

Lo primero que debemos saber es que **nunca tenemos que subestimar el poder de un buscador**. No se trata de buscar un nombre y ya está, *Google*, por ejemplo, tiene muchas opciones de búsqueda:



Aunque buscar el nombre de alguien por Internet sea básico, podemos mejorar estas búsquedas con trucos sencillos:

- Buscar el nombre completo si lo tenemos disponible.
- Entrecomillar el nombre completo para buscar exactamente dicho nombre. Hacerlo con el formato `<Nombre> <Apellidos>` y `<Apellidos>, <Nombre>`
- Añadir la profesión, localización, hobby, ...o cualquier otro dato relevante que sepas tras el nombre para mejorar la precisión de la búsqueda.
- Usar ***Google Advanced Search*** (https://www.google.com/advanced_search) para incluir o excluir ciertas palabras, frases o webs de los resultados de búsqueda.
- Buscar el nombre en redes sociales, especialmente en *LinkedIn* por ser la que más datos potenciales puede dar.

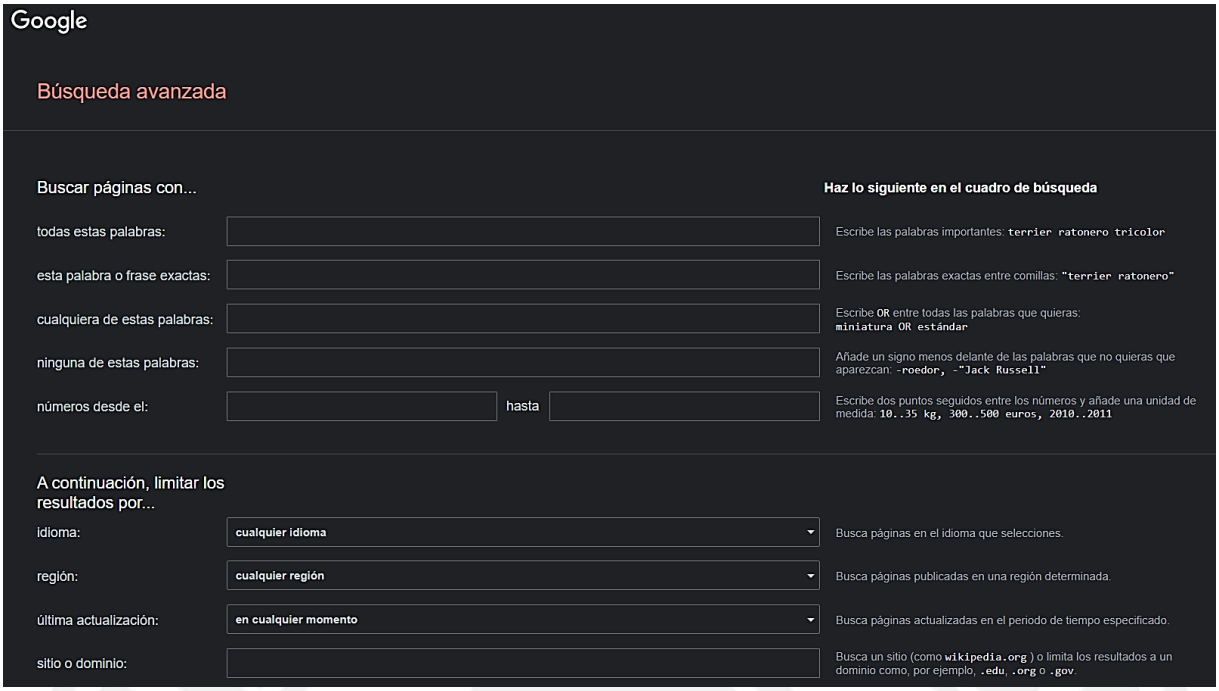


Figura 39. *Google Advanced Search*

Si no encontramos la información que estamos buscando en un momento dado, también podemos usar ***Google Alerts*** (https://www.google.com/alerts?hl=es) para recibir una notificación cuando *Google* indexe ciertos términos de búsqueda que pueden ser interesantes para nosotros y que podemos configurar, de modo que podamos saber cuándo aparecen ciertos datos en Internet. Esto también es recomendable para saber cuándo aparecemos **nosotros mismos** en los datos de búsqueda de *Google*, para evitar fugas de datos, menciones no deseadas o situaciones similares.

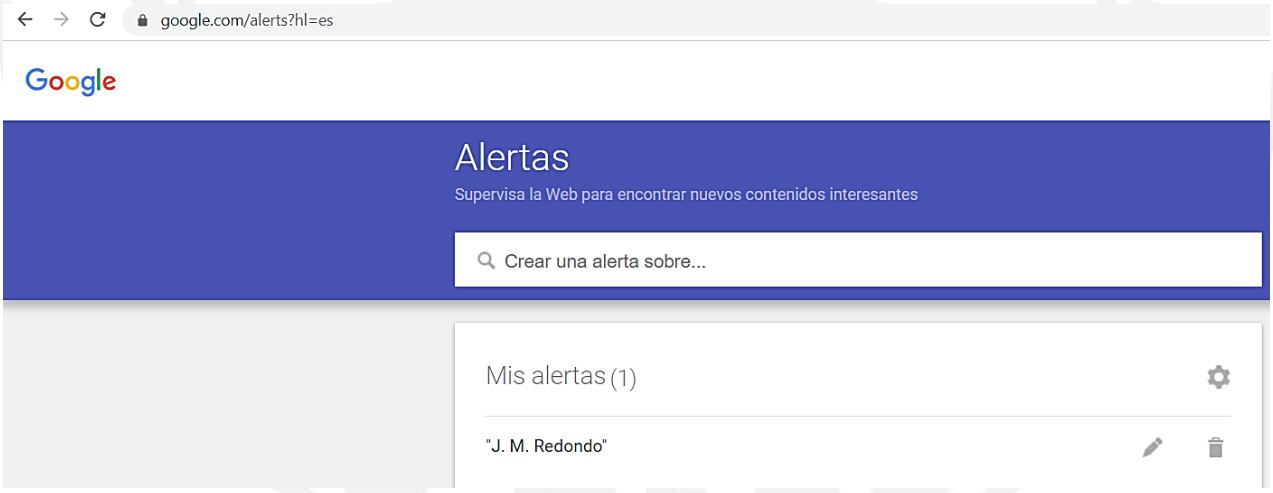


Figura 40. Servicio *Google Alerts*

Si por el contrario no contamos con ningún dato de un objetivo más allá de una foto, la mejor forma de encontrar más información es recurrir a una **búsqueda inversa de imágenes con *Google*** (https://www.google.com/imghp). Simplemente descarga o recorta la imagen que te interese (aplicación "***Recortes***" de *Windows*, por ejemplo), guárdala y arrástrala y suéltala al cuadro de búsqueda.

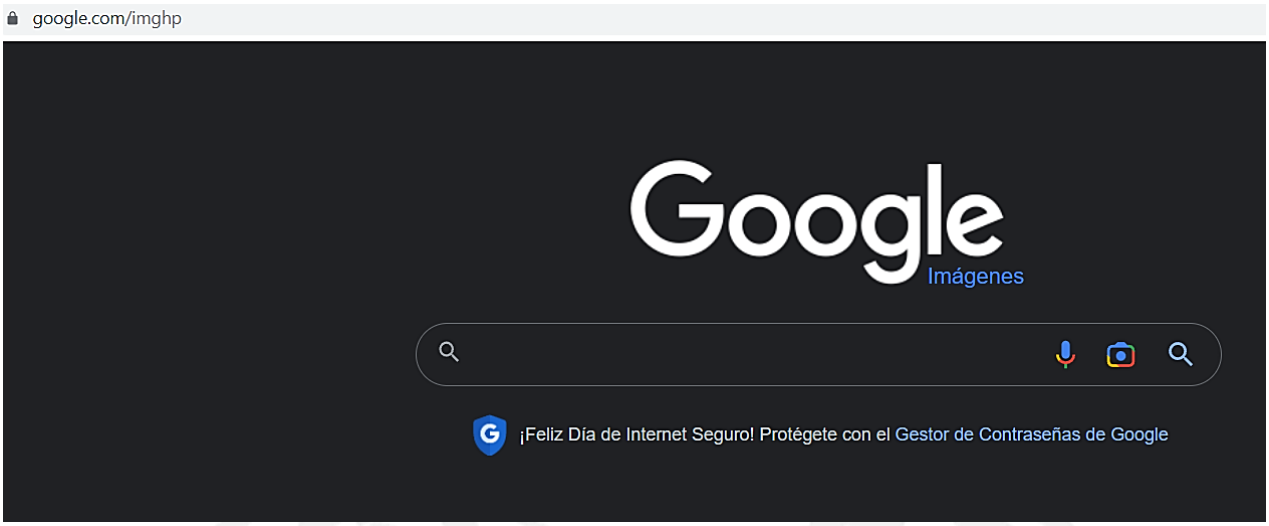


Figura 41. Búsqueda de imágenes de *Google*

Con un poco de suerte podremos tener un nombre y la combinación de este con la imagen probablemente nos permita encontrar **redes sociales** del objetivo. Con ello, es frecuente que aparezcan más datos interesantes como nombres de personas asociadas al investigado (familiares, compañeros de trabajo, pareja(s), descendientes, etc.) de los que "tirar del hilo" para seguir con la investigación. Toda persona encontrada así es investigable para saber qué podemos averiguar a partir de ella. Otras herramientas de OSINT pueden hacer **reconocimiento facial** vía fotos en redes sociales, pero se debe localizar alguna que siga mantenida en el repositorio de ***CiberPatrulla***.

Una foto es algo que puede ofrecernos mucha más información:

- Por ejemplo, salvo que esté subida a una red social típica (ya que los eliminan) es posible examinar sus **metadatos** en busca de más información (lo vimos en el **Laboratorio 2**, *¿recuerdas?*). Aunque las redes sociales eliminen los metadatos, también es posible que una misma foto presente en una red social del objetivo se haya subido a distintos sitios (es decir, hay que usar la búsqueda de imágenes en *Google* de nuevo). *¿Por qué es eso interesante?* Pues porque algunos de esos sitios pueden no quitan los metadatos, y si el objetivo no lo hace por su cuenta, ahí estarán. Por tanto, en Internet puede haber una misma foto con y sin metadatos, dependiendo de dónde se esté mostrando. Así puede darse el caso paradójico de que se puede saber mucho de los metadatos de una foto en una RRSS por una copia de esta fuera de esa RRSS.
- Aunque una foto reciente no ofrezca metadatos, quizá **una foto del pasado** sí lo haga (la preocupación por eliminar metadatos es algo que se popularizó recientemente). *¿Qué conoces que pueda "rebuscar" en el pasado de una web en Internet y, dentro de ella, en sus archivos individuales?* Exacto, la web ***Wayback Machine*** del **Laboratorio 2** 😊.
- Por otro lado, **detalles** de la foto como **lugares reconocibles**, **fecha de la foto** (¡a veces aparece impresa en la propia foto!), su nombre (pero ojo, muchos sitios como las redes sociales modifican su nombre por algo aleatorio), **matriculas** de coche, números de **paradas** de autobús/tren/taxi, **calles** o **plazas** concretos, **monumentos** y un largo etc. pueden ayudarnos, con un hábil uso de *Google Maps / Earth*) a **geolocalizar** a una persona en un momento dado en el tiempo (o aproximado) o cualquier otra información relevante.

  Esto no es descabellado, si quieres un ejemplo puedes buscar por qué un conocido *streamer* quedó en *sock* (*guiño* *guiño* *codo* *codo*) cuando se supo la dirección de su nueva casa por un detalle que aparecía en una ventana en el fondo de una foto que subió a sus RRSS. A veces los investigadores usan una identidad anónima (como las que se crean en la sección anterior) para **preguntar por distintos foros a gente** (también anónima) si conocen el lugar al que pertenece una

foto. Las probabilidades de encontrar a alguien que viva cerca o pase por allí son altas, especialmente en foros grandes. Por otro lado, imagina que en la foto aparece el modelo de cámara con la que se hizo, y que esta es una cámara profesional muy cara. *¿Cuánta gente puede tener esa cámara entre una lista de posibles "sospechosos"?* Y si es una cámara muy cara, es muy probable que "vacile" de ella en RRSS... *¿entiendes un poco como va la cosa?* Sí, si se parece a un juego de detectives es poque a veces es un "juego" de detectives 😊.

¿Y tú para que quieres saber eso? pues porque cuanta más información tengas, mejor será tu investigación. Una persona que le guste "fotografiar" su vida más o menos intensamente está exponiendo, quizá sin saberlo, una enorme porción de su vida. Si quieres ejemplos reales, te puedo hablar de unos cuantos 😊

- Si acostumbra a hacerse fotos en **diferentes lugares y momentos con alta frecuencia**, es técnicamente viable hacerse un "mapa de su rutina" con *Google Maps*/*Earth*. Aparte de los riesgos que puede tener para la persona tener lugares donde es más posible encontrarla en un momento dado (*creeps*, *stalkers*...), incluyendo hacer "espionaje en vivo" si el objetivo es suficientemente importante, establecer una rutina en un mapa puede revelar también lugares de trabajo, de recreo, gimnasios o incluso la vivienda de uno, incluso sin tener activada la geolocalización. Obviamente con la geolocalización activada esto es mucho más trivial 😊
- Si la persona acostumbra a poner **fotos de vacaciones**, esquiando en lugares conocidos, de monte, etc., no solo está revelando sus gustos, sino también a dónde suele ir cuando llegan periodos vacacionales largos, nuevamente incrementando las opciones de encontrarla en un sitio en el mundo real en ciertas fechas...o dónde no encontrarla.
- Este tipo de fotos sociales **pueden involucrar a más personas**. Cada una de ellas es un objetivo adicional a investigar siguiendo las mismas técnicas que se hacen para la "persona primaria". *¿Y para qué quieres hacer eso?* Pues porque no pocas veces esas personas allegadas escriben cosas relativas a la primera, a sus actividades o su trabajo que resultan útiles para la investigación. *¿No quieres decir que vas de boda?* Mala suerte, tu amigo te ha hecho una foto en la que sales de fondo en el banquete y la ha subido a *Instagram*. *¿Quieres mantener la privacidad de tus hijos?* Lo siento, la madre de su mejor amigo les ha hecho una foto en el parque y la ha puesto en *Twitter*...etc. Lo peor es que las "víctimas" de este tipo de problemas tienen muy difícil esquivarlos o defenderse de ellos. **A nivel social hay muy poca conciencia de los peligros de este tipo de actividades**.

  Ten en cuenta que hasta anécdotas laborales se pueden usar para un ataque de ***Spear phising***, por ejemplo. Por supuesto, esto nos permite construir una red de allegados como decíamos antes, que no deja de ser otra poderosa fuente de información. Esa red se puede construir e ir enriqueciendo con el programa *Maltego* que veremos después.
- Una foto también puede sacar **características físicas**. Por ejemplo, postear un foto de cuerpo entero con algo al lado de tamaño reconocible permite calcular la altura con una sencilla regla de 3, como el señor que sujeta un folio A4 de la siguiente foto, que medirá en torno al 1,77m. Incluso también hay herramientas de IA a partir de estas fotos que permiten calcular el peso aproximado (en este caso entre 84-87Kg).

Figura 42. Averiguando la altura de una persona con una foto. Fuente: https://ciso2ciso.com/introduction-to-doxing-osint-methods-for-information-gathering-by-hadess/

Y ahora te preguntarás. "*Vale, muy bien. Ahora dime para que me vale eso*". Pues fíjate, es cuestión de dejar volar tu imaginación...te cuento como dejo yo volar la mía, que te puede servir de pista 😉

- Si la **altura varía ligeramente en el tiempo**, puede que detectes que esa persona tiene un problema de espalda, como una compresión vertebral o una lumbalgia que te obliga a caminar ligeramente encorvado.
- Si la persona es **muy alta o baja**, puede que tenga complejo o problemas para encontrar tallas de ropa, lo cual puede servir de dato para una operación de *phishing* o saber el tipo de tiendas que frecuenta en casos muy concretos.
- Una **variación de peso acusada** en un periodo de tiempo relativamente corto, tanto hacia arriba como hacia abajo puede ser indicio de:
  - Un **trabajo más bien sedentario** (nuevo o actual)
  - Haber **dejado el gimnasio**
  - Estar sometido a determinada **medicación** donde el aumento o pérdida de peso sea un efecto secundario claro
  - Una **enfermedad** (normalmente grave: anorexia, bulimia, depresiones y otras enfermedades mentales (la medicación contra ellas suele provocar esos efectos secundarios))
  - Un **cambio de estilo de vida a mejor o peor** (preparar una carrera exigente, un "susto" de salud que le he hecho "ponerse las pilas", una etapa compleja de la vida ("crisis de los 40"), unas oposiciones o incluso una ruptura sentimental

Aunque son posibilidades a introducir en el "*mix*" de información acerca de una persona, estoy de acuerdo que **son suposiciones y algunas muy arriesgadas**. Pero al menos, son temas a introducir en conversaciones posteriores con el investigado que podamos tener (recuerda el apartado de identidades falsas en casos donde haya *engagement*), que usar en una campaña de *phishing* luego, o que contrastar con otra información que obtengamos a ver si alguna tiene posibilidades de ser correcta. Como digo, el límite está en tu imaginación 😊.

Deportes de interés

Las fotos con banderas, bufandas, gimnasios, etc. de equipos de futbol / baloncesto / etc. dicen mucho de una persona. Si una persona se fotografía haciendo distintos deportes o en distintos lugares, no solo tienes

un posible sitio más donde encontrarlo, sino que tienes un tema de conversación más o herramientas para usar en *phishing*. Por supuesto, también salen otras personas a las que investigar a raíz de ese tipo de fotos.

Figura 43. Persona que publica su vida deportiva en RRSS. Fuente: https://ciso2ciso.com/introduction-to-doxing-osint-methods-for-information-gathering-by-hadess/

Esto es más importante de lo que parece realmente. Si eres un "fan fan", por ejemplo, del *Sporting* o del *Oviedo*, te sacas fotos en el campo, yendo a los partidos, etc. en realidad estás diciendo **dónde vas a estar** (y, ojo, también **dónde NO vas a estar**) con una gran probabilidad en momentos de tiempo públicos y conocidos. Si te sacas fotos siguiendo al equipo amateur de tu hijo/a en sus partidos, esos calendarios de partidos son fáciles de encontrar, por lo que ya sabrán donde encontrarte regularmente. La composición de los equipos también suele ser pública (frecuentemente, hasta con fotos), por lo que ya tienen a tus hijos y sus amigos. El personal que forma parte del equipo es también fácil de obtener, así que ya tienen allegados con los que dar contigo "saltando" a través de perfiles de RRSS si tu foto no les dice nada. Y un gran etc., si lo piensas bien...

Coches (u otras posesiones)

¿Tienes un coche chulo? Me alegro mucho, en serio. Pero si haces una foto y se ve la **matrícula** pueden usarla para, a través de ella, saber quién eres y más datos de ti. Aparte de, por supuesto, estimar **tu poder adquisitivo**. Esto es así salvo que la persona sea un "*poser*" que los alquila para "vacilar". En ese caso la matricula, pegatinas o características distintivas del coche (que lo hagan reconocible en fotos de otras personas) puede servir para saber que la persona es precisamente eso :)

Figura 44. Bonito coche. Este tipo de coche me dice que no tienes hijos... Fuente: https://ciso2ciso.com/introduction-to-doxing-osint-methods-for-information-gathering-by-hadess/

Obtener los datos de un propietario de un automóvil por su matrícula no obstante **no es gratuito y requiere aportar una causa justificada**, como se indica en estos dos enlaces.

- https://tramicar.es/blog/como-identificar-al-propietario-de-un-coche-por-la-matricula
- https://blog.reale.es/averiguar-propietario-coche-por-matricula/

Tampoco sería de extrañar que un día (si no ha ocurrido ya...) **se filtre una base de datos de automóviles registrados en un país** donde directamente se puedan buscar datos tuyos por matrículas (recuerda: en realidad estamos pensando cómo actúan los delincuentes, si teniendo límites éticos y morales ya puedes averiguar mucho, imagina si no los tienes...). No obstante, muchas veces **ni siquiera hace falta**, porque es la propia persona la que, a base de analizar todas las fotos que sube a las redes, da él mismo esta información. Esto es un aspecto importante también: una sola foto podría ser inofensiva, pero **la combinación** de dicha foto con otras puede darnos datos interesantes. Y normalmente los investigados no tienen en cuenta el contexto de sus fotos pasadas antes de subir una nueva...

Comida

Si te haces foto de lo que comes, es posible que revele no solo qué restaurantes frecuentas, sino que se puedan hacer suposiciones acerca de tu país de procedencia o raza, intolerancias o si eres o no vegano,

por ejemplo. Aparte de otros elementos (personas, objetos...) reveladores que pueda haber en la foto, claro 😊

Colores / Estilo de ropa / Cuidados personales

Tu color favorito puede ser usado en temas de *spear phishing* (hay toda una disciplina de *marketing* que usa datos sutiles de la mente humana para llegar a las personas...), además de que el **tipo de ropa** puede decir mucho de **tu personalidad**, **nivel económico** o incluso tiendas en las que frecuentas comprar (físicas u online) si son de marcas concretas. Como todo lo que uno revela, no sabes cuando va a ser un dato útil...

Si la persona lleva peinados que se ve a la legua que llevan trabajo hacer, cambia de estilo de corte frecuentemente, lleva barbas arregladas, maquillajes de cierto estilo, etc. También está revelando datos sobre sus gustos, personalidad, posible trabajo, intereses o intenciones y, quizá, lugares que frecuenta.

Figura 45. Persona que se gusta mucho a sí mismo. Fuente: https://ciso2ciso.com/introduction-to-doxing-osint-methods-for-information-gathering-by-hadess/

Libros / Películas / Canciones favoritas

El revelar públicamente tus libros / películas / canciones favoritas, incluido los comentarios que puedas hacer de ellas o de sus personajes en foros públicos, puede ser muy revelador de cosas como:

- *Background* educativo
- Inclinaciones políticas
- Fe religiosa
- Valores vitales o pensamientos acerca de temas que puedan resultar polémicos para ciertos sectores de la población si estás a favor o en contra de ellos.

En otras palabras, **que es muy fácil que te puedas "catalogar socialmente"** de esta forma. Si bien es cierto que es muy frecuente que la gente dé esa clase de información "gratis" en redes sociales. Por ejemplo, hay gente que habla de su militancia en un partido político determinado como si fuera el escudo de armas de su familia...

Aunque pueda parecer una tontería, revelar publica y repetidamente listas de *Spotify*, la canción que estás escuchando en estos momentos (y los comentarios que hagas) puede usarse para inferir tu estado de ánimo, eventos importantes próximos de tu vida, enfermedades que estés pasando (depresión, ansiedad....),

y rasgos de tu personalidad. Estoy hablando en serio. No es la primera vez que veo listas como "`Canciones para superar una ruptura`", "`Temas para la boda`", "`Música de fondo para el MIR`" o cosas así. Aparte, sabes que la propia aplicación te clasifica canciones en función del estado de ánimo que le pongas, por lo que tú también puedes hacer ese tipo de suposiciones. Todo ello puede ser información útil en un *spear phishing*, por ejemplo.

Fotos de tu casa

Como pasa habitualmente (y ha sido noticia recientemente con algún *streamer*, como decíamos antes) hacer fotos de tu casa puede revelar fácilmente **dónde vives** por el fondo de éstas (si es reconocible) u otros detalles de tu vida social y personal. Una foto de fondo que muestre **la cara de alguien**, **animales de compañía** (placas identificativas, nombres que pueden ser indicios de claves que tú uses en tus webs...), **tipo de dispositivos** que tienes (que pueden dar pistas de como atacarlos desde fuera simplemente, por ejemplo, averiguando su marca y modelo por lo que sale en la foto), o **cualquier objeto que salga por accidente** en la foto y revele algo de ti (no es la primera vez que aparecen documentos confidenciales, DNIs, tarjetas de crédito...), los tuyos o de algo que tienes. Por ejemplo, hace poco una empresa vacilaba en redes sociales de sus nuevos *routers* instalados en su centro de datos. En las fotos se veía el modelo y marca claramente, y ese modelo tenía una vulnerabilidad conocida reciente. *¿Los habrán parcheado?* Ya me entiendes...

El caso más llamativo es el de "fondo de la ventana". Piensa que aunque el fondo en sí no sea fácilmente reconocible, si unes ese dato a lo que ya hayas podido averiguar (rutas, rutinas, etc.) haces que localizar tu casa sea mucho más sencillo. Ahora pon que vas de vacaciones en una fecha concreta, y ya entiendes porque esto es peligroso....

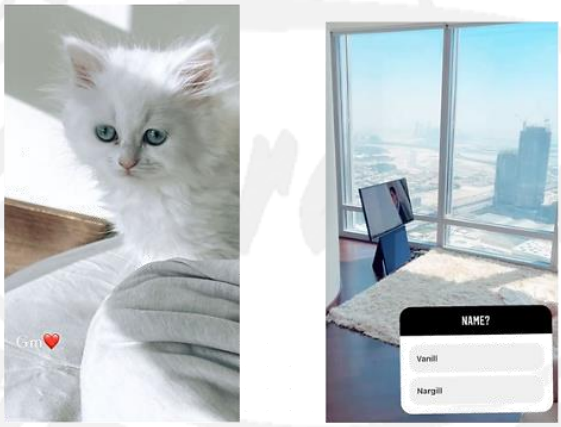


Figura 46. Foto de fondo de ventana que puede dar lugar a reconocer dónde vives. Fuente: https://ciso2ciso.com/introduction-to-doxing-osint-methods-for-information-gathering-by-hadess/

Fotos de las llaves de tu casa / otras propiedades

Publicar una foto de las llaves, por ejemplo de tu nueva casa comprada o alquilada puede parecer algo muy bonito (de hecho lo es), y es algo que es frecuente en las redes sociales. **Desgraciadamente, es una muy mala idea**. Estas fotos **se pueden usar para hacer una copia física funcional** de las llaves que acabas de fotografiar, con el problema de seguridad para ti y los tuyos que ello supone. Si además tienes en cuenta que las fotos pueden estar etiquetadas con tu localización, pueden contener algún metadato o puede examinarse el contexto de tu RRSS para deducir donde vives, comprenderás porque esto puede ser un problema grave.

Una impresora 3D moderna puede copiar una llave. El motivo es que las fotos de los móviles hoy en día tienen una resolución muy alta, y por tanto es viable traducirlas a un esquema digital imprimible de la misma. Si luego se usan elementos de la foto como manos, anillos u otros elementos de fondo para tener idea de una escala y dimensiones de la llave, entonces eliminas otro problema. Aunque las llaves se diseñan con distintas longitudes y profundidades de crestas, la sombra de estas puede usarse para estimar su tamaño. Además, si van a por ti por algún motivo, dado que la impresión 3D hoy día es muy barata, se pueden hacer varias versiones de la llave hasta dar con la que tiene la profundidad y longitud adecuada.

Muchas llaves tienen además **un código escrito**, que se puede usar para obtener una copia exacta de la misma del fabricante. Si alguien es capaz de ver dicho código por una foto, pueden usarla para obtener una copia física de la llave también. En otras palabras, **cualquier persona que ponga una copia de sus llaves en redes sociales con suficiente resolución tiene un problema de seguridad que debe eliminar antes de que sea demasiado tarde**.

Figura 47. Bonita llave...pero yo no debería estar viéndola

Direcciones de email / Nombres de dominio

Saber a quién pertenece un dominio o un email de contacto puede servir no solo para identificarle, sino para buscar filtraciones de datos relacionadas con dicho email (o con la empresa a la que pertenece el dominio). Y claro, *¿cuánta gente usa la misma clave para varias cosas? ¿y/o no usa MFA?* Pues eso....en el **Seminario 1** y en el **Laboratorio 2** vimos herramientas para poder ver **quién es el propietario de un dominio** (servicios `WHOIS`) y buscar esta información. Contamos también con la ayuda de los motores de búsqueda mencionados en la sección correspondiente de este documento.

Los emails de cualquier persona son también muy interesantes porque se pueden usar como una fuente de contenido para una investigación muy interesante. El siguiente esquema de ***OSINT Dojo*** (https://github.com/sinwindie/OSINT), es un muy buen resumen de lo que se puede llegar a sacar a partir de un email que recibas de alguien (ahora entiendes porque tener cuentas falsas que reciban emails de los objetivos investigados es más importante de lo que crees 😊). Tampoco está de más buscar trozos del contenido en *Google* por si se tratase de **emails genéricos** copiados de una plantilla (cosa que es típica de programas como SET, como veremos luego), o **bien *scams* conocidos**.

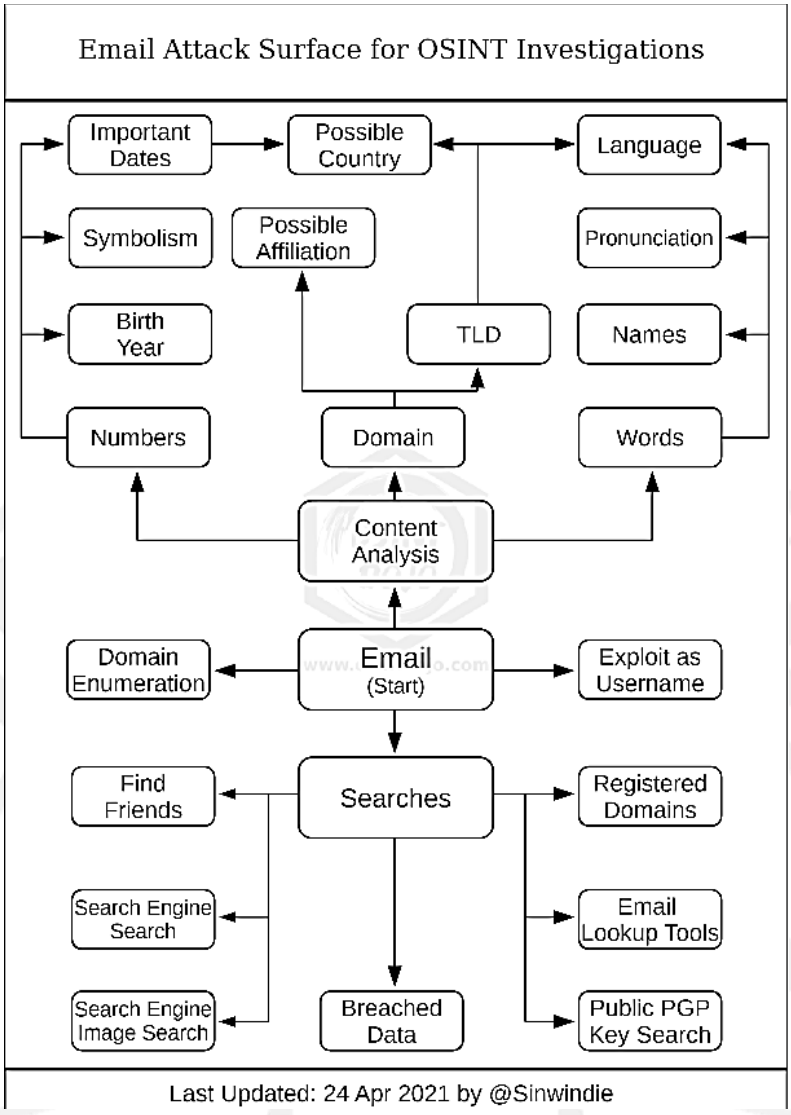


Figura 48. Superficie de ataque de un email

Otro esquema alternativo que puedes seguir es este, hecho por Michael Bazzell (https://www.linkedin.com/in/michael-bazzell-a83572122/). Date cuenta de que en este diagrama aparecen mencionados servicios que te pueden ayudar en esta tarea:

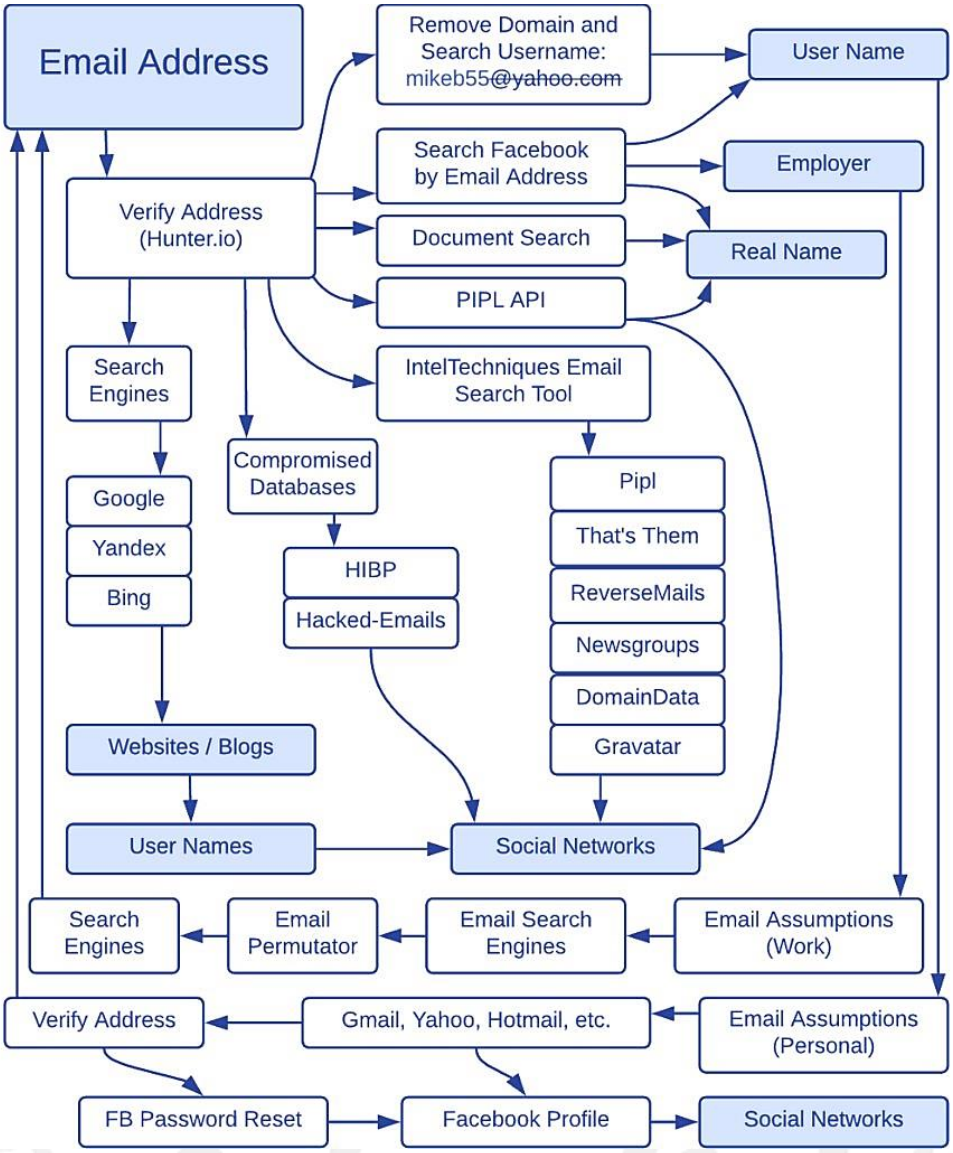


Figura 49. Superficie de ataque de un email, pero siguiendo una secuencia alternativa

Truco para averiguar el nombre detrás de una cuenta de Gmail

A veces podemos necesitar saber **el nombre de quien está detrás de una cuenta de *Gmail*** que nos ha escrito por algún motivo, y sin alertar de ello al emisor. Se puede hacer de forma sencilla creando un documento en *Google Documents* cualquiera y compartiendo ese documento con esa cuenta de correo (pero acordándose de **desmarcar la opción de notificar al usuario**). Con ello *Google* nos mostrará el nombre que luego podemos usar para hacer más investigaciones. Hay que acordarse de descompartir luego el documento para no dejar rastro alguno de esta operación. Esta técnica se puede usar, por ejemplo, para desenmascarar quién está detrás de cuentas de correo de *Gmail* (aunque quizá sirva con otros proveedores equivalentes, como *Outlook*) que nos envíen correos inapropiados.

NOTA: Existe otra técnica similar que se hace compartiendo un calendario de *Google Calendar*, pero esto si notifica al usuario de que hemos compartido algo con él queramos o no...

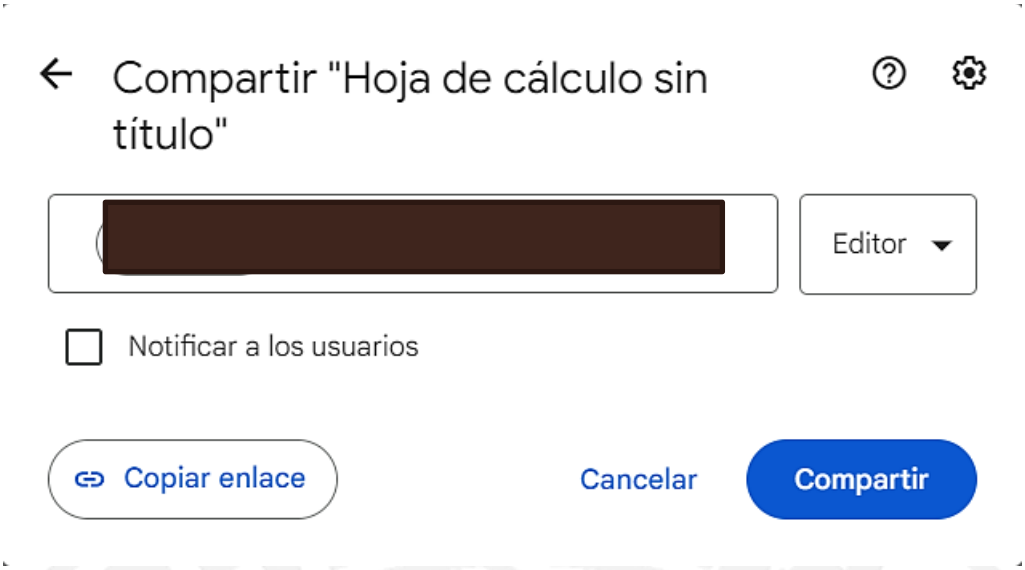


Figura 50. Compartiendo un documento con una cuenta de *Gmail* desconocida

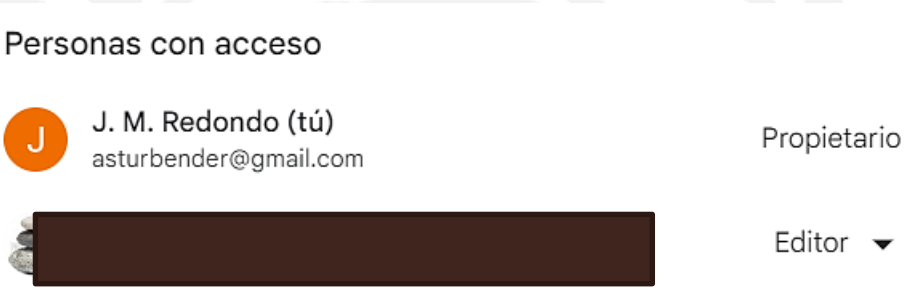


Figura 51. Nombre del propietario de la cuenta en la lista de usuarios con acceso al documento

Un truco parecido, aunque solo funciona si el receptor tiene la configuración de manera que lo muestre, es hacer una transferencia por una cantidad muy pequeña (lo mínimo que *PayPal* permita) a **una cuenta de *PayPal* que pueda estar asociada a ese nombre**. En algunos casos puede revelar el nombre de la persona detrás de una dirección de correo si la tiene registrada en *PayPal* a la hora de crear su cuenta. Más información: https://www.quora.com/Does-PayPal-show-your-full-name-to-the-sender

Redes sociales

Como dijimos, **cada foto que uno ponga de alguien lo convierte en un posible objetivo a estudiar**. Pero no solo es eso. Piensa además que si antes aparecías mucho con una persona y ahora ya no lo haces, puedes estar revelando un **alejamiento / ruptura** con dicha persona, y que por tanto puede estar molesto contigo. Lo mismo si de repente borras fotos tuyas con esa persona, o de la persona en sí, le dejas de seguir, o comentas cosas indebidas acerca de ella. Hay formas de que esta persona pueda ser engañada para que revele información importante sobre ti, y es más fácil si está enfadada o molesta contigo (obviamente, lo mismo puede pasar con cualquiera de tus amigos...). Es un poco lo mismo que pasa si se localiza a un **exempleado** de una empresa (por ejemplo, usando ***LinkedIn***): es posible que sea más propenso a revelar secretos de ésta, que luego se necesiten para otro tipo de ataques.

Figura 52. Una persona y sus amigos / objetivos de estudio. Fuente: https://ciso2ciso.com/introduction-to-doxing-osint-methods-for-information-gathering-by-hadess/

Finalmente, nunca desprecies la posibilidad de que una misma persona tenga una red social más cerrada que otra, y que mientras en una no tenga apenas información relevante, en otra sea mucho más "hablador". Para localizar estas redes sociales alternativas, muchas veces simplemente hay que mirar la cabecera de la red de la persona, ya que da esa información directamente, o usar herramientas como ***Project Sherlock*** (**Seminario 1**), que localiza cuando **una persona usa el mismo identificador en distintas redes sociales**, algo que es muy común. Yo mismo he "destapado" la hipocresía de un par de personas que negaban haber hecho algo en una red mientras alardeaban de haberlo hecho en otra, pensando que no era posible establecer una relación entre ellas (algo muy frecuente en personas de perfil no técnico).

¿Como estudio manualmente una red social?

El mayor problema de una red social no es que estudiarla sea en sí difícil, es hacerse una idea de la información que contiene y como extraerla correctamente, ya que hay una **gran cantidad de información** y **hacerlo de forma organizada puede ser realmente muy complejo**. Por suerte el repositorio ***OSINT Dojo*** (https://github.com/sinwindie/OSINT) tiene una serie de esquemas llamados "*attack surface*" que son excelentes para hacer esta clase de investigaciones. Si bien **el repositorio contiene muchas más**, reproducimos aquí los esquemas de las redes que hemos considerado más importantes de cara a hacer una investigación OSINT para que los podáis usar más fácilmente.

Discord

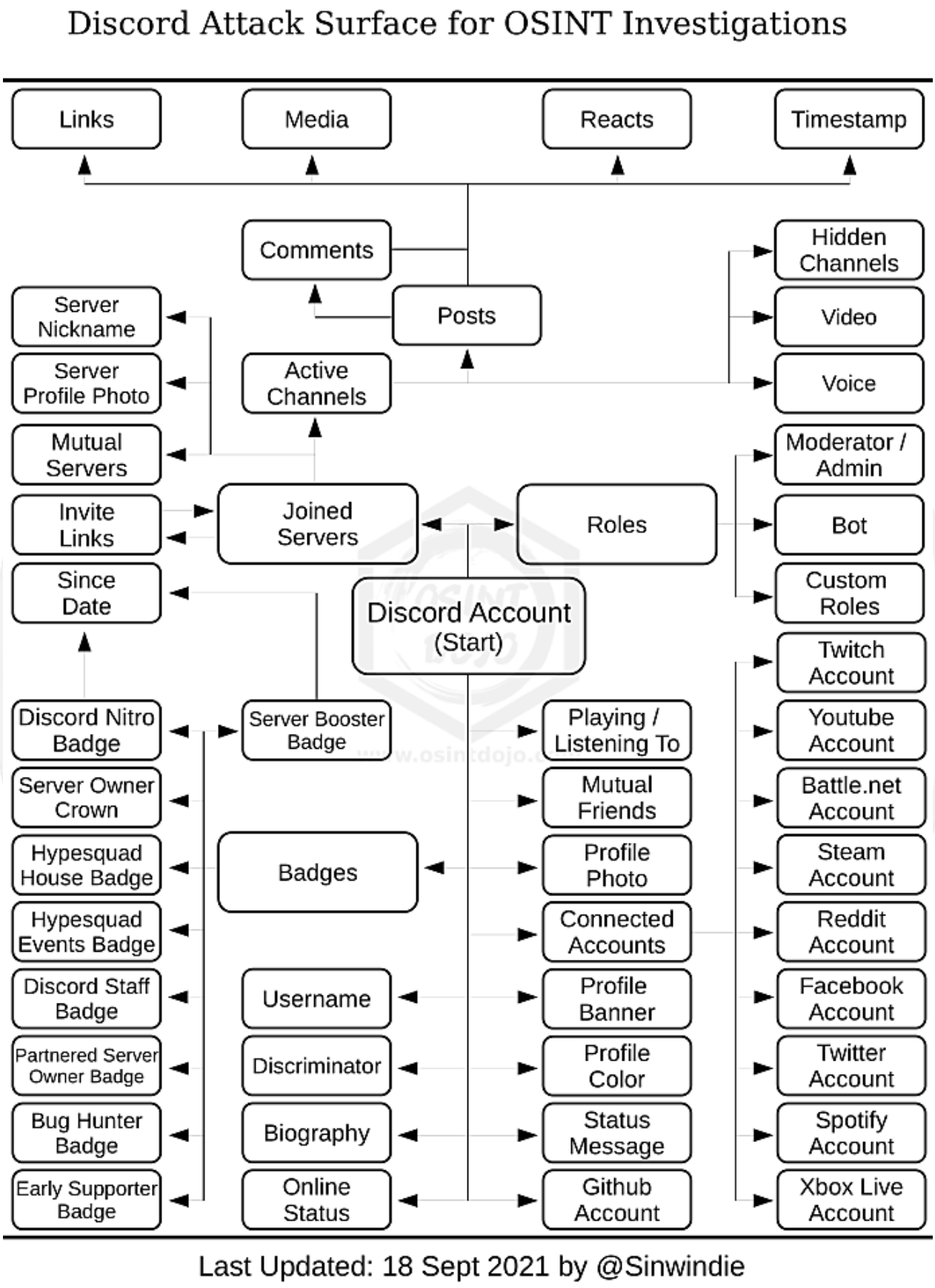


Figura 53. Superficie de ataque de *Discord*

YouTube

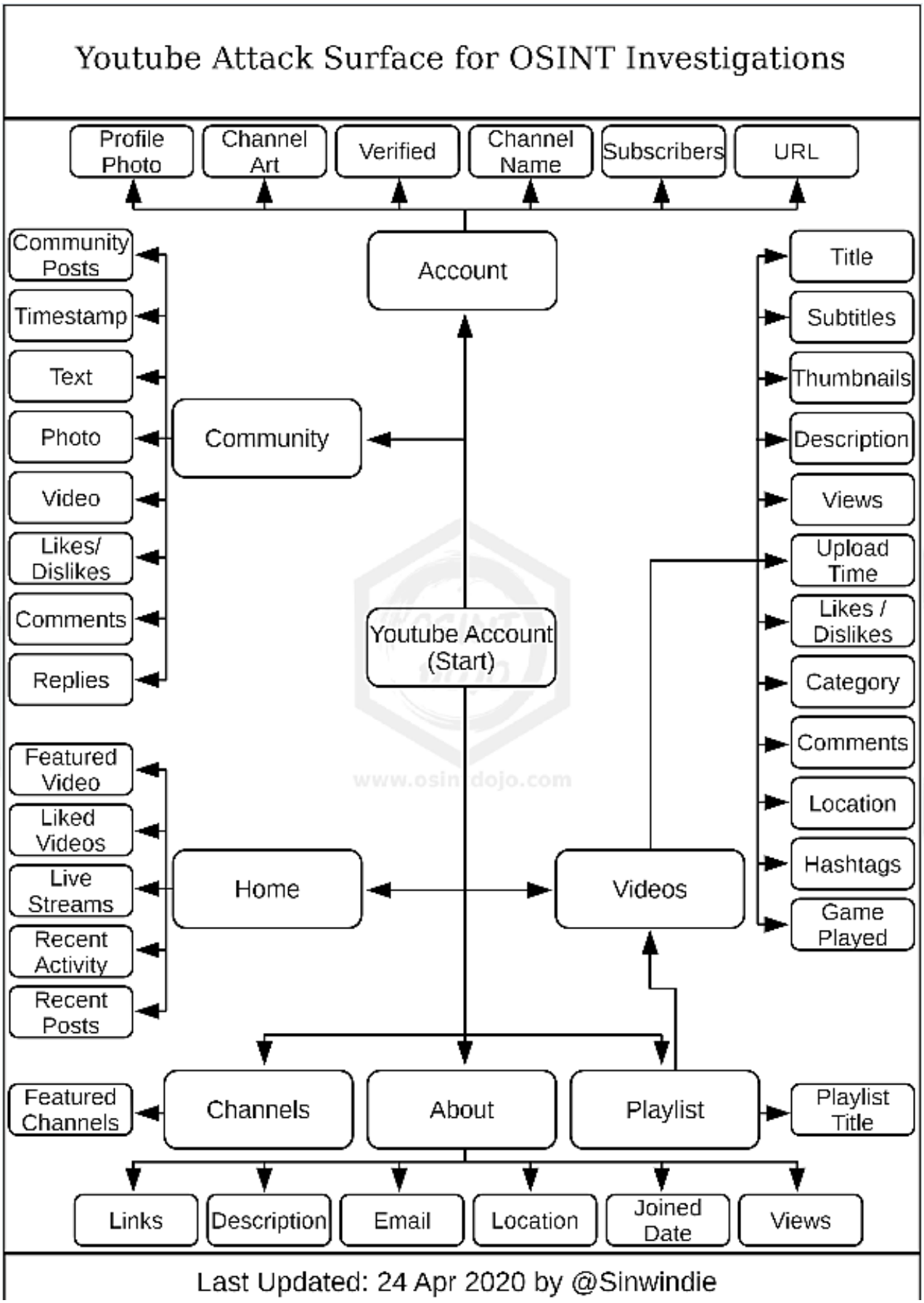


Figura 54. Superficie de ataque de *YouTube*

Twitter (¿Alguien lo llama X realmente? 😊)

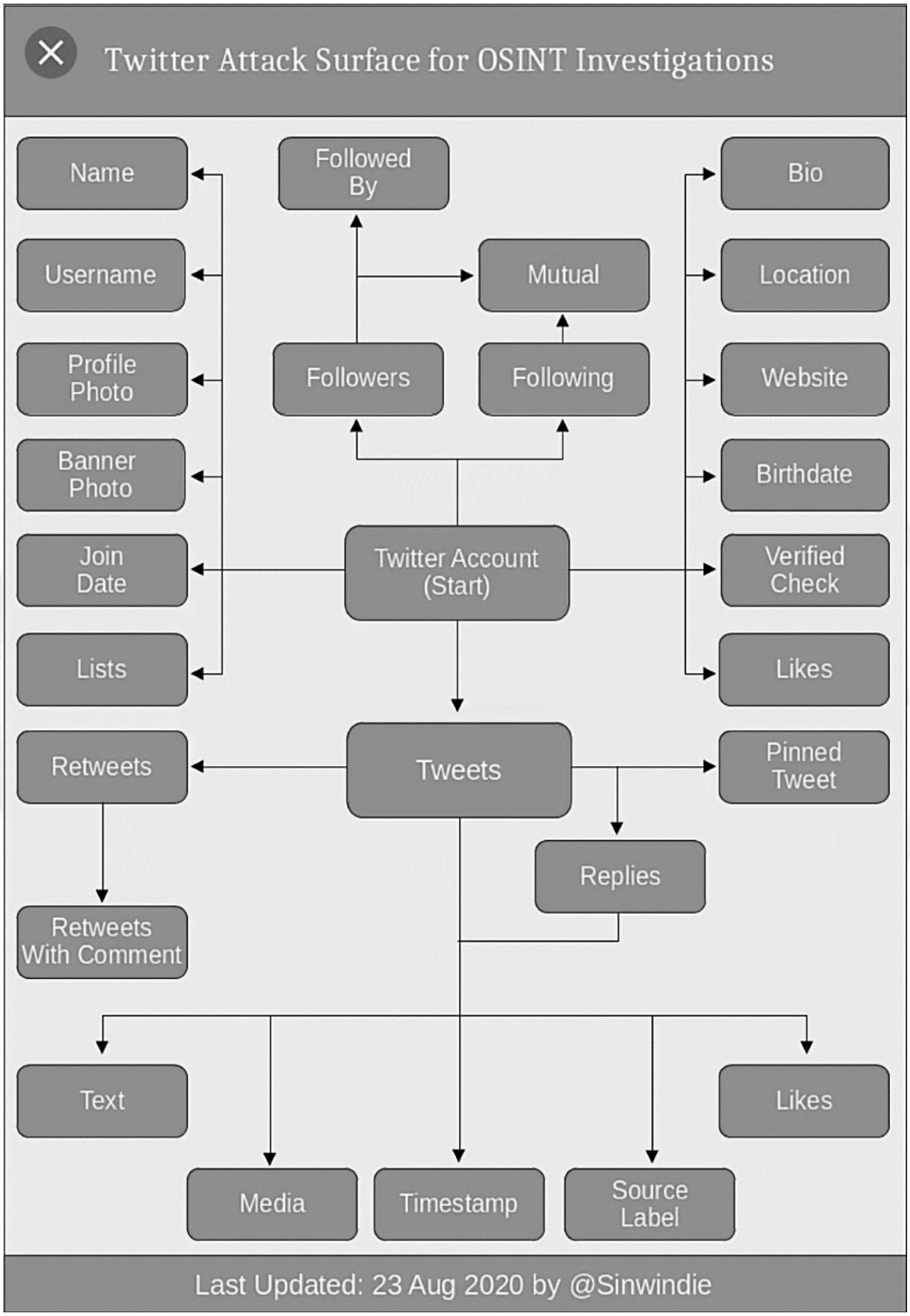


Figura 55. Superficie de ataque de *Twitter*

Instagram

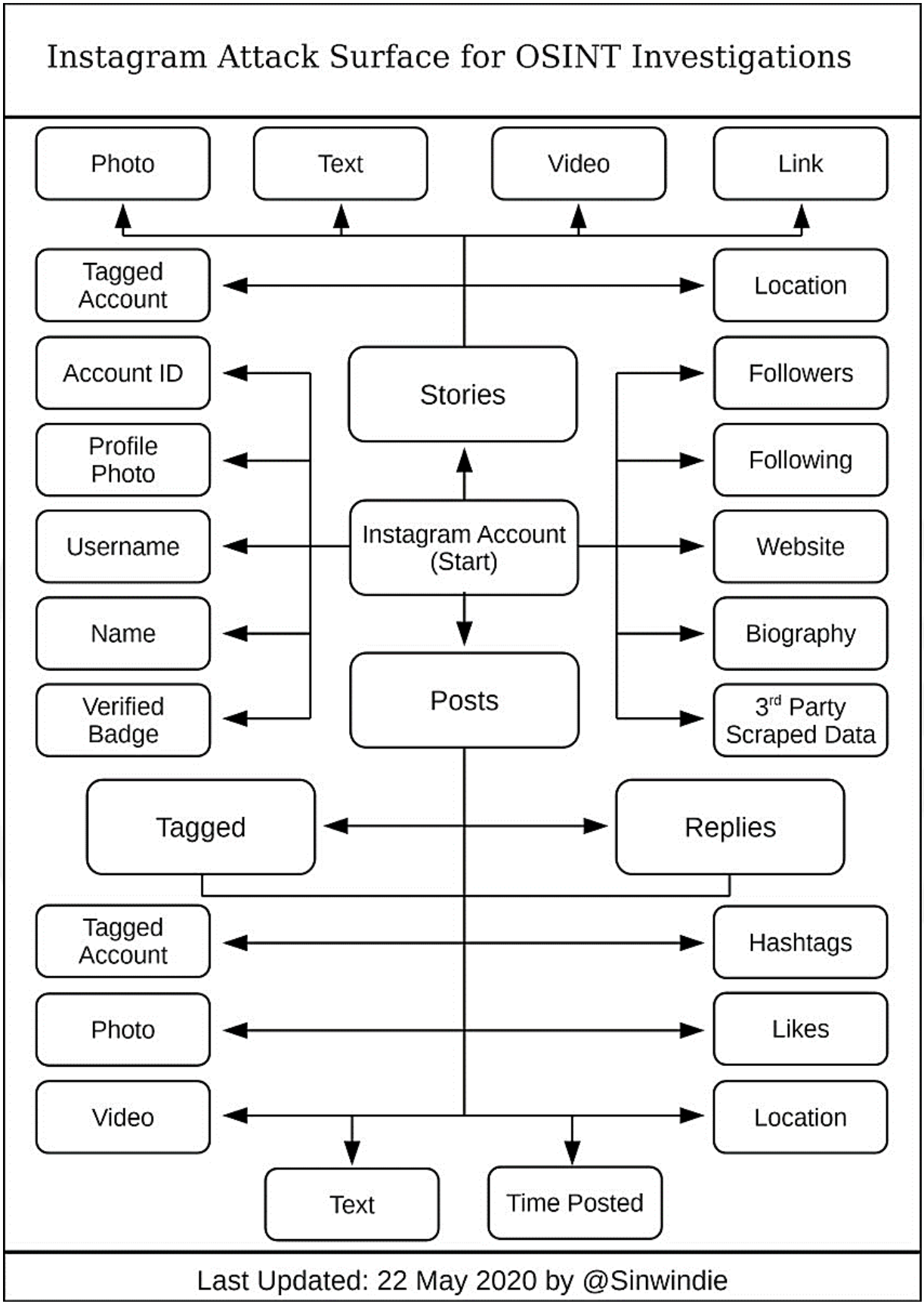


Figura 56. Superficie de ataque de *Instagram*

LinkedIn

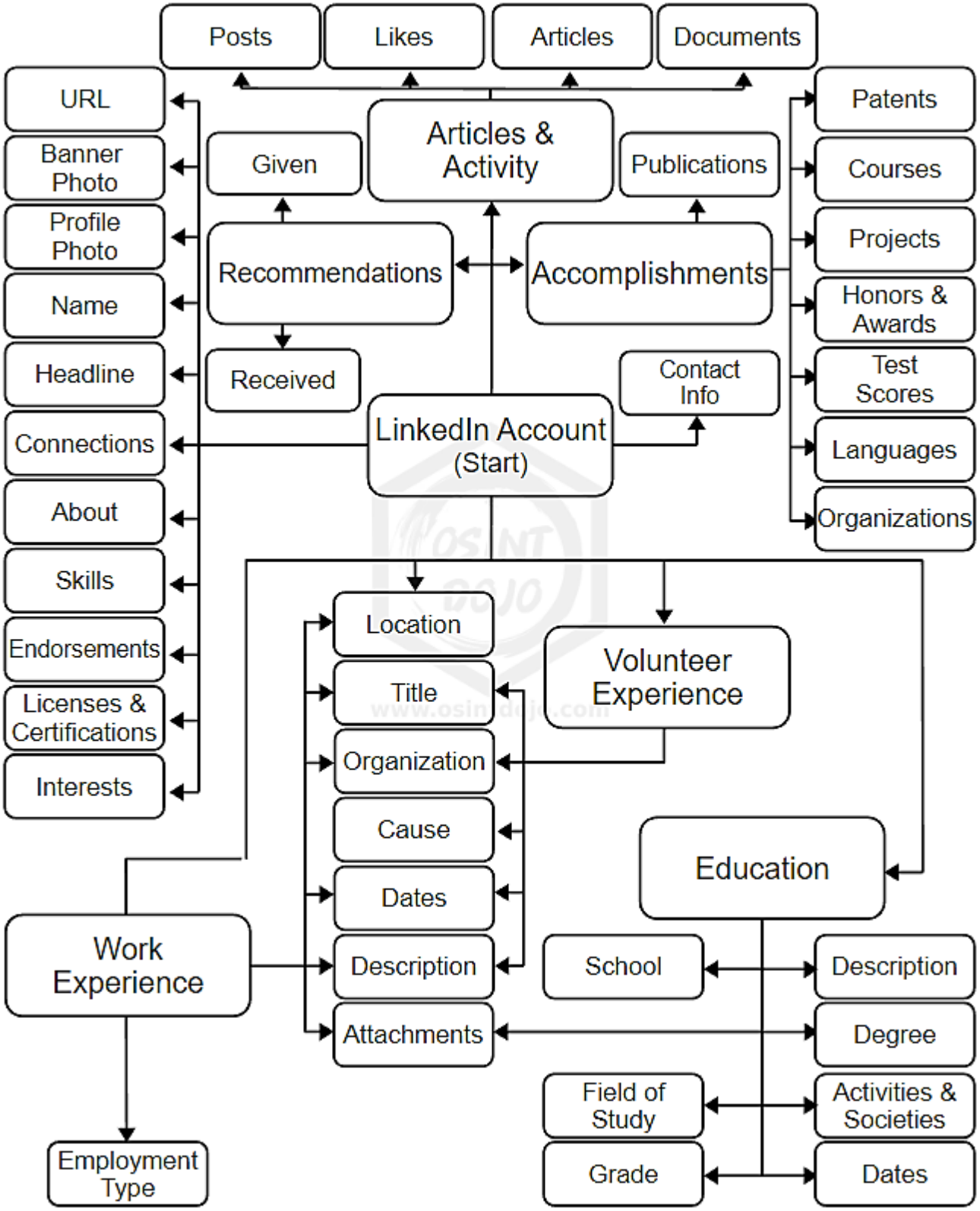


Figura 57. Superficie de ataque de *LinkedIn*

Snapchat

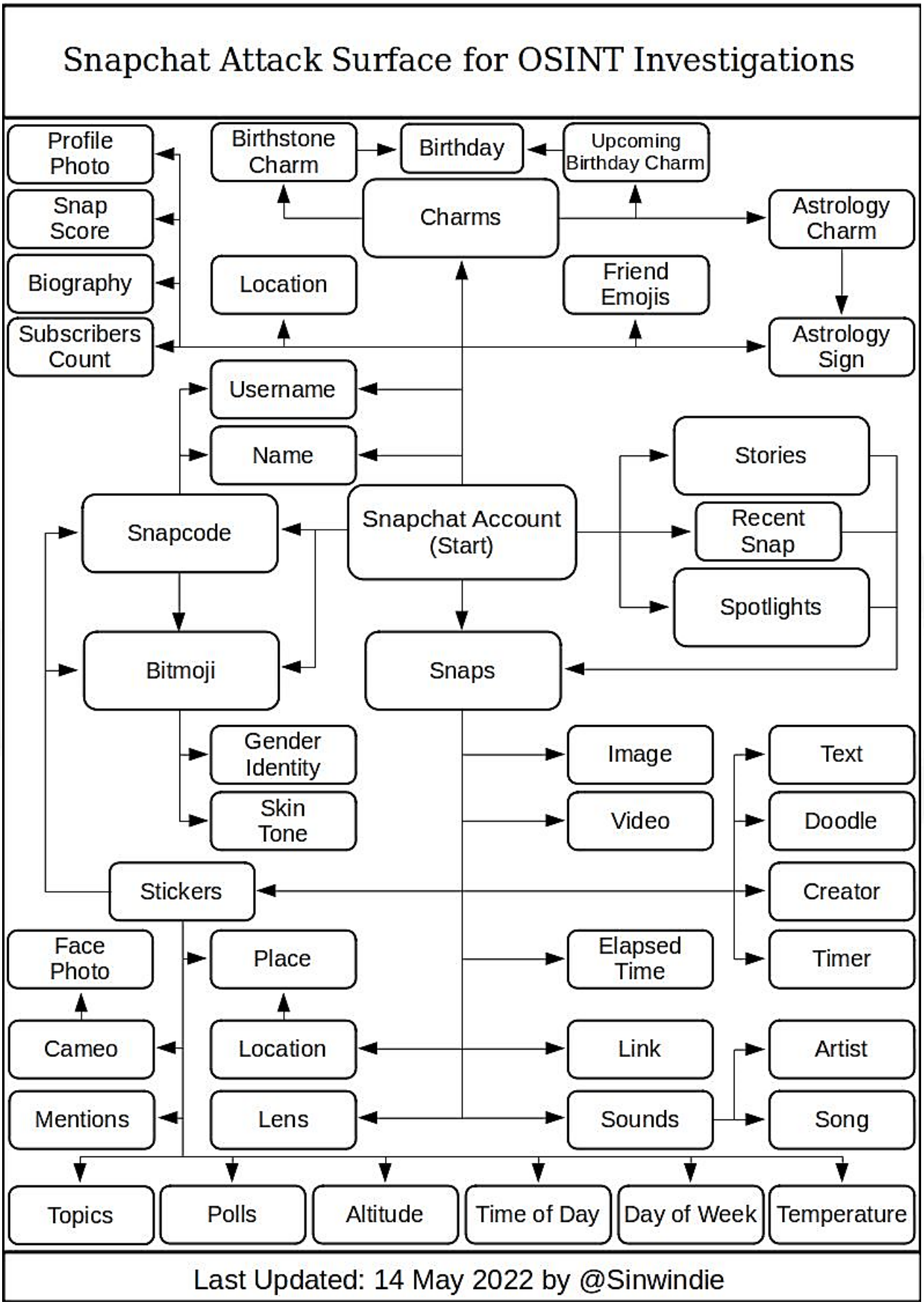


Figura 58. Superficie de ataque de *Snapchat*

TikTok

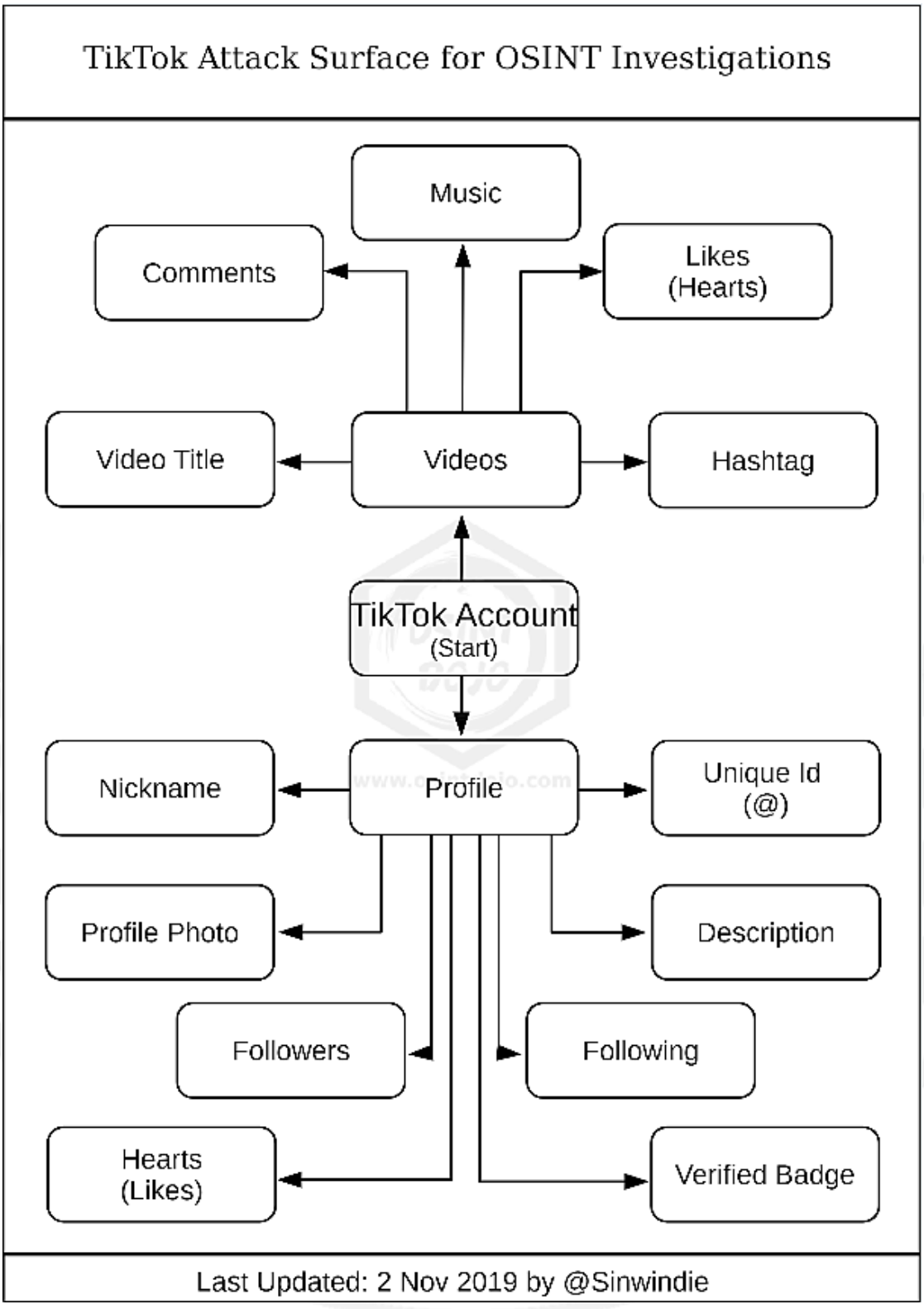


Figura 59. Superficie de ataque de TikTok

GitHub

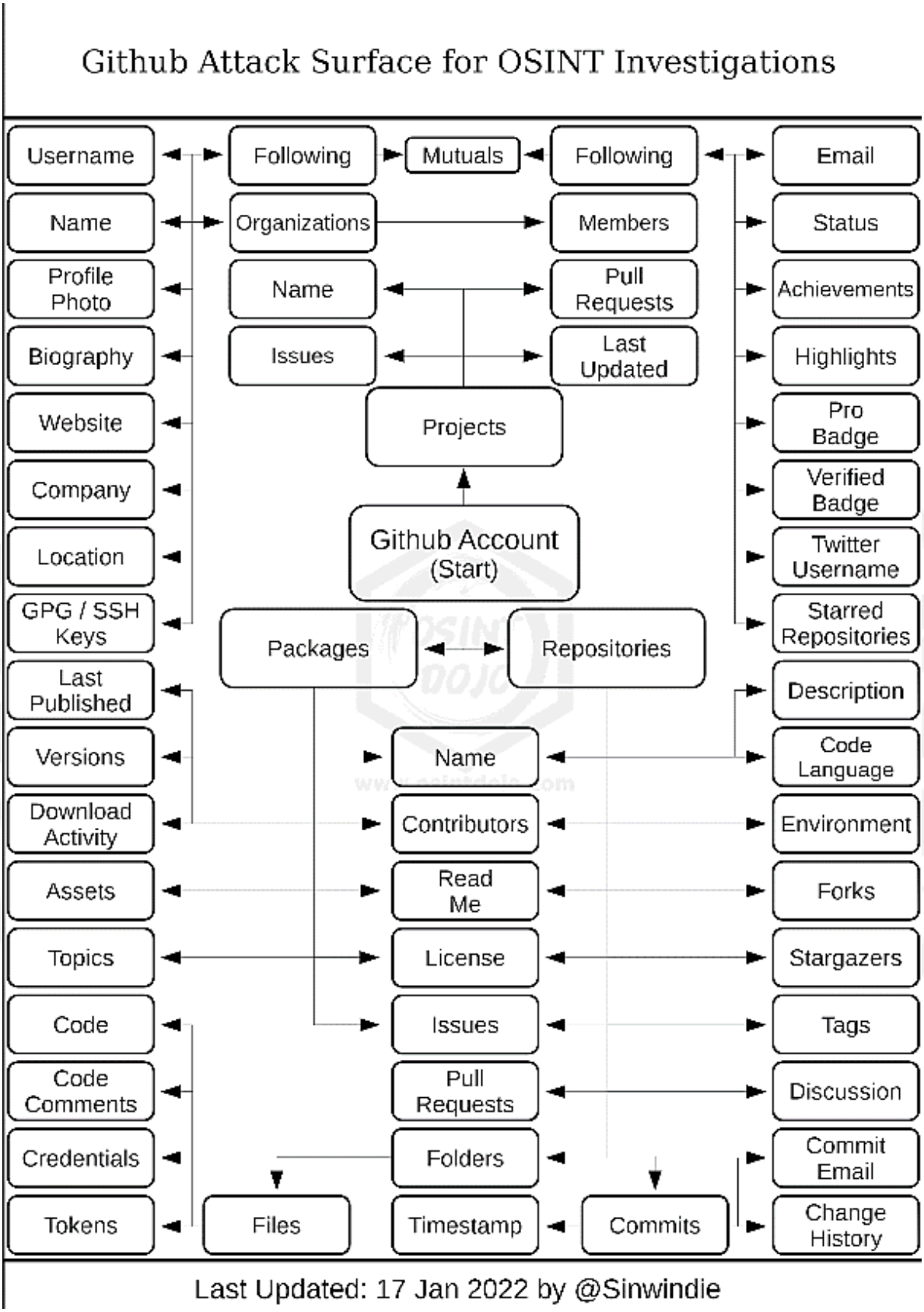


Figura 60. Superficie de ataque de *GitHub*

¿Y si quiero estudiarlas con ayuda de herramientas?

En ese caso vas a tener una ventaja y una desventaja. La ventaja es que habitualmente hay muchas herramientas que trabajan sobre diversas redes sociales en un momento dado, y muchas de ellas están

especializadas en obtener información concreta de cada una de ellas (imágenes, contactos, relaciones...). El problema de estas herramientas se puede resumir en una serie de puntos:

- **Están mantenidas por personas anónimas,** y debes fiarte de que sus intenciones **no son maliciosas** (es decir, que por usar una herramienta no te instalar un *malware*). Hay que recurrir a un entorno de ejecución aislado, mirar reputación del autor y la herramienta, etc.
- **A las redes sociales habitualmente no les gustan esta clase de herramientas**, por lo que podrían dejar de funcionar sin previo aviso. El caso más representativo es *Twitter* y el cierre de su API para convertirla en una de pago.
- En consonancia con lo anterior, **los autores pueden dejar de mantenerlas en cualquier momento** (al fin y al cabo, no les pagan por ello), y una herramienta que una vez te daba una funcionalidad puede ser inútil en poco tiempo...

Si eres consciente de ello, herramientas puedes encontrar muchas a poco que busques. No solo en el repositorio de ***Ciberpatrulla*** mencionado antes, sino también en ***Social Media OSINT***, especializado en herramientas de esta clase: https://github.com/cqcore/Social-Media-OSINT

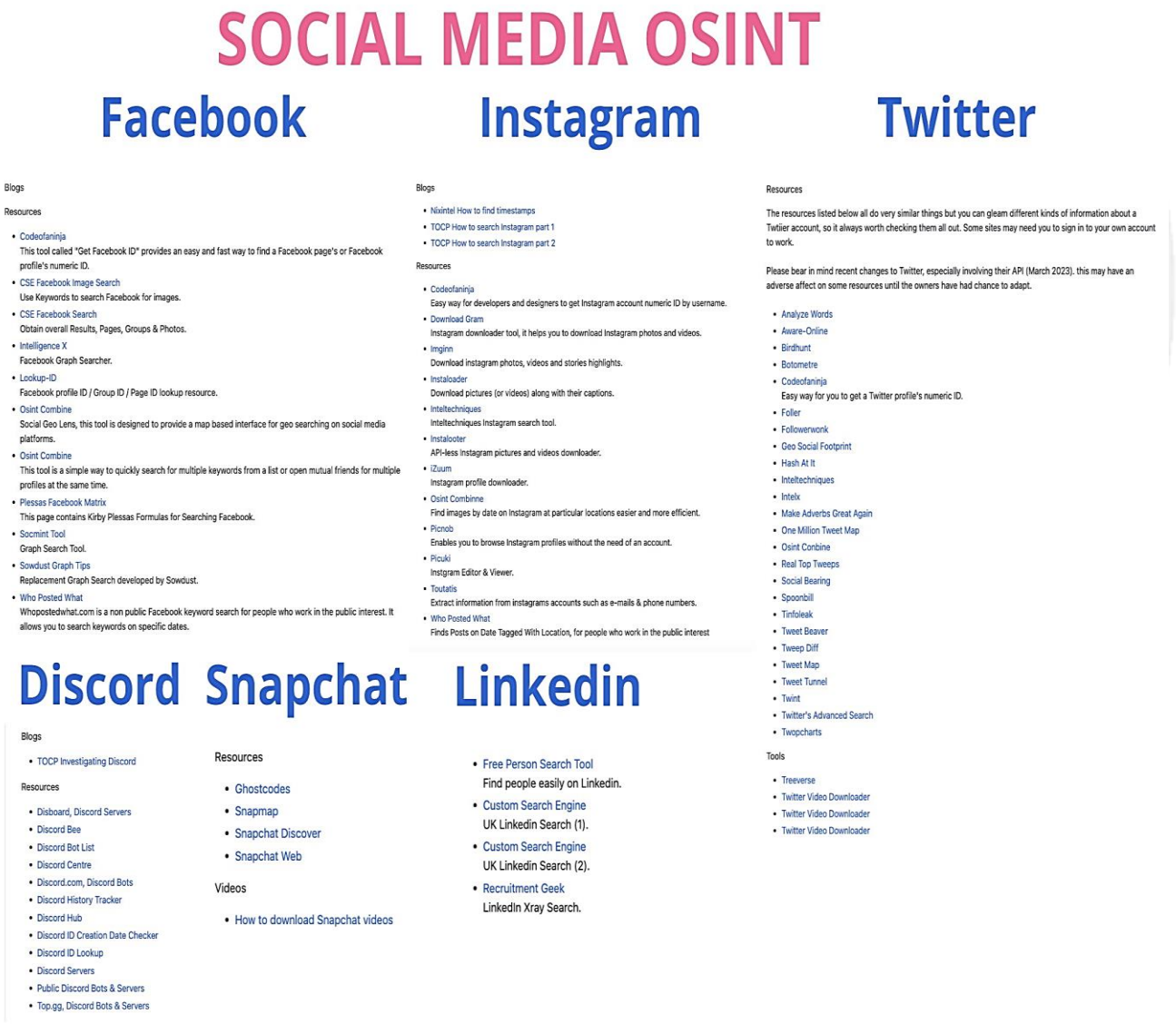


Figura 61. Resumen del contenido del repositorio Social Media OSINT. Fuente: https://twitter.com/cyb_detective/status/1647588776762388480

Sobre las "Cuenta B" en redes sociales

Hay personas que usan cuentas alternativas a la suya principal en redes sociales. Muchas lo hacen para mantener el anonimato, de manera que en la "Cuenta B" pueda **decir u opinar cosas que, de hacerlo en su cuenta oficial, podría traerle problemas ante ciertos colectivos**, perder reputación o cualquier otro efecto. O simplemente desea ejercer su derecho a ser anónimo en Internet. Eso es un tema aparte del que tratamos en este documento, que es saber cuándo una cuenta es una "Cuenta B" de otra porque de esta forma sabremos la identidad de la persona que está detrás de ella.

Aunque esto no es una ciencia exacta, hay varios indicios que delatan a una "Cuenta B" de esta clase y que puedes usar para ver si estas ante una o no:

- La cuenta A tiene muchos seguidores pero sigue a poca gente. Entre ellos, está la "Cuenta B" candidata que quieres investigar.
- La "Cuenta B" siempre repostea a la "Cuenta A" ponga lo que ponga, y realmente tiene pocas interacciones propias.
- La "Cuenta B" siempre defiende todas las posturas de la A, incluso de forma muy vehemente.
- La "Cuenta B" se dedica a insultar y faltar al respeto a todo el mundo que tenga una opinión en contra de la A, o rebata cualquiera de sus argumentos.

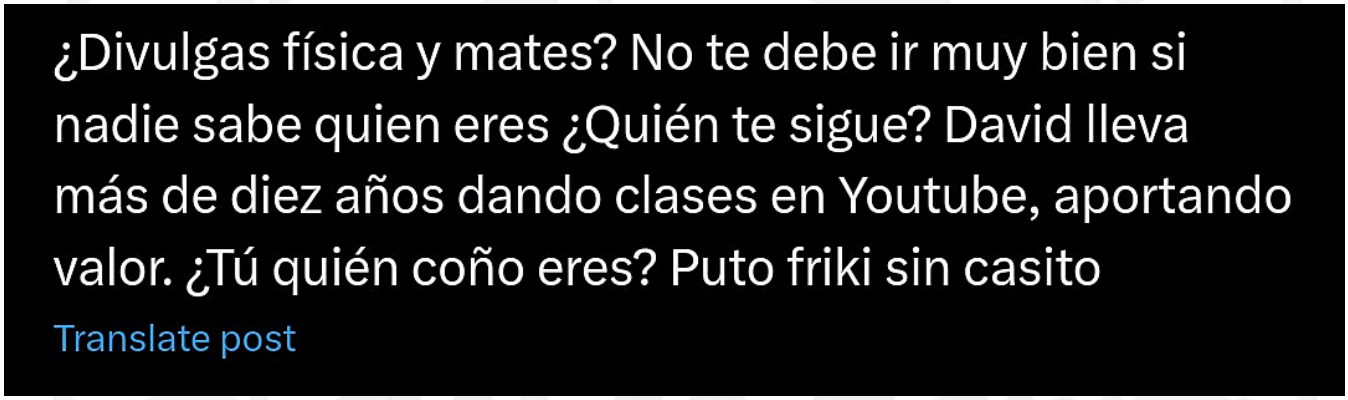


Figura 62. Este post es supuestamente de la cuenta B de un conocido divulgador

Redes de búsqueda de personas

Hemos hablado de redes sociales, que también son una red de búsqueda de personas en general. También con los motores de búsqueda "típicos" se puede conseguir el mismo efecto. Si eso no es suficiente, hay **motores de búsqueda de personas** que se dedican expresamente a eso, y que podemos usar como complemento.

Puedes encontrar mención a algunos de ellos en este artículo, aunque solo permiten hacer búsquedas a gente de EEUU, y probablemente tu objetivo viva en otro país: https://www.lifewire.com/search-engines-that-top-the-web-3482269. Un ejemplo de motor de búsqueda precisamente **para gente en EEUU** (ya te estarás dando cuenta que la privacidad en EEUU no se cuida tanto como en Europa) es ***Intelius***: https://www.intelius.com/.

Si una empresa necesita hacer algo como esto en plan serio debe plantearse opciones profesionales de pago, como ***Pipl***: https://pipl.com/product/pipl-search. Como no es nuestro caso, en la asignatura podemos usar la funcionalidad que tiene para ello ***Facebook*** o (mejor) ***LinkedIn***, como se muestra en el artículo referenciado.

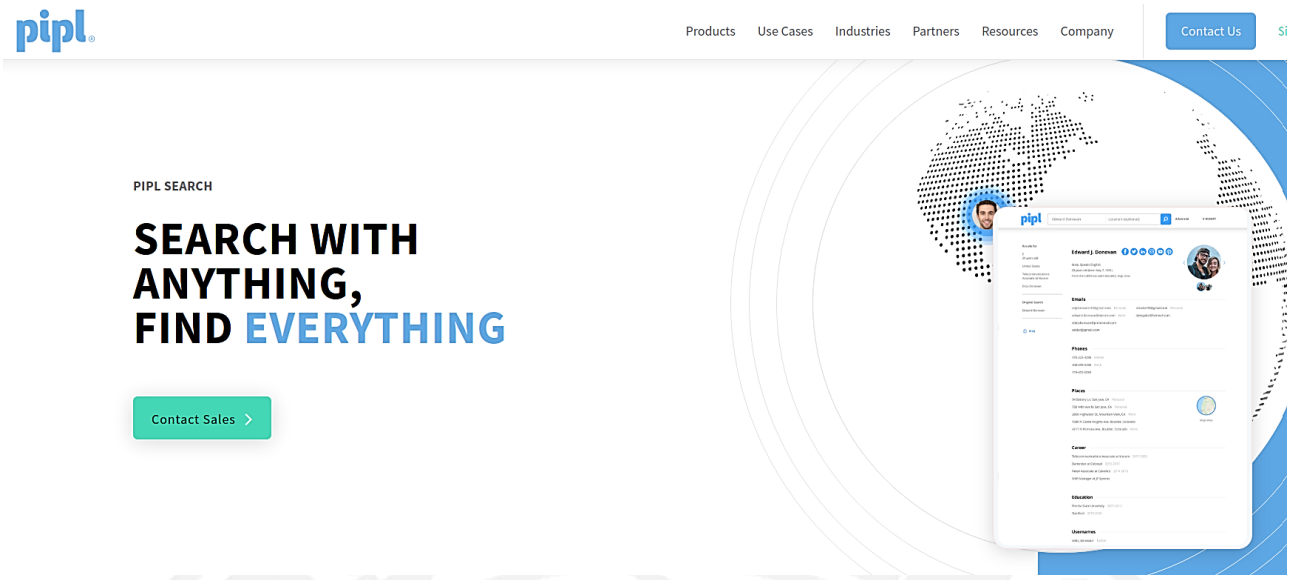


Figura 63. Motor y API de búsqueda de personas *Pipl*

Si lo de usar motores de búsqueda web no es lo tuyo, también hay **herramientas instalables en tu PC que hacen ese tipo de trabajo**. Aquí tienes un par de ellas:

- ***Mr. Holmes*** (https://github.com/Lucksi/Mr.Holmes): Una aplicación que busca información de dominios, nombre de usuario y nºs de teléfono con varias técnicas.
- ***Alfred*** (https://github.com/Alfredredbird/alfred): Una herramienta para buscar información acerca de nombres de usuario por Internet.

Aplicaciones de citas

Aunque esto te pueda parecer una broma, no lo es en absoluto. Una vez que tienes información básica de una persona, la probabilidad de que puedas encontrarla en una aplicación de citas estilo ***Tinder***, ***Grindr***, etc. no es cero. Si además la persona que investiga se crea una identidad falsa para darse de alta en esta red, como vimos en una sección anterior, puede ser, en manos de alguien especialmente entrenado en como hablar con otras personas para intentar sonsacarles información, una vía muy potente **para enterarse de datos o secretos que hagan otros ataques posibles**. Obviamente, esto es una investigación que requiere "*engagement*" con el objetivo, por lo que **está fuera del ámbito de esta asignatura**, pero nuevamente se comenta no para lo que hagáis, sino para que lo tengáis en cuenta **por si algún día sois vosotros los que estáis en el lado del investigado**, y no del investigador. Hacer legalmente este tipo de investigaciones es trabajo de las fuerzas del orden.

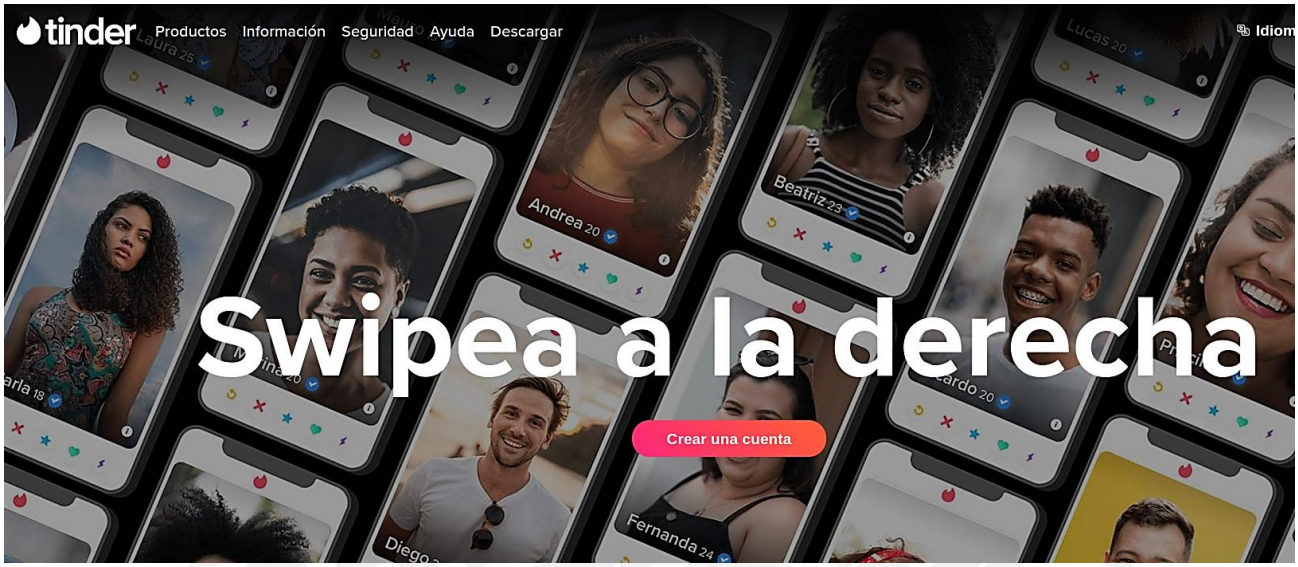


Figura 64. Con una foto, un nombre y una localización geográfica aproximada de un objetivo, sitios como *Tinder* pueden ser un punto del que sonsacar más información a los investigados

Números de teléfono

En el año 2012 moría una de las mayores filtraciones legales de datos de personas: **las páginas blancas**. Un enorme libro que llegaba a muchas casas gratuitamente y que listaba todos los teléfonos de España, ordenados por apellidos y mostrando la dirección de cada persona. Obviamente, de seguir existiendo facilitaría mucho el trabajo de investigación y la obtención de datos para hacer ataques de *spear phishing* o similares. Lo que no quiere decir que a lo mejor en tu casa quede alguna por ahí olvidada y resulte que localices el teléfono (y la dirección) de alguien que no lo haya cambiado en estos años...

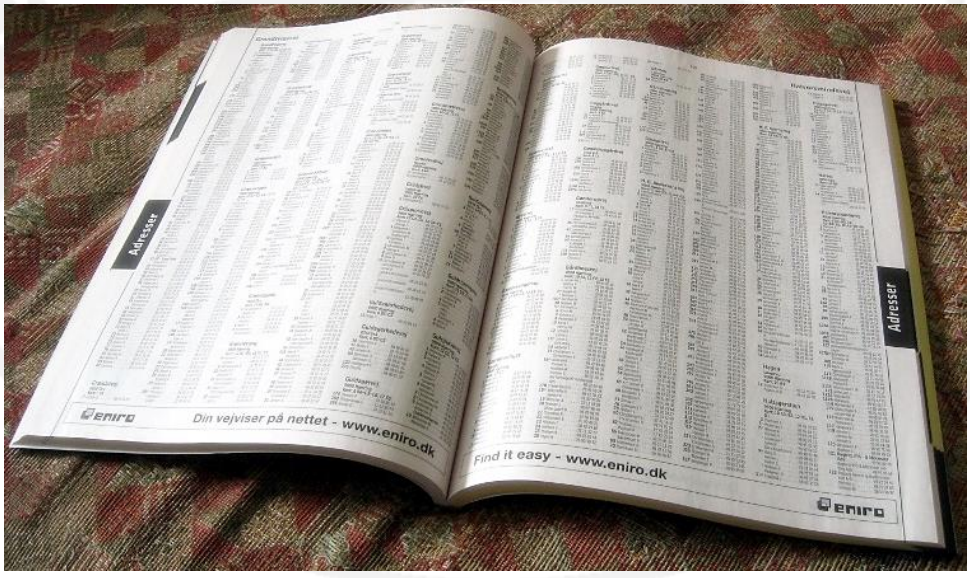


Figura 65. Una guía de Paginas Blancas física, que es como imprimir un robo de datos moderno. Los años 2000 fueron salvajes en más de un sentido...Fuente: *Wikipedia*

Esto es el pasado, lo mismo que las **páginas amarillas** (un directorio telefónico nacional de **empresas y negocios** solamente), que dejaron de imprimirse también 9 años después que las blancas. Lo que pasa en este caso es que sí que hay una versión online disponible aquí: https://www.paginasamarillas.es/ , lo cual posibilita **encontrar más información de negocios**, si son nuestros objetivos de estudio.

Figura 66. Páginas amarillas buscando información de negocios

Realmente si estamos buscando el de alguien de España o cualquier país de la UE solo podemos confiar en encontrar su teléfono filtrado en Internet por accidente (o no, ofrecido por la persona directamente), ya que **las búsquedas telefónicas**, ya sean directas (de persona a teléfono) como inversa (de teléfono a persona) **están prohibidas por la ley de protección de datos**. Solo es posible usar webs que recojan nºs de teléfono que sean identificados y valorados por particulares, lo que normalmente se hace cuando corresponden a negocios o teléfonos usados para prácticas agresivas o no convenientes de telemarketing. Una de las páginas más conocidas para eso es esta: https://www.listaspam.com/

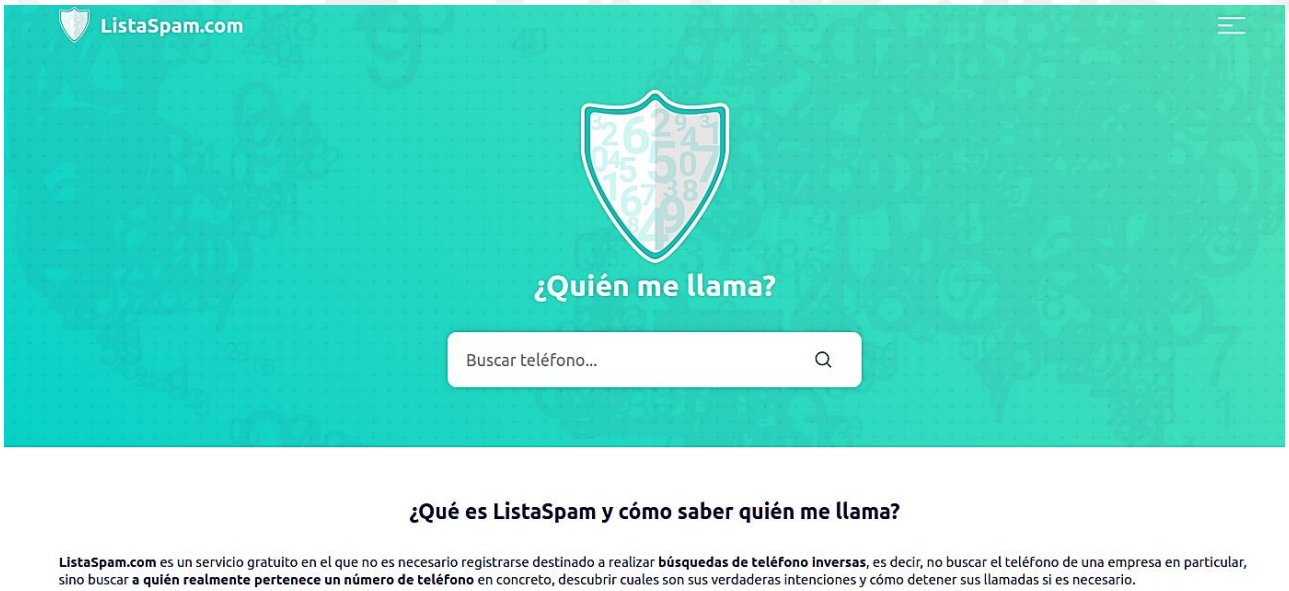


Figura 67. *ListaSpam* y similares al menos te permite saber si un objetivo está haciendo prácticas comerciales agresivas o consideradas molestas

También es cierto que algunas empresas u organizaciones (privadas o públicas) podrían mantener un **directorio de sus empleados** que puede dar más información sobre los mismos, y que merece la pena

consultar. Es por ejemplo el caso de la Universidad de Oviedo, que permite **búsquedas anónimas** (limitadas): https://sies.uniovi.es/paginasblancas/web/busqueda-anonima/personal.faces

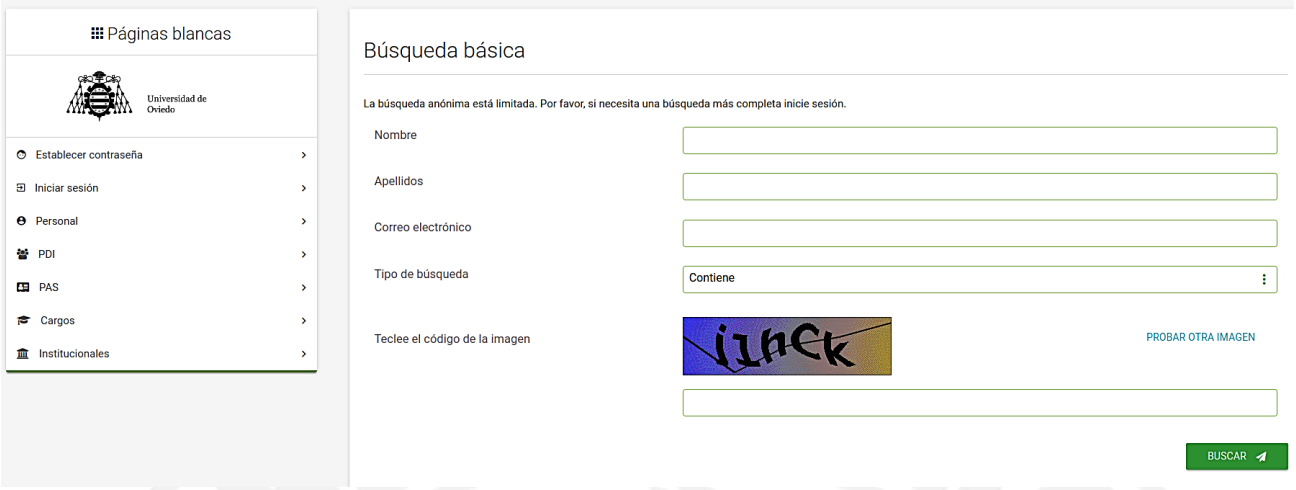


Figura 68. Páginas blancas de empleados de la Universidad de Oviedo

Cabe decir que **estas restricciones no se aplican en otros países fuera de la UE**, como por ejemplo EEUU, por lo que si la persona que estás buscando es de allí, puedes hacer una búsqueda inversa con servicios como por ejemplo ***Spokeo***: https://www.spokeo.com/

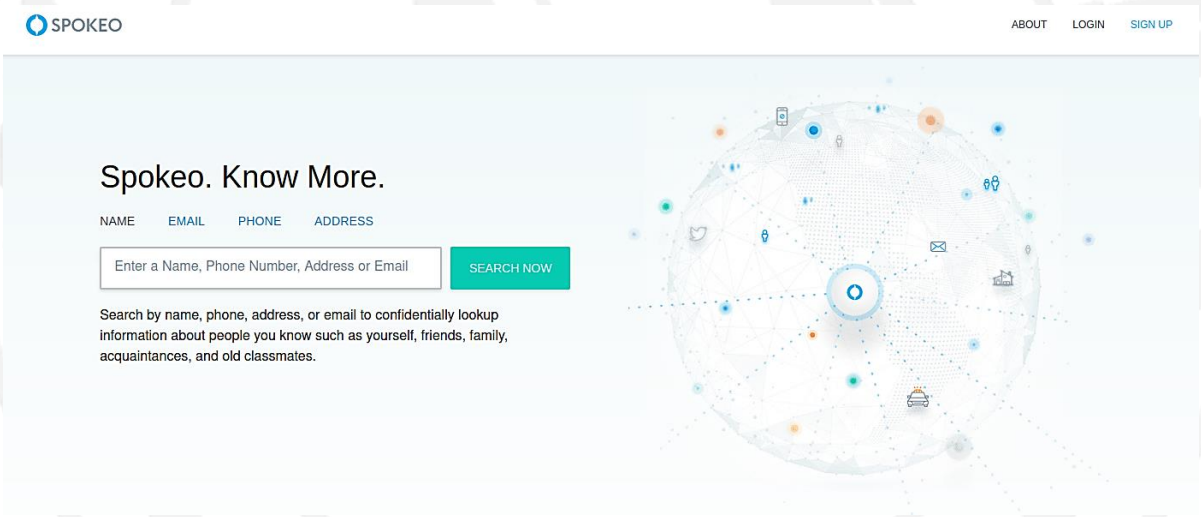


Figura 69. En EEUU el tema de la privacidad de los nº de teléfono es otro rollo. Incluso hay buscadores de antecedentes penales también....

Número de teléfono parcial a partir de PayPal

A veces algunos servicios nos muestran parte de un nº de teléfono del propietario de una cuenta cuando hacemos determinadas operaciones, **como recordar contraseñas o similares**. En estos momentos es lo que pasa con *PayPal*: es posible saber **5 dígitos del número de teléfono** del propietario de una cuenta cualquiera (y su prefijo) o su dirección de correo. A veces incluso parte de su tarjeta de crédito y el tipo de tarjeta. Esta información puede usarse para facilitar búsquedas que localicen datos del objetivo o, en combinación con otros dígitos que revelen otros servicios ante una operación similar, incluso tener todo o gran parte de algo tan privado como tu teléfono o la tarjeta. **Conviene no obstante hacer una prueba con nuestra propio cuenta** del servicio, por si el propietario recibiese un aviso de que se están haciendo estas pruebas (en el caso de *PayPal*, en Marzo de 2023 no lo hacía). Ahora imagina que es alguien que conoces

y que tienes en tus contactos telefónicos. Comprobar estos nºs con tus contactos reduce mucho la cantidad de objetivos.

Figura 70. *PayPal* revelando mucha información de un propietario de una cuenta

Puede que esta información incompleta creas que no vale para mucho, y lo cierto es que usarla de forma aislada efectivamente no lo vale. Pero ahora imagina que esa persona usa *PayPal* para algo que te perjudica y sospechas que es un conocido: ¡con estos datos es muy fácil que solo haya una persona en tu lista de contactos que encaje con ellos!

Otras formas de obtener el propietario de un nº de teléfono

Si tenemos un nº de teléfono pero no tenemos ni idea de quien puede estar detrás, hay dos cosas más que podemos hacer para ver si nos ofrecen más datos de este:

- **Mandar un nuevo mensaje al teléfono por *Telegram*** añadiendo en el proceso un nuevo contacto (Ej.: https://www.makeuseof.com/how-to-add-a-new-contact-on-telegram/). Si esa persona tiene un perfil medianamente público, podremos saber cosas de el o ella, como una foto, nombre real y nombre de usuario. Con ese nombre de usuario si hace falta podremos buscar por distintas redes sociales si hay alguien que usa el mismo y, aunque no es 100% seguro, si que es muy probable que con eso saquemos más información de la persona que estamos buscando.
- **Iniciar el proceso de mandar un *Bizum*** por una cantidad ridícula al nº de teléfono (por ejemplo, 0,50 euros) en nuestro banco. Podremos saber al menos parte del nombre del destinatario (depende de cómo lo tenga configurado) **sin necesidad de confirmar el envío de dinero**. Por lo general, antes de realizar la transferencia, te debería aparecer una confirmación con el nombre de la persona destinataria.

En cualquier caso, puedes seguir este esquema general a la hora de encontrar información de nºs de teléfono:

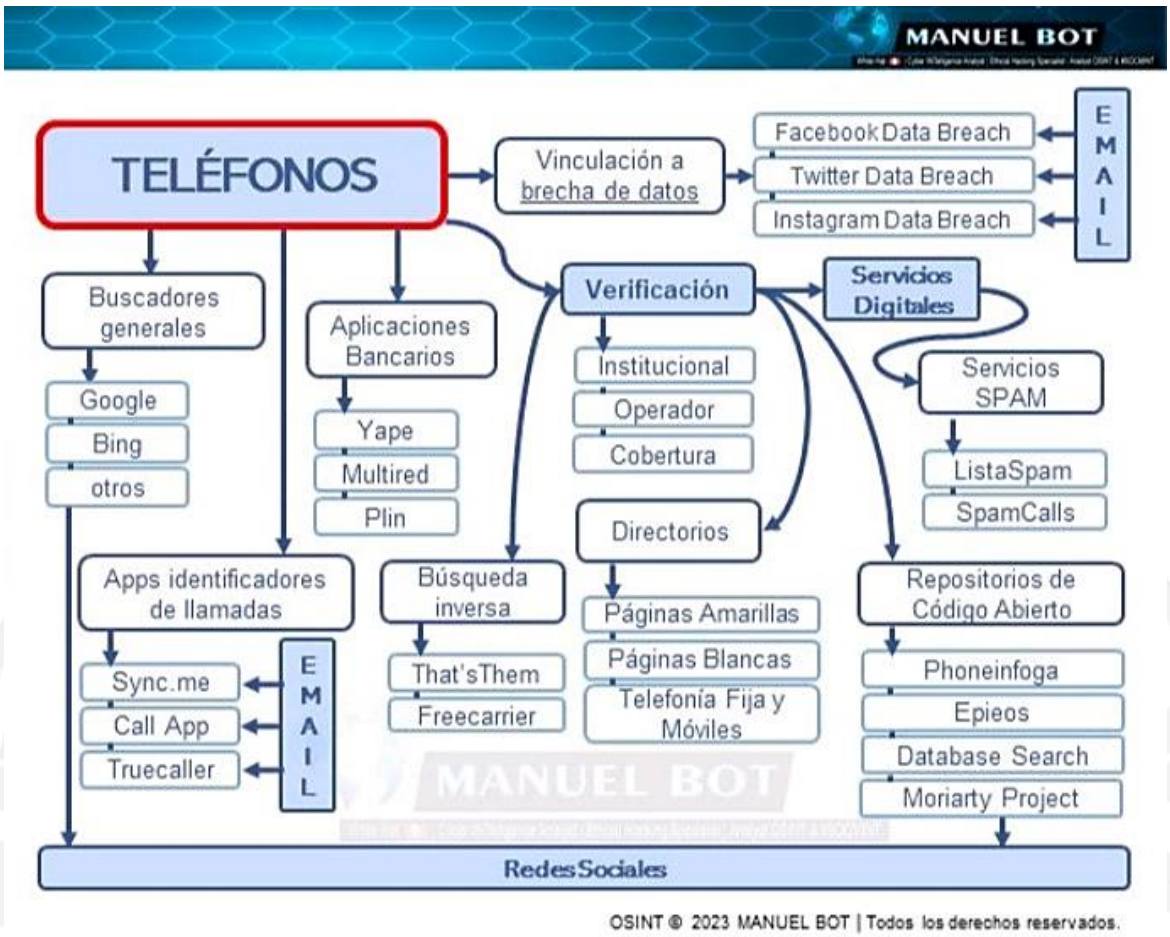


Figura 71. Donde encontrar información de un nº de teléfono, según Manuel Travezaño (https://www.linkedin.com/posts/manuelbot59_osint-workflow-saberespoder-activity-7142187376654643200-ws0-/?originalSubdomain=es)

Registros públicos

Hay varios lugares públicos donde es posible encontrar información variada de personas, aunque cabe destacar que es posible que el acceso a la misma **tenga un coste**, se tenga que **aportar una razón plausible para hacerlo** y/o **se alerte al titular de que se ha hecho esa búsqueda**. Antes de usar estos servicios hay que mirar muy bien las condiciones en las que se da el mismo, ya que muy fácilmente puede caer en la categoría de "*engagement*" **que NO estamos buscando** hacer en esta asignatura. Obviamente, si en el futuro alguno de estos registros públicos se filtrara por un ataque, la información privada de todo un país estaría seriamente comprometida (no se habla mucho de la filtración del SEPE al principio de la pandemia, pero miedo da pensar qué se pudo exfiltrar entonces...).

- **Registro Mercantil** (https://www.registradores.org/el-colegio/registro-mercantil): Trámites para crear una empresa en España, Sociedades Anónimas, empresas con contrataciones y compras del estado, directorio de registros empresariales, impuestos, datos económicos y financieros, datos jurídicos y económicos de las sociedades y personas, quién es su representante de una empresa o sociedad...se puede buscar por nombre o número de registro. También puede obtenerse información básica de forma gratuita de las bases de datos de empresas privadas como ***Axesor***: https://www.axesor.es/

Figura 72. Página web de *Axesor* con información gratuita básica de cualquier empresa de España

- **Registro de la Propiedad** (https://www.registradores.org/el-colegio/registro-de-la-propiedad): Datos públicos de propiedades y bienes inmuebles para conocer quiénes son los propietarios.
- **Registro Civil** (https://www.mjusticia.gob.es/es/ciudadania/estado-civil/registro-civil) para nacimientos, matrimonios, divorcios, defunciones, fecha de nacimiento, etc.
- **Censo Electoral**: Consultar si alguien está o no en el censo implica autenticarse previamente (vamos, que solo te dejan consultar esos datos de ti mismo), por lo que no es viable hacerlo anónimamente para un tercero desde la web habilitada para ello por el ministerio correspondiente: https://sede.ine.gob.es/ss/Satellite?c=SETramite_C&cid=1254734728608&lang=es_ES&p=1254734719739&pagename=SedeElectronica%2FSELayout. No obstante, este tipo de listados **se publica en papel** en los ayuntamientos en épocas electorales habitualmente, y esas copias físicas sí nos pueden servir para extraer información de otras personas sin que nadie se percate de ello.
- **Vehículos** (Consulta por número de placa): Como se dijo anteriormente, es posible, previo pago y autenticación y aportando una razón válida, saber quién es el propietario de un coche, si tienen al día el pago de impuestos, multas y las Infracciones por número de placa o nombre en el Ministerio de Transporte. Puede parecer excesivo, pero es algo necesario si vas a comprar un coche de segunda mano, para asegurarte de que está "libre de cargas", por lo que los datos son accesibles.
- **Catastro** (https://www.sedecatastro.gob.es): Es una forma de obtener información de fincas, mapas, planos, bienes raíces, inmobiliarias, créditos, compras, ventas, prestamos, gravámenes, hipotecas de fincas y condominios. Mediante la solicitud y pago de una **nota simple** en el catastro se puede averiguar quién es el dueño de una propiedad y dónde está ubicada, o incluso ver los planos de una casa. De forma gratuita y vía web en el Catastro se pueden buscar **mapas por satélite**.

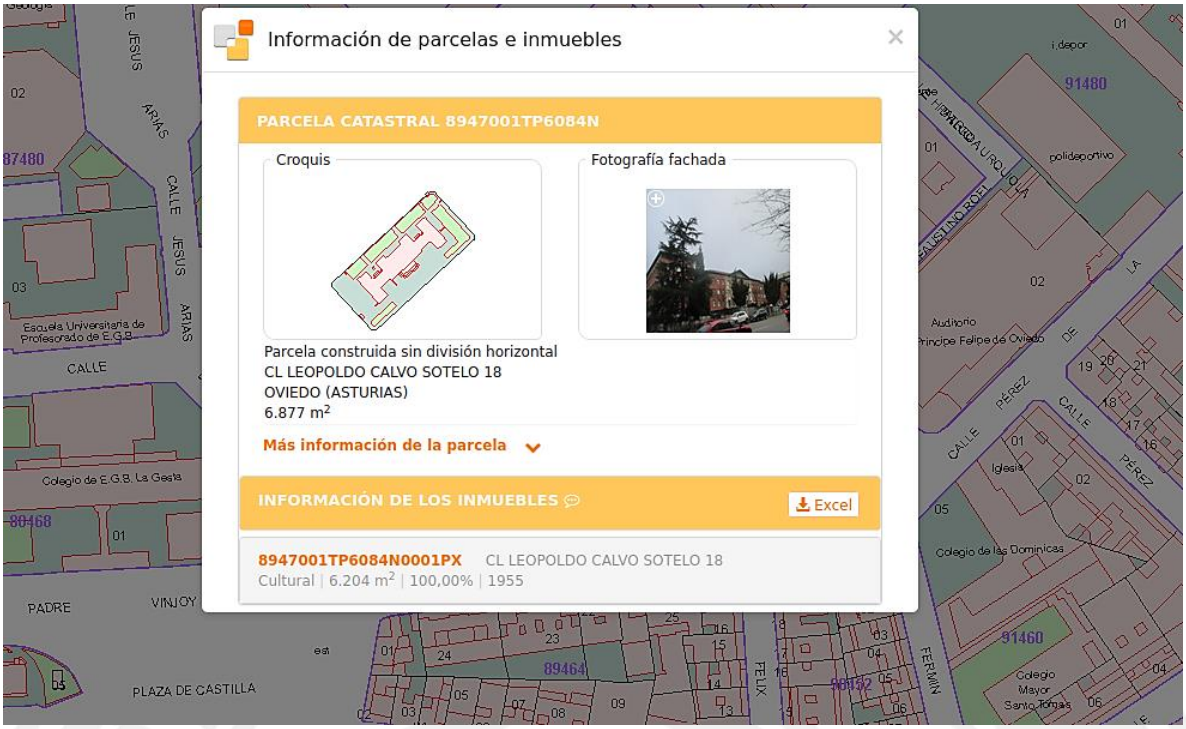


Figura 73. Mapa por satélite del catastro Español

Para acceder a los datos de un inmueble lo primero que necesitamos es **la referencia catastral**. Pero en el mapa de la imagen anterior podemos acceder fácilmente a dicho código (de 20 caracteres), por lo que no debería ser un problema. No debemos confundir este número con el **IDUFIR** (IDentificador Único de FInca Registral) de una finca, ya que estos términos no tienen relación entre sí. Este documento también sirve para hacer la valoración de una vivienda en portales online gratuitos. Con dicho código podremos solicitar la nota simple de la que hablábamos antes, bien de forma digital o en persona en las oficinas del catastro. La solicitud cuesta entre 3€ y 10€ y no es inmediata (pueden tardar unos 3 días). En dicha nota simple aparecerán desde los datos de la persona titular del inmueble hasta las posibles cargas que tiene la misma, lo cual de nuevo es necesario para comprar un inmueble de segunda mano con la certeza de que no tiene cargas. Otras opciones de búsqueda son:

  - **Por datos registrales**. Para este tipo de búsqueda es necesario tener el número registral de la finca, y la sección y el ordinal de la subfinca.
  - **A través de IDUFIR**. Como su nombre indica, permite encontrar una finca concreta dentro del territorio nacional.
  - **Por titular**. Para ello necesitamos su nombre completo o parcial, o su NIF (ver punto siguiente).
  - **Con otros datos**. Este tipo de consultas sirve para cuando solo disponemos de la localización de la finca.
  - **Mediante geolocalización**. A partir de las coordenadas del punto seleccionado en el visor, podemos identificar el IDUFIR asociado a él, siempre que la finca haya sido geolocalizada en el registro.

- **DNI / NIF**: El nº de DNI/NIF de alguien puede resultar sorprendentemente fácil de encontrar ya que puede aparecer directamente listado en Internet, muchas veces sin que la persona sea consciente de ello. Si solo tenemos el DNI, la letra correspondiente se puede calcular por un algoritmo que se puede encontrar fácilmente por Internet. Para localizar DNI/NIFs basta con hacer búsquedas muy sencillas en *Google* como `"<nombre y apellidos completo de la persona>" DNI` o `"<apellidos, nombre completo de la persona>" DNI`

Es común, por ejemplo, que la gente **cuelgue sus currículum vitae en Internet/LinkedIn** (o redes similares)**,** incluyendo mucha información personal y laboral (más incluso que la que dan en redes sociales), y que podemos usar para completar el perfil del investigado. Entre esa información no es raro que se encuentre el nº de DNI, fotos o incluso domicilios. Esto es común en investigadores o profesores de universidad, donde crear un CVN (***Curriculum Vitae Normalizado***) y exponerlo al público es muy común para poder acceder a proyectos u otros usos académicos. En estos documentos se suele dar una gran cantidad de datos de la persona y/o de las personas y empresas relacionadas con el mismo (proyectos de investigación llevados a cabo con empresas, dirección de tesis...), que se pueden usar fácilmente para "tirar del hilo" de nuestra investigación.

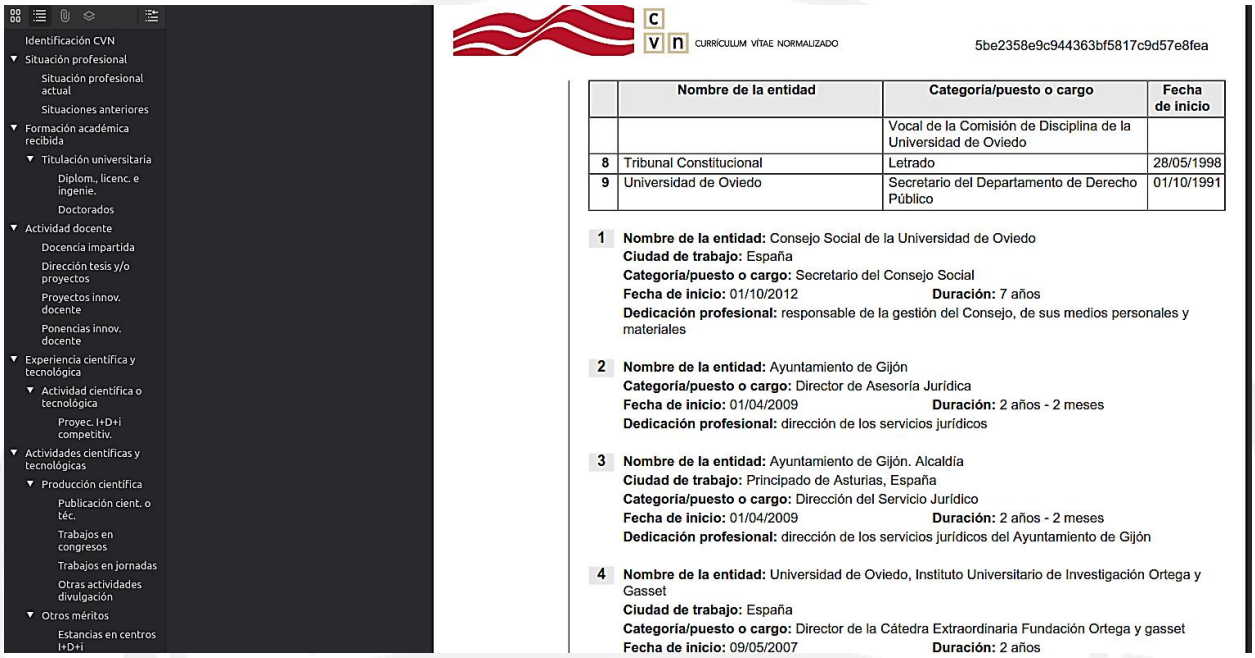
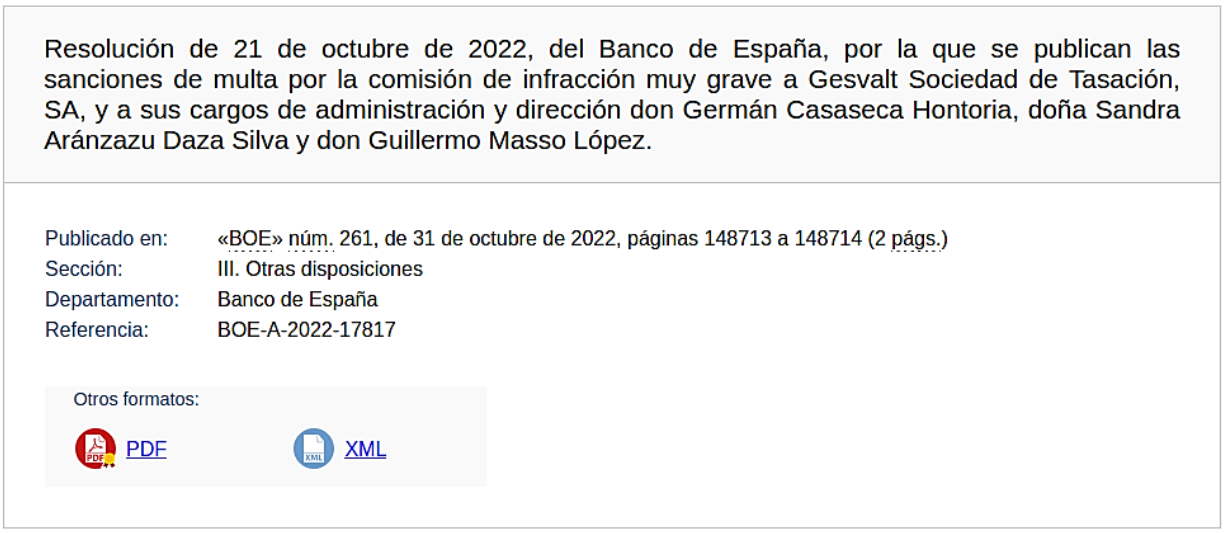


Figura 74. CVN público de un miembro de la Universidad de Oviedo

No debes tampoco descartar la presencia de **documentos antiguos "olvidados"** en la red donde se exponga esta información porque no se han retirado del acceso público, bien a través de búsquedas directas como en la ***Wayback Machine***. Hace no tanto la privacidad en Internet no era una preocupación y no había casi medidas contra ella. Esa información puede seguir ahí, y ni el propio titular podría saberlo. Yo he llegado incluso a localizar notas mías de la carrera 15 años más tarde...

- **BOE / o boletines autonómicos como el BOPA**: Si alguien ha sido **multado, o no ha pagado una multa de tráfico** o de otro tipo, es posible que salga publicado en el BOE u otro boletín autonómico. No solo se revela el nombre de la persona, sino el motivo y la cuantía de la multa, lo que puede ser usado para recabar más información de este, para spear phishing, y también para saber si hay alguna actividad que haya hecho que pueda resultar sospechosa o no fiable, que se pueda usar en algún contexto. Recuerdo una vez que un conocido mío localizó que uno de sus potenciales clientes tenía una serie de notificaciones de esta clase derivadas de múltiples impagos...



Resolución de 21 de octubre de 2022, del Banco de España, por la que se publican las sanciones de multa por la comisión de infracción muy grave a Gesvalt Sociedad de Tasación, SA, y a sus cargos de administración y dirección don Germán Casaseca Hontoria, doña Sandra Aránzazu Daza Silva y don Guillermo Masso López.

Publicado en: «BOE» núm. 261, de 31 de octubre de 2022, páginas 148713 a 148714 (2 págs.)
Sección: III. Otras disposiciones
Departamento: Banco de España
Referencia: BOE-A-2022-17817

Otros formatos:
PDF XML

TEXTO ORIGINAL

Con fecha 27 de septiembre de 2022 el Consejo de Gobierno del Banco de España dictó la Resolución sancionadora por la que finalizó el expediente de referencia IE/ST-2/2021, incoado a Gesvalt Sociedad de Tasación, SA, y a sus cargos de administración y dirección, don Germán Casaseca Hontoria, doña Sandra Aránzazu Daza Silva y don Guillermo Masso López por Acuerdo de la Comisión Ejecutiva del Banco de España de fecha 25 de noviembre de 2021.

Por otra parte, el artículo 115.5 de la Ley 10/2014, de 26 de junio, de ordenación, supervisión y solvencia de entidades de crédito, establece que las sanciones por infracciones muy graves serán publicadas en el «Boletín Oficial del Estado» una vez que sean firmes en vía administrativa.

Toda vez que la citada resolución ha adquirido firmeza en vía administrativa, procede, al amparo de lo establecido en el mencionado artículo 115.5, dar publicidad en el «Boletín Oficial del Estado» a las siguientes sanciones, impuestas a los citados expedientados en dicha Resolución sancionadora del Consejo de Gobierno del Banco de España:

– Sanción de multa impuesta a Gesvalt Sociedad de Tasación, SA, por importe de ciento treinta y seis mil euros (136.000 €) –tras la aplicación de la reducción del 20 % por pago voluntario contemplada en el artículo 85 de la Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas–, prevista en el artículo 97.1.a) de la Ley 10/2014, de 26 de

Figura 75. BOE notificando multas a personas

Algunas veces uno puede encontrarse en los ficheros `robots.txt` de las webs que mantienen estos documentos accesibles enlaces a PDFs concretos. Esto puede hacerse para que su contenido no salga listado en buscadores como *Google*. Uno de los motivos por los que esto ocurre es precisamente porque se estén revelando datos de personas, y hay una legislación relativa al "derecho al olvido". No obstante, sigue siendo información pública, por lo que en caso de que sospechemos que algo que necesitemos está ahí, **puede desarrollarse un *script*** para descargar esos ficheros y buscar contenidos entre ellos en caso necesario. Puedes pensar que esto es algo bastante extraño, pero si lo piensas es una forma de mantener a la vez un cierto "derecho al olvido" y que la información sea pública...

Otra cosa que también es localizable vía BOE o similar es cuando una administración u organismo públicos saca un **pliego de condiciones** para que otras empresas sean capaces de concursar y servir lo que requiere dicha administración. Si dicho pliego de condiciones incluye servir unos determinados **modelos o tipos de *hardware/software*** podremos entonces identificar qué máquinas están usando y quizá con ello también sistemas operativos y software por estar vinculados a ellas. Si localizamos quien ha ganado el concurso (debería ser información pública), ya es un dato para añadir que puede ser usado en timos de *spear phishing*. Hay varios ejemplos de fraudes del CEO que usan esta clase de información...

Teléfonos y tarjetas de crédito (u otros documentos de los que aparezcan fotografías en Internet)

Esto es un poco lo mismo que lo anterior: en Internet es típico encontrar **filtraciones de nºs de teléfono** que se hayan usado para darse de alta en un servicio, ya que los robos de información a diversas empresas o RRSS son desgraciadamente comunes. Con ellos es más fácil ser víctima de *phishing* o *spam*, etc. y habitualmente se venden (o se regalan) en enormes packs de nºs en formato CSV o texto, muchas veces acompañados del nombre del titular o más datos. Esta información no suele estar publicada en motores de búsqueda típicos no obstante, sino en la *Dark Web*.

El **número de tarjeta de crédito** es una información más sensible que el **número de cuenta o IBAN**. Pero sólo con este dato no se puede sacar dinero de un cajero. Sólo se podrá operar con ella (ej.: realizar compras por internet) en caso de que, **además, se haya obtenido también la fecha de caducidad, número de seguridad o PIN**. Pero claro, uno puede dejar sin querer foto de ambas cosas en una red social por descuido o accidente...o que alguien obtenga el número y luego el PIN/fecha por *phishing,* o simplemente **mirando lo que tecleas en un cajero** (de ahí que sea importante lo de no revelar por dónde acostumbras a ir diariamente). En otras ocasiones, gente (sobre todo joven y/o inexperta) **directamente publica estos datos en Internet**, y el objetivo investigado (o alguno de sus familiares) puede ser uno de ellos. No estamos hablando, como antes, de que aparezca "de fondo" en una foto (que también), sino **directamente publicar esta información**. Es más común de lo que parece verlo en *Twitter*, aunque muchas veces son bromistas.

Figura 76. Sí, esto pasó realmente....y continúa pasando de vez en cuando

¿Qué pasa si alguien tiene un número de tarjeta ajeno? La respuesta es **nada**, compartir el número de tarjeta no representa ningún peligro, ya que, como decíamos, para realizar una compra en línea o hacer algún pago es necesario poner también la fecha de vencimiento y el código CVV. El problema es, como decíamos, que se saque ese CVV de otras formas...

Si alguien se hace con un **número de cuenta o IBAN** de un tercero **TAMPOCO le puede robar** (¡imagina entonces lo que pasaría con tantas campañas de donación y ayuda que hay dando esos datos para hacer ingresos!). Solo con esa numeración nadie puede sacar el dinero de tu cuenta. El problema es si **además**

conocen el nº de DNI del titular (de las formas que vimos antes). En ese caso sí existe la posibilidad de que en algunos casos se pueda realizar una domiciliación de recibos con estos datos.

Uso de medios de transporte varios

¿Qué pasa si usas medios de transporte y lo publicas? De nuevo, que estás dando demasiada información. Eso además incluye los **billetes** del viaje, porque permiten averiguar el trayecto a seguir, tiempos, paradas, escalas, rutas...etc. en no pocas ocasiones. Es decir, de nuevo permiten ver dónde vas a estar en un momento dado y si es un viaje que haces periódicamente, es aún peor...

A modo de ejemplo, fíjate en este **billete de avión**. Si se te ocurre ponerlo en una red social, solo con la fecha y con el identificador del avión (en este caso AF112) hay páginas web que te informan del trayecto, escalas, etc. (las propias compañías tienen esa información publicada de vuelos recientes) e incluso pueden seguirlo en tiempo real, como ***FlightAware***, por ejemplo: https://es.flightaware.com/.

Figura 77. Ejemplo de típico billete de avión

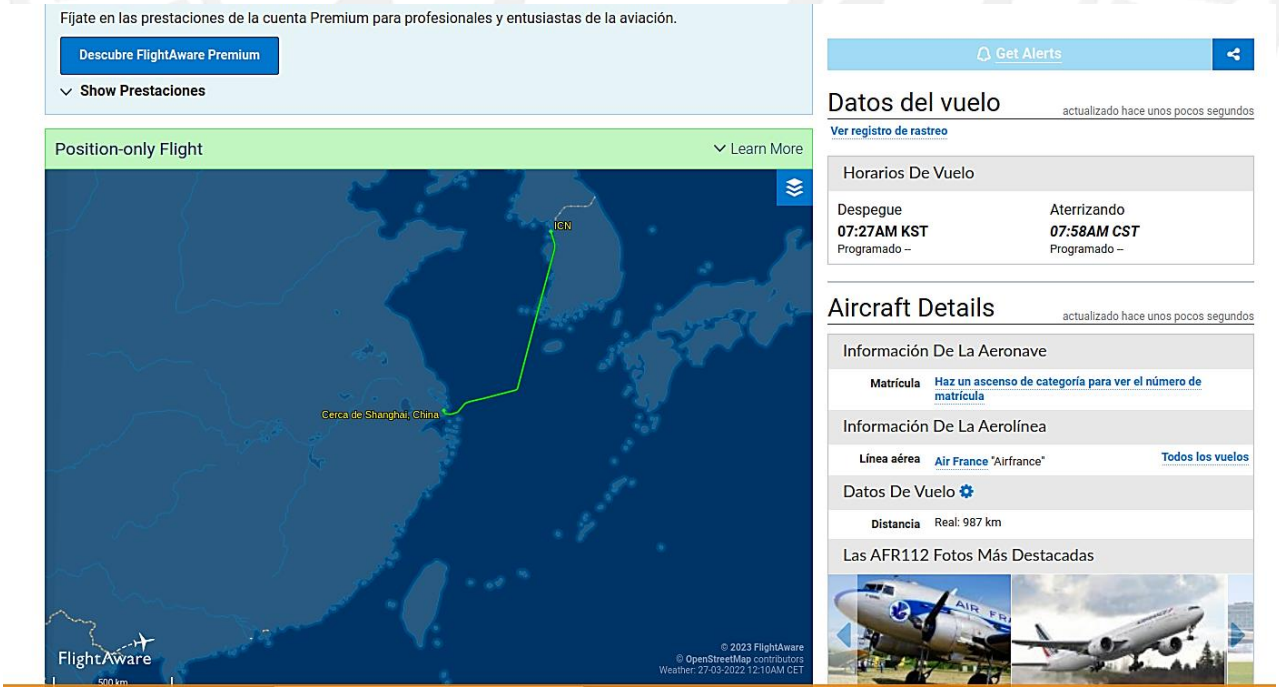


Figura 78. Vuelo hecho por el avión del billete anterior en una fecha posterior gracias a *FlightAware*

Geolocalización

En la asignatura hemos visto que es viable **geolocalizar una IP** o unas coordenadas que encontremos en metadatos, pero muchas veces eso ni siquiera es necesario porque la propia persona investigada da dicha localización en la propia imagen que sube a la red. No pierdas de vista esto, porque aunque parece demasiado evidente, la gente lo hace por desconocimiento u olvido. Esto por ejemplo es muy típico en **posts de *Instagram***. Aquí no hay coordenadas, directamente te dicen el nombre del sitio...

Otra técnica para tener una estimación de la geolocalización de un investigado es usar **una baliza digital** como las que páginas como ***IPLogger*** pueden proporcionar (https://iplogger.org/es/). No obstante, esto requiere *engagement* (el objetivo debe hacer clic en el enlace) y queda fuera de la asignatura.

Con el tema de direcciones IP y geolocalización hay muchas herramientas para sacar información asociada a la misma. Por ejemplo, ***Ghost-Track*** (https://github.com/HunxByts/GhostTrack) es una herramienta que permite obtener una geolocalización de cualquier IP desde línea de comandos **y además trabaja con teléfonos**.

Figura 79. Desde luego no faltan opciones para obtener información de geolocalización de IPs...

Pero vamos a detenernos un poco más en este apartado para insistir en que muchas veces, con un poco de habilidad investigadora y sabiendo **trucos habituales en este tipo de investigaciones**, es posible encontrar

la geolocalización de una imagen aunque inicialmente no tengamos pistas acerca de la misma., Esto es un "arte" muy estudiado que merece la pena leer y practicar porque obtiene resultados más prometedores de lo que parece. Puedes encontrar las técnicas usadas, e incluso un pequeño tutorial de entrenamiento a la hora de aplicarlas, en estos enlaces:

- https://www.osintcombine.com/post/image-analysis-and-verification
- https://infosecwriteups.com/image-geospatial-osint-7cf8560b54f6
- https://medium.com/@ashfquejahan5/image-analysis-osint-tools-2bad341a3d29

Con técnicas como esta puedes encontrar información incluso en fotogramas de series famosas (cuando en lugar de fotos tienes videos se analizan fotogramas individualmente). Por ejemplo:

Figura 80. Donde tiene lugar *Big Bang Theory* solo con un fotograma de la serie

El boom de la **inteligencia artificial** también está teniendo su influencia en el mundo OSINT, y prueba de ello son las IAs de geolocalización como ***Geospy AI***. https://geospy.web.app/. Esta IA analiza la vegetación y otros elementos de una foto que le pases para decirte dónde cree que se ha tomado la foto, dándote unas coordenadas que puedes usar en *Google Maps* para validar la información. Es un proyecto en beta y **sus resultados son aproximados**, pero para lograr una primera pista de la zona donde una foto pudo ser tomada, cumple con su función perfectamente (además de ser gratis 😊). Este es un ejemplo con una foto que encontré en redes sociales donde se ve un fondo mínimo:

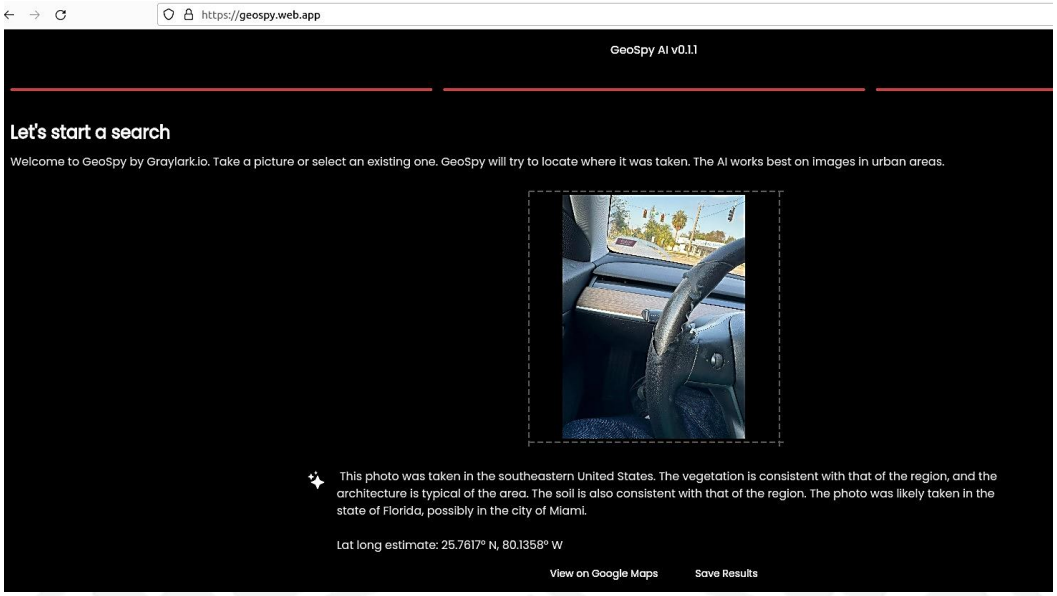


Figura 81. Reconozco que hace bastante con tan poco como muestra la foto...

Y esta es una foto de Adrián Barbón en una localidad Asturiana. Está claro que no ha acertado, pero por lo menos sabemos que se tomó en el norte de España (ten en cuenta que si alguien extranjero ve esta foto no sabría ni por dónde empezar, ya que ni siquiera sabría quién es el de la foto 😊):

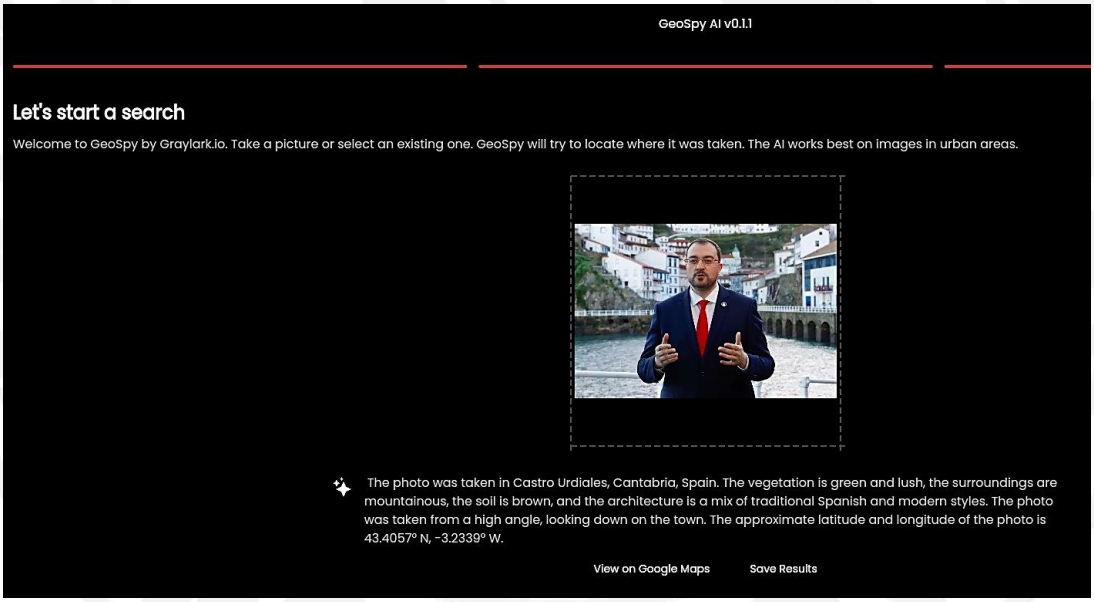


Figura 82. Los resultados son aproximados...pero al menos ahora sabes por dónde empezar a mirar

Foros / Blogs / IRC...

Como decíamos al principio, nuestros objetivos pueden escribir datos relevantes o interesantes para determinadas investigaciones en foros, blogs, canales IRC, *Discord*, comentarios de *YouTube*, etc. Es común que una persona concreta use el mismo ID en muchas redes sociales o páginas, por lo que localizarlo es viable en el momento que usemos herramientas como ***Project Sherlock*** (**Seminario 1**) o porque simplemente esa información la estén dando en otras webs sin darse cuenta.

Hay también que tener en cuenta que **mucha gente se "suelta" en sitios que no son redes sociales** (es decir, opina o dice cosas que no se atrevería a decir en foros más públicos, demostrando una tremenda hipocresía), amparados en ese **supuesto anonimato** que tienen en esos sitios, pensando que nadie va a ser

capaz de averiguar su identidad real. No es infrecuente, tanto por delincuentes como por cuerpos de seguridad del estado, usar una identidad falsa, como las que vimos antes, para hacerse un usuario en estos sitios y "ciber-infiltrarse" simplemente a leer que se comenta en esos foros en busca de información válida para algo.

Además de la **coincidencia de IDs** usados en distintos sitios y/o posibles metadatos que haya en cosas que suban, también a veces es viable **identificarlos por lo que escriben y cómo lo escriben**: expresiones concretas que sean propias de la persona, faltas de ortografía "raras" que suela cometer esa persona al escribir, uso de palabras particulares, narrar la misma historia personal (o copiar párrafos) en distintos sitios, ...un análisis de lo que una persona escribe puede revelar más de lo que crees...**Moraleja**: Nunca subestimes el poder de un **lingüista forense**, como por ejemplo Sheila Queralt: https://es.wikipedia.org/wiki/Sheila_Queralt_Estevez.

Como anécdota ilustrativa de este tipo de técnicas, una vez identificamos a un troll en *Internet* porque, para intentar alejar las sospechas de su identidad real, un día decidió criticar a una persona concreta e identificable en sus redes sociales. El fallo fue que de toda la gente a la que criticaba desde su "anonimato", esta persona era un "***outlier***", es decir, que no tenía rasgos en común (por ejemplo, puestos o categoría laboral) con las otras a las que criticaba. Con esa información, y otras más que fue revelando a medida que hacía sus campañas de troleo (siempre siguiendo los 4 mandamientos del OSINT que vimos antes), pudimos hacer un perfil que redujo el nº de "sospechosos" a 1. Puesto en conocimiento de las autoridades todo esto, finalmente resultó que era el verdadero culpable y la campaña se paró. Efectivamente, al final el culpable era el "*outlier*" al que había criticado muy duramente en redes (es decir, que se criticó a si mismo 😊). Quizá eso de criticarse a uno mismo para apartar sospechas te pueda parecer una táctica rara, pero no es tan infrecuente como puedas creer...

Anuncios

Buscar a una persona en un **portal de anuncios o ventas** (a través de fotos de propiedades o automóviles por ejemplo que haya subido a la red, o un ID de usuario común con sus redes sociales) es una forma alternativa localizar y obtener información adicional del objetivo. También debemos incluir en este supuesto las **empresas que buscan empleados de un determinado perfil** o con conocimiento de algunas tecnologías, ya que eso puede usarse para **identificar software o hardware** que la empresa esté usando en ese momento. Si la empresa por ejemplo pide mucha gente con un determinado perfil, querrá decir o que va a abrir una nueva línea de negocio, o que las personas que estaban ocupándola han sido despedidas por algún motivo, con lo cual se abre la posibilidad de buscar exempleados para seguir con nuestra investigación. Otro caso que he visto en mi experiencia son **empresas que piden trabajadores de un determinado perfil demasiado frecuentemente**. Esto muchas veces revela un entorno de trabajo malo o tóxico, y que genera un gran nº de extrabajadores "quemados", con todo lo que ello implica.

Por tanto, no debe subestimarse la información que se pueda encontrar en portales de venta de segunda mano, propiedades, inmobiliarias, empleo o similares porque, si podemos encontrar un vínculo con la persona que estamos investigando, podremos abrir incluso **un canal de comunicación** bajo una identidad falsa como las que hemos mencionado anteriormente para extraer más información de esta. No obstante esto es una técnica que requiere "*engagement*" y por tanto queda fuera de los objetivos de esta asignatura. nuevamente solo se menciona por si acaso algún día tú eres el objetivo. **Las estafas en páginas de venta son demasiado comunes**...

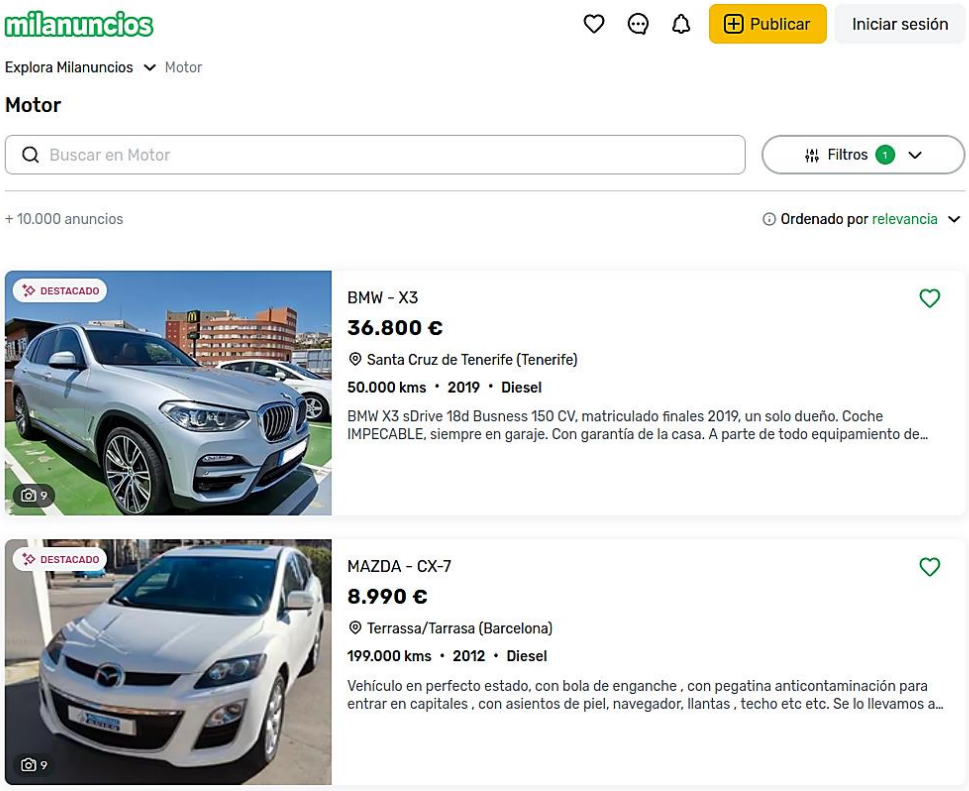


Figura 83. Nunca subestimes la información que se puede encontrar en un portal de venta de segunda mano o similar...

OSINT sobre máquinas y código

Poco puedo decirte aquí nuevo, puesto que se basa en aplicar las técnicas que se vieron en el **Seminario 1** de la asignatura y también se probaron en las **Prácticas 1, 2**, y **8-9**. *Shodan*, *Censys* y *Google Dorking* sobre **máquinas** son formas pasivas típicas de hacerlo y `nmap` la forma activa preferida en estos casos, tanto con escaneos estándar como a través de *scripts*. Para **código** en *GitHub*, en el **Seminario 1** se vieron herramientas y técnicas.

Se recomienda mirar estos temas de la asignatura para hacer este tipo de investigación, **priorizando siempre hacerla de forma pasiva**. También debe consultarse la sección de este documento donde se describen motores de búsqueda que ofrecen diversa información, que pueden ser de aplicación en este caso. Si lo que vimos de la GHDB te parece complejo, **puedes usar una IA que trabaja con este concepto** (hacer consultas de *Google* que encuentran información interesante de objetivos) y/o quieres plantear algunas personalizadas, puedes usar *DorkGPT* (https://www.dorkgpt.com/), que diciendo lo que quieres encontrar te genera consultas que lo buscan.

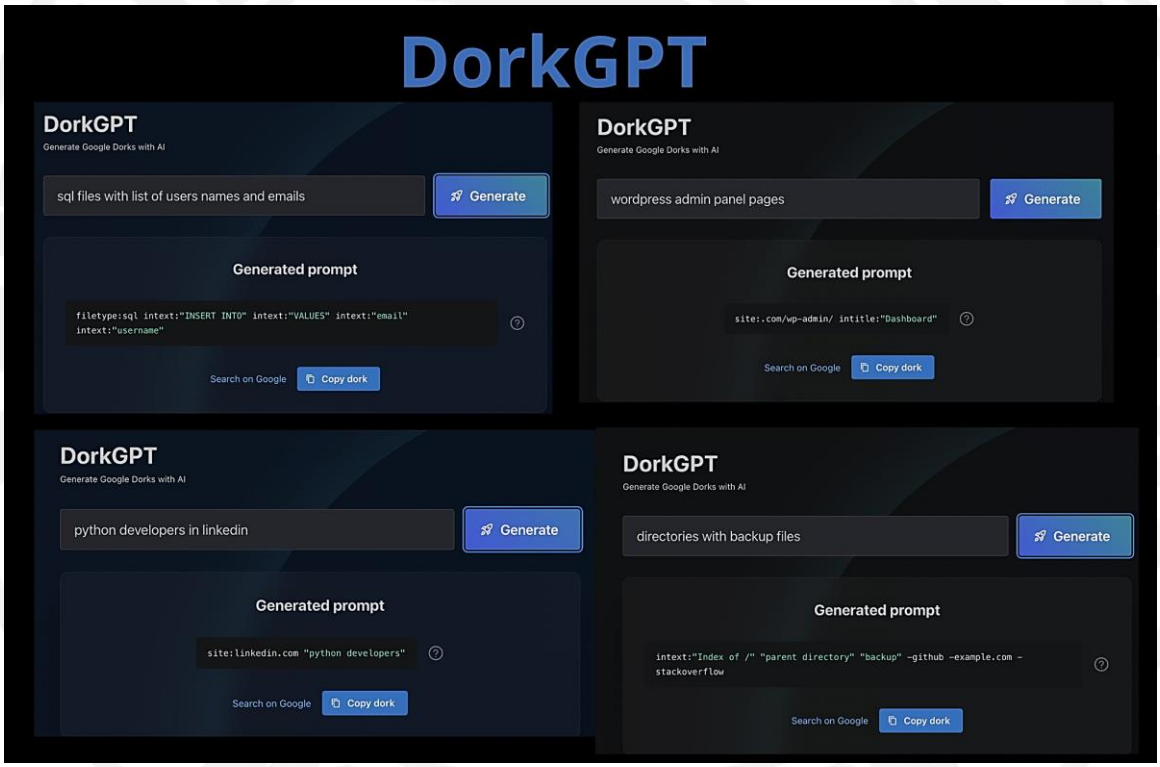


Figura 84. *DorkGPT*, usando IA con *Google Hacking*. Fuente: https://twitter.com/cyb_detective/status/1654835052260335618

Obviamente, si la información de máquinas y código se obtiene a partir de, o se vincula con la que hemos obtenido **examinando la de personas**, el estudio tendrá muchas más posibilidades de éxito (y calidad). Sobre este apartado concreto vamos a destacar solo unas pocas cosas en las siguientes subsecciones.

Emulación de terminales móviles

Quizá puedas pensar que no tiene sentido acceder a una web como si fuera un móvil para temas de investigación OSINT. Pero no te creas: **es frecuente que una web para móviles no tenga las mismas protecciones ni mecanismos de seguridad** que una para PCs. Además, también es posible que la web para dispositivos móviles **sea independiente** de la otra, con lo cual pueda presentar vulnerabilidades o información extra (como documentos) que la otra no presenta. **Tampoco hace falta usar un móvil para ver**

esas versiones de una web, hay extensiones de navegadores (o las propias funcionalidades integradas en los mismos) que permiten verlas desde un PC: https://stackoverflow.com/questions/16651911/how-can-i-simulate-mobile-devices-and-debug-in-firefox-browser

Direcciones IP / dominios

Ya sabemos que con la IP de un objetivo tenemos forma de averiguar fácilmente los servicios que ofrece y su versión (nmap, **Laboratorio 7-8**) y, por tanto, si son o no vulnerables (**Tema 1**). Los motores de búsqueda vistos en la sección correspondiente de este documento también pueden ser de mucha ayuda para encontrar más información de los objetivos cuyas IPs conozcamos o averigüemos, así como la web ***IPInfo*** (https://ipinfo.io/). Como vemos, la IP de un objetivo puede llegar a ser una "mina" de información.

Otras de las cosas que podemos mirar de las IPs / dominios son:

- Si aparecen en **listas de bloqueo** (*blacklists*), lo que querría decir que el objetivo podría haber sido comprometido o sus actividades no son muy "normales". Hay muchas listas de bloqueo disponibles en Internet, descargables o buscables online. Una muy popular y grande, clasificada por el tipo de actividad que lleva a cabo cada IP, la mantiene Steven Black https://github.com/StevenBlack/hosts
- Si están **protegidas por una CDN** (Ej.: *CloudFlare, TransparentEdge*), lo que implica que probablemente haya una inversión detrás y la seguridad se toma "algo" en serio.
- Si pertenecen a un **proveedor de nube** / hosting grande o pequeño. Esto les podría dotar de una mayor seguridad inicial, pero también de **volatilidad**. La volatilidad implica que son dominios que se crean solo durante espacios pequeños de tiempo, normalmente porque son para estafas o similares. Se compra el dominio (normalmente muy barato), se le pone un nombre que "suene como" el original (o se manda a las víctimas un enlace acortado), se clona una página a suplantar, se lanza la campaña y en unos días se borra todo para dificultar las investigaciones (***TA0042 Resource Development*** del MITRE ATT&CK). Si en tus investigaciones das con una máquina en la nube de, por ejemplo, Amazon, con una copia de una página conocida, has dado con una estafa.

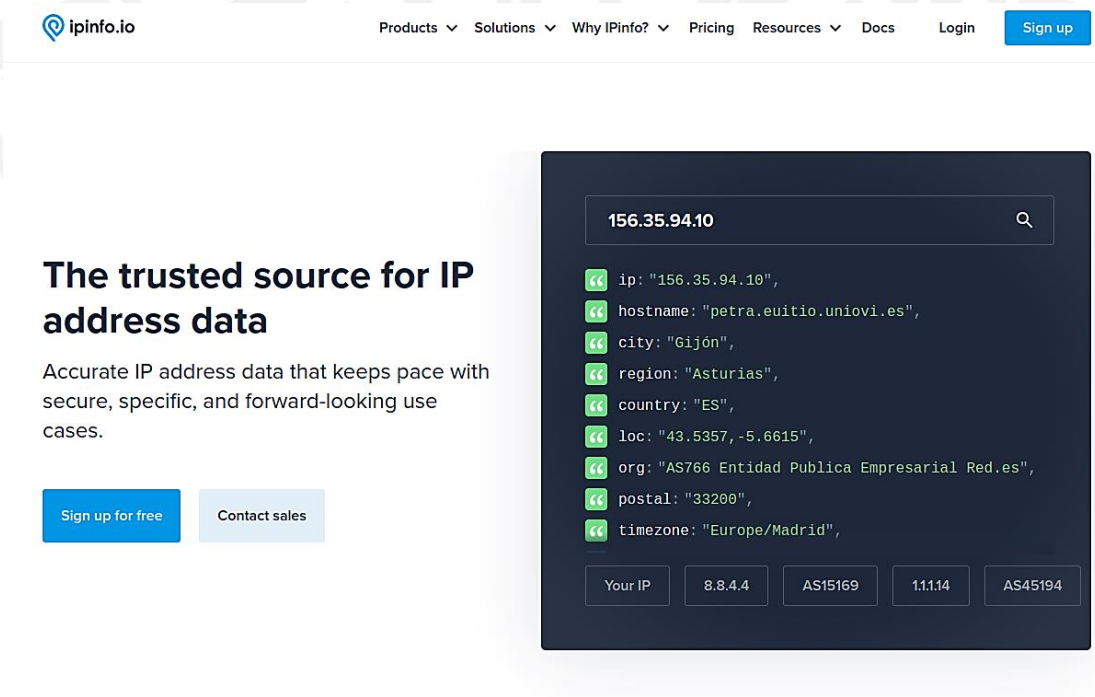


Figura 85. Web IPInfo

Si quieres **saber más sobre un dominio y los sitios "secundarios" a los que accede o desde los que lee información** sin necesidad de visitarlo por si tuviera algún tipo de *malware*, siempre puedes recurrir a los servicios de ***Lookyloo***: https://lookyloo.circl.lu/

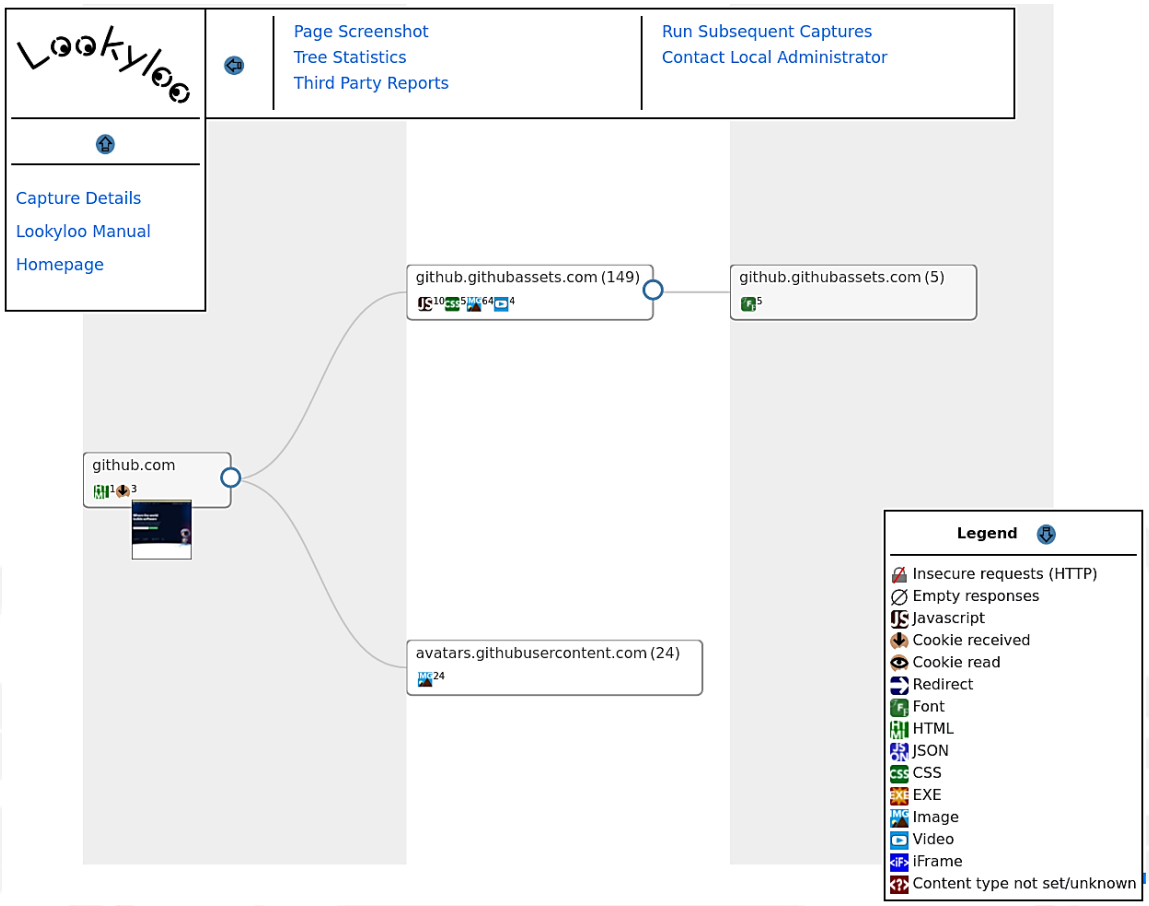


Figura 86. *Lookyloo* analizando lo que hace un dominio. Fuente: https://github.com/Lookyloo/lookyloo

Análisis offline de archivos en la red

Examinar los datos de los archivos que contiene una página es obviamente una fuente de información que ya hemos visto en la asignatura. Si lo incluimos aquí es simplemente para recordar que muchas veces es una investigación que **lleva tiempo y varias sesiones de trabajo**, y en esos casos quizá lo mejor sea hacerse **una copia de una web en local** para mirar lo publicado con calma Esto es posible gracias a herramientas como ***Burp Suite*** (instalada por defecto en *Kali Linux*), que permiten hacer **"*mirrors*" locales** de cualquier web con un solo clic: https://portswigger.net/burp/documentation/desktop/getting-started/running-your-first-scan

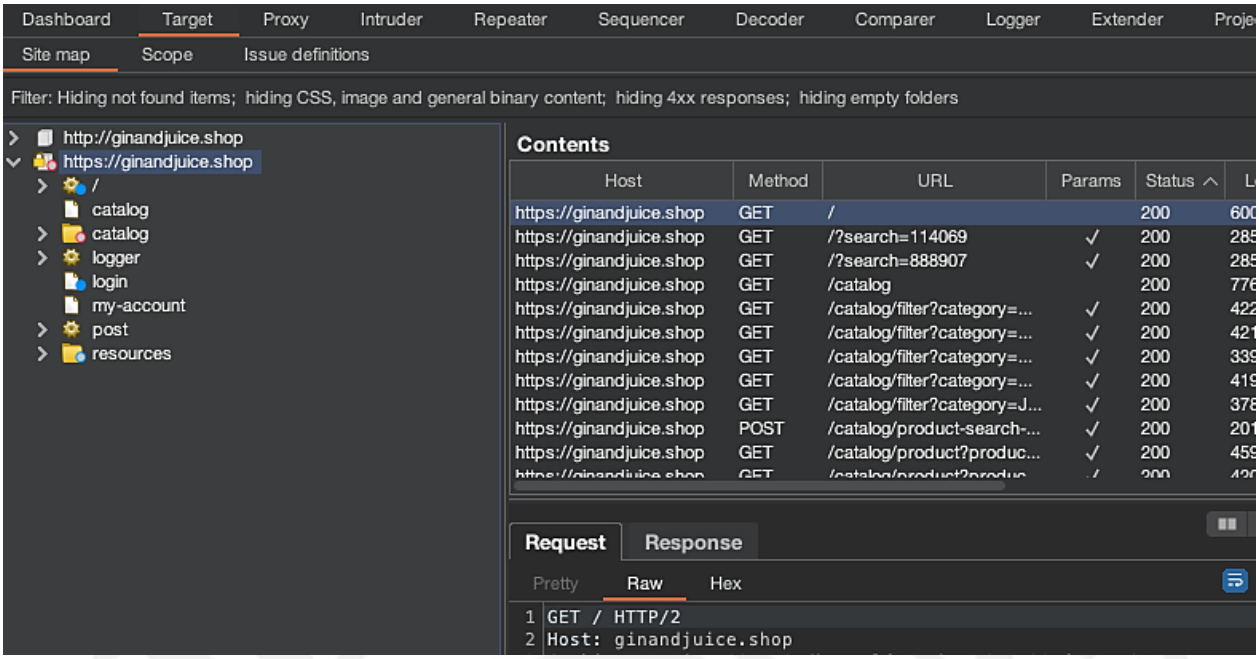


Figura 87. *Burp Suite* haciendo una copia local de una web

Con la web en nuestro poder ya podemos hacer todas las búsquedas que queramos en nuestra casa sin prisas y a nuestro propio ritmo 😊.

Además de datos que se puedan encontrar en comentarios de la web descargada, es posible encontrar otra información más sutil si sabemos exactamente qué mirar. Por ejemplo, en la siguiente imagen podemos ver un ejemplo analizando el código que tiene la web asociado a ***Google Analytics*** (una herramienta que se usa para analizar el comportamiento de los usuarios en la web). Con la herramienta *NerdyData* (https://www.nerdydata.com/) mencionada en dicha imagen puedes encontrar otras webs vinculadas a la que estás visitando...y puede que en esas otras webs haya más información de su propietario, empresas asociadas, etc.

En otras palabras, estamos usando el identificador de análisis de *Google* para encontrar **webs vinculadas entre ellas**. *¿Por qué estarán vinculadas?* Te corresponde a ti saberlo 😉. Además, ahora tienes otra herramienta (*NerdyData*, que puede probarse de forma gratuita haciéndose una cuenta) para **relacionar webs que usen las mismas tecnologías SaaS**.

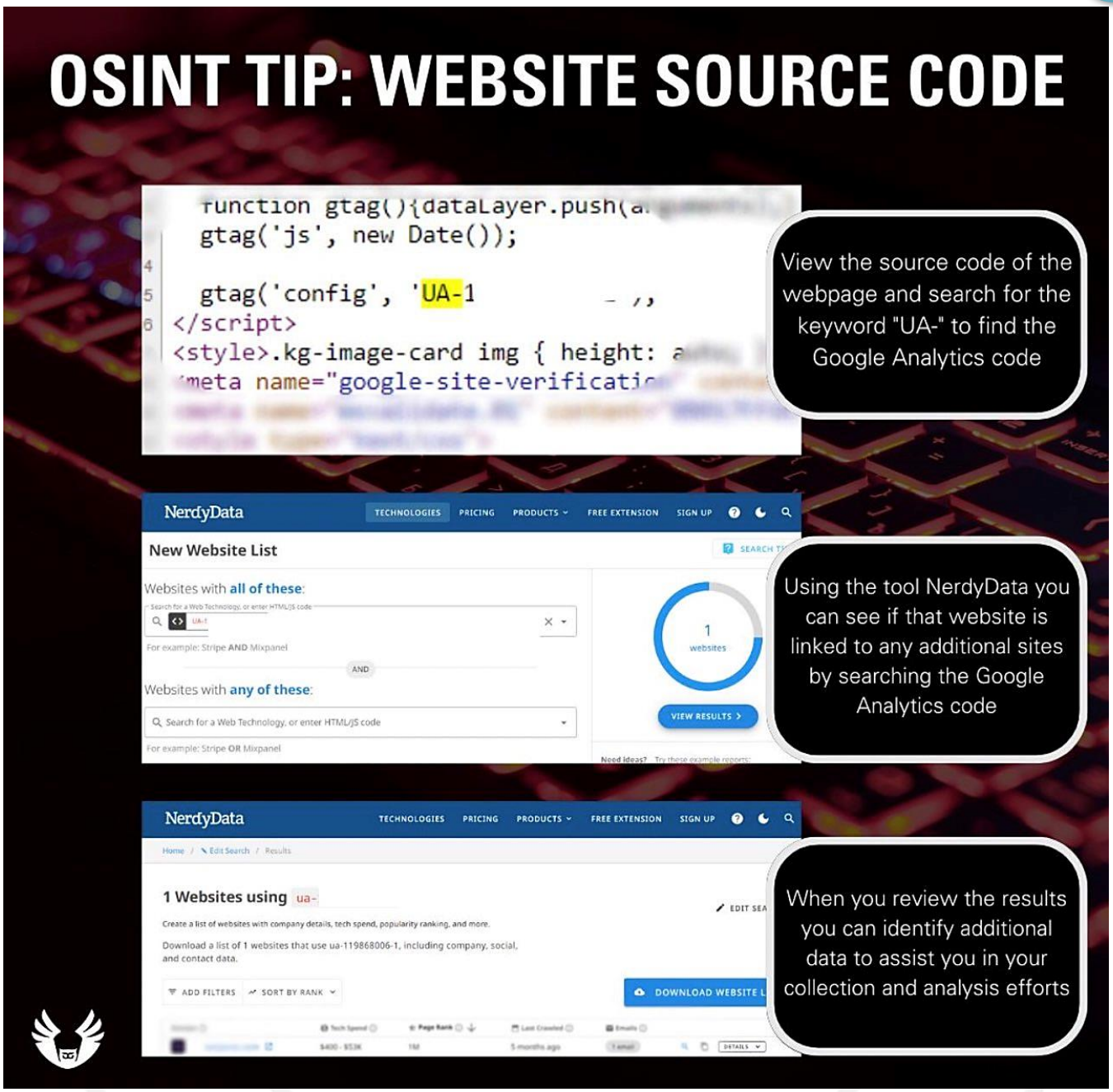


Figura 88. El código de una web puede ser una fuente de datos inesperada...a veces de manera sorprendente y sin que se requiera una negligencia de sus autores. Fuente: https://twitter.com/cybernitewatch/status/1664975350999252992

Qué pasa si te dejas tus dispositivos desbloqueados

Todo el mundo se imagina que si un *malware* es capaz de entrar en nuestro sistema tenemos un problema grave, si bien mucha gente no tiene una idea clara de **exactamente qué se puede robar de allí**. El motivo es que no saben exactamente qué tipo de información tiene cualquiera de nuestros dispositivos, y esto queda claro en la imagen siguiente.

No obstante, antes de mostrarla es importante recordar una cosa: el mismo tipo de acceso es posible **si no se bloquea un dispositivo de trabajo** cuando nos ausentamos de él, y no será la primera ni la última vez que alguien saca partido de eso para sacar información de alguien. Esto no es OSINT, ni algo que vayamos a hacer en la asignatura, pero es importante conocerlo porque es un riesgo muy real y frecuentemente bastante ignorado (es decir, **puedes ser una víctima**).

La imagen siguiente se aplica a ciudadanos de los Estados Unidos, pero muchos de los elementos que aparecen en ella son aplicables a todo el mundo. Piensa en la cantidad de claves recordadas que puedes tener, tu historial, documentos personales, etc. Si ya es malo que te lo roben usando *malware*, no permitas encima que lo hagan por un descuido: **no es una broma**.

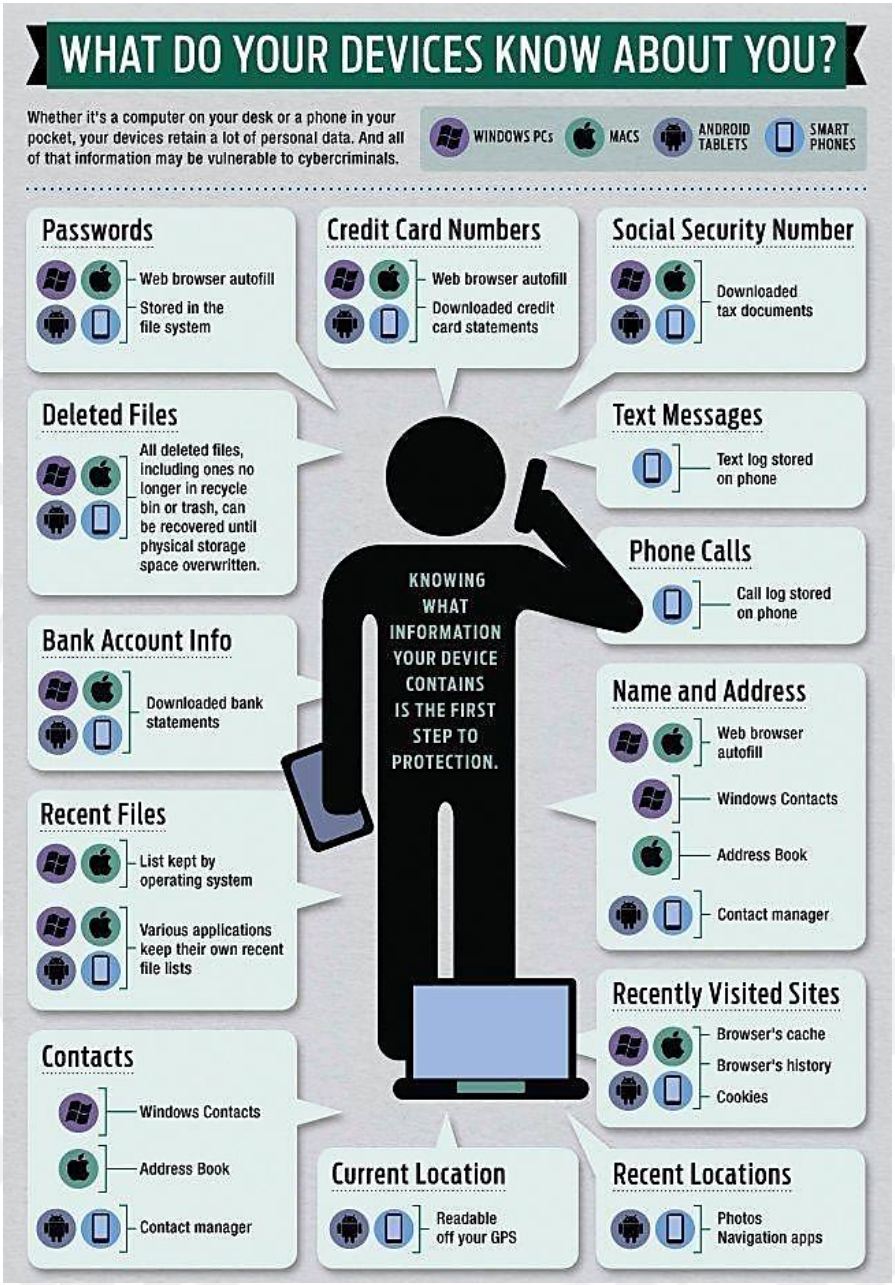


Figura 89. Cosas que se pueden averiguar si alguien tiene acceso a tu dispositivo. Fuente: *Cyber Crime watch*

OSINT "Automático". OSINT Ex Machina

OSINT sobre personas

Antes mencionamos redes sociales, pero el análisis que se puede hacer de lo que contienen **puede llegar a ser muy amplio y revelador**, más cuantas más cosas comparta el investigado en ellas. Por eso es oportuno complementar este análisis con herramientas automáticas. Para ello debemos usar **herramientas específicas para la red social concreta** que estemos investigando, como las vistas en el **Seminario 1** y las listadas en la web de ***Ciberpatrulla*** (https://ciberpatrulla.com/), que nos permiten sacar un listado histórico de información subida, geolocalizaciones, buscar entre el texto escrito y otros datos. Nos permiten por tanto un análisis más en profundidad de lo que aparece en cada red social, puede revelar también muchas cosas.

No obstante, conviene aclarar que aunque **hay herramientas automáticas para todas las redes sociales**, muchas de ellas dependen del acceso al API de cada una de ellas, lo que nos obliga a crear una cuenta en las mismas y puede tener posibles limitaciones de uso / ratio que puede interferir con nuestras operaciones. Además, como con el caso de ***Twitter***, el acceso al API **puede ser de pago** e impedir su uso totalmente si no contamos con financiación para ello. En este caso concreto, necesitamos herramientas que no usen el API para extraer la información, sino que sean ***scrappers***: navegan de forma automatizada la interfaz web de la aplicación y extraen información leyendo lo que sale en dichas páginas web, estructurándola adecuadamente para que luego podamos usarla con otros programas o analizarla más detenidamente. Un buen ejemplo de una herramienta *scrapper* que funciona con varias redes sociales es `snscrape`: https://github.com/JustAnotherArchivist/snscrape

Como muestra de herramientas que trabajan con personas, la herramienta ***Email Info*** (https://epieos.com/) nos puede ayudar a localizar información asociada a un email que pongamos, lo que se puede usar también para dar pie a investigaciones más en profundidad.

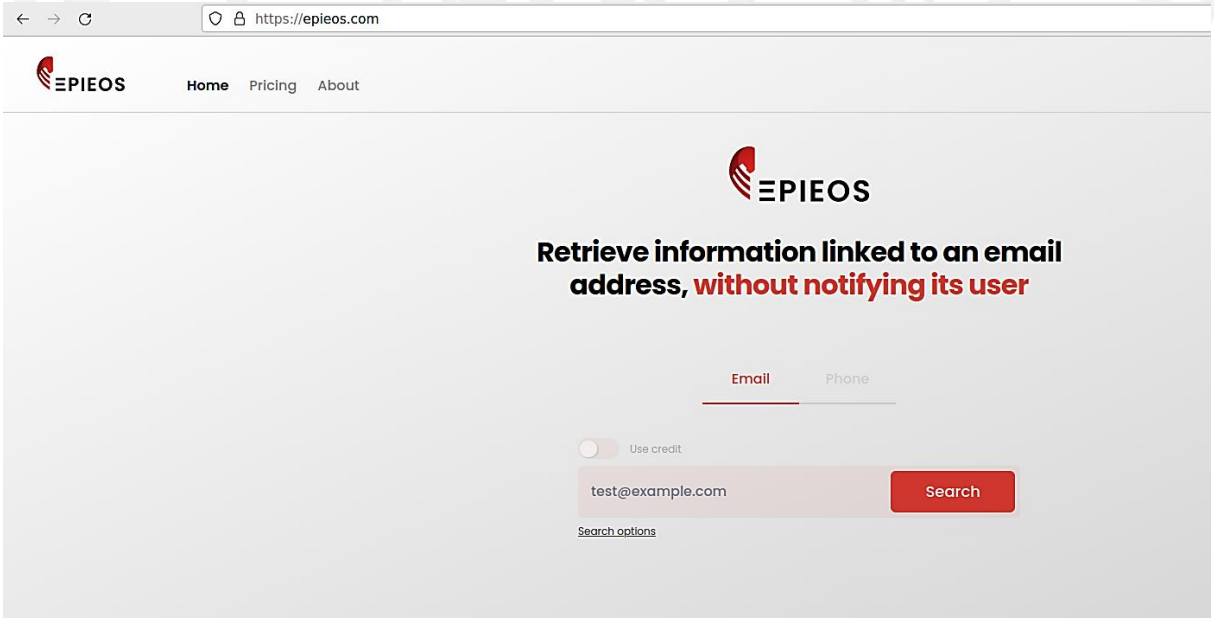


Figura 90. Herramienta *Email Info*

Por otro lado tenemos ***WhatsMyName*** (https://whatsmyname.app/), que hace lo mismo pero con nombres de usuario.

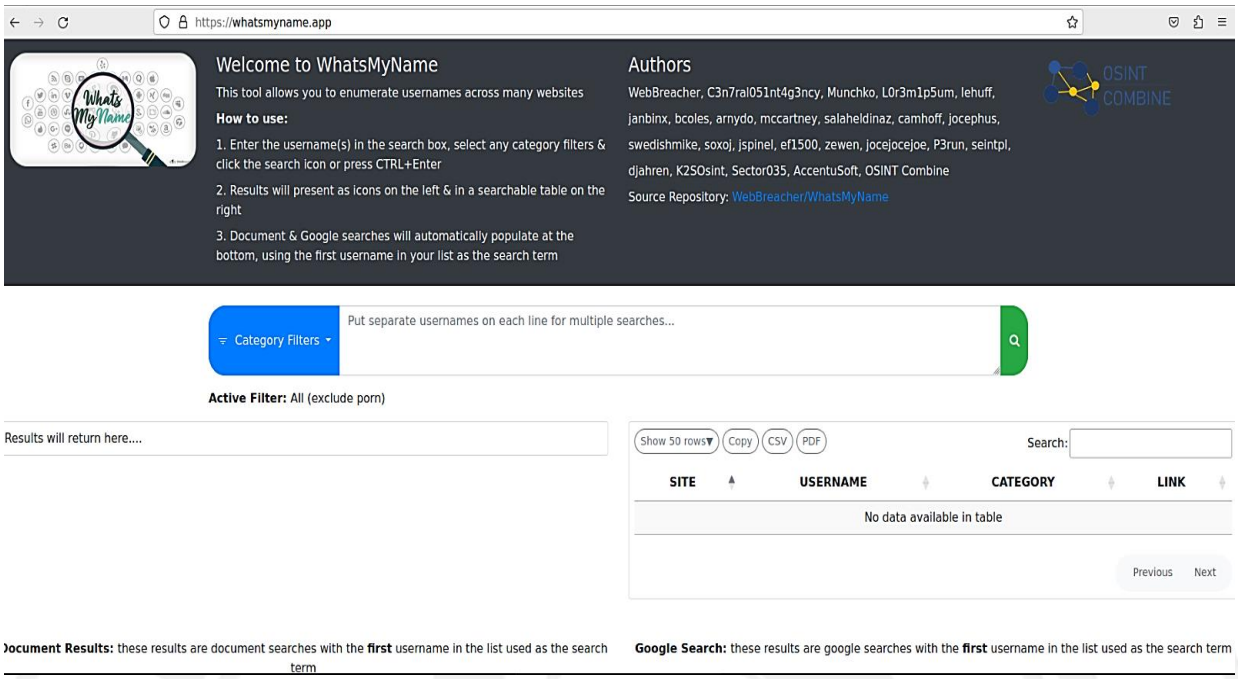


Figura 91. Herramienta *What is my name*

Finalmente, hay una forma automática de averiguar datos de una **persona abusando de la funcionalidad de contactos** de nuestro móvil en combinación con aplicaciones como *Snapchat*, *WhatsApp*, *Threema*, *Strava*, *FitBit*, *Duolingo*, *TikTok* o *Vivino*. Consiste en hacer lo siguiente:

- Si tienes un teléfono o email de la persona que quieres investigar, **añádelo como un contacto** de tu teléfono
- Usa una de las apps que mencionamos antes y añade este nuevo contacto a dicha app si no lo hace ya automáticamente.
- La app mostrará toda la información de dicho contacto, que muchas veces incluye una foto, nombre completo y localización.

Estás técnicas y más técnicas automáticas para investigar personas puedes verlas documentadas **en este video,** que incluye la **obtención de información de la *dark web***: https://www.youtube.com/watch?v=ywUxkvCK96w . Aunque esto pueda parecer que tiene un único propósito delictivo, no es así: muchas organizaciones lo hacen porque así saben cuándo **se están filtrando datos robados de ellos mismos**.

OSINT sobre máquinas: Metadatos masivos con FOCA

Introducción

Dada la enorme importancia de la información que puede obtenerse con metadatos de los recursos de una web, vamos a usar **una herramienta que permite obtener y relacionar automáticamente los metadatos de toda una web completa**. ***FOCA*** (*Fingerprinting Organizations with Collected Archives*) es una herramienta utilizada para encontrar metadatos e información oculta en los documentos que examina, que normalmente se encuentran en páginas web y que *FOCA* puede descargar y analizar. Permite analizar muchos tipos de documentos, comúnmente ficheros *Office*, *Open Office*, o PDF, aunque también analiza ficheros de *Adobe InDesign*, o SVG por ejemplo.

Estos documentos se buscan utilizando buscadores como *Google*, *Bing*, *Shodan* y *DuckDuckGo*. La integración de la información encontrada por estos buscadores hace que se puedan conseguir muchos datos. También existe la posibilidad de añadir ficheros locales para extraer la **información EXIF de archivos gráficos**, y hacer un análisis completo de la información descubierta a partir de una URL. **FOCA unifica la información de los metadatos extraídos de todos los ficheros**, tratando de reconocer qué documentos han sido creados desde la misma máquina, qué servidores y clientes los alojan, las personas que han editado los documentos, etc. En otras palabras: a medida que va estudiando los documentos de una página web va haciéndose un "mapa" de todo lo que puede averiguar a partir de ellos.

¿Qué diferencia hay entre esto y lo que hemos hecho en los laboratorios con `exiftool`*?* Precisamente que puede relacionar automáticamente los metadatos encontrados en todos y cada uno de los documentos de la web. Por ello, si te encuentras con muchos documentos con unos metadatos como los de este ejemplo:

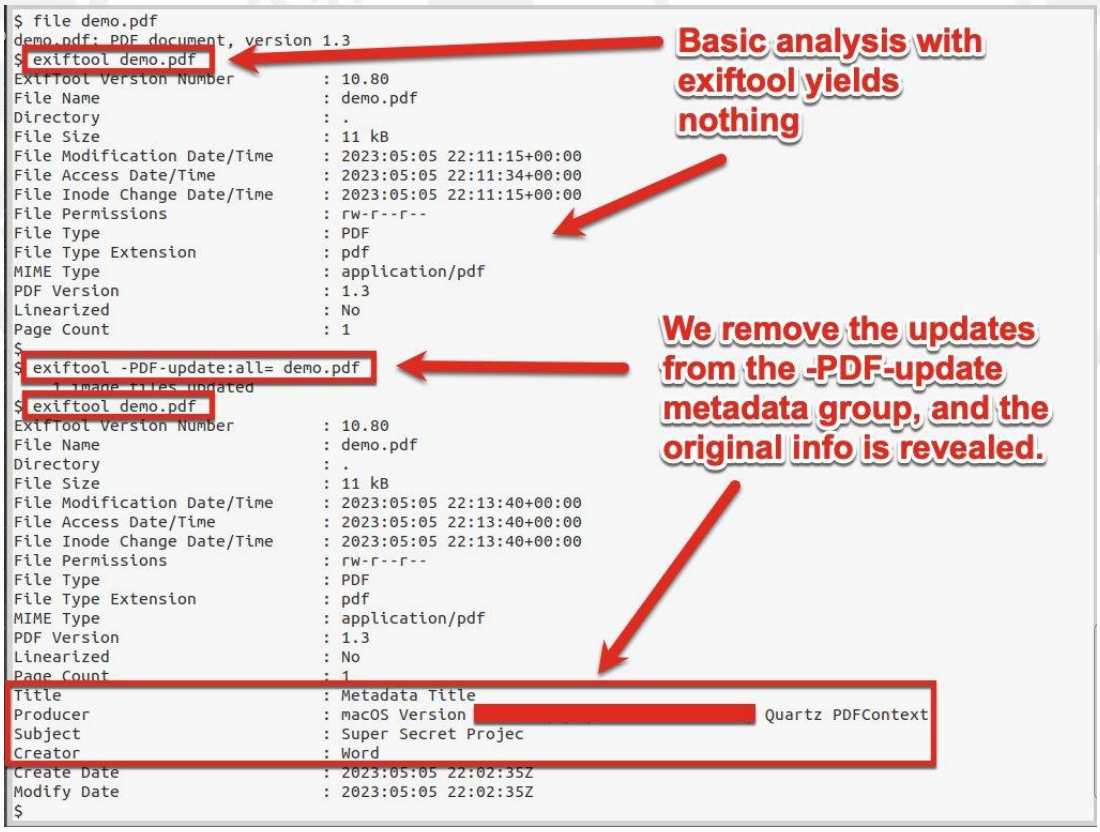


Figura 92. Extrayendo metadatos de una imagen individual de forma avanzada. Fuente: https://twitter.com/cyb_detective/status/1705261248512712781

La cantidad de información que podríamos encontrar puede llegar a ser enorme.

Instalación

Para usar *FOCA* **necesitamos usar una máquina virtual o sistema *Windows 10*** donde instalarla. Además, antes de instalarla es necesario tener una instalación operativa de *SQL Server*. Sirve perfectamente la *SQL Server* de evaluación descargable gratuitamente de aquí (opción "*on premises*"): https://www.microsoft.com/es-es/sql-server/sql-server-downloads

Hecho esto, ya podemos proceder a la descarga, descompresión y ejecución del instalador desde el repositorio oficial de la herramienta: https://github.com/ElevenPaths/FOCA. En el primer arranque la herramienta preguntará por unas credenciales para acceder a la base de datos instalada. Las correctas son:

- ***Server name***: *localhost*
- ***Authentication***: *Windows Authentication*

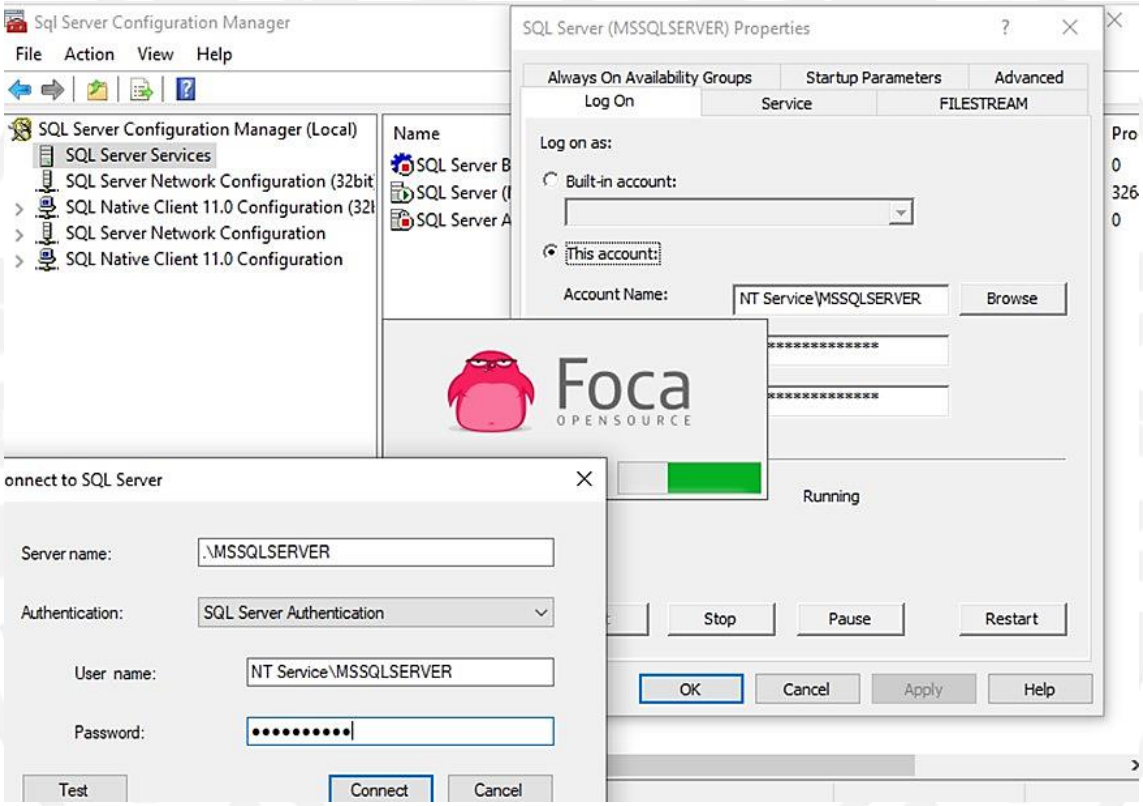


Figura 93. Arranque inicial de FOCA

Con esto ya podríamos entrar a la aplicación y comenzar a usarla.

Uso de FOCA

Lo primero que debemos saber es que *FOCA* funciona a base de **proyectos**, cada uno de los cuales asociaremos a una web que queramos analizar. Para ello, al principio creamos uno de la siguiente forma.

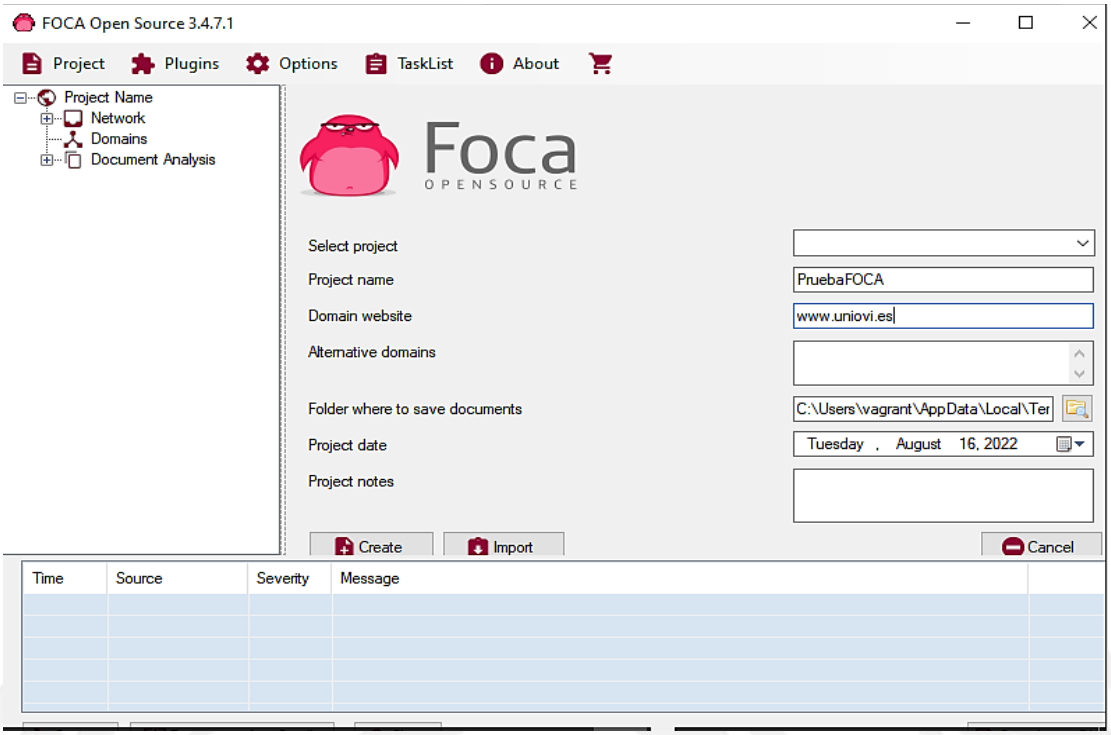


Figura 94. Nuevo proyecto de FOCA

Hecho esto, podemos usar **la opción *Search*** para que la herramienta analice los documentos de la web.

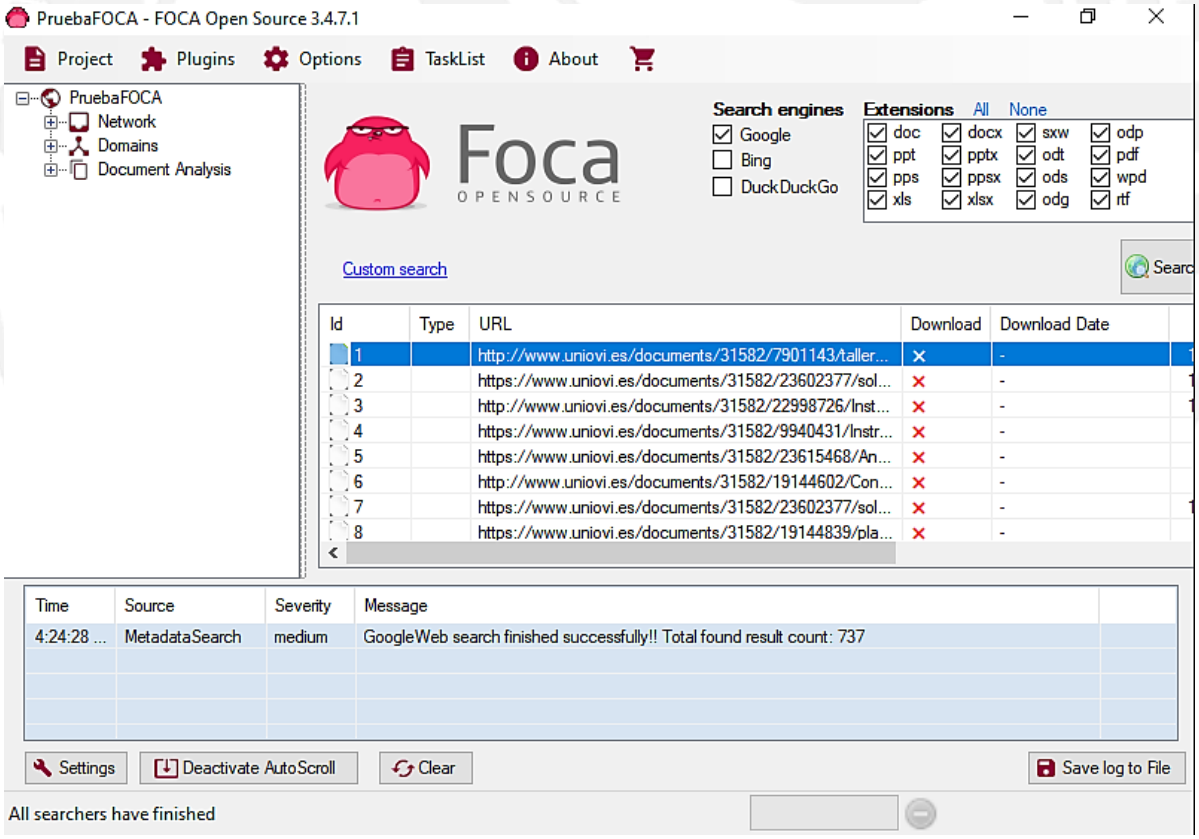


Figura 95. Buscando ficheros de una web con FOCA

Cada documento obtenido puede ser analizado fácilmente **con su menú contextual con el botón derecho**, para lo cual podemos primero **descargarlo** (*Download*) y luego **extraer sus metadatos** (*Extract Metadata*). Date cuenta de que justo debajo tenemos **la opción de hacer lo propio con todos los documentos encontrados en la web**, lo cual es adecuado si no queremos hacer un análisis en detalle de cada uno.

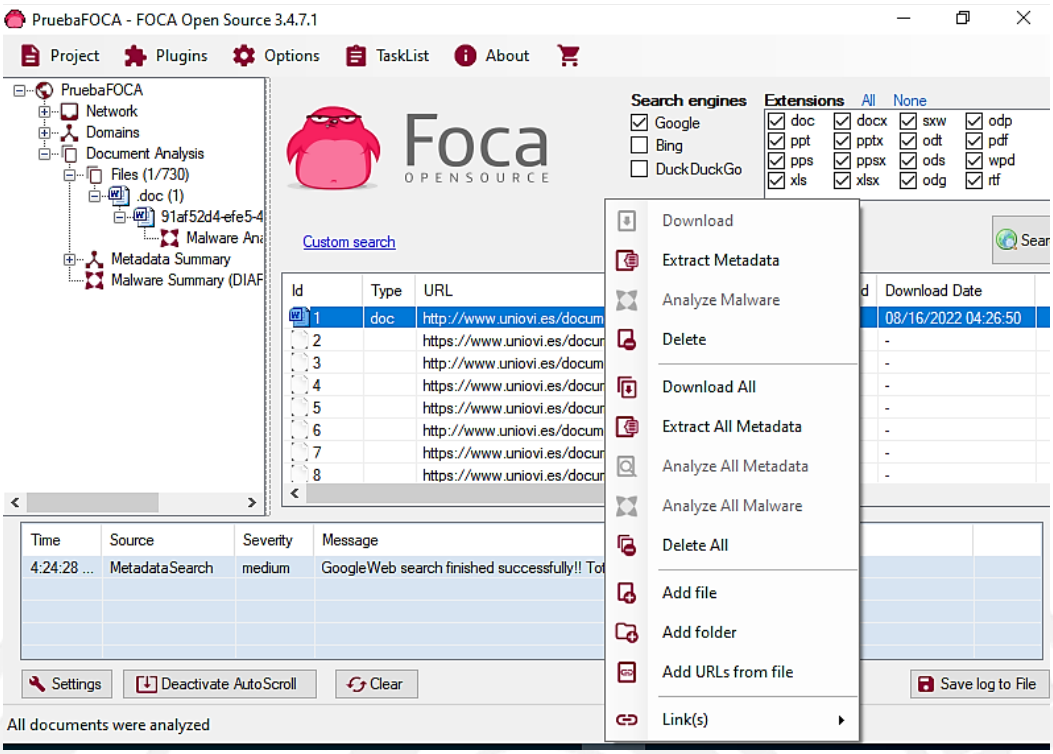


Figura 96. Extracción de metadatos con FOCA

Al extraer los metadatos vemos como ahora nos da toda la información que se ha podido encontrar en el mismo que hemos mencionado anteriormente.

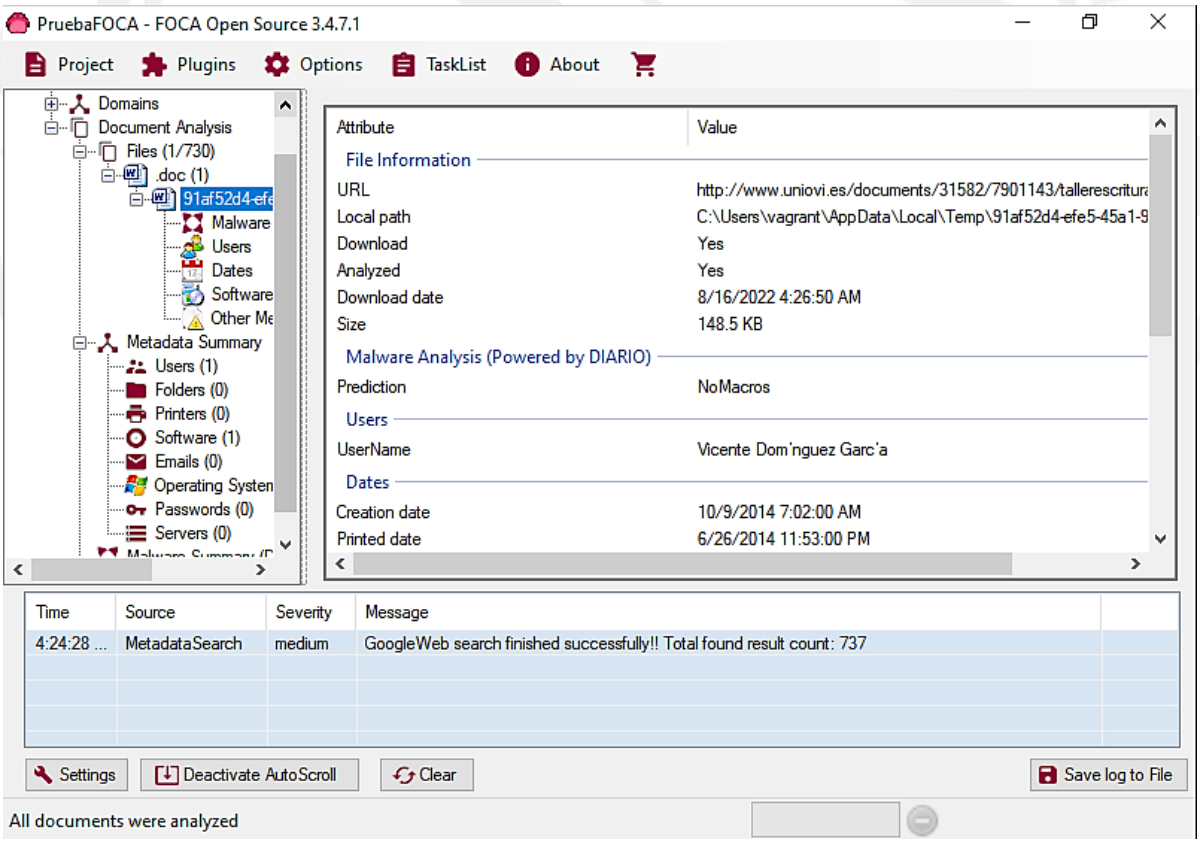


Figura 97. Metadatos de un fichero extraídos con FOCA

Uno de los aspectos más interesantes es que **tiene un sistema *antimalware* integrado** para poder comprobar la posible peligrosidad de cada archivo. También vemos apartados para otros metadatos comunes como usuarios, fechas, *software*, etc.

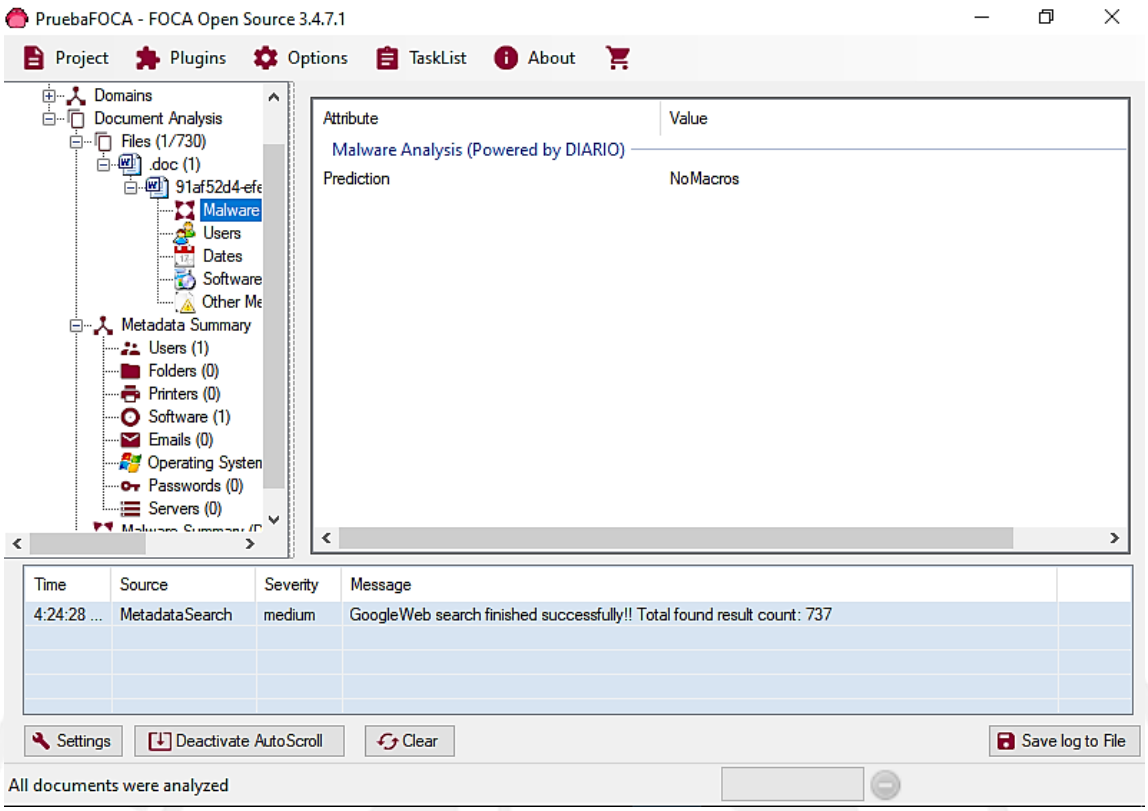


Figura 98. Antimalware de FOCA

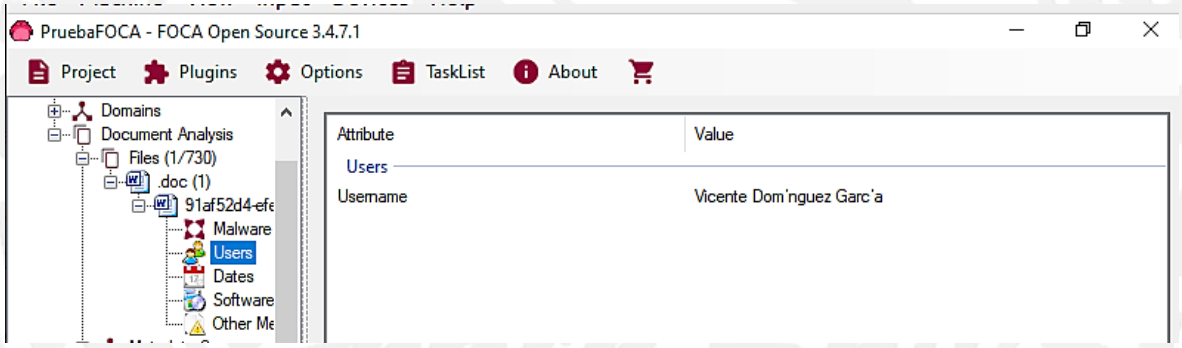


Figura 99. Metadatos de FOCA por grupos (I). Usuarios creadores de documentos

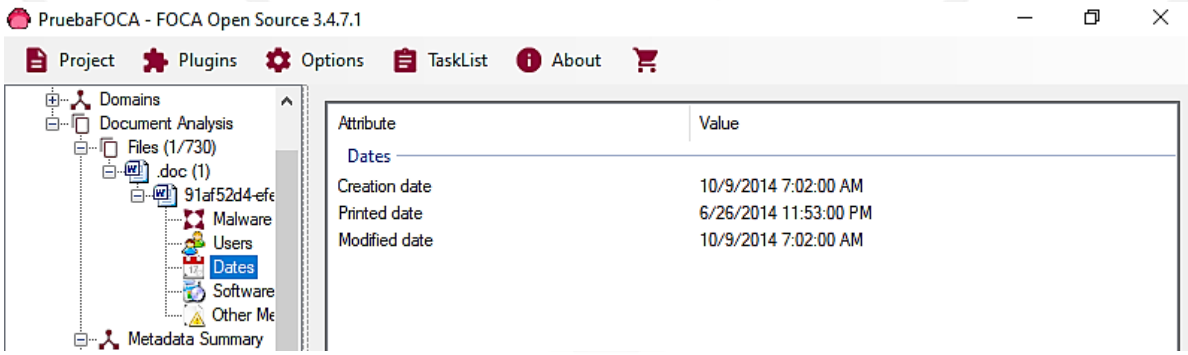


Figura 100. Metadatos de FOCA por grupos (II). Fechas de creación

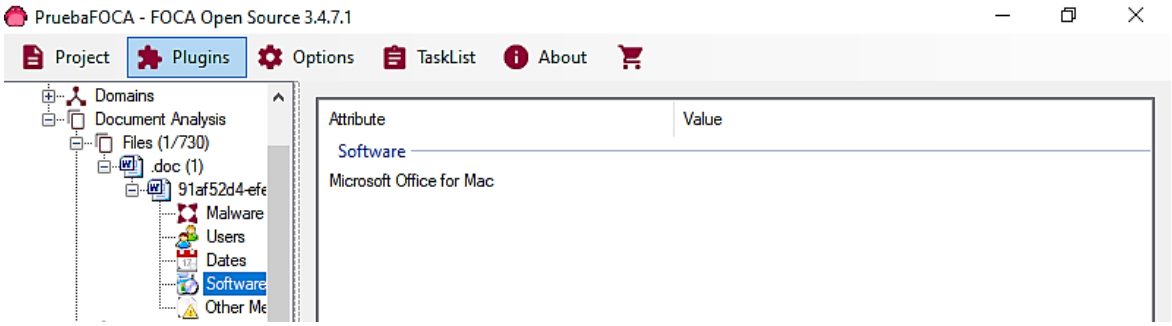


Figura 101. Metadatos de FOCA por grupos (III). Programas usados para crear los documentos

Si ahora descargamos otro documento distinto, vemos como tenemos la misma información **y la herramienta va recopilando y agrupando la de toda la web** en el apartado "*Metadata Summary*" en función de lo que vaya encontrando.

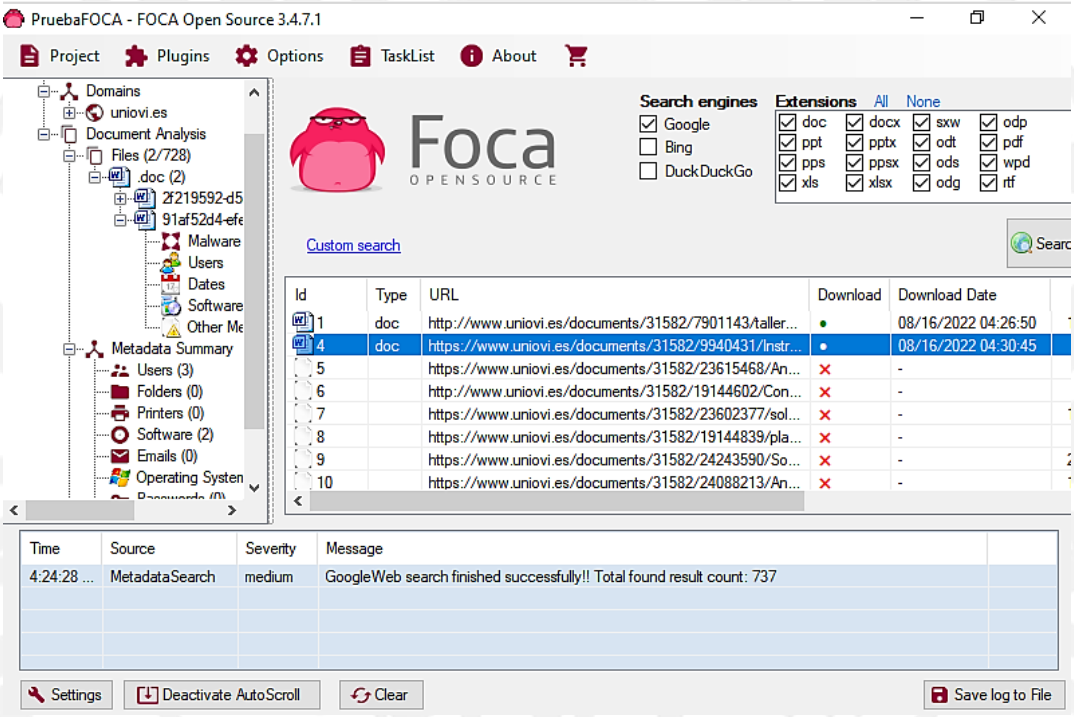


Figura 102. Recopilación de Metadatos de FOCA

En las entradas del resumen podemos ver la lista de usuarios, carpetas, emails, SO, etc. que se vayan encontrando gracias a los metadatos de la web analizada.

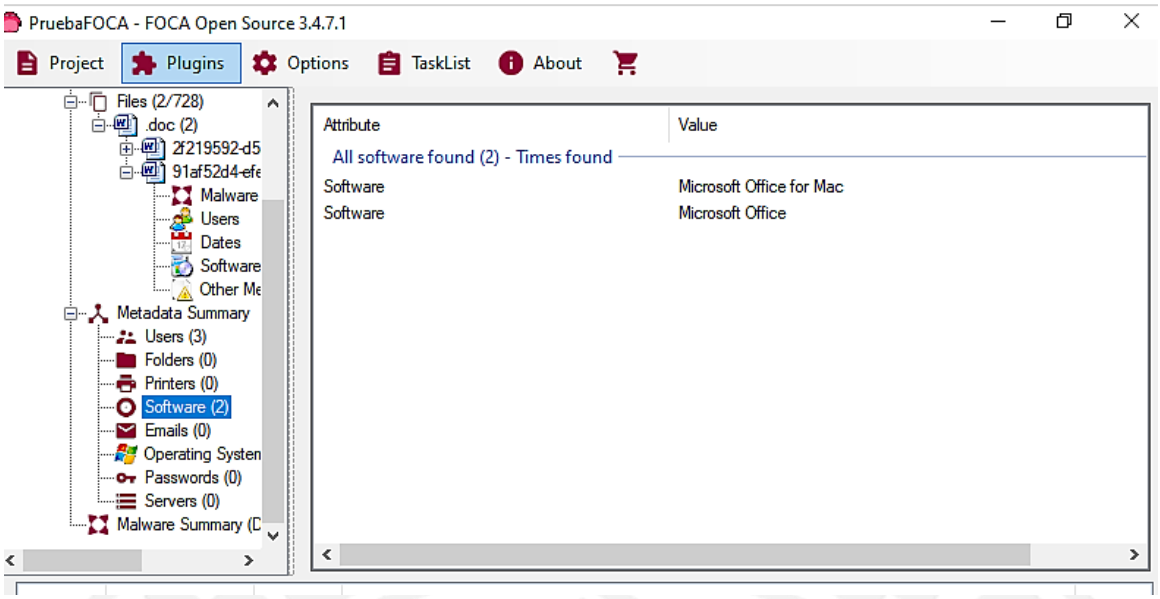


Figura 103. Datos agregados obtenidos por FOCA (I)

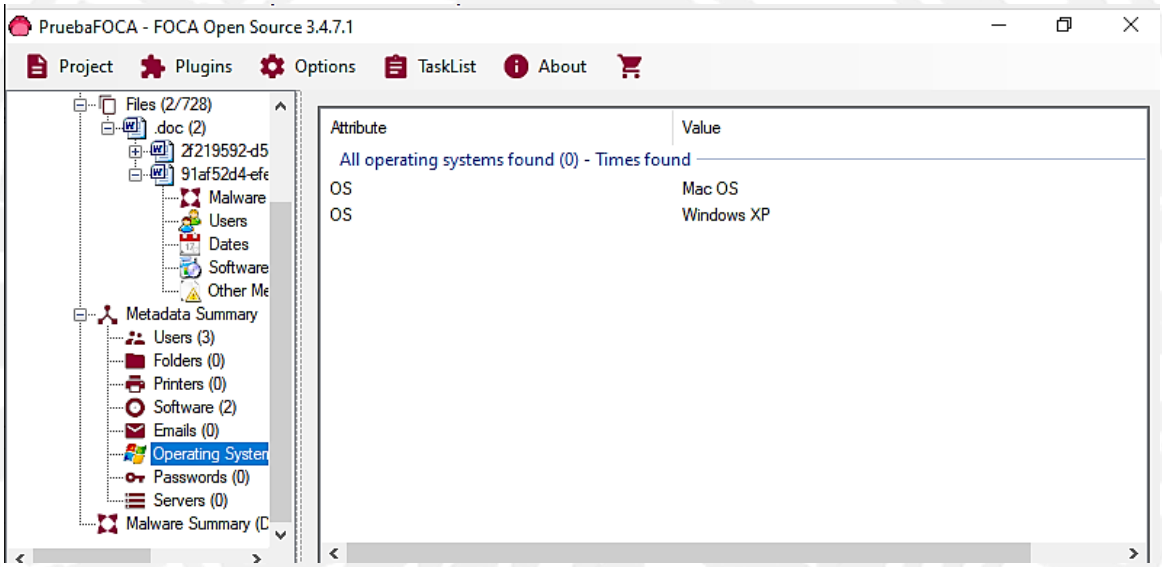


Figura 104. Datos agregados obtenidos por FOCA (II)

En las **opciones** finalmente podemos introducir las claves de las API de otros servicios externos de buscadores que podamos usar para mejorar nuestros análisis, como algunos de los buscadores mencionados.

Minería de datos para OSINT: Maltego

Maltego es una herramienta de **minería de datos** para OSINT muy conocida descargable de https://www.paterva.com/web7/downloads.php. El objetivo de *Maltego* es **representar la información que sabes de una entidad** (persona, máquina, red...) de una forma visual y fácilmente estudiable, usando un grafo de relaciones entre distintos nodos. **Cada nodo representa a un elemento descubierto de un tipo determinado** (hay muchos posibles tipos: dominios, IPs, ….) y cada tipo está representado por una imagen descriptiva. La información de ese grafo se puede construir por medios automáticos, manuales o una combinación de ambos.

Con *Maltego* es fácil por ejemplo presentar a cualquiera **un gráfico** que vincule un resultado con otro, como por ejemplo los perfiles de RRSS asociados a los trabajadores de una empresa (representada por su dominio), sus teléfonos, imágenes, direcciones de correo, etc. Para ello podemos ir añadiendo nosotros la información a medida que la vamos descubriendo con nuestras propias investigaciones, o bien dejar a la herramienta que haga una serie de consultas automáticas contra distintas fuentes de información o datos. Concretamente podemos identificar relaciones entre estas entidades:

- Personas
- Grupos de personas (redes sociales)
- Compañías
- Organizaciones
- Sitios web
- Infraestructura de internet
- Afiliaciones a distintos sitios, eventos, grupos u organizaciones
- Documentos y archivos
- Más entidades, según sea necesario en cada contexto (creadas por nosotros a base de texto o incorporadas de terceros ya que, como veremos, es una herramienta modular)

Maltego tiene versiones de pago y una ***Community Edition*** (CE) que es la única que podemos usar en esta asignatura por la política de cero gasto que llevamos. No obstante, la versión CE es muy potente y más que suficiente para esta asignatura y para todo estudio OSINT que hagamos a baja/media escala.

Instalación

Instalar *Maltego* es muy sencillo, y simplemente hay que descargarlo de la página oficial. Como ves en la siguiente imagen, está disponible para *Windows*, *Linux* y *Mac*.

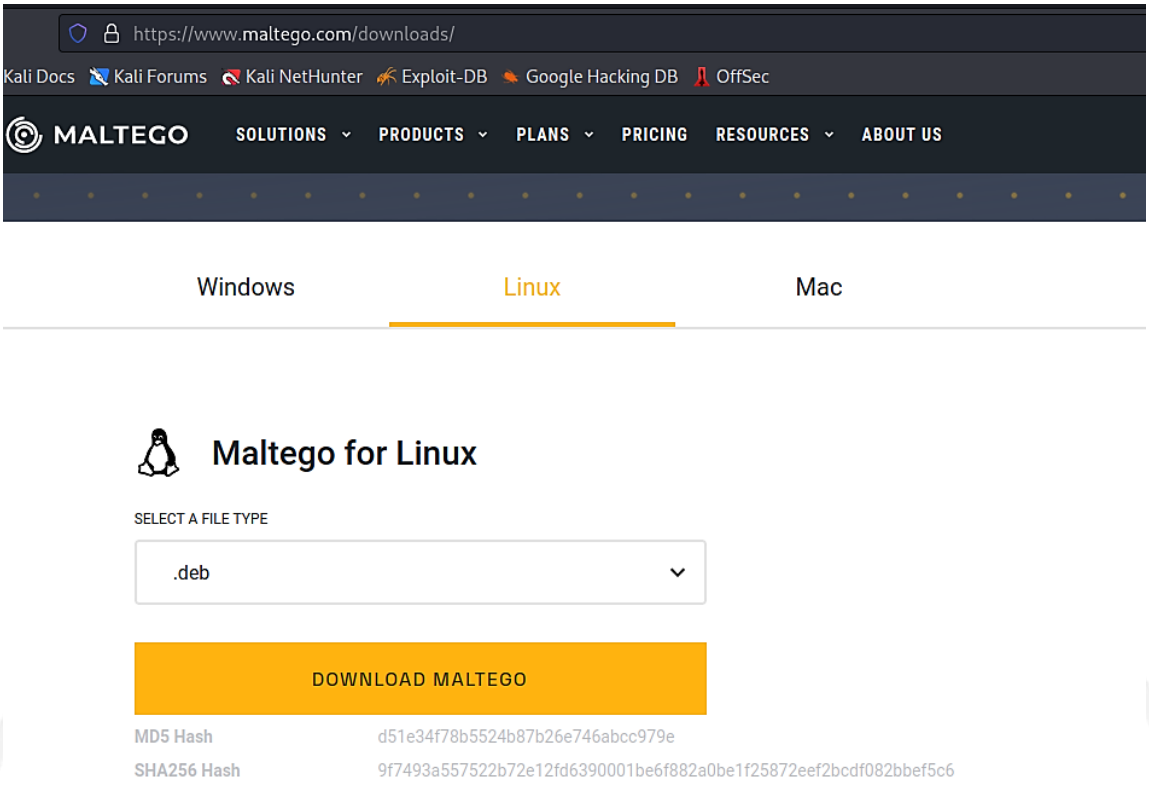


Figura 105. Descarga de *Maltego*

No obstante, si cuentas con sistema operativo ***Kali Linux*** moderno con entorno gráfico, se puede instalar muy fácilmente **de su repositorio de paquetes oficial**, método que es muy recomendable porque tendrás el programa operativo de manera muy rápida y sencilla.

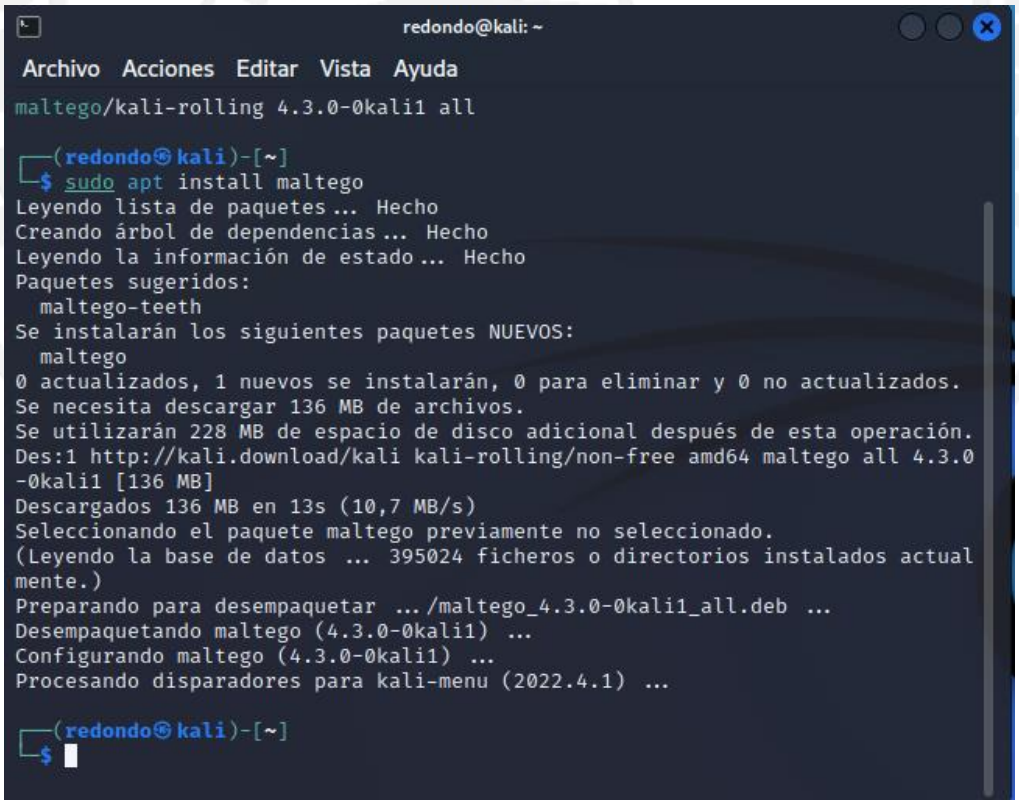


Figura 106. Instalación de *Maltego* desde los repositorios de *Kali Linux*

Para poder usar *Maltego* (incluso su versión CE gratuita, *Community Edition*) **es necesario crear una cuenta en la web de su fabricante** (*Paterva*) a través de este enlace:

https://www.paterva.com/web7/community/community.php. La cuenta es **gratuita** y una vez hecha podremos acceder a la herramienta con las credenciales que creemos. En la pantalla de registro podéis poner que sois estudiantes y que vuestra empresa es la Universidad de Oviedo.

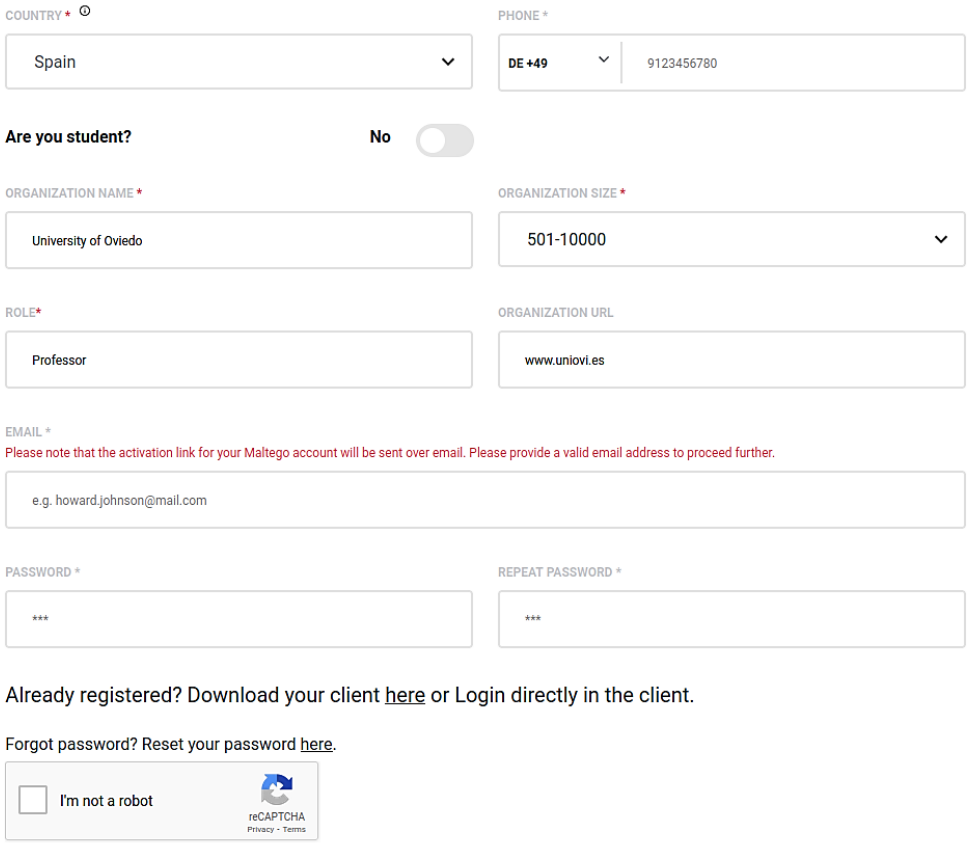


Figura 107. Registro en la web de *Paterva*

El registro mandará un enlace de activación de la cuenta al correo que aportéis, que podría ser uno desechable como los vistos que *Paterva* no prohíba, aunque en este caso tampoco es necesario ser tan precavido porque lo que estamos haciendo es algo perfectamente legal y normal. En cualquier caso debéis aseguraos de que puede leerse, porque es necesario hacer clic en el para que la cuenta se considere correctamente creada.

Primer arranque

Una vez creada la cuenta podemos arrancar la herramienta y **seleccionar la versión CE** (ver imagen) cuando se nos pregunte en el primer arranque cual queremos usar. También podría usarle la versión ***CaseFile***, que está destinada solo a crear gráficos manualmente, sin toda la parte automatizada que vamos a ver. En este caso no nos conviene, pero puede que para otros usos sí (ya que instala más tipos de entidades de base).

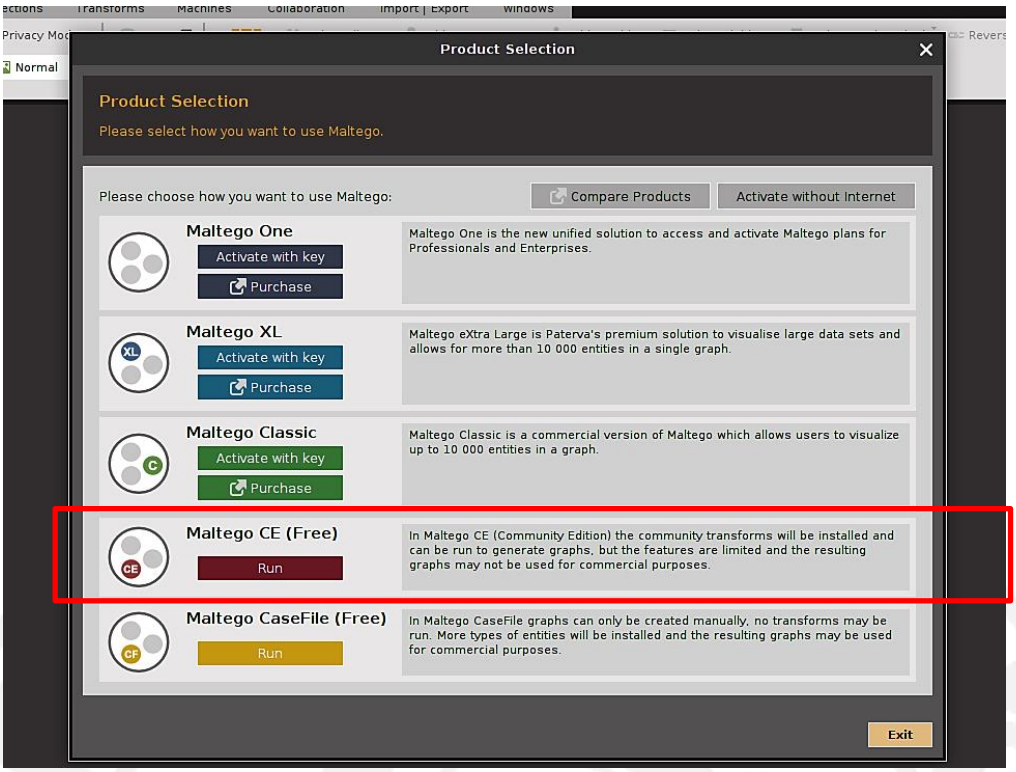


Figura 108. Versiones de *Maltego*, gratuitas y de pago

Una vez seleccionada la versión es hora de introducir los datos de la cuenta de *Paterva* que acabamos de crear, algo que solo deberíamos tener que hacer una vez.

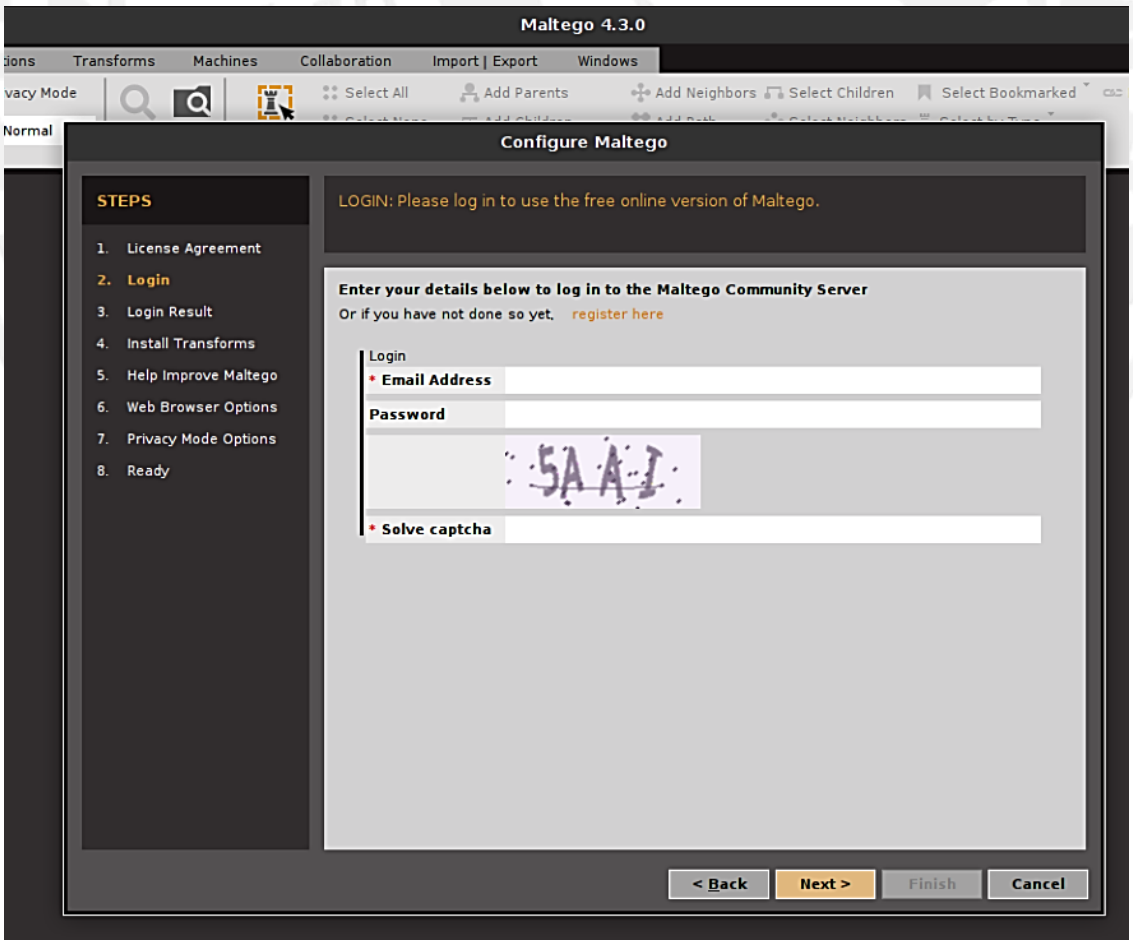


Figura 109. Datos de la cuenta de *Paterva* en *Maltego* en el primer arranque de la herramienta

Si todo es correcto *Maltego* terminará su instalación instalando los módulos necesarios y ya podemos avanzar para terminar su configuración.

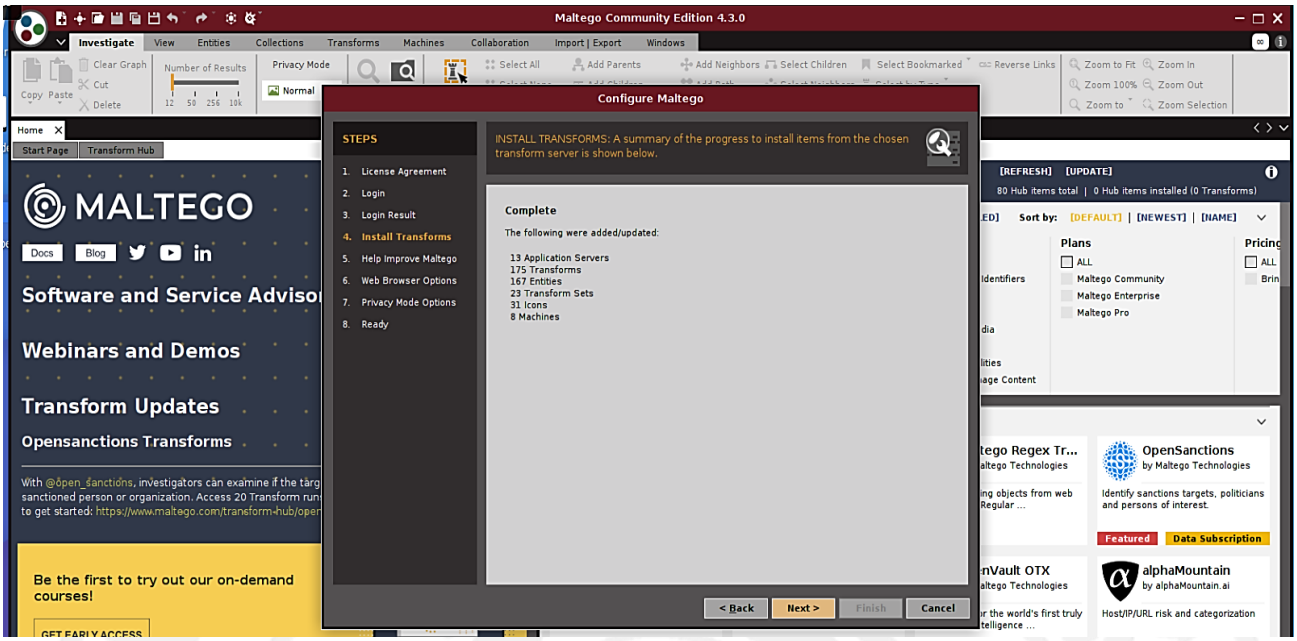


Figura 110. Fin de la instalación de *Maltego*

Ahora tenemos una decisión que tomar. *Maltego* puede funcionar en dos modos: **Privacidad normal** y ***Stealth***.

- En el primer modo nuestro ordenador **se conectará** con los objetivos investigados para descargar ciertas cosas, como imágenes y otros datos de sus URLs. Esto hace que se genere tráfico hacia el investigado desde nuestro PC, que puede analizarse para saber quiénes somos en caso de que el objetivo haga este tipo de pesquisas.
- El otro modo **reduce al mínimo el tráfico que nuestro PC genera al exterior** (no genera tráfico hacia el objetivo, salvo en casos muy concretos), lo cual es importante en escenarios de investigaciones reales serias (como las que puede hacer la policía).

En los ejemplos que voy a mostrar **he elegido el modo normal** porque solo voy a hacer uso de información pública y no voy a hacer nada malo contra ninguna entidad sensible (vamos, no soy un servicio secreto ni un atacante 😊). **El modo normal garantiza también que se obtengan el mayor nº de datos posible.**

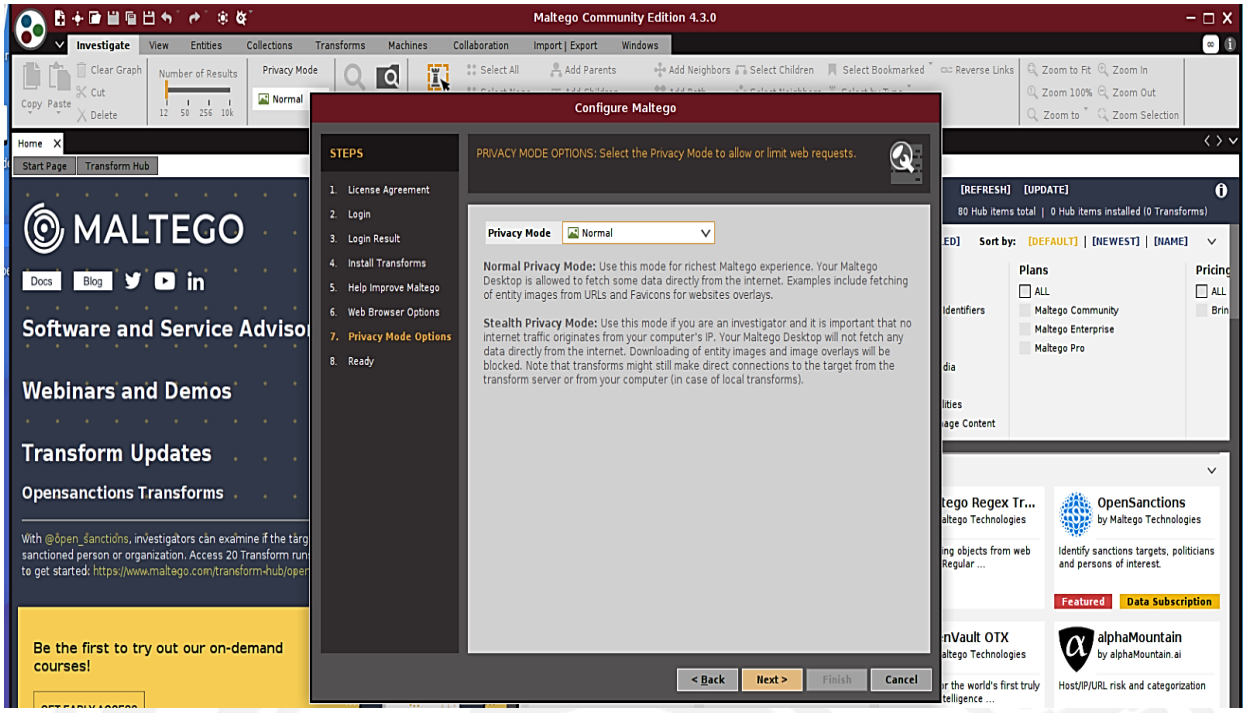


Figura 111. Modos de funcionamiento de *Maltego*. *Are you human or are you hacker?* 😊

Hecha esta configuración inicial ya podemos usar la herramienta para hacer nuestras investigaciones. Para empezar a hacerlo optamos por la última opción y nos quedamos con la pantalla de bienvenida.

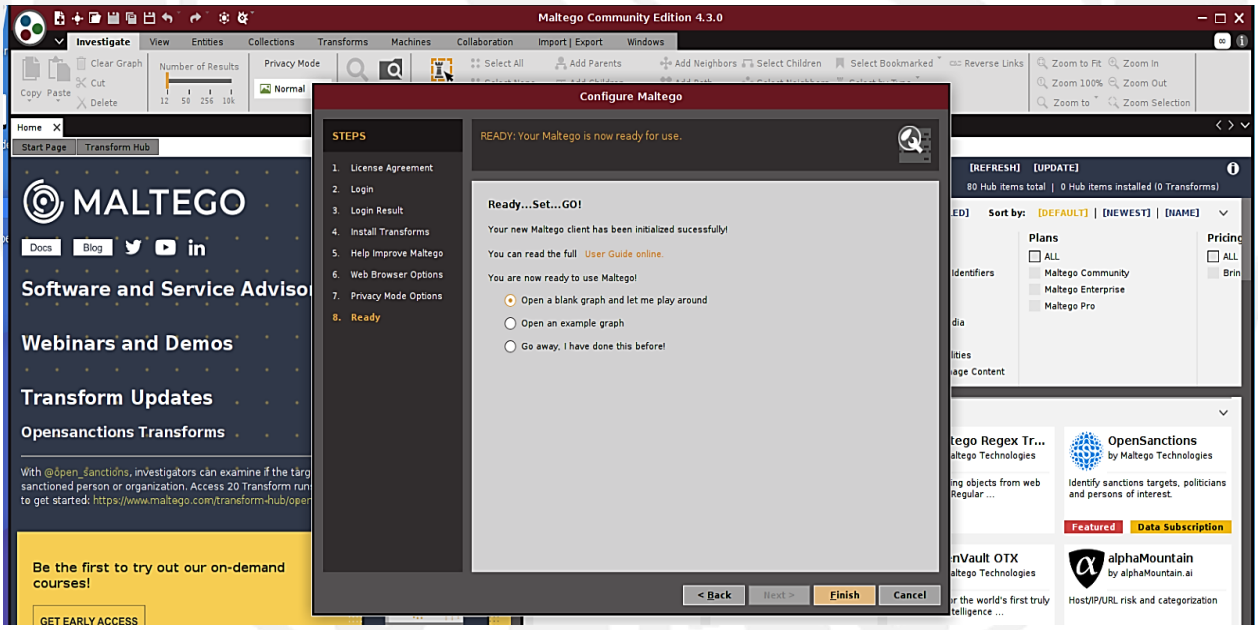


Figura 112. Fin de la configuración de *Maltego*

La pantalla inicial muestra el siguiente contenido, que muestra las distintas operaciones que podemos hacer con la información de las que dispone la versión comunitaria. Estas operaciones podrían ir cambiando con la versión que instalemos.

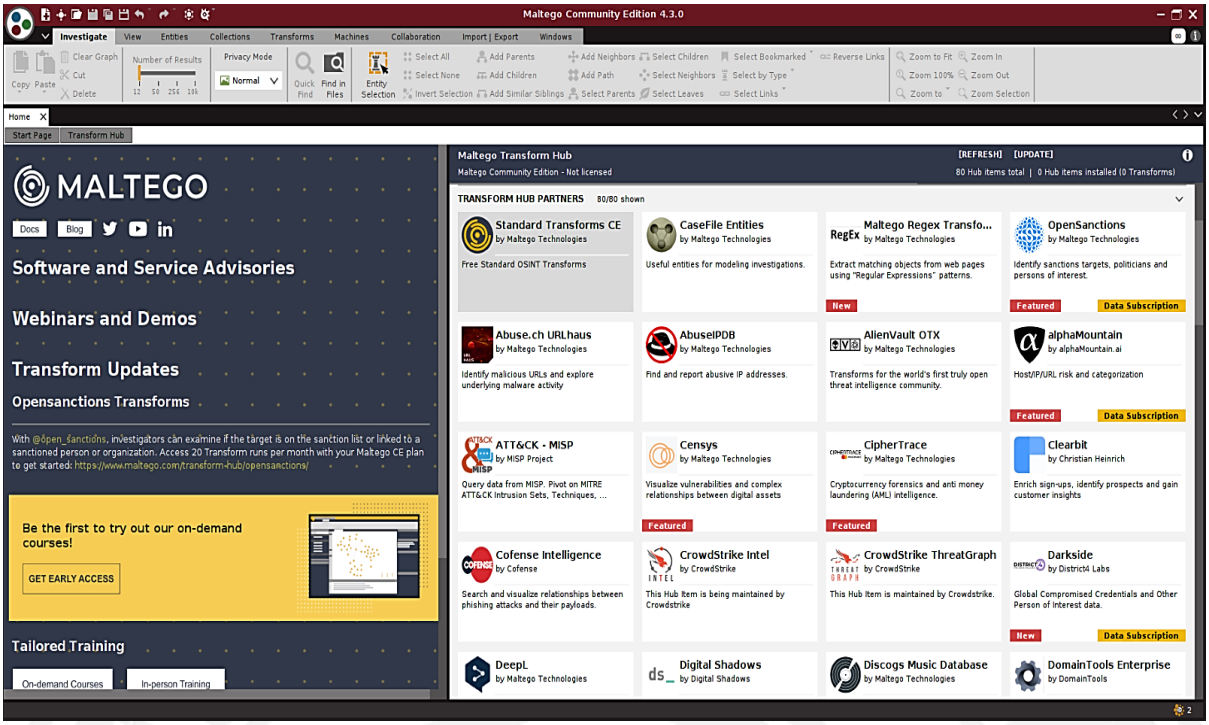


Figura 113. Pantalla de bienvenida de *Maltego*

Creación de un proyecto

***Maltego* funciona con proyectos**. Cada proyecto guarda las relaciones que podemos establecer sobre una entidad, y de esta forma podemos mantener múltiples proyectos para varias entidades bajo estudio. Por ahora crearemos un nuevo proyecto (llamado "*Graph*" en la herramienta) de la siguiente forma:

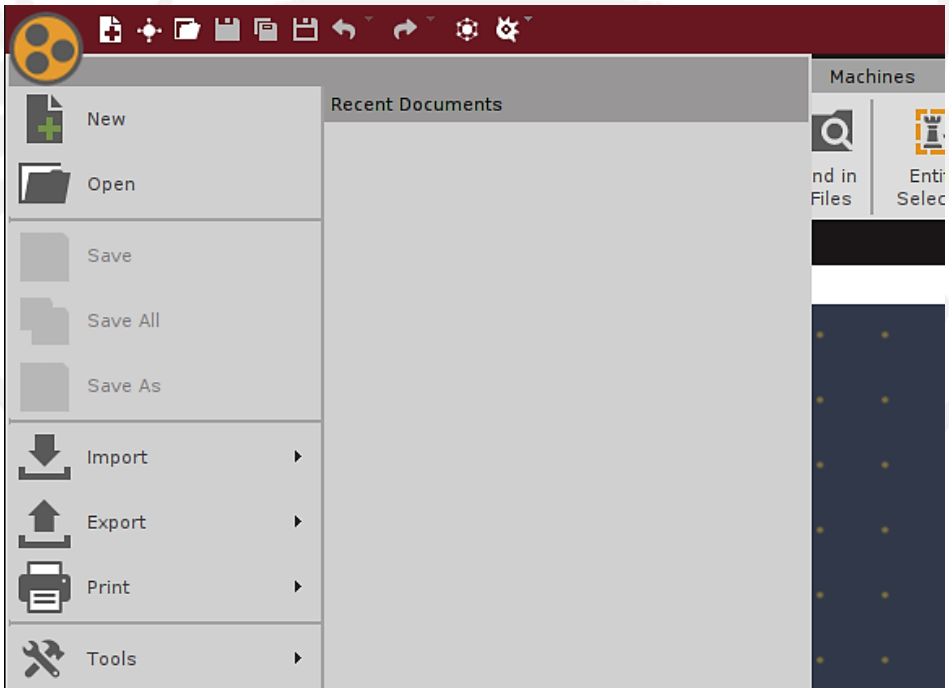


Figura 114. Creación de un proyecto en *Maltego*

Una vez creado, nos abrirá una ventana en blanco y en la barra izquierda nos encontramos con una ventana llamada ***Entity Palette*** (Paleta de Entidades).

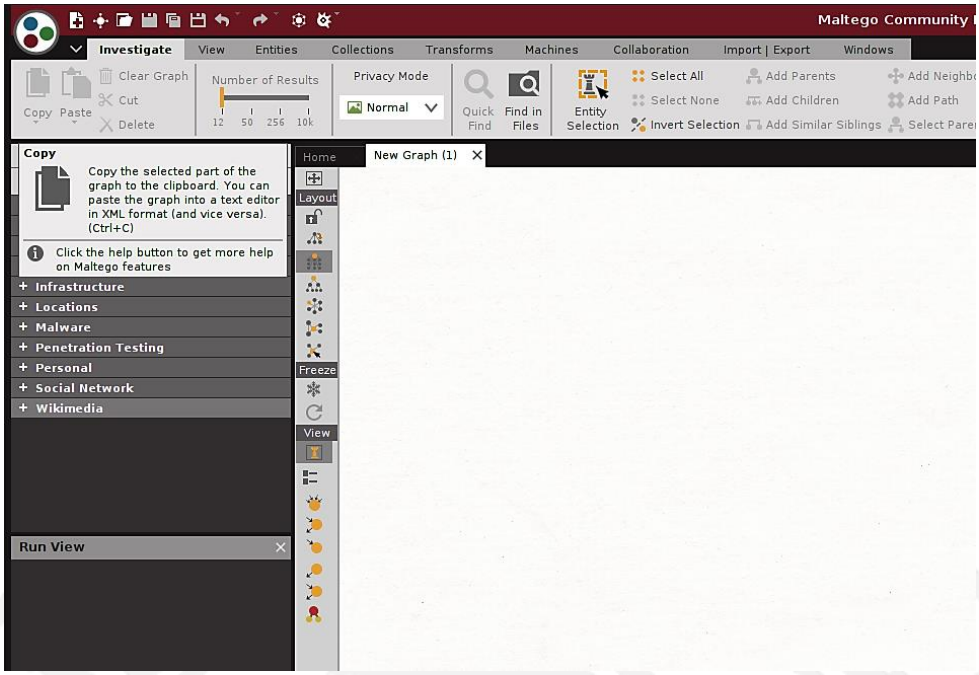


Figura 115. Paleta de entidades de *Maltego*

Para empezar a trabajar lo único que tenemos que hacer es arrastrar una entidad y soltarla en el ***Graph*** en blanco. En este caso escogemos una entidad de tipo "*Domain*", que representará a un dominio cualquiera de Internet.

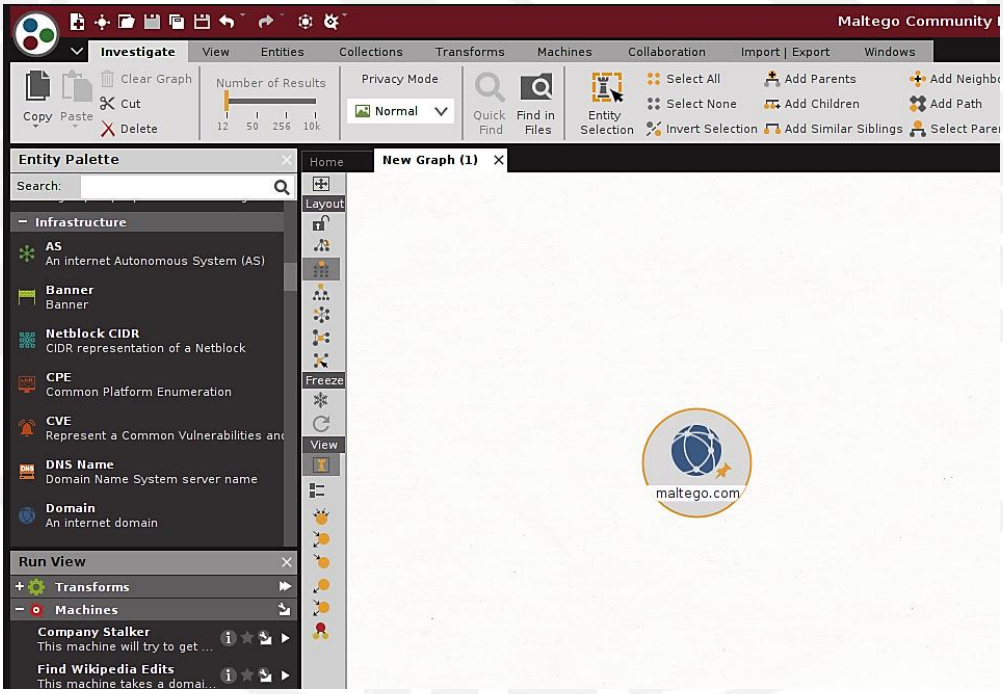


Figura 116. *Graph* de *Maltego* con una entidad añadida

Para **editar** cualquier entidad podemos hacer **doble *clic*** sobre ella o ir al panel derecho en el apartado "*Property view*" y editar el contenido que aparece de ejemplo, en este caso `maltego.com`. Es muy parecido a un editor visual para crear GUIs de aplicaciones. Ahora ya podemos cambiar ese dominio por el que queramos y ¡comienza el juego! 😊

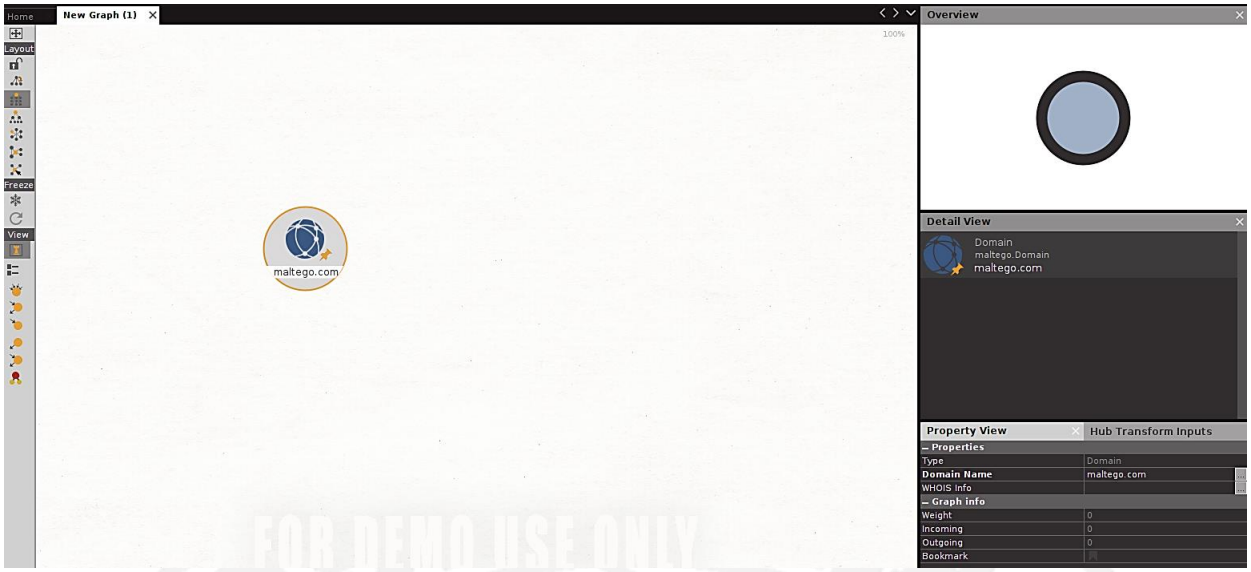


Figura 117. *Property view* de *Maltego*

Búsqueda automatizada de información

Cómo ejemplo vamos a cambiar el dominio anterior por `uniovi.es` y le vamos a decir a *Maltego* que haga automáticamente **una búsqueda de información pública** sobre él. Para hacer eso damos al botón derecho sobre la entidad y vemos las opciones de búsqueda de información de distinto tipo que salen. A cada una de estas opciones se las llama **transformaciones (*Transforms*)**, que representan las distintas búsquedas soportadas por *Maltego* en estos momentos para el tipo de entidad seleccionada. **Obviamente, cada entidad tiene un conjunto de opciones diferente**. Para un dominio tenemos estas:

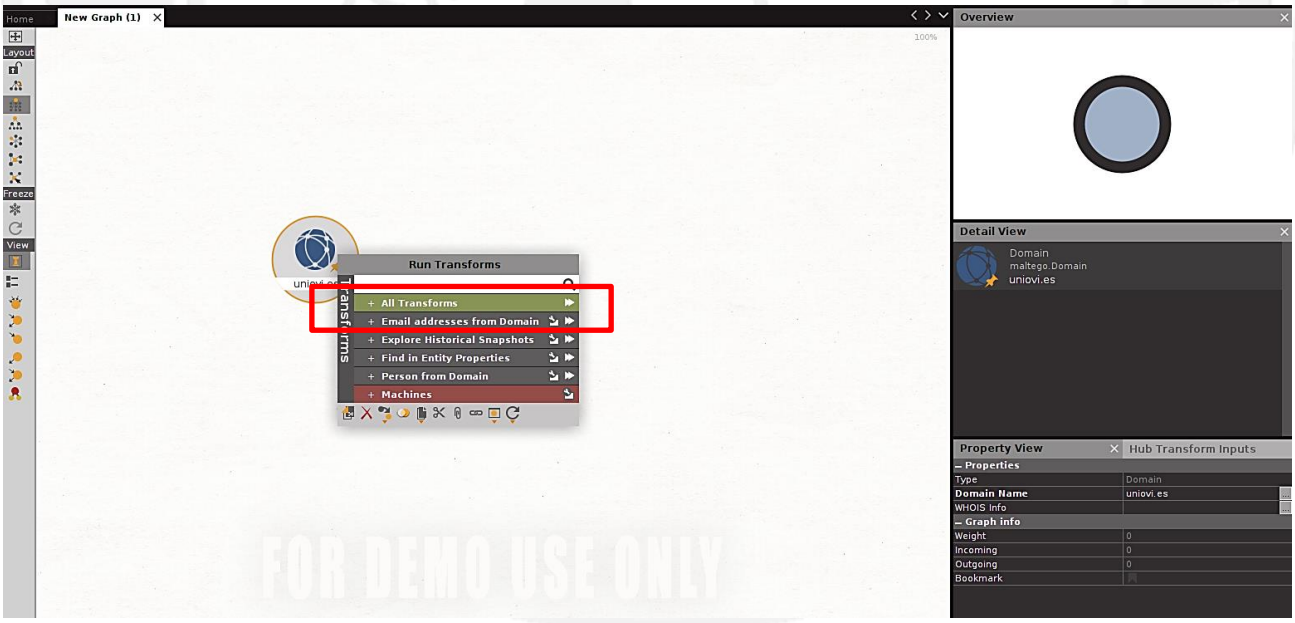


Figura 118. *Transforms* de una entidad "*Domain*" de *Maltego*

Si pulsamos en el nombre de cada ***transform*** podemos ver en detalle cada una de las operaciones que hace la misma. Esto nos permite aplicarlas una a una (pulsando el icono de la flecha que tiene al lado del nombre) e ir haciendo una búsqueda detallada y ordenada de aquellos aspectos que más nos interesen de cada entidad, para así tener un gráfico sólo de aquellas cosas que nos interesen (de la lista de cosas que vimos en el apartado "OSINT *Manual*"). No obstante, también podemos ir "a lo bestia" y pulsar en la flecha de la categoría "***All transforms***", poniendo la herramienta a trabajar contra el ***Domain*** (o la entidad que hayamos

puesto) con todas las transformaciones que conozca sobre ella, a ver qué sale. Un amigo mío llama a esto "a mi señal, ira y *Maltego*" 😊.

Figura 119. Arrancando todas las transformaciones de una entidad en *Maltego*

Por tanto, **las búsquedas en *Maltego* se llaman transformaciones** y la cantidad de resultados que obtendremos va a depender de la versión de la herramienta que usemos. Nuestra versión CE tiene la capacidad de realizar un análisis a 10.000 entidades en un gráfico y puede devolver hasta 12 entidades por transformación. Como decíamos, para hacer una primera prueba usamos "***All transforms***" al dominio `uniovi.es`, sacando un numero de resultados enorme que podemos pasar a analizar. Eso sí, algunas de estas transformaciones necesitan que les especifiques o les confirmes valores de parámetros que controlan su funcionamiento, con lo cual nos los pedirá antes de hacer la búsqueda. En este caso por ejemplo se están pidiendo fechas de principio y fin sobre la búsqueda del dominio sobre la ***Wayback Machine***, que podemos dejar con sus valores por defecto.

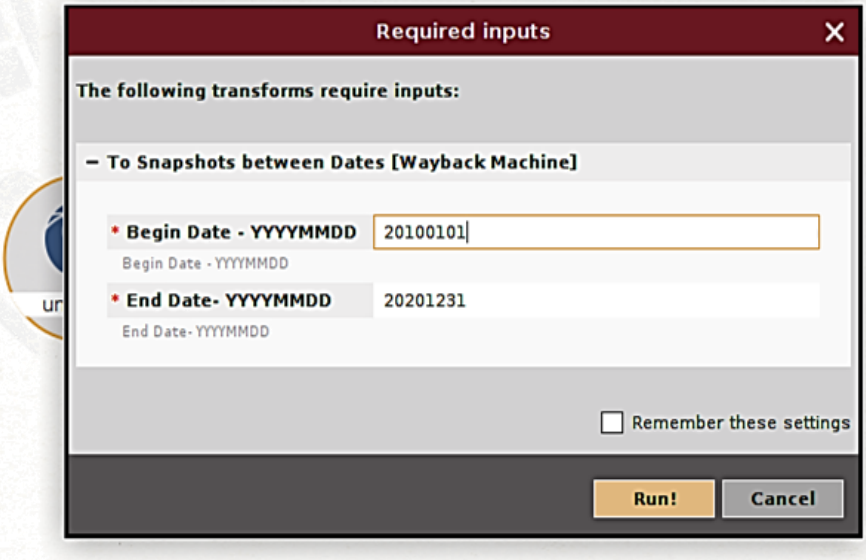


Figura 120. Petición de parámetros de una *Transform*

A veces nos podemos encontrar con que **alguna de las *Transforms* falla** por alguna razón en tiempo de ejecución (por *bugs* que puedan tener), pero normalmente son entidades independientes unas de otras y el hecho de que una falle no implica que lo hagan las demás: tendremos resultados independientemente de ello.

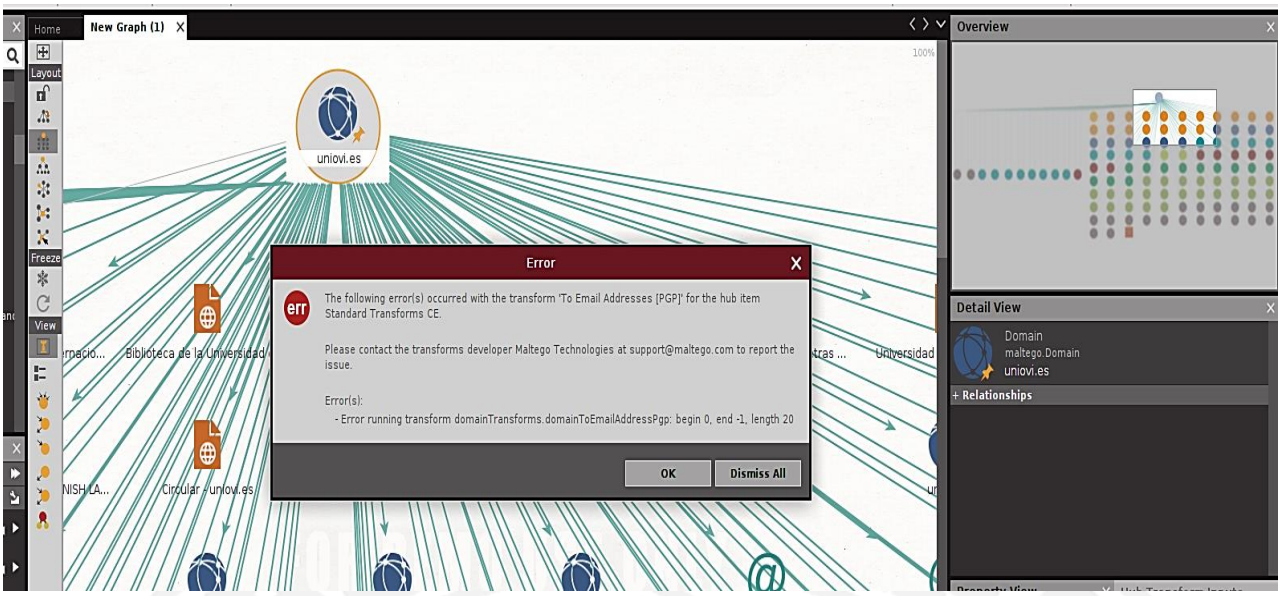


Figura 121. Fallo en tiempo de ejecución de una *Transform* de *Maltego*

Al acabar la búsqueda, vemos que aparecen asociadas al dominio `uniovi.es` **una cantidad enorme de entidades de otros tipos**: documentos, máquinas, IPs, e-mails, personas...y todas ellas pueden ser investigadas individualmente usando la misma técnica de **clic derecho + *Transform*** una a una, con lo que podéis comprobar que las posibilidades de "exploración" de esta herramienta son enormes.

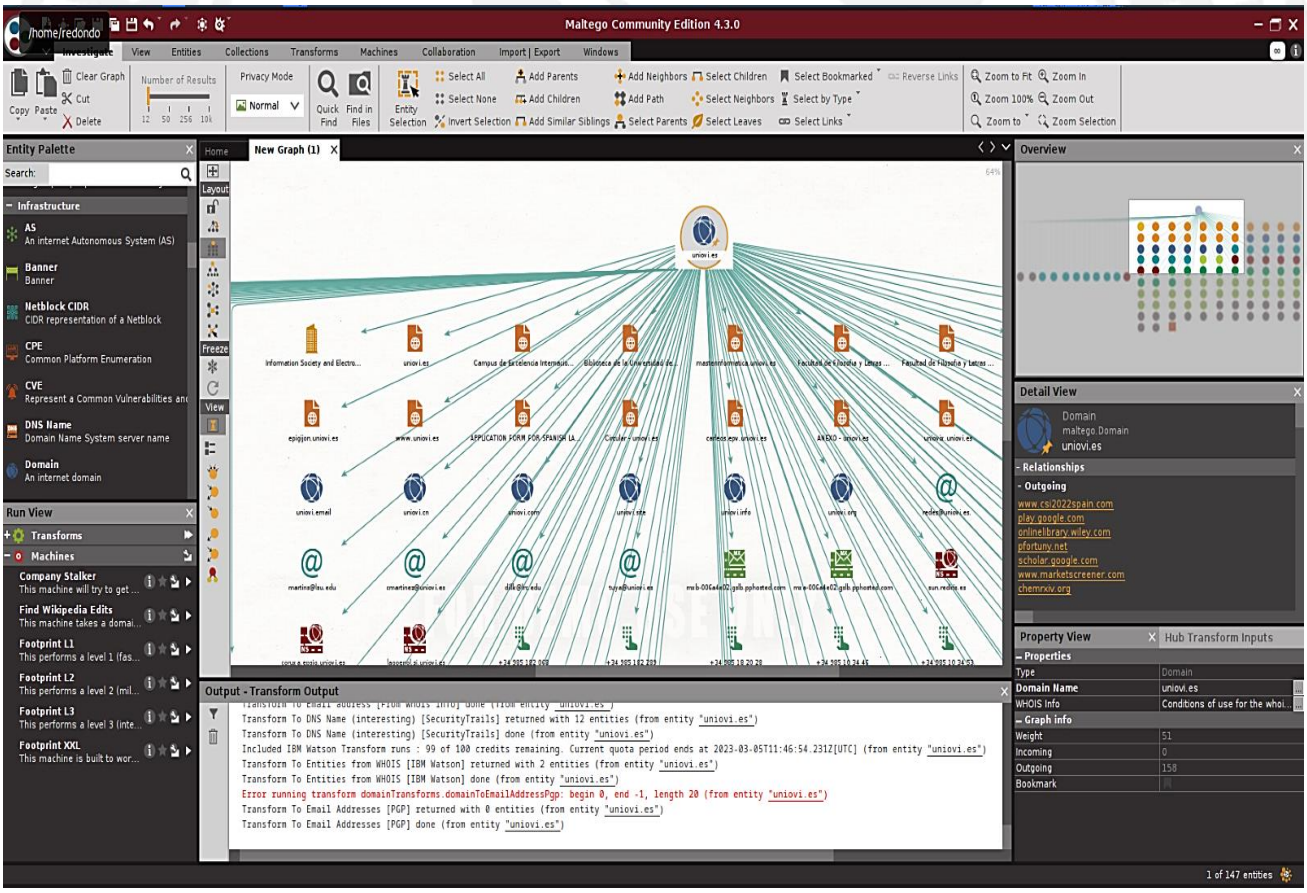


Figura 122. Despliegue de información de un dominio gracias a *Transforms* automatizadas

Estas búsquedas son posibles gracias a las **distintas técnicas que usa *Maltego* para recopilar información**, que nos libera de hacerlas nosotros manualmente como vimos en la sección anterior. Estas técnicas no son distintas a las que se mencionan en otras partes del documento: hacer un `whois` a un dominio, preguntar a servidores DNS, usar diversos motores de búsquedas, buscando en distintas redes sociales como

Facebook, *Twitter*, *Instagram*, etc., pero la herramienta lo hace por nosotros. Es una forma de obtener información "masiva" cuando no tenemos que "hilar fino" y necesitamos tener toda la información posible para empezar a estudiar el objetivo.

El grafo, que puede ser enorme, es navegable y debemos recorrerlo poco a poco en busca de cosas interesantes, como por ejemplo teléfonos y otras webs asociadas al dominio investigado, como vemos en la siguiente imagen

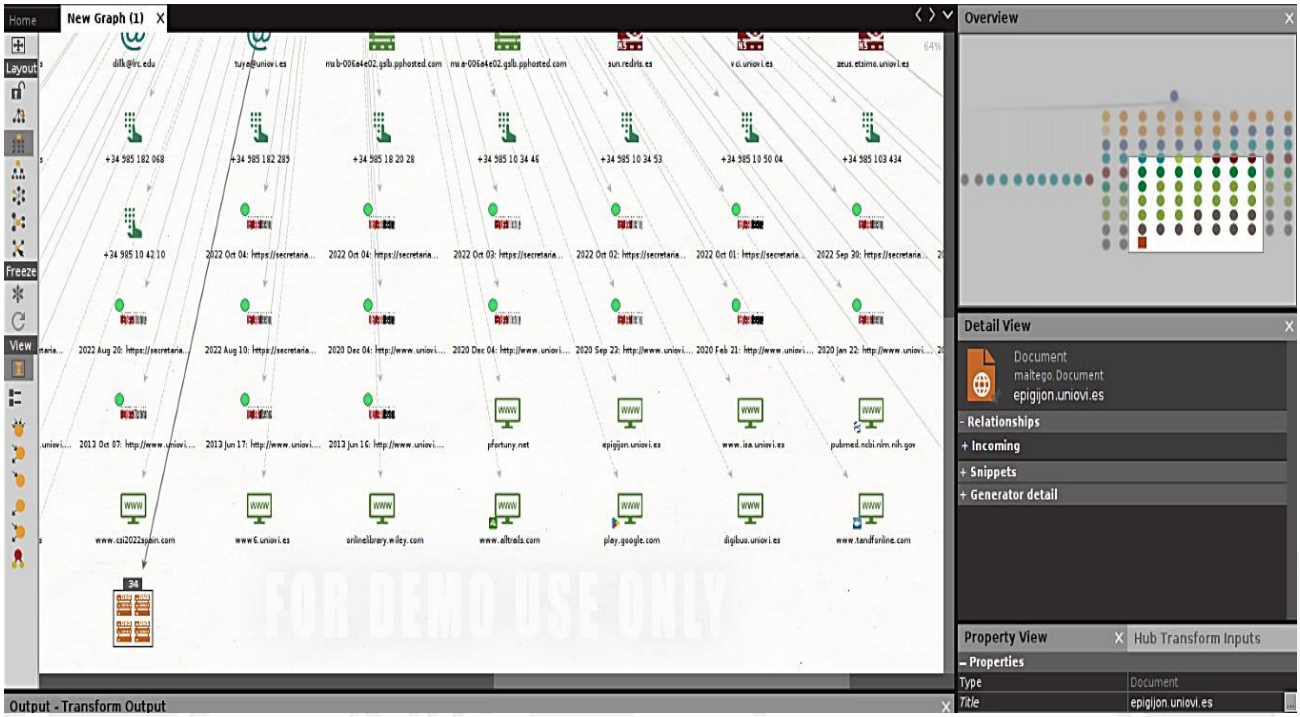


Figura 123. Entidades varias descubiertas por las *Transforms* de *Maltego*

Buscar entre lo buscado 😊

Como dijimos antes, una vez obtenido un conjunto de información, podemos seguir realizando transformaciones a cualquiera de las entidades que representen datos descubiertos por *Maltego*. Vemos por tanto como **el proceso de descubrimiento es completamente recursivo**. Antes lo hicimos sobre un nombre de dominio, que nos derivó en descubrir un subdominio asociado a dicho nombre. Vamos a usar esta entidad y analizarla de nuevo otra vez con "***All Transforms***". Insisto en que esta es una aproximación "basta" al problema, normalmente uno va más "fino" y el detalle, pero aquí hemos venido a jug...experimentar 😊:

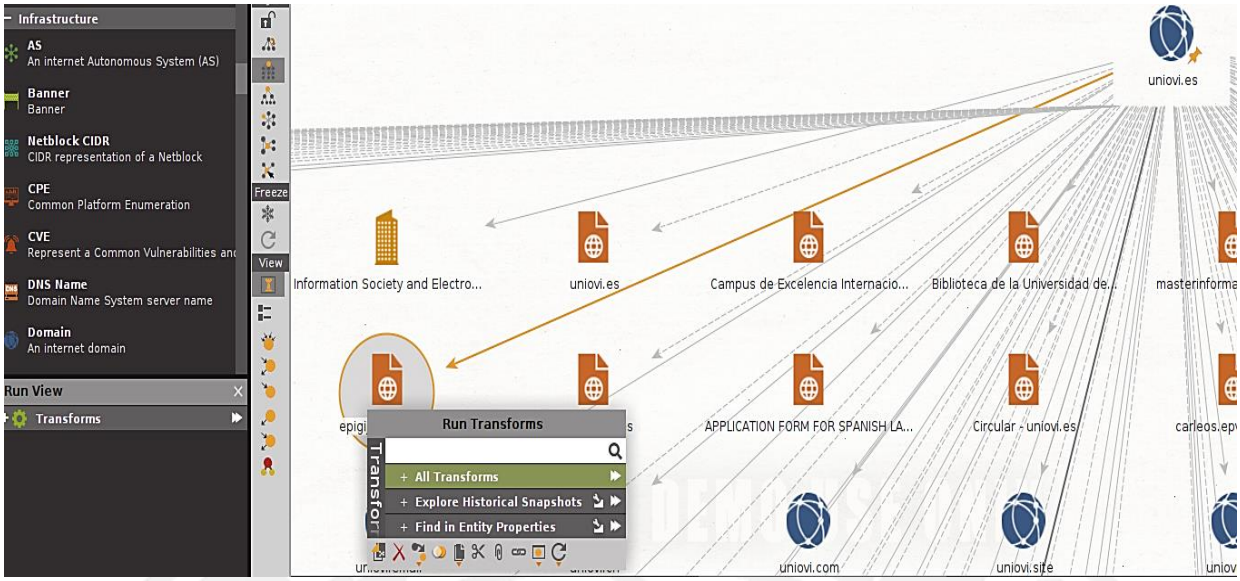


Figura 124. Ejecutando todas las *Transforms* sobre una entidad de la primera búsqueda de *Maltego*

Esto nos da un nuevo conjunto de resultados:

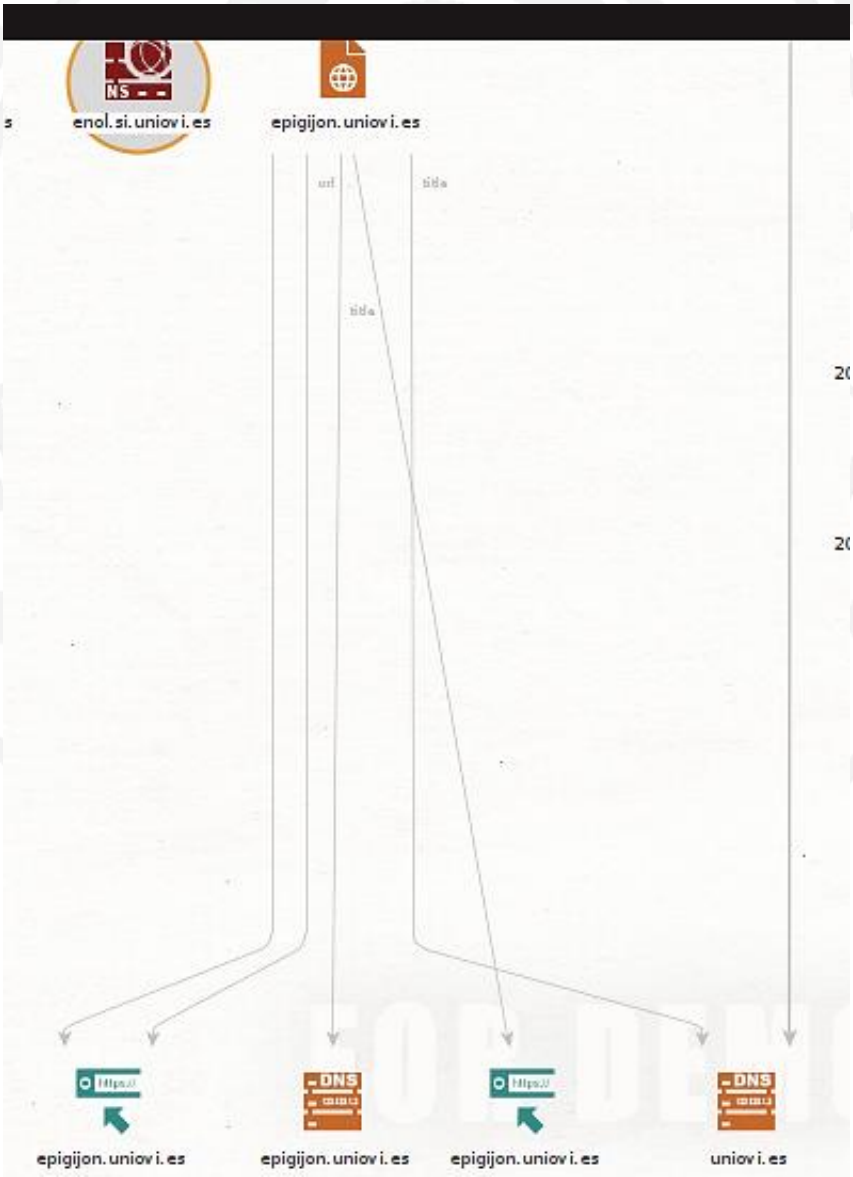


Figura 125. Resultados de *Maltego* tras ejecutar las *Transforms* de un subdominio

Ahora nos sale una entidad sitio web (***Website***) asociada a este dominio, sobre la que procedemos a hacer lo mismo. Nuevamente se nos piden parámetros para las *Transforms* que lo necesiten, y ahora nos aparecen más que antes puesto que las *Transforms* que se ejecutan ahora son diferentes a las de la primera vez (entidad ***Website*** vs entidad ***Domain***). En esta prueba volvemos a dejar los valores por defecto.

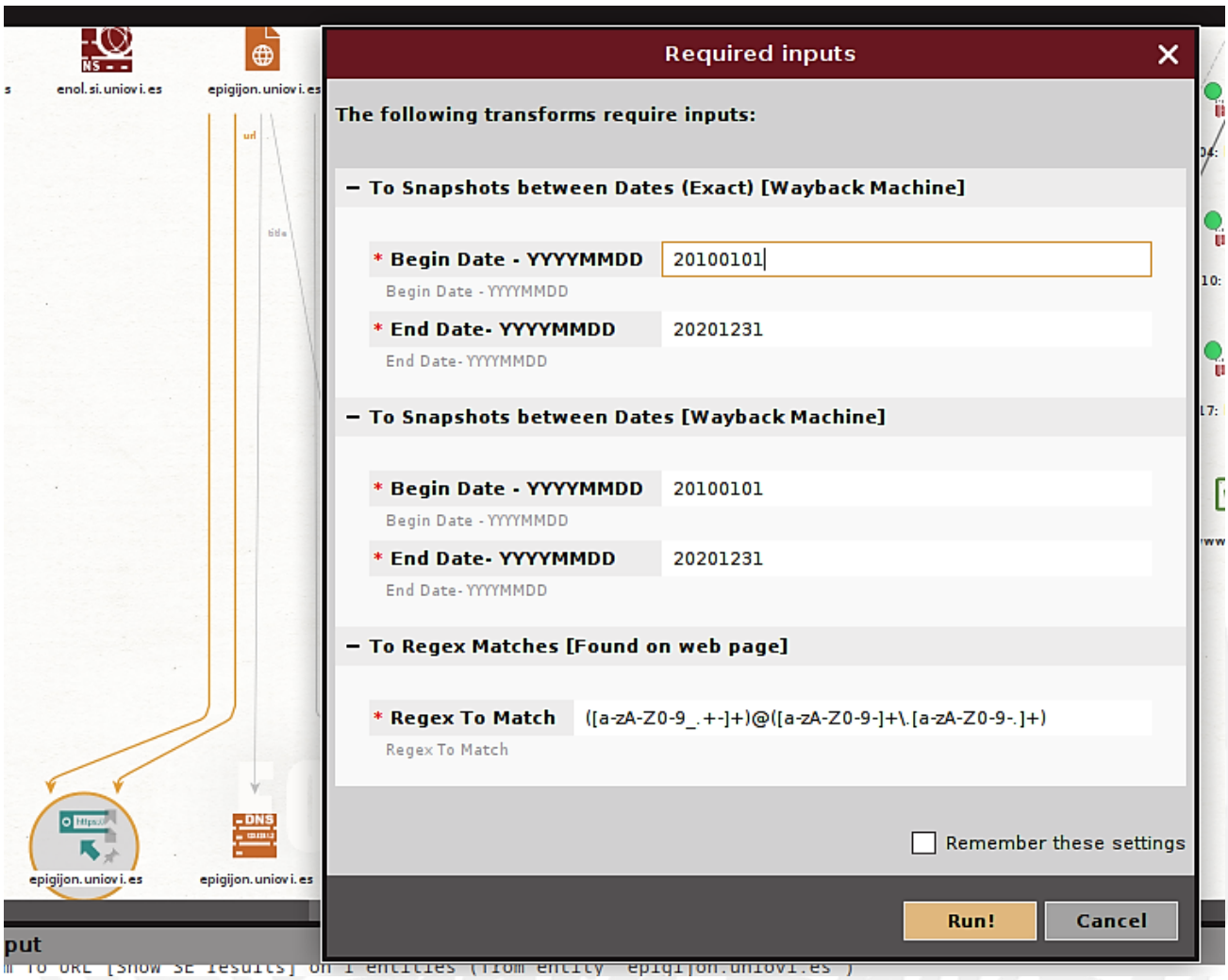


Figura 126. Parámetros de las *Transforms* de una entidad *Website*

Vemos que ahora *Maltego* obtiene otras entidades, como el ***framework* usado para crear la web** y **algunas otras de dudosa fiabilidad** como el uso de pagos en libras o en yenes. Como diremos posteriormente, **no todo lo que *Maltego* obtiene es fiable**, se necesita analizar cada entidad manualmente (con las técnicas de la sección anterior) para ver si es o no fiable dicha información y si realmente nos aporta algo. Por supuesto, **podemos borrar todo aquello que resulte inútil para hacer nuestro gráfico más manejable**.

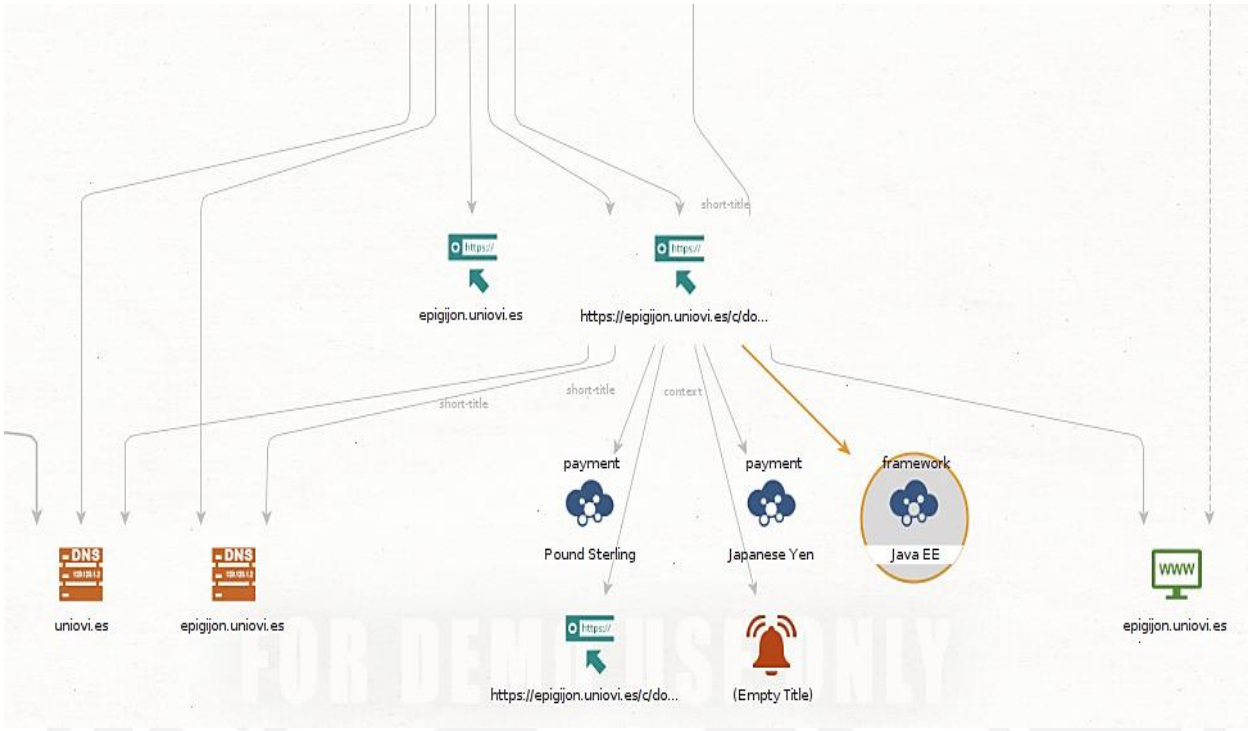


Figura 127. Análisis de una *Website* en *Maltego*

Si seguimos analizando podemos llegar hasta el servidor DNS que se encarga de servir la IP de la web explorada y una IP de este, que luego puede ser explorada por técnicas activas como las que vemos en la asignatura (Ej.: `nmap`). Como vemos, podemos llegar a saber muchas cosas (de alto y de bajo nivel) de cualquier entidad explorada si insistimos lo suficiente 😊.

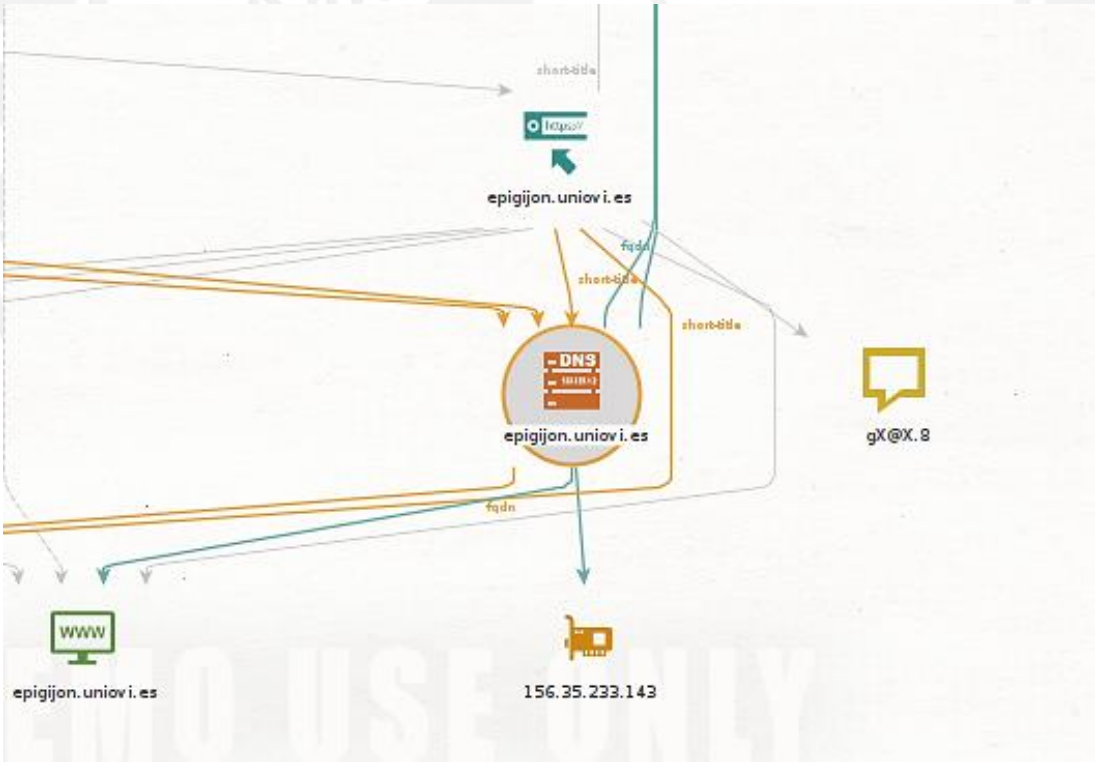


Figura 128. DNS e IP de una web obtenidas con *Maltego*

Vemos por tanto como se puede extender este proceso todo lo necesario, siendo una forma de coordinar toda la información descubierta de una entidad mediante transformaciones "a lo bestia" (como ahora), más específicas (si sabemos lo que estamos buscando) o bien **aportando datos que podamos ir descubriendo**

nosotros mediante nuestras investigaciones más "a mano" con otras herramientas dadas en este documento en la sección anterior o con transformaciones de terceros que veremos luego.

Antes de continuar, es importante destacar una cosa. *Maltego* es una herramienta automática muy potente para minería de datos, pero hemos visto que, como toda herramienta automática, **puede devolver información que no sea del todo fidedigna o directamente falsa**. Por eso tenemos que conocer cómo sacar información con otras herramientas más "manuales": no solo es posible que no tengamos a mano *Maltego*, sino que además tenemos que confirmar que la información que nos devuelve es cierta en caso de que nos sea útil para algo. Tiene por tanto todo el sentido que si *Maltego* descubre una IP por ejemplo, luego vayamos a sitios como ***IPInfo*** (https://ipinfo.io/) para confirmar que esos datos son 100% precisos. **No confíes por defecto en todo lo que *Maltego* te dice** (la probabilidad de que sea correcto es alta, pero no 100%).

Guardar las búsquedas

Como obviamente toda esta información de búsquedas puede ser demasiada para "digerirla" en una sola sesión de trabajo, podemos guardar el gráfico en un fichero para examinarlo en varias veces de esta forma:

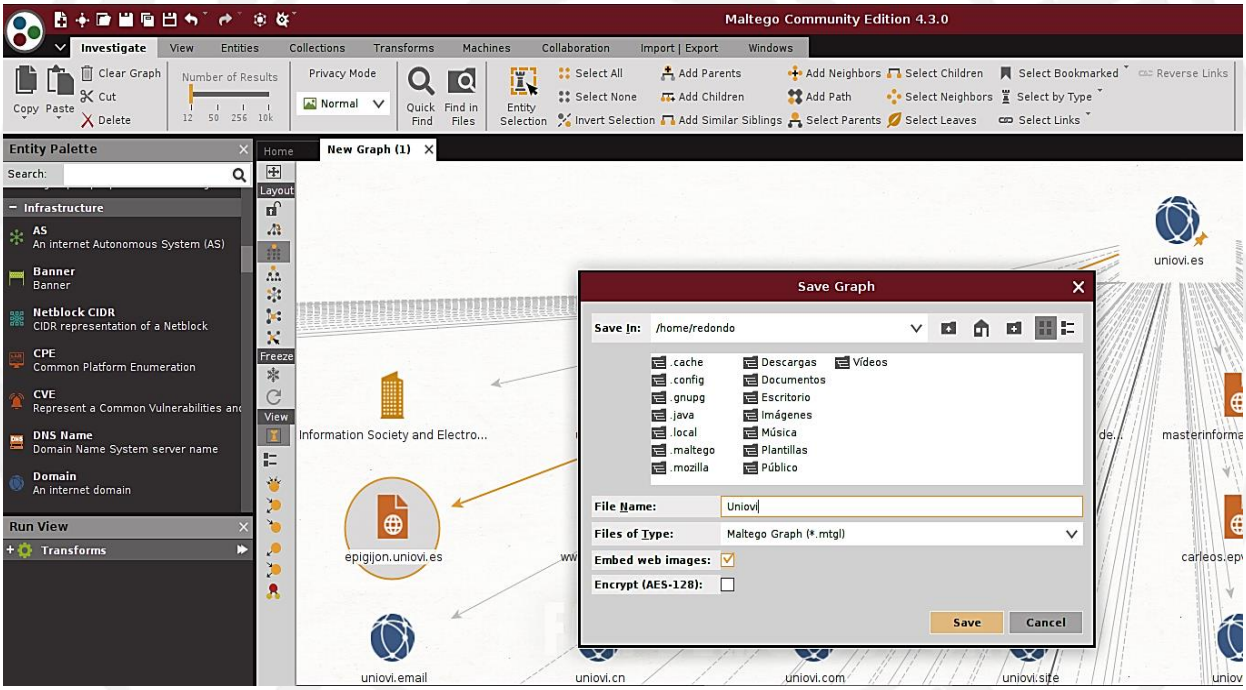


Figura 129. Guardar un gráfico con búsquedas de *Maltego*

Detalle de una entidad de *Maltego*

Hecha esta exploración general el siguiente objetivo es **revisar cada uno de estos resultados y sus entidades**, elegir las que más nos interesen para lo que necesitamos hacer, y **hacer doble clic en ellos para ver sus detalles**. La siguiente pantalla es una típica ficha de una de esas entidades:

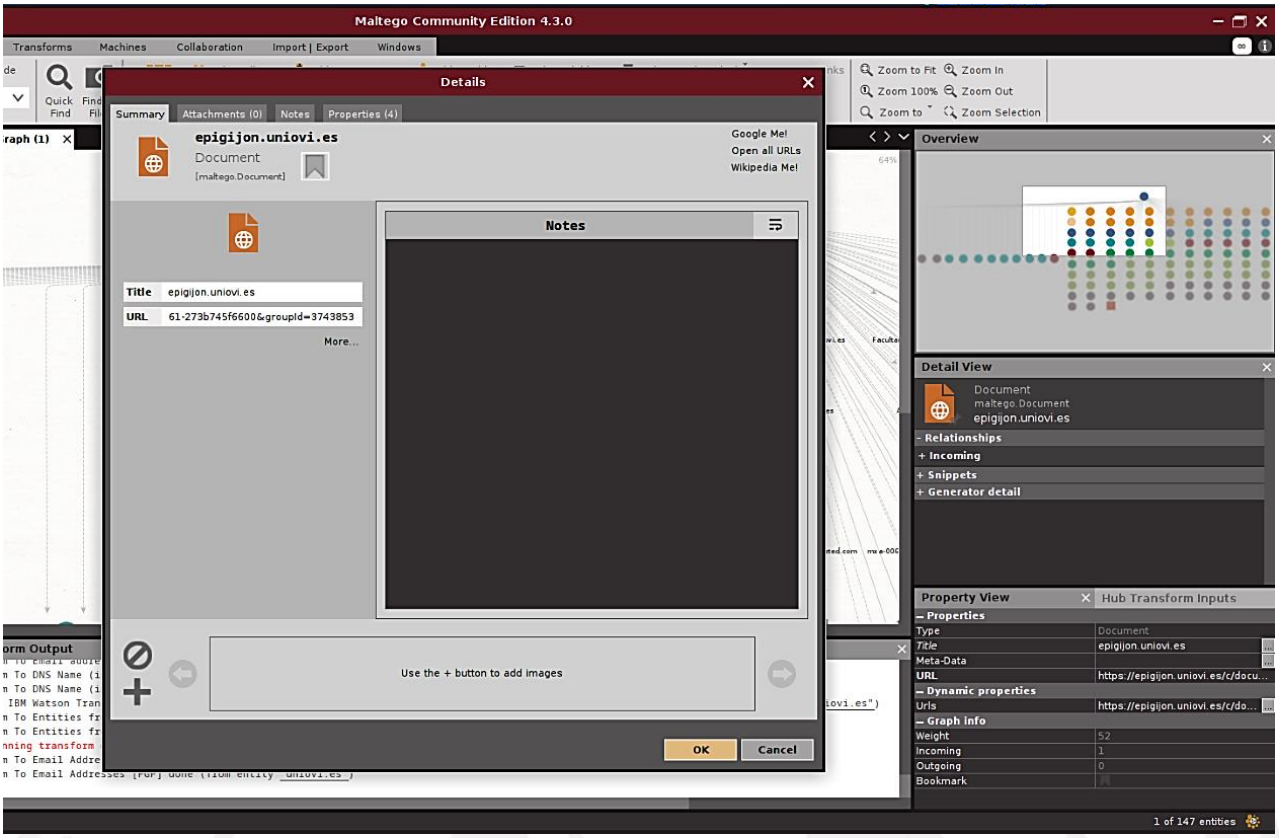


Figura 130. Detalles de una entidad de *Maltego*

Los detalles de una entidad se dividen en:

- **Primera pestaña**
  - Un **resumen** (en la parte superior)
  - La opción de hacer una **búsqueda en *Google***, abrir en el navegador (aunque solo porque la entidad analizada es un dominio) o buscar información de este en la ***Wikipedia***, si no sabemos exactamente qué es la entidad que hemos encontrado.
  - Opción de **añadir imágenes** asociadas a la entidad para usar luego en nuestra investigación (Ej. Captura de pantalla de una web, foto de un perfil...). Ten en cuenta que este grafo **en situaciones reales lo tendrá que ver mucha gente** y cuanta más información podamos poner de cada entidad más fácil será a la gente luego entender la misma o comprobar qué es a simple vista. Cuanto más pienses en los demás que vayan a usar este "mapa" que estás construyendo, mejor. ¡La usabilidad también es importante aquí! 😊
  - Un **cuadro de notas**, para escribir lo que estimemos oportuno a medida que investigamos la entidad, especialmente algo importante que queramos recordar luego (a nosotros o a los demás).
- **La segunda pestaña** es la de archivos adjuntos. Esto se usa si queremos adjuntar algo, como informes de la entidad o persona, ficheros que ha generado en otras partes, capturas de *Tweets* o mensajes en RRSS generales, por ejemplo. Dicho de otra forma, **pruebas y evidencias** sobre la entidad que luego nos sean útiles para algo.
- **La tercera pestaña** es para agregar notas extendidas sobre la entidad. A veces encuentras cosas que requieren una descripción detallada, o dejar notas / preguntas para otros miembros del equipo que estén examinando las entidades del gráfico. **Recuerda**: esto es una herramienta de investigación.

- **La última pestaña** es la de propiedades, donde muchas veces hay información sobre puertos, imágenes, URLs, metadatos u otros antecedentes relacionados con la búsqueda que *Maltego* haya encontrado.

Como ves, una investigación profesional a conciencia con *Maltego* requiere paciencia, conocimientos técnicos (aunque no especialmente avanzados), atención al detalle, documentar adecuadamente los hallazgos combinando elementos obtenidos automáticamente con otros obtenidos de manera manual y, seguramente, la colaboración de varias personas con perfiles diferentes. Pero es, sin duda, apasionante 😊. Cibermaruja++.

Creando gráficos manualmente

Antes hemos dicho que con *Maltego* podemos representar la información que vayamos encontrando en análisis "manuales" que hagamos con procedimientos de OSINT. Para ello la herramienta soporta la creación de gráficos manualmente. *¿Para qué queremos hacer eso?* Pues porque la información también la podemos **encontrar o inferir** de forma manual (como vimos en la sección anterior), podemos hacer suposiciones (que viene bien documentar en el detalle de la entidad), o bien mediante otras herramientas.

Recuerda que en el **Seminario 1** se vieron muchas herramientas que se usan para obtener de forma más o menos automática información de personas, máquinas y de código que se uniría a la que tenemos y que probablemente *Maltego* no pueda encontrar. Todos esos datos tenemos que añadirlos al gráfico que tenemos, arrastrando, soltando y vinculando las entidades encontradas de forma "manual" con las "automáticas" y **enriquecer así el resultado final**.

Para ello debemos arrastrar las entidades que necesitemos, editarlas con la información que hayamos descubierto, dirigir el cursor hacia el objeto que queramos enlazar y mover hacia otro generando un "vínculo" estableciendo algún tipo de identificación de esa relación, como se ve en este ejemplo:

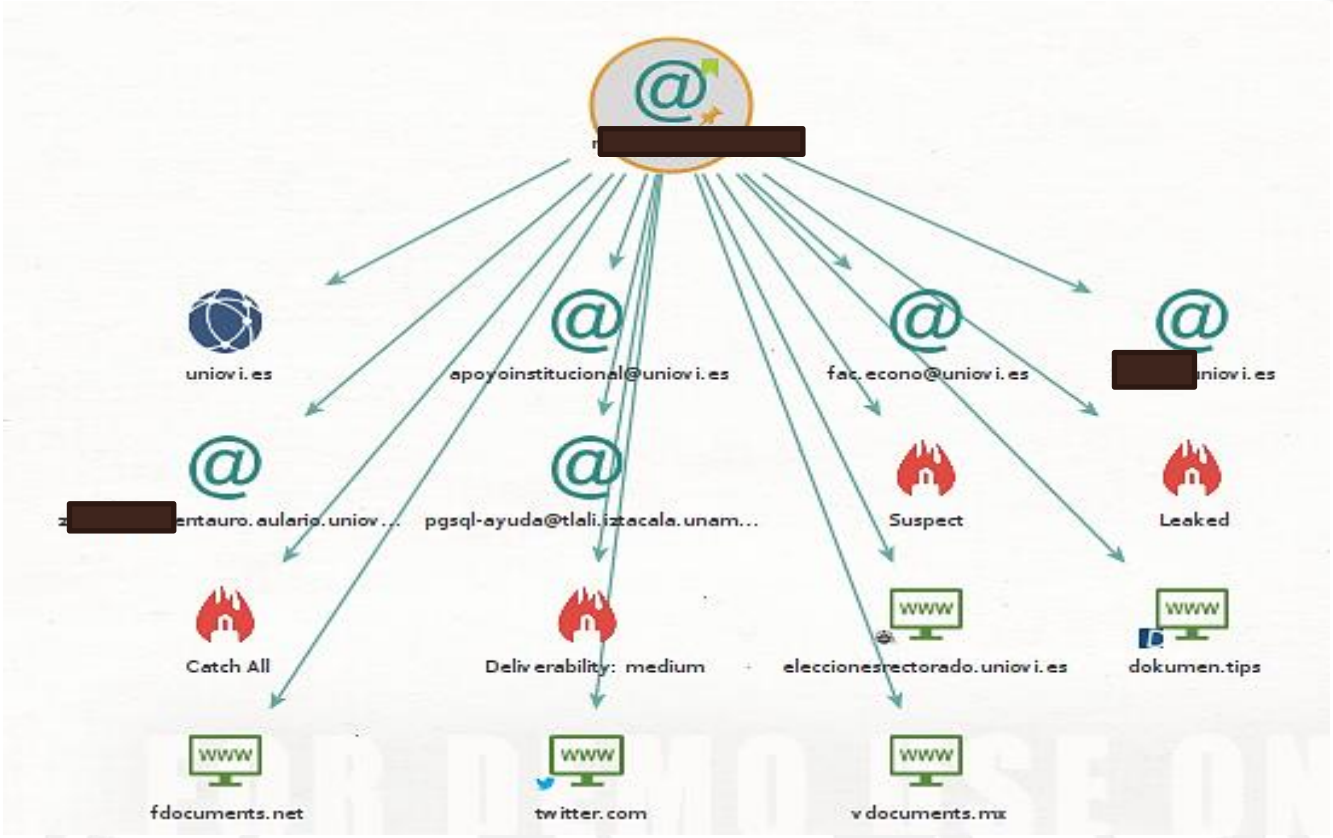


Figura 131. Conjunto de información de una persona combinando las transformaciones automáticas con la información obtenida de manera manual

Al hacer esto podremos visualizar la información obtenida de forma ordenada e ir completando nuestra investigación con lo que vayamos recopilando (y verificando). Además, todo lo añadido manualmente puede combinarse con la parte automatizada de la herramienta: si seleccionamos alguna(s) de las entidades añadidas (**pueden seleccionarse varias pulsando *mayúsculas***), podemos aplicar de nuevo "*All transforms*" sobre todas las seleccionadas y pueden salirnos más datos. Es posible que en estos casos la herramienta nos pregunte otra vez por parámetros a aplicar a las transformaciones antes de ejecutarlas. Por ejemplo, si se va a usar un motor de búsqueda pueden preguntarnos por uno concreto que queramos usar (Ej.: www.google.com). En estos casos puede salirnos información interesante nueva, como **si una dirección de email se encuentra filtrada** en una base de datos de emails robada conocida, como se ve en la siguiente imagen.

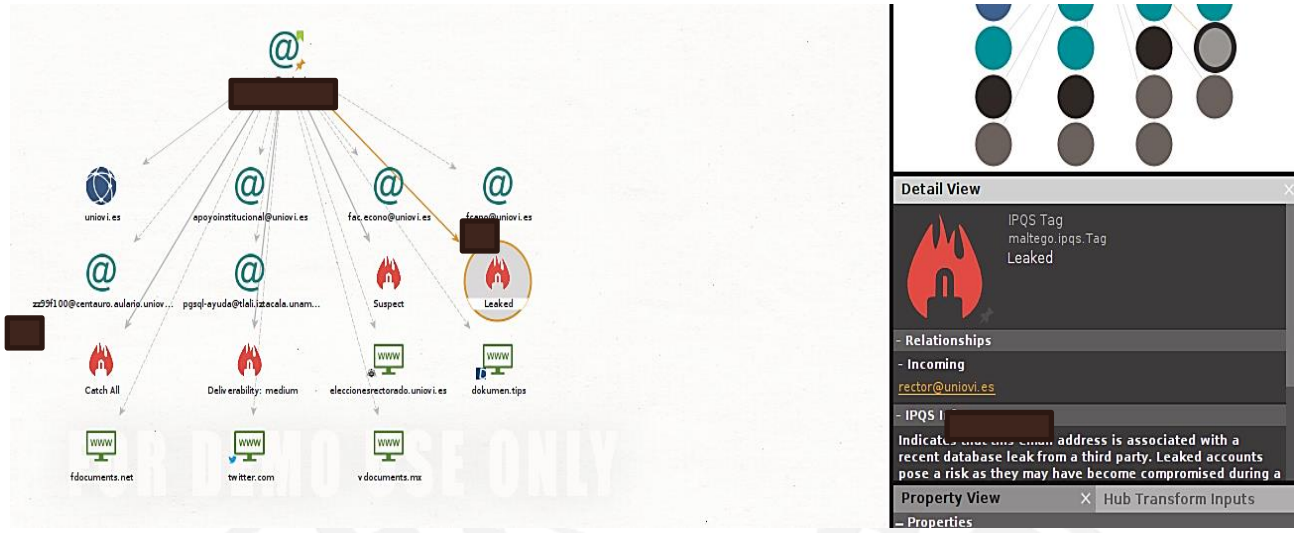


Figura 132. Dirección de email filtrada en una base de datos detectada por *Maltego* (tranquilos, no es una filtración peligrosa, ya se cambió la clave desde entonces 😊)

Vemos como la herramienta es capaz de detectar cosas importantes, como si la dirección es sospechosa de enviar *spam*, si maneja errores de forma inadecuada...como decía, un grafo de *Maltego* puede contar una historia muy interesante que merece la pena leer 😉.

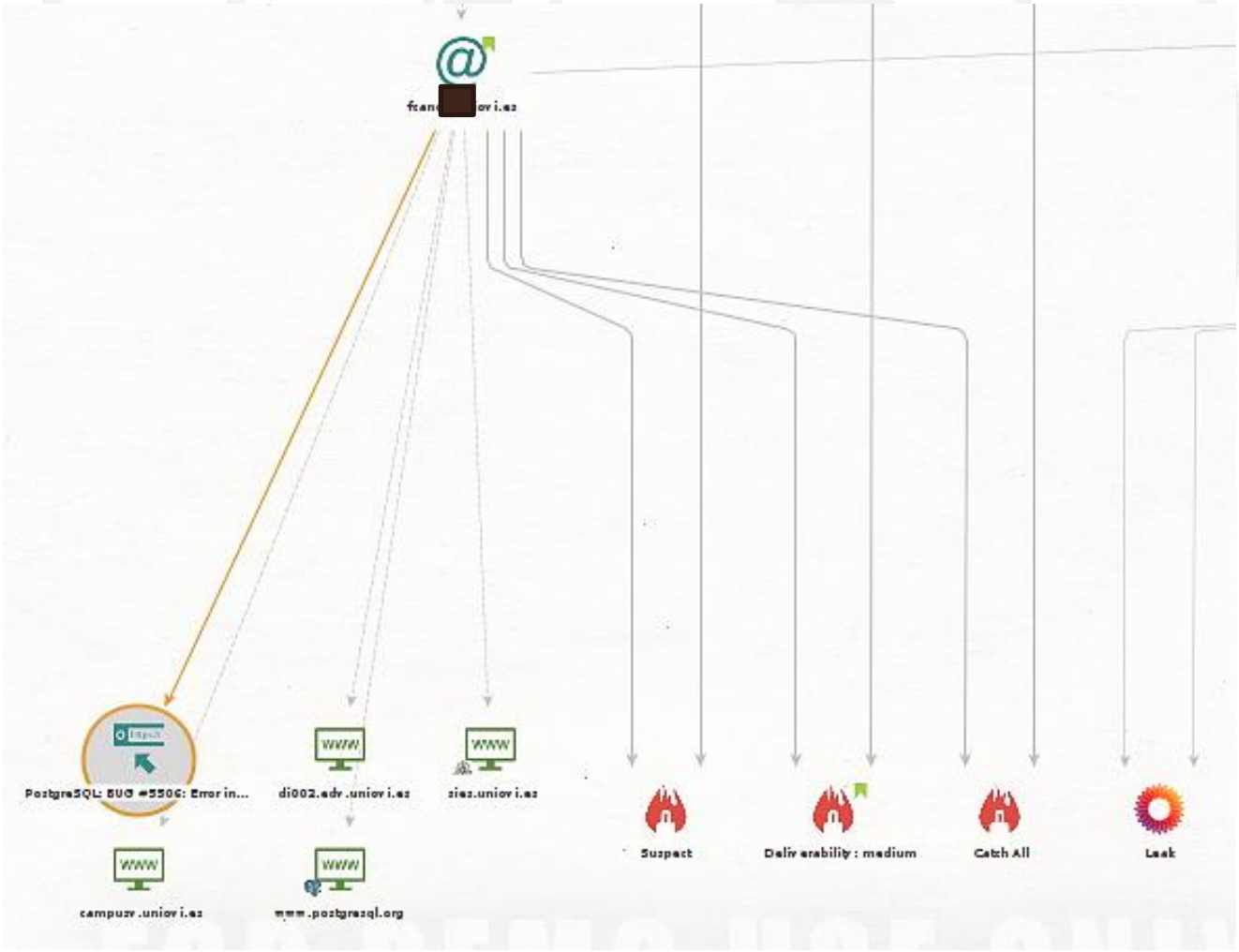


Figura 133. *Maltego* descubrimiento cosas interesantes de una dirección de email

En este ejemplo añadimos manualmente una entidad ***Person***, ponemos un nombre conocido y hacemos un análisis automático para sacar más correos electrónicos, ubicación, DNS y alias, entre otros. También tenemos perfiles de redes sociales, sobre los que hablaremos en la siguiente sección.

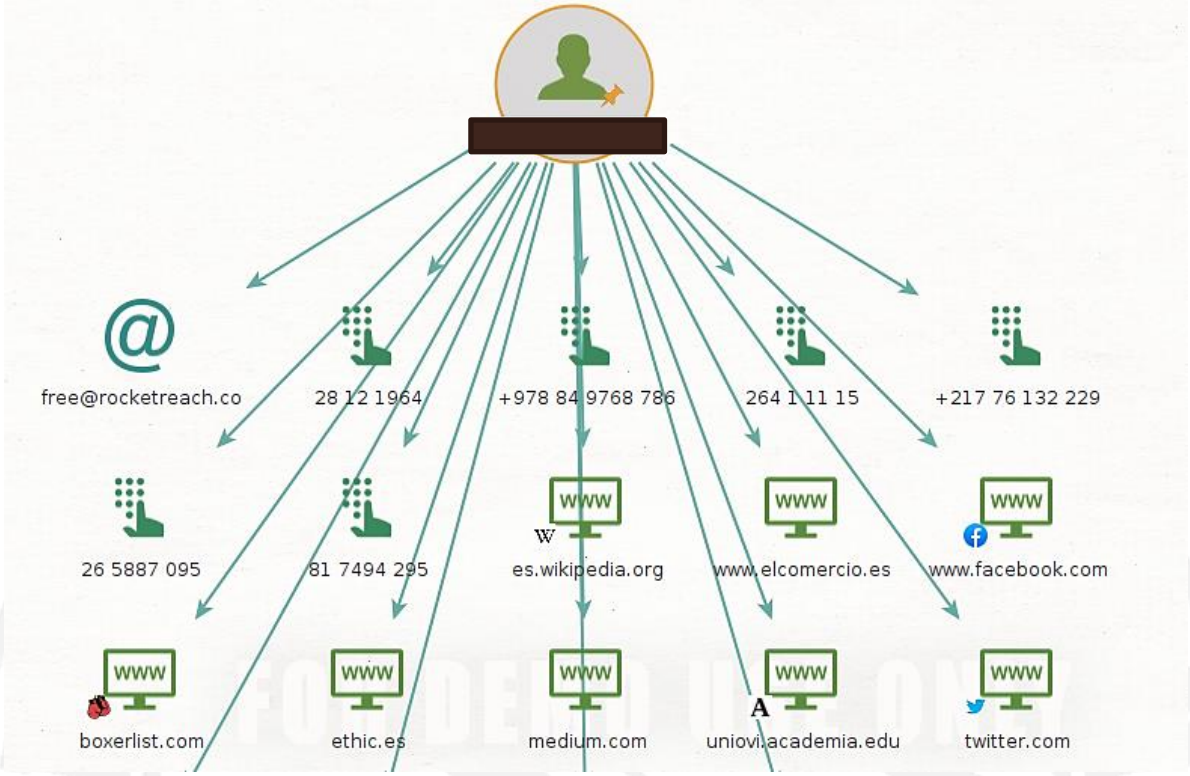


Figura 134. Típico análisis de un nombre de una persona hecho por *Maltego*

Por supuesto, hacer un poco de ***egosurfing*** (buscarse a uno mismo) es bueno para revelar si estamos dando demasiada información pública. Este es mi caso, donde solo aparece a) información falsa (o de otras personas que tienen mi nombre) o información que tengo pública por mi trabajo. Vamos, todo bien 😊.

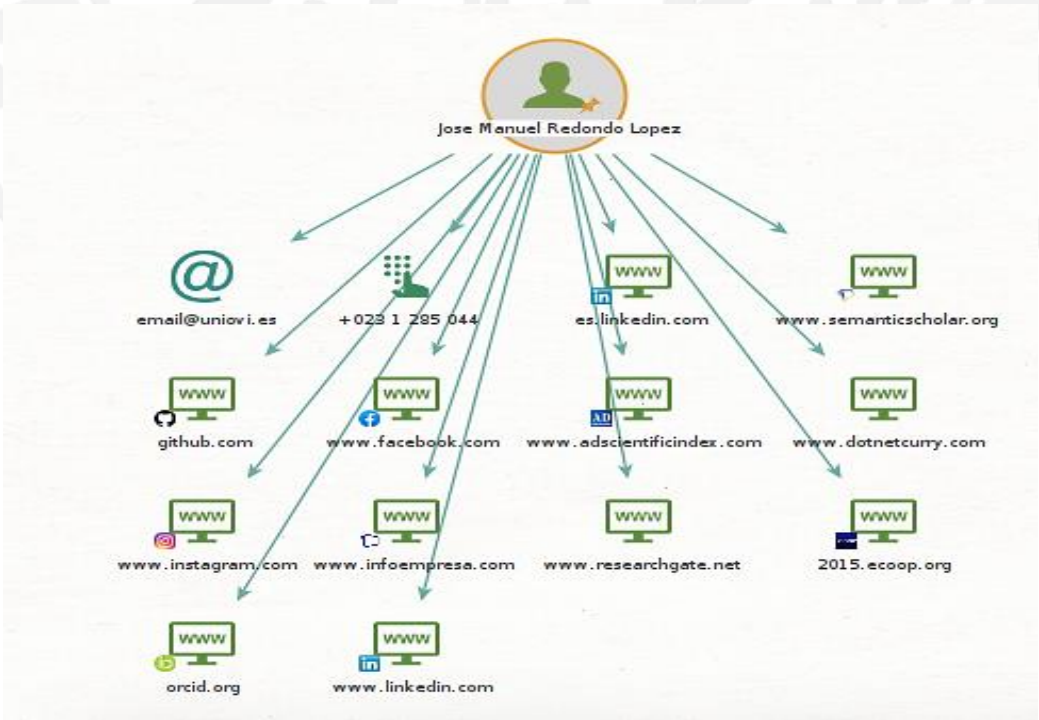


Figura 135. Información sobre mí en *Maltego* (el nº de teléfono no tengo ni idea de donde sale :D)

Buscando en redes sociales

En este punto, una vez obtenidas cuentas de redes sociales de las personas investigadas, es posible usarlas para seguir buscando información de forma automatizada también. Normalmente para hacer eso es necesario tener una cuenta de la red social que queremos investigar, ya que la **API Key necesaria** para hacer las consultas solo la tendremos si tenemos una cuenta activa en ella. La siguiente imagen muestra por ejemplo el resultado de usar una entidad que analiza *Twitter*, algo que ya va a ser imposible de hacer gratis por culpa de haber hecho el uso de la API de pago a partir del 09/02/2023[1].

Por otro lado, hay *Transforms* asociadas con *Facebook* tanto en la paleta como por parte de terceros, y en el caso de terceros también las hay de muchas otras redes sociales (Ej.: *LinkedIn*, *Tiktok*...). Gracias al uso de estas ***Transforms* de terceros** (que veremos en el apartado de "***Centro de Transformaciones***"), la potencia de la herramienta puede expandirse enormemente.

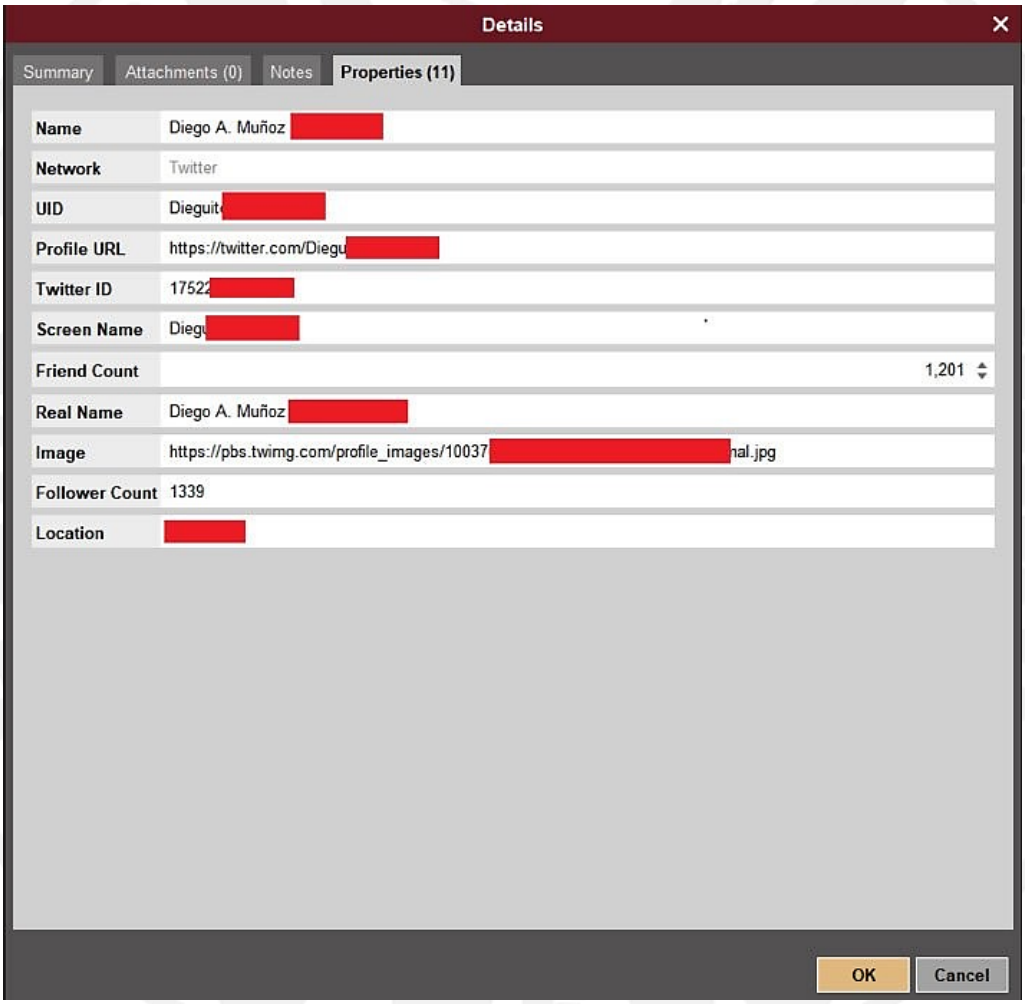


Figura 136. Ficha de una cuenta de Twitter de alguien. Fuente: https://ciberpatrulla.com/maltego/

Esta información va variando de acuerdo con lo recopilado por *Maltego*. Algunas veces encontraras mucha información y otras veces nada, pero esa es la gracia. Ninguna buena investigación es fácil ni inmediata, necesita trabajo, astucia y constancia para obtener buenos resultados, combinar investigaciones manuales con automáticas recursivas (afinar), tomar notas, subir capturas...y echar tiempo. En este documento no

[1] He oído un rumor que dice que hay por ahí un *GitHub* que tiene *API Keys* válidas de *Twitter* asociadas a las aplicaciones de la red social para teléfonos (por ejemplo *IPhone*), entre otras. Si eso fuera cierto, esas API Keys podrían funcionar en *Maltego*, ya que para la API de *Twitter* serían peticiones que hace una aplicación de *IPhone* legítima y no hay razón para prohibirlas. Pero claro, sólo es un rumor...

obstante estamos enseñando las bases para que os hagáis una idea de la potencia de la minería de datos asociada a OSINT contra entidades que, como veis, es mucha 😊.

Otras entidades de Maltego

Maltego trae muchas entidades usables de serie, y cada una de ellas puede ser arrastrada y utilizada para buscar diferentes tipos de datos. La cantidad de entidades disponible **depende de la versión de la herramienta que usemos**. Esta imagen muestra algunas de la versión CE que hemos usado hasta ahora.

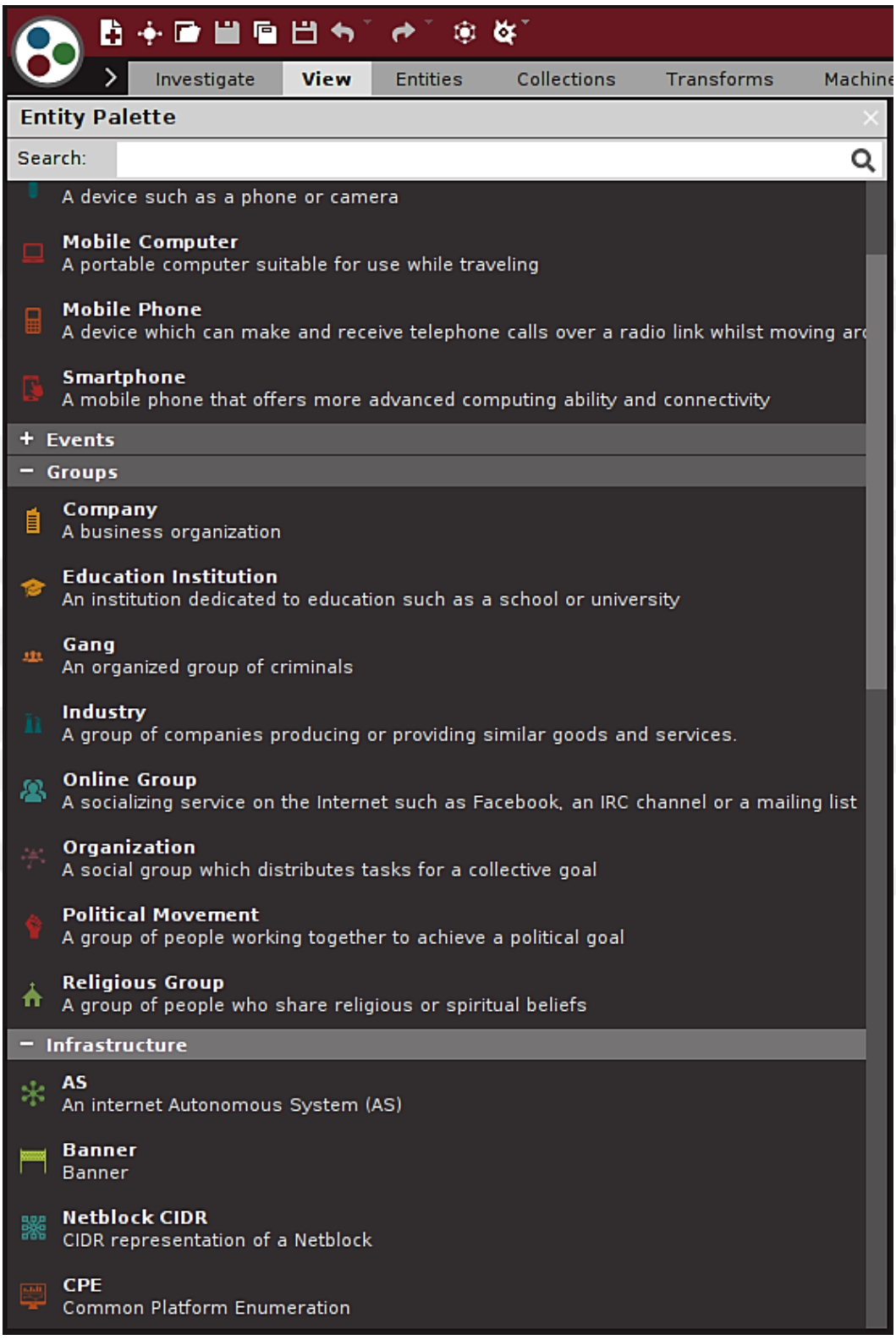


Figura 137. Entidades de la versión CE de *Maltego* (algunas)

En relación con la sección anterior, también tenemos de serie varias relacionadas con redes sociales conocidas, en este caso se muestran las de *Twitter*.

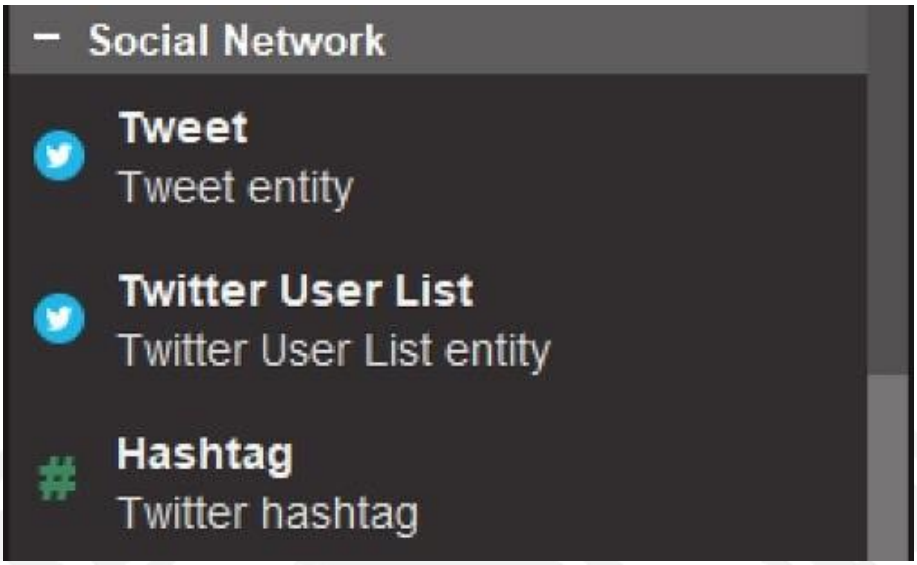


Figura 138. Entidades relacionadas con *Twitter* incluidas en *Maltego*

Como vemos, herramientas de este tipo nos facilitan cierto tipo de búsquedas de información, ya que puede llegar a identificar una gran cantidad de antecedentes. No obstante, quizá en la versión CE más moderna falten entidades para ciertas redes sociales. Pero siempre podremos encontrar alguna hecha por terceros que se pueda importar. Todas ellas se pueden controlar con el centro de transformaciones que veremos a continuación.

El *Centro de Transformaciones*

El centro de transformaciones (siguiente imagen) es lo que nos permite el uso de transformaciones externas, haciendo la herramienta increíblemente ampliable y potente con los desarrollos de colaboradores de *Paterva*.

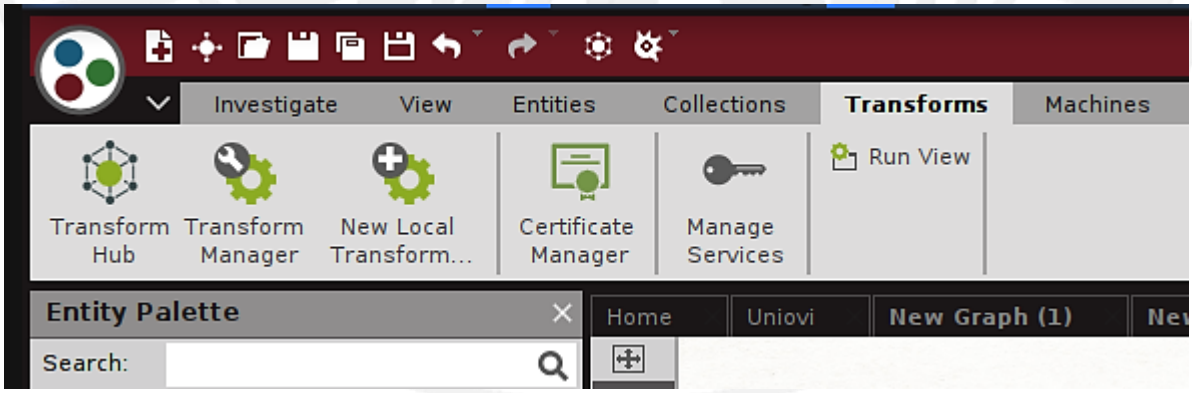


Figura 139. Centro de transformaciones de *Maltego*

Una vez abierto el "***Transform Hub***" vemos un listado de transformaciones disponibles para instalar en nuestro *Maltego* base, con su descripción para saber si nos va a ser o no de utilidad. Como vemos, hay una gran cantidad de ellas disponibles directamente en la GUI de la herramienta.

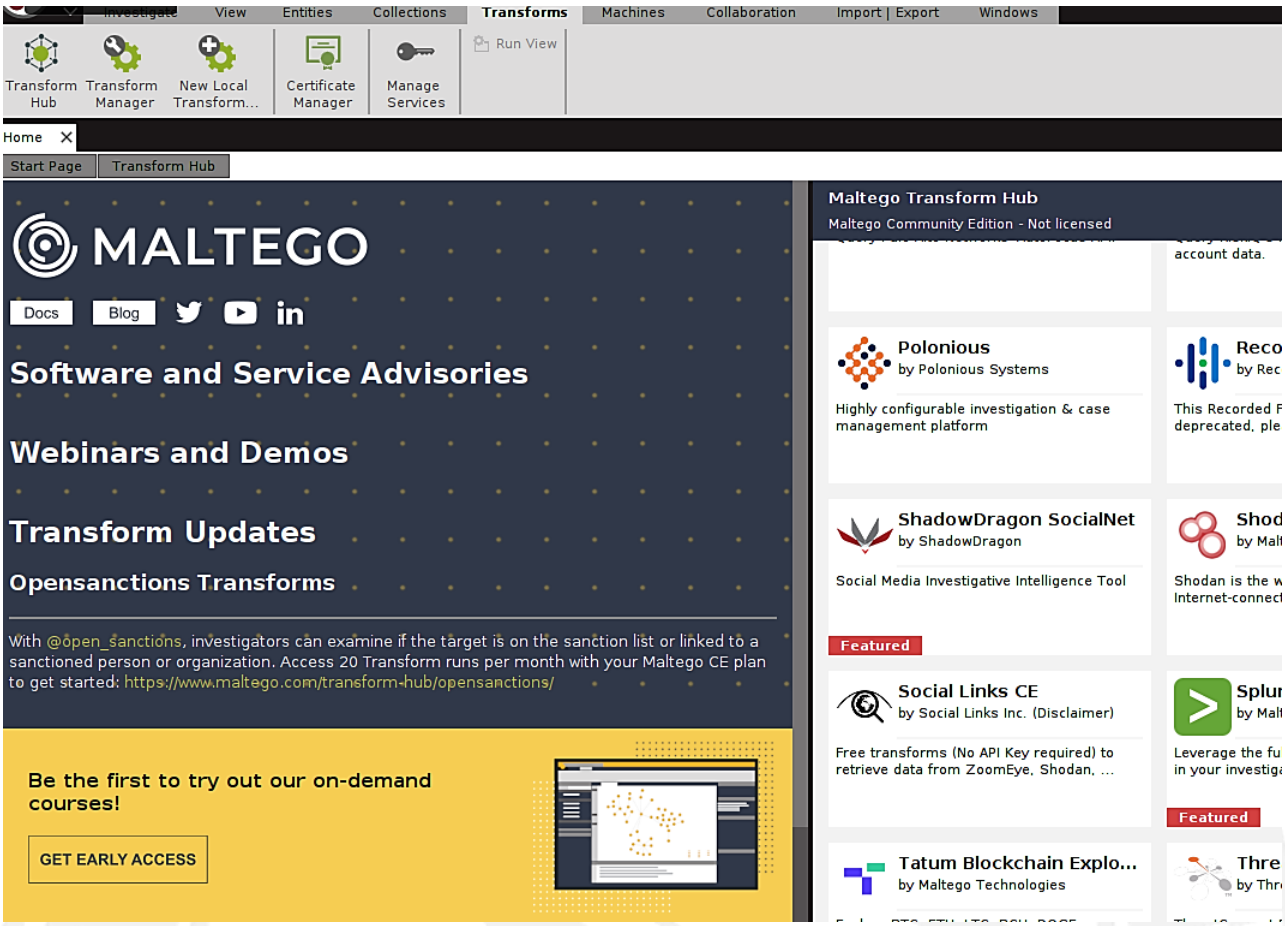


Figura 140. GUI del *Transform Hub* de *Maltego*

Este "*Centro de Transformaciones*" permite a *Maltego* acceder a datos de diferentes herramientas o proveedores de algún servicio utilizando su *hub* oficial: https://www.maltego.com/transform-hub/. Como vemos, **no tendremos a nuestro alcance todos ellos**, ya que muchos son de pago o para otras versiones de la herramienta. Por suerte para nosotros, podemos filtrar por cualquier criterio para quedarnos solo con aquellos a los que realmente podemos acceder.

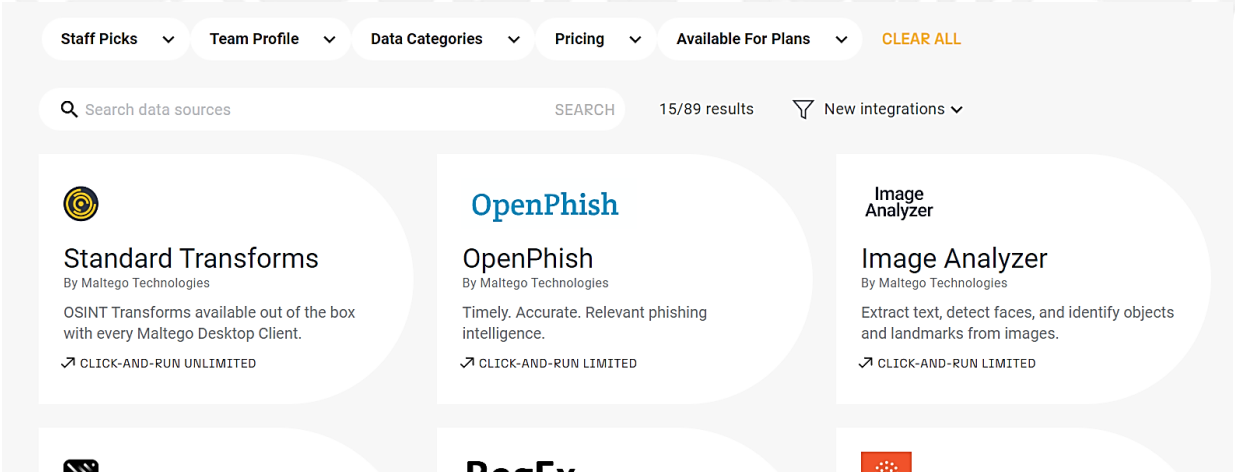


Figura 141. Web del *Transform Hub* de *Maltego*

A modo de ejemplo instalaremos el hub de ***SocialLinks***, el cual nos dará algunas entidades adicionales que podemos usar en cualquiera de los análisis vistos. Para ello simplemente buscamos el hub en la GUI y hacemos clic en instalar, que como vemos añade muchas más transformaciones, entidades e iconos para enriquecer nuestros gráficos.

Figura 142. Instalación de *SocialLinks* para *Maltego*

Ahora, si creamos un *Graph* nuevo y nos vamos al apartado de entidades, podemos ver que ya se encuentra la sección de ***Social Links***, la cual nos permite en este momento y con la versión actual del *hub*, buscar información utilizando una entidad "***Pipl***" (la red de búsqueda de personas que vimos antes), una entidad de "***Company***" y antecedentes relacionados con ***GitHub***. Estas entidades se van modificando a medida que la empresa que lo mantiene desarrolla este módulo (*Maltego* (*Paterva*), solo las integra, no las controla). Nuevamente para hacer uso de alguna de ellas necesitamos ***API Keys*** de los servicios relacionados, como por ejemplo esta entidad de *GitHub* que nos pide el *API Key*, que podemos obtener si tenemos una cuenta en *GitHub* activa. Como vemos, de aportar una podemos usarla para sacar información de los **miembros de una empresa u organización**, sus **repositorios, detalles de su código** y otros detalles que se encuentren en *GitHub* sobre la entidad analizada. Por tanto, gracias a *Social Links* se puede cubrir también otro de los aspectos de búsqueda en Internet que mencionamos en el **Seminario 1**, que hasta ahora no habíamos podido cubrir 😊.

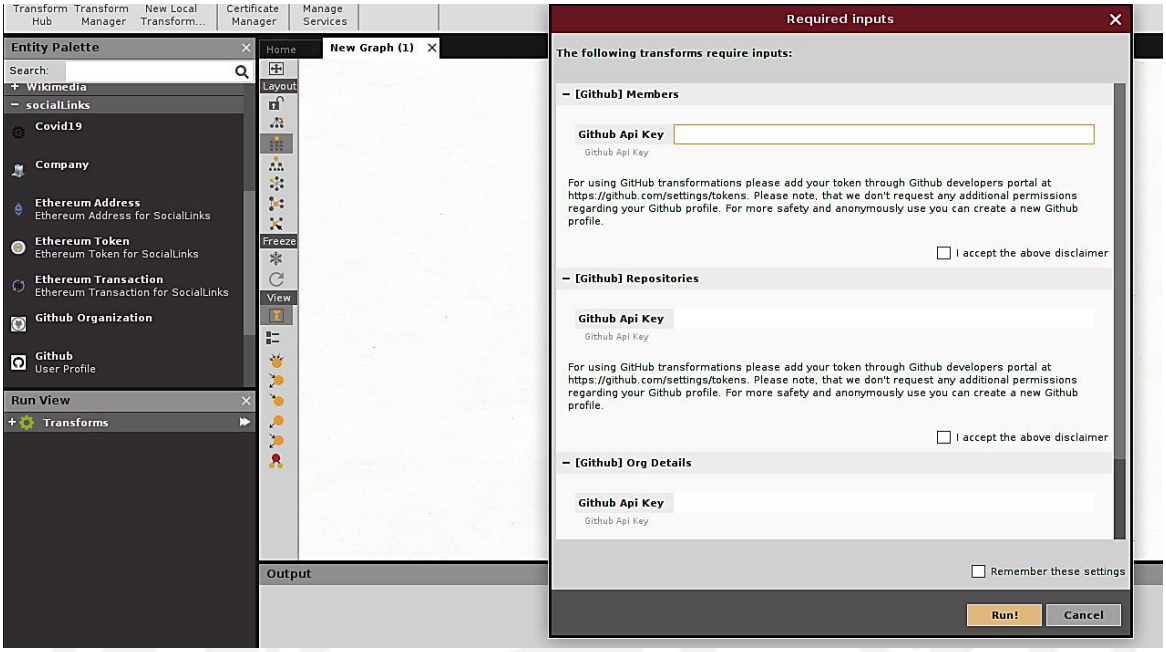


Figura 143. Entidad *GitHub* de *SocialLinks* de *Maltego*

Al añadir más entidades a la paleta de *Maltego*, repetir búsquedas anteriores **nos saca más resultados**. Por ejemplo buscarme a mí mismo de nuevo ahora no solo saca sitios web asociados conmigo, sino supuestas personas y documentos que también lo están. Curiosamente, **en este caso son todo falsos positivos** (no conozco a esas personas de nada...), aunque obviamente no siempre es así. En la imagen también vemos como **un documento ofrece muchas propiedades que determinan qué es sin falta de abrirlo**.

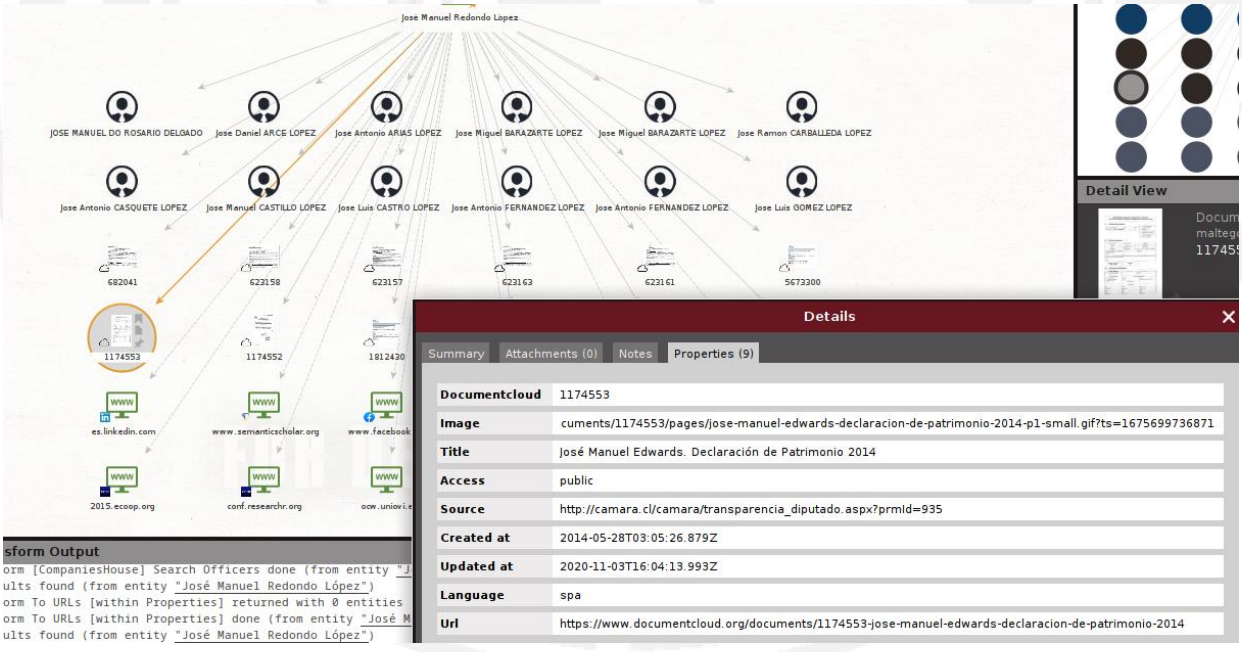


Figura 144. Búsqueda "enriquecida" con entidades de *SocialLinks* de *Maltego*

Para agregar nuevos *hub*, basta con seleccionar la opción de agregar y seguir los pasos que se detallan.

Figura 145. Agregar nuevo *hub*. Fuente: https://ciberpatrulla.com/maltego/

En el siguiente ejemplo, instalamos el *hub* de ***Shodan*** y tenemos que introducir el API que conseguimos por crear nuestra cuenta como se vio en el **Seminario 1**.

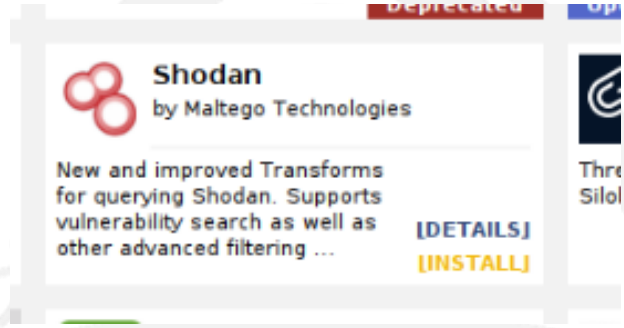


Figura 146. Hub de *Shodan* para *Maltego*

Figura 147. Introduciendo el API Key de *Shodan* cuando nos lo solicite el nuevo *hub*. Fuente: https://ciberpatrulla.com/maltego/

Una vez hecho esto podríamos consultar una entidad tipo ***IP*** y al ver sus opciones de transformación comprobaremos que se ha añadido la de hacer un escaneo *Shodan*. **Date cuenta lo que esto quiere decir**: puedes integrar la enorme cantidad de información que *Shodan* te da sobre máquinas, con la que *Maltego* te da de personas (y máquinas también) por otro lado, con lo cual **el "mapa" de tu objetivo** será mucho más completo. Ahora imagina que sigues esta misma filosofía con otros servicios (que es lo que se hace). Comprobarás que esta herramienta, en su versión de pago sin limitaciones y en manos de alguien con experiencia, puede ser una autentica bestia a la hora de trabajar con datos de fuentes abiertas 😊.

Figura 148. El poder de *Shodan* al alcance de un clic (nótese que también hay *Transforms* de otros *hubs* como *Censys* y otros buscadores 😉)

En este repositorio de *GitHub* hay enlazadas una **gran cantidad de transformaciones** para *Maltego* que pueden ser útiles en distintos escenarios, algunas de ellas interactuando con motores y *frameworks* que también se ven en este documento: https://github.com/cipher387/maltego-transforms-list/

Ahora solo te queda seguir interactuando con *Maltego* y probar cada una de sus entidades y funcionalidades, integrar nuevas APIs para búsquedas específicas e ir desarrollando una metodología propia para nuestras investigaciones. *To the moon!* :D. Obviamente con la versión CE estarás limitado, pero aun así, el poder que tendrás es mucho mayor que el no usarla 😊.

Exportar la información

Maltego tiene la opción de exportar sus resultados para poder incluirlos en una documentación y/o presentárselos a un cliente que nos haya contratado para hacer una auditoría. Para ello solo es necesario ir al menú y seleccionar la pestaña "*Import | Export*" para finalmente usar "*Generate Report*"

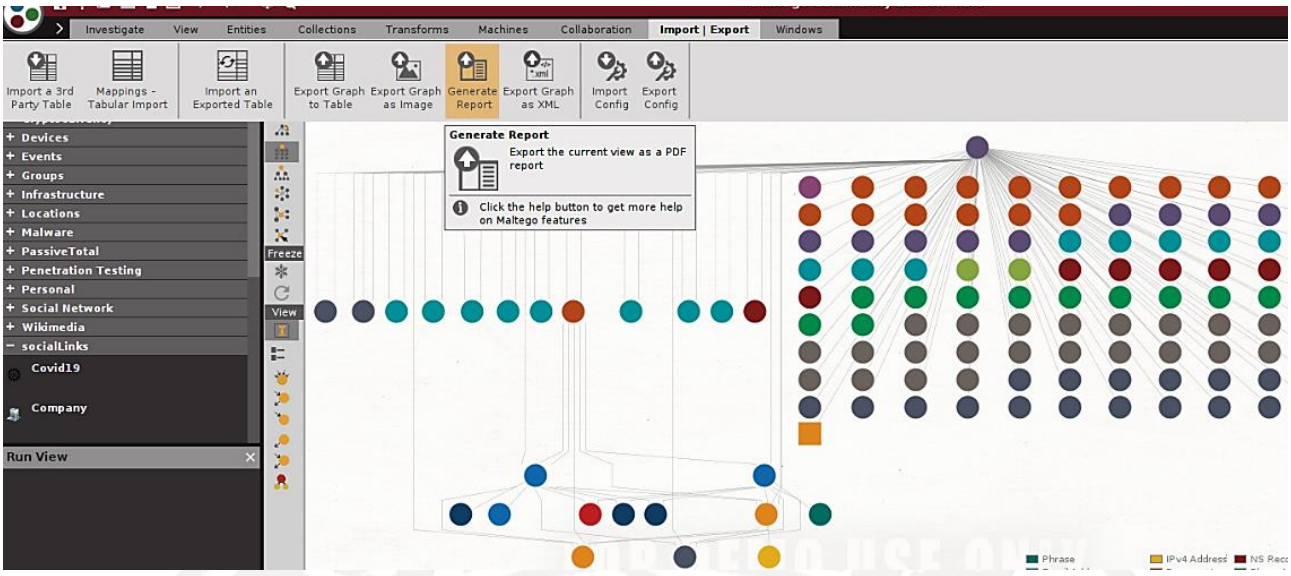


Figura 149. Generación de informes en *Maltego*

Generamos un informe y guardamos los resultados en PDF.

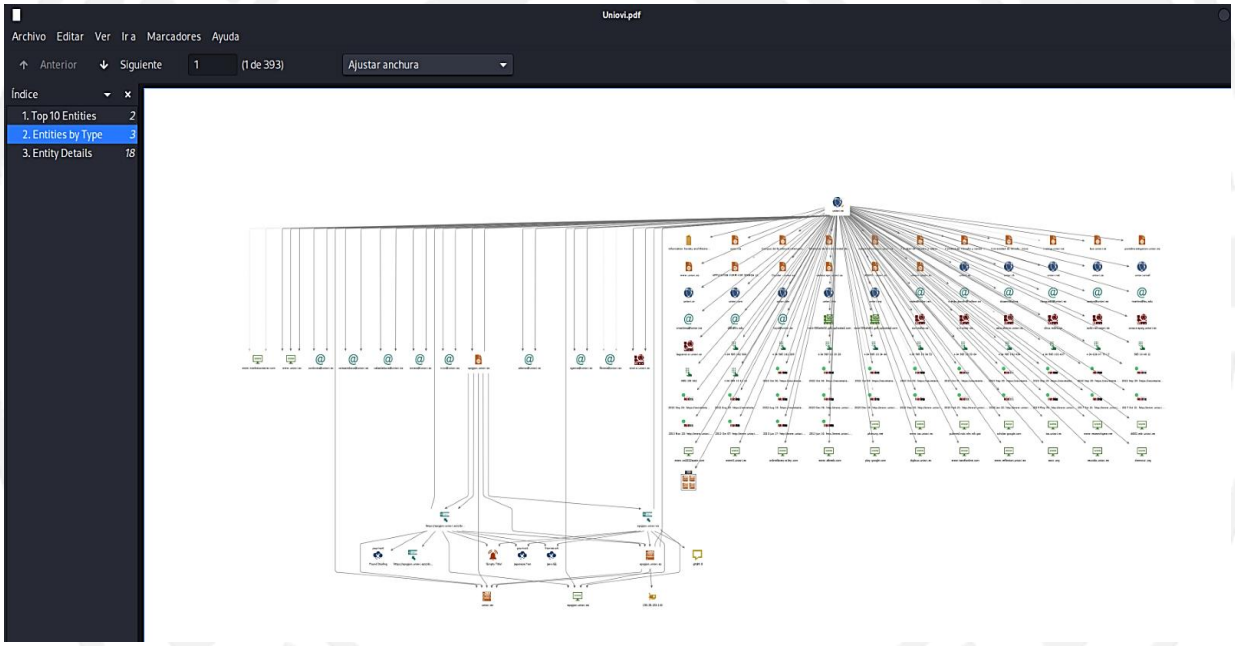


Figura 150. Informe en PDF de *Maltego*

También tenemos la opción de importar como tabla o imagen, según necesites o más te guste. Ten en cuenta que **este informe puede ser muy extenso**. No se limita sólo a mostrar el gráfico, sino tablas de datos e información sobre cada entidad encontrada y analizada de forma exhaustiva. Esto es así porque es el tipo de información que necesitan las personas que analizan esta información. Para que te hagas una idea, solo el informe resultante de analizar `uniovi.es` y `epigijon.uniovi.es` y sus entidades que vimos al principio de esta sección de OSINT automático casi llega al as 400 páginas. Lo bueno es que no queda nada sin documentar 😊.

¡Quiero saber más!

En este documento solo hemos visto un uso típico común de *Maltego* para hacer investigaciones OSINT orientado a esta asignatura. Pero la herramienta ofrece mucho más 😊. Uno de los desafíos que puede ser más complejo al principio es como hacer una investigación una PoI (***Person of Interest***, es decir, a quien

estamos investigando) de una forma coherente y ordenada, dada la gran cantidad de información que podemos obtener a poco que expandamos las entidades y hagamos transformaciones. Este diagrama puede ayudarte con ello:

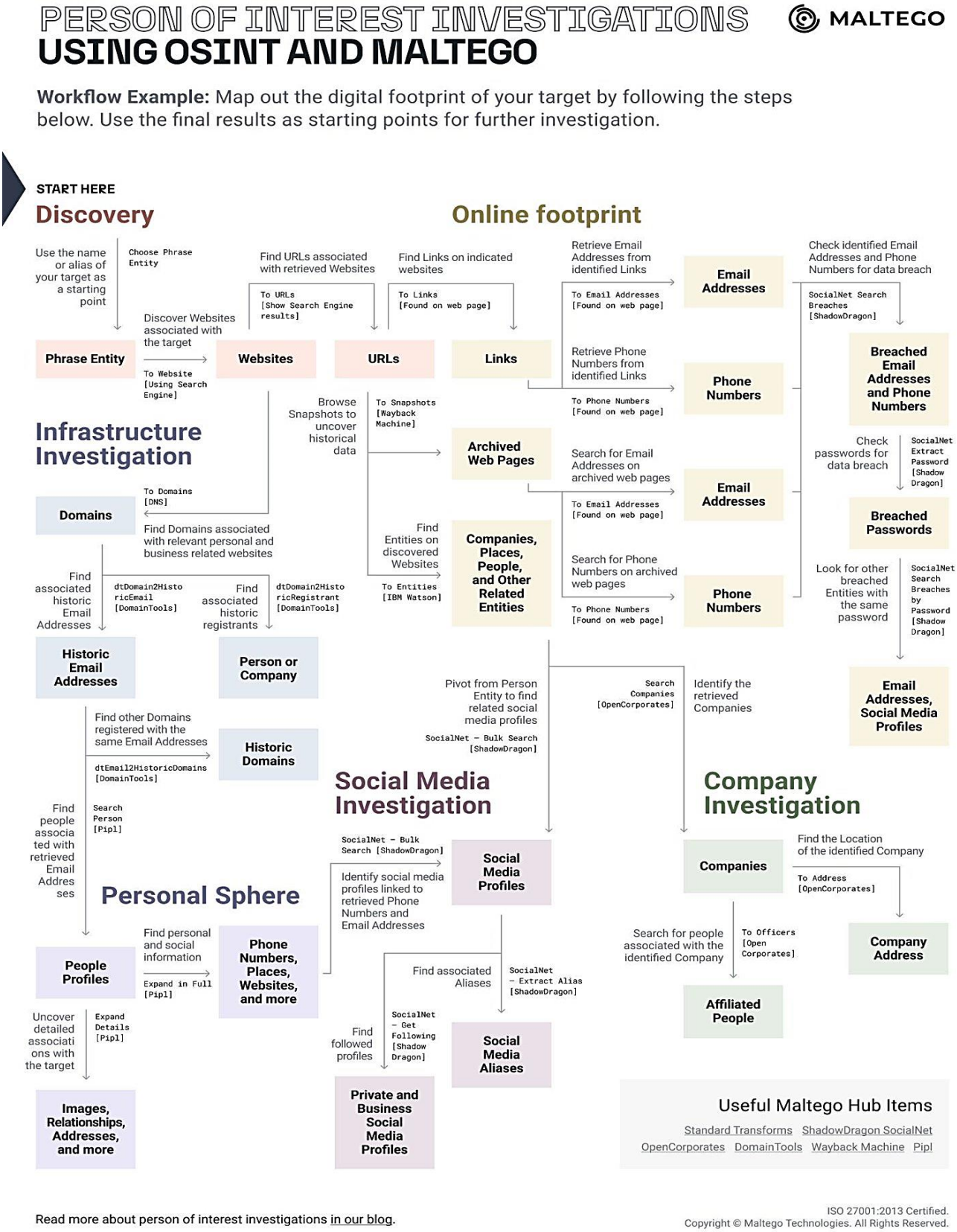


Figura 151. Usar la potencia de *Maltego* correctamente conlleva planificar. Este esquema puede ayudarte con ello

Si quieres aprender más sobre *Maltego*, puedes consultar estos enlaces

- **Sitio Oficial de *Maltego***: https://www.paterva.com/web7/
- **Documentación**: https://www.paterva.com/web7/docs.php
- ***Maltego User Guide (Version 3)***: https://www.paterva.com/web7/docs/M3GuideGUI.pdf
- **Canal de *YouTube* de *Paterva***: https://www.youtube.com/user/PatervaMaltego
- **Canal de YouTube *Cylon Null***: https://www.youtube.com/channel/UCseLiPDQKBrKJfqtiHuH07w/videos

OSINT "Automático": Motores de búsqueda y *Threat Intelligence*

Introducción

La **inteligencia de amenazas cibernéticas** (*Cyber Threat Intelligence* o CTI) es información basada en conocimientos, habilidades y experiencia sobre amenazas cibernéticas y físicas y actores que hacen uso de esas amenazas. Incluye tanto determinar su existencia en un contexto determinado como evaluar la gravedad de estas. Tiene como objetivo ayudar a mitigar posibles ataques y eventos dañinos que ocurren en el ciberespacio. Hablamos de este término aquí porque las fuentes de inteligencia de amenazas de este tipo incluyen:

- Información obtenida vía **OSINT**
- Información obtenida de **redes sociales**
- Información obtenida de **actores humanos**
- **Información técnica**: archivos de *log* de dispositivos, datos adquiridos forensemente o tráfico de Internet, datos obtenidos o derivados de lo que se obtenga de *Deep Web* o la *Dark Web*...

Es decir, este documento cubre una buena parte de estas fuentes 😊. Si lo de ser "cibermaruja" no te convence, puedes decir que estás "obteniendo inteligencia de amenazas" 😉. Eso sí, para decir eso obviamente tienes que estudiar actores (humanos o máquinas) **que puedan ser amenazas para las infraestructuras o clientes que protejas o audites**, no a cualquiera por Internet (no seas tramposo 😊).

En los últimos años, la inteligencia de amenazas se ha convertido en una parte crucial de la estrategia de seguridad cibernética de las empresas, ya que permite a las empresas ser más proactivas en su enfoque y determinar qué amenazas representan los mayores riesgos para un negocio. Esto **hace que las empresas busquen activamente sus vulnerabilidades** y evita ataques antes de que ocurran. Dado que este método está ganando importancia en la actualidad, conocer técnicas como las que os he escrito os da un plus laboral 😊.

Hasta ahora hemos visto formas de obtener inteligencia de amenazas por medios manuales y con herramientas que automatizan parte de ellas. Pero en esta sección vamos a presentar **motores de búsqueda**, en su inmensa mayoría gratuitos, que permiten expandir la información de los objetivos **de forma pasiva** ("*do not engage*") y completar nuestro estudio.

El objetivo de estos motores es múltiple:

- Puedes usarlos para hacer una **evaluación rápida** sin coste y anónima de un posible del objetivo de estudio. En ese caso, **recomiendo usar varios**. No solo porque posiblemente se complementen entre ellos, sino porque un objetivo podría bloquear unos sí y otros no. Por ejemplo, es típico que un objetivo bloquee *Shodan* porque es quizá el más popular. Pero luego sin embargo la misma información la encuentras en otros que hacen cosas muy similares pero son menos conocidos.
- Puedes **completar tu gráfico de *Maltego*** con los datos que obtengas de varios de ellos. Es más, seguramente *Maltego* tenga un módulo que directamente te permita usarlos y que esto se haga automáticamente. Con suerte el módulo será gratuito y compatible con la versión CE 😉.

- Si estás desde el punto de vista "defensor", **puedes ver claramente lo que "los malos" pueden saber de ti** mirando Internet. Esto es tal y como suena. Esta información está ahí, al alcance de quien sepa como buscarla, y se puede usar contra ti. Así que mejor "tapa" tú el "agujero" antes de que se te cuele quien no debe.

Si tu empresa es grande e importante, es muy conveniente plantearse hacer un gasto para suscribirse a algunos de estos motores de búsqueda y monitorizar sus resultados para encontrar posibles amenazas que la afecten. No es broma, es algo preventivo y muy efectivo.

A continuación vamos a ver un conjunto de motores de búsqueda para CTI muy conocidos y usados a nivel mundial, clasificados por el tipo de información que son capaces de encontrar primariamente.

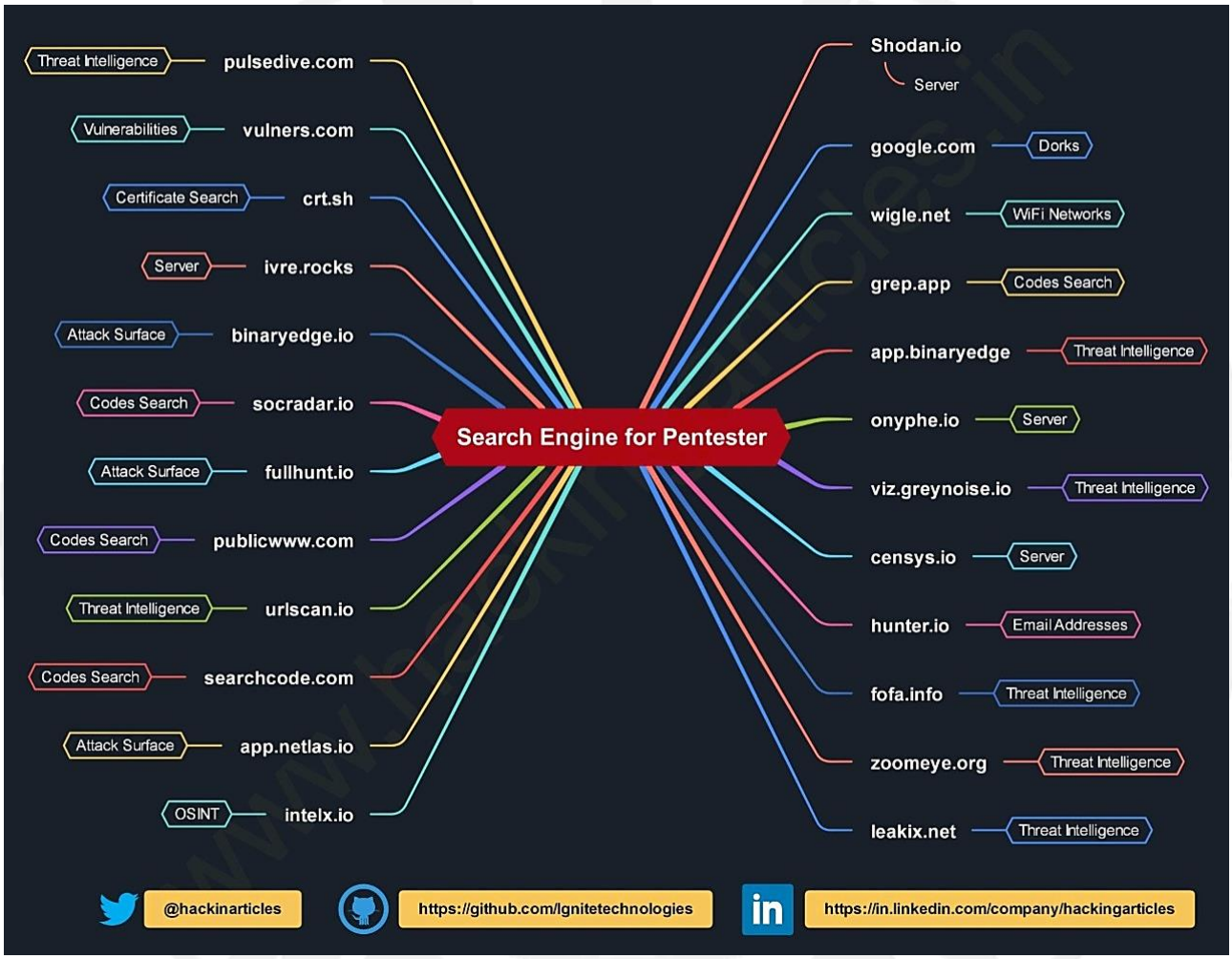


Figura 152. Esquema general de motores útiles para distintos propósitos e información buscada

APIs para OSINT: ¡Programas que usen OSINT!

El OSINT no es solo una técnica que se haga de forma manual o automatizada con herramientas específicas que descubran información sobre distintas fuentes de datos. **También existen APIs** que permiten encontrar información diversa sobre multitud de entidades y, como APIs que son, **pueden permitirnos diseñar programas que hagan uso de dicha información** para alguna de nuestras necesidades de investigación, combinarlas o hacer inferencias por programa a partir de ellas. Efectivamente, como ingenieros del software ¡ahora podremos crear programas que usen datos OSINT! 😊

Las APIs suelen requerir una API Key para poder usarlas y estas no son una excepción. Algunas la ofrecen gratuitamente con una ratio de peticiones / tiempo limitado (aunque suficiente para pruebas esporádicas). Otras son de pago, por lo que solo tienen sentido si nos dedicamos profesionalmente a ello. También debemos asegurarnos de que las APIs incluyen los países que nos interesan, pues no es infrecuente encontrarnos APIs **que solo ofrecen datos de Estados Unidos**, cuya política de tratamiento y publicación de datos es muy distinta a la de, por ejemplo, Europa. Antes de pagar, ¡debemos estar seguros de que los resultados nos van a servir!

Un sitio donde se proporciona acceso a distintas APIs es este: https://inteltechniques.com/tools/API.html

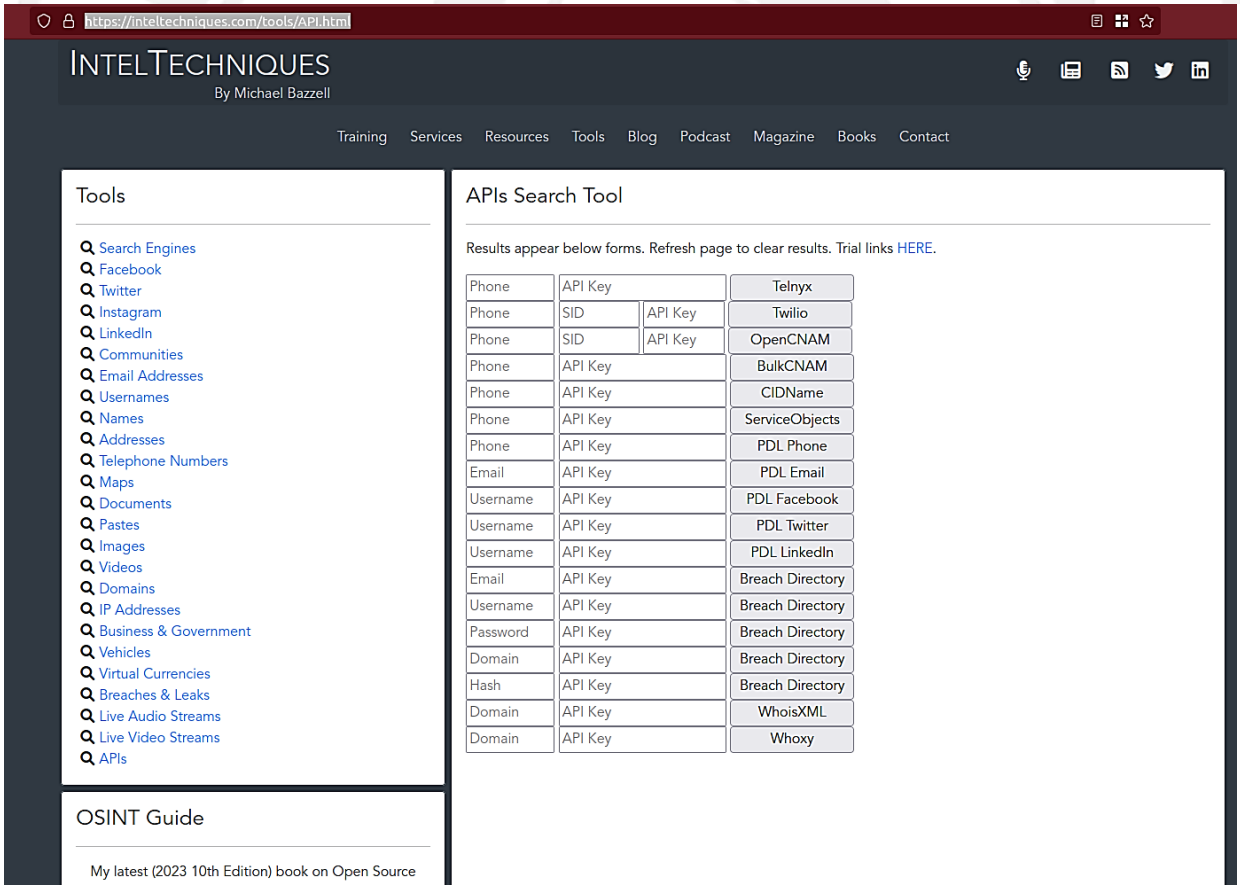


Figura 153. APIs para OSINT de *Intel Techniques*

También tenemos el siguiente repositorio de *GitHub* con una gran cantidad de referencias a APIs para obtener cualquier clase de dato que se nos pueda ocurrir: https://github.com/cipher387/API-s-for-OSINT

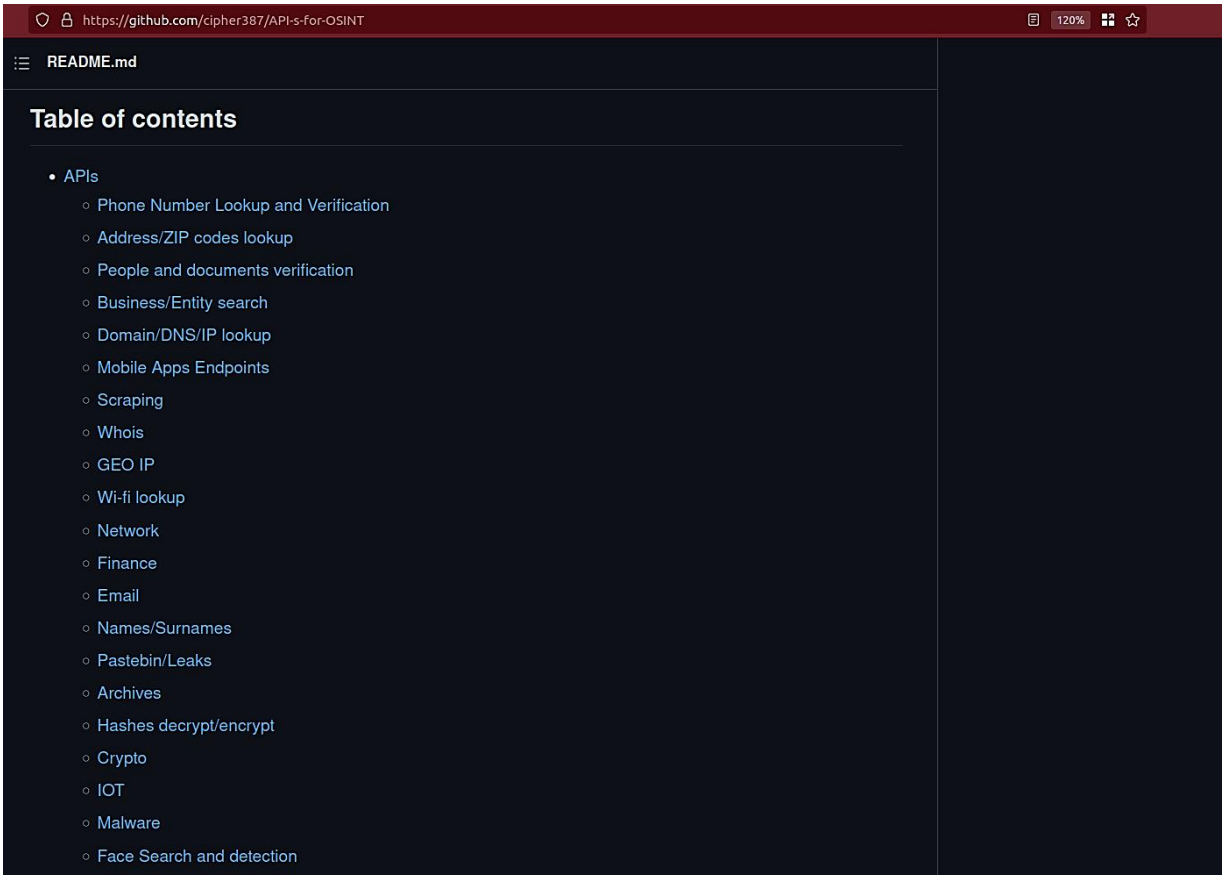


Figura 154. Repositorio de *GitHub* con APIs para OSINT

Árboles interactivos de herramientas OSINT clasificadas por tarea a realizar

Con la cantidad de herramientas OSINT existentes, muchas veces es muy difícil saber qué herramienta usar en cada caso, bien porque no sabemos de una o simplemente porque no la recordamos. Este tipo de herramientas en forma de árbol interactivo de herramientas OSINT nos puede ser de mucha ayuda para ello, y por ese motivo en esta sección presentamos algunos ejemplos

Malfrat's OSINT (https://map.malfrats.industries/)

Mapa que da acceso a distintas herramientas en función de lo que se vaya buscando en el árbol desplegable que va mostrando. Como se ve en la imagen, cubre una enorme cantidad de aplicaciones y usos

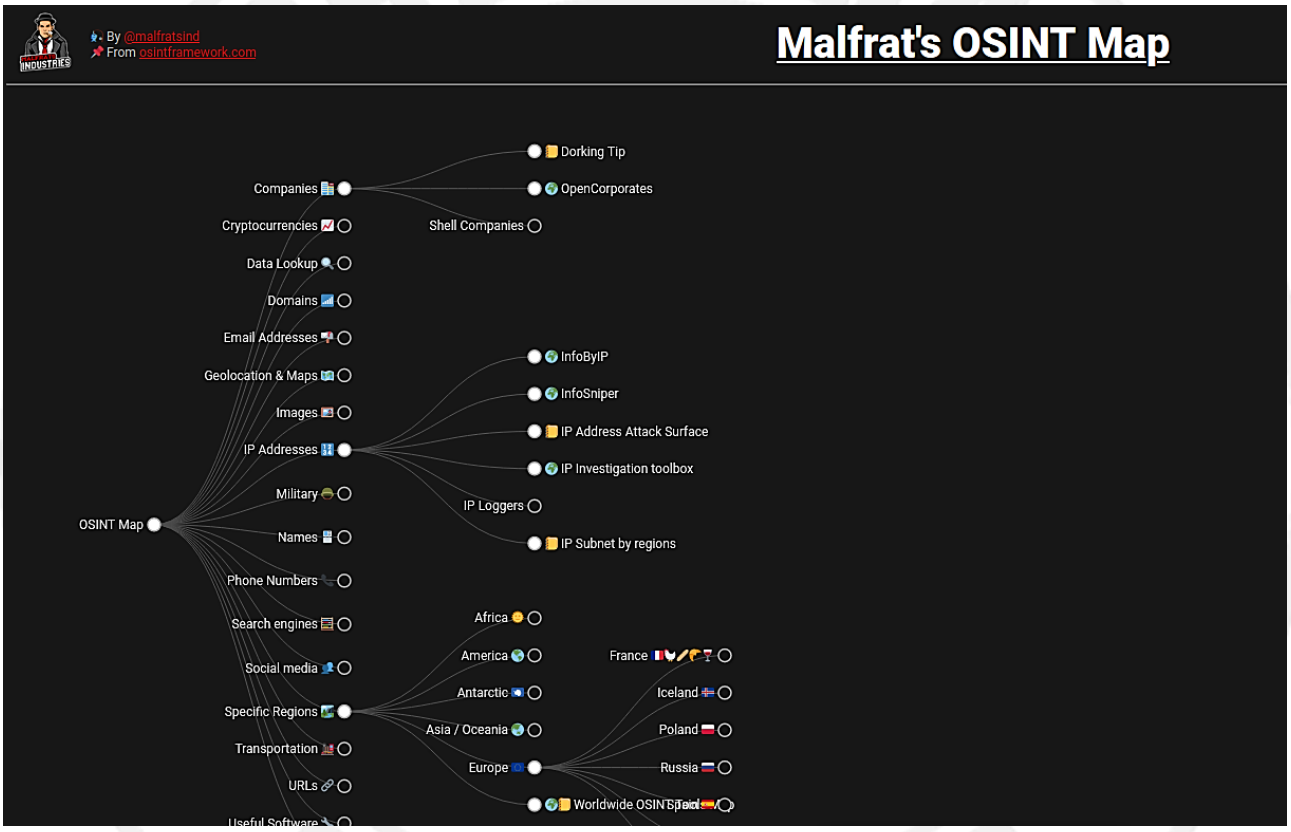


Figura 155. Malfrat's OSINT

OSINT Framework (https://osintframework.com/)

Otro mapa que da acceso a distintas herramientas en función de lo que se vaya buscando en el árbol desplegable mostrado, pero enfocándose más en información de personas. *¿Necesitas encontrar algo concreto de una persona?* Aquí tienes las herramientas que puedes usar más adecuadas para este tipo de información. Es una guía perfecta para cuanto no sabes exactamente cómo proceder a la hora de encontrar información.

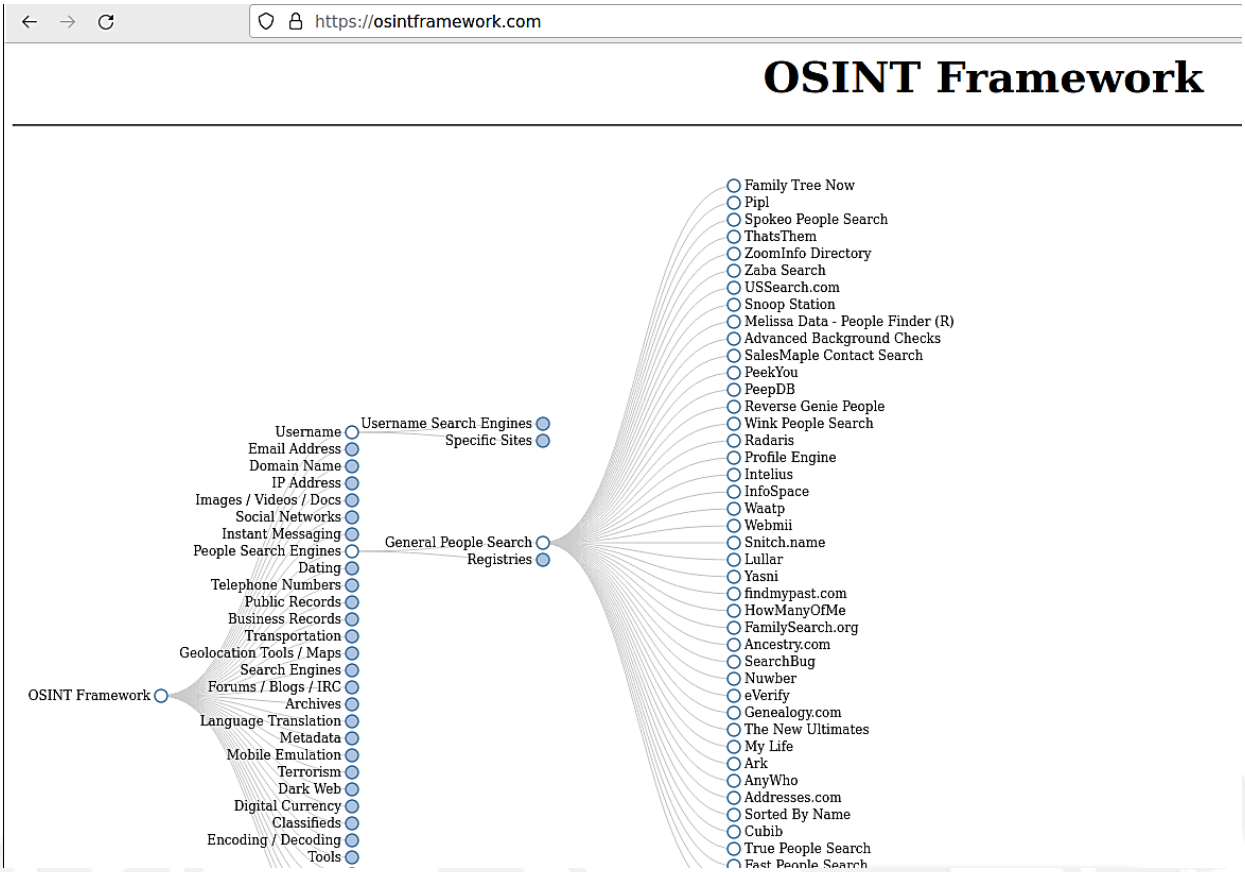


Figura 156. OSINT Framework

Motores de búsqueda para máquinas y redes

Existen una gran cantidad de motores de búsqueda para la investigación por Internet y en esta sección vamos a ver varios de ellos. No obstante, **no son todos los que hay**, y en esta imagen podrás encontrar más que los que incluimos en esta sección y en el documento en general. Si crees que alguno te puede venir bien, ¡no dudes en usar sus servicios!

Figura 157. Motores de búsqueda para investigadores hay muchísimos, y cada uno con su especialidad

Binary Edge (https://www.binaryedge.io/)

Recopila datos de Internet para alimentar su propio motor de búsqueda de manera muy similar a *Shodan*, *Censys* y otros que hemos visto o veremos. Permite hasta 250 consultas por mes gratuitas y desde 10 dólares al mes para 5000 consultas. Puedes encontrar muchos **ejemplos de consultas** útiles en: https://github.com/balgan/binaryedge-cheatsheet

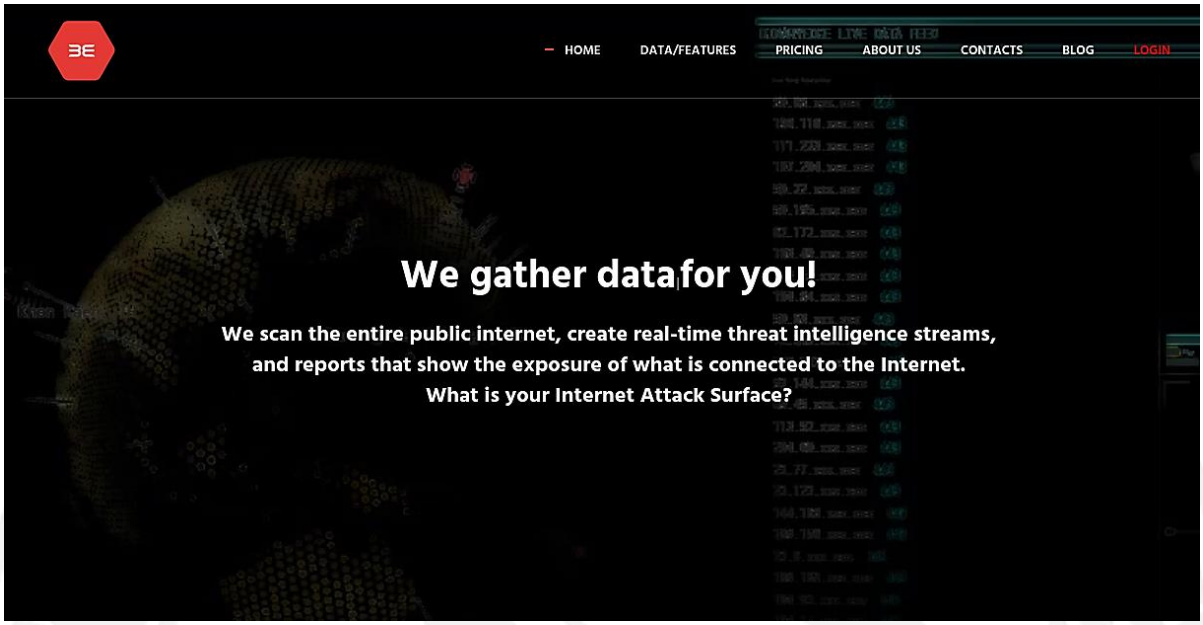


Figura 158. *BinaryEdge*

Censys (https://search.censys.io/)

Es el motor que vimos en el **Seminario 1** y que hace una evaluación de la superficie de ataque para dispositivos conectados a Internet. Como vimos, este motor de búsqueda encuentra y detecta versiones de servidores web, servidores de bases de datos, vulnerabilidades sin parches, *routers*, versiones del sistema operativo, servicios inseguros, firewalls de aplicaciones web...y, en general, mucha información que permite reconocer un sitio cualquiera para ver su posibles problemas de seguridad.

Puedes encontrar muchas **consultas de ejemplo** para trabajar con él aquí: https://cheatography.com/himajedi/cheat-sheets/censys-search/. El siguiente *cheatsheet* resume las **consultas más frecuentes** que puedes hacerle.

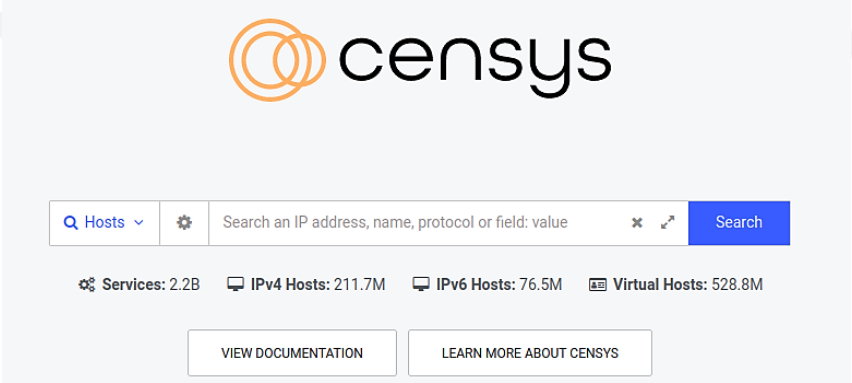
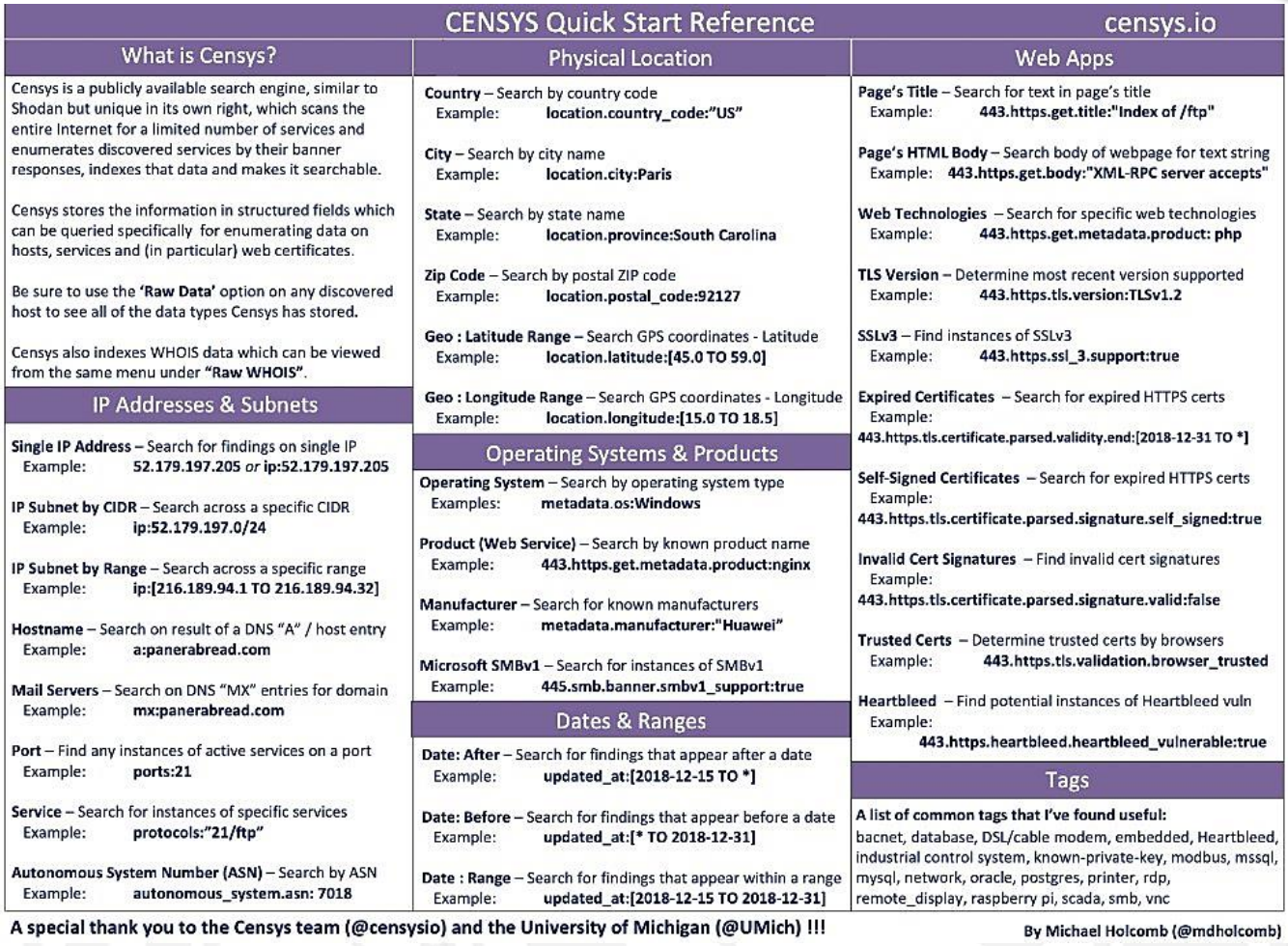


Figura 159. Censys



CENSYS Quick Start Reference — censys.io

What is Censys?

Censys is a publicly available search engine, similar to Shodan but unique in its own right, which scans the entire Internet for a limited number of services and enumerates discovered services by their banner responses, indexes that data and makes it searchable.

Censys stores the information in structured fields which can be queried specifically for enumerating data on hosts, services and (in particular) web certificates.

Be sure to use the **'Raw Data'** option on any discovered host to see all of the data types Censys has stored.

Censys also indexes WHOIS data which can be viewed from the same menu under **"Raw WHOIS"**.

IP Addresses & Subnets

Single IP Address – Search for findings on single IP
Example: **52.179.197.205** *or* **ip:52.179.197.205**

IP Subnet by CIDR – Search across a specific CIDR
Example: **ip:52.179.197.0/24**

IP Subnet by Range – Search across a specific range
Example: **ip:[216.189.94.1 TO 216.189.94.32]**

Hostname – Search on result of a DNS "A" / host entry
Example: **a:panerabread.com**

Mail Servers – Search on DNS "MX" entries for domain
Example: **mx:panerabread.com**

Port – Find any instances of active services on a port
Example: **ports:21**

Service – Search for instances of specific services
Example: **protocols:"21/ftp"**

Autonomous System Number (ASN) – Search by ASN
Example: **autonomous_system.asn: 7018**

Physical Location

Country – Search by country code
Example: **location.country_code:"US"**

City – Search by city name
Example: **location.city:Paris**

State – Search by state name
Example: **location.province:South Carolina**

Zip Code – Search by postal ZIP code
Example: **location.postal_code:92127**

Geo : Latitude Range – Search GPS coordinates - Latitude
Example: **location.latitude:[45.0 TO 59.0]**

Geo : Longitude Range – Search GPS coordinates - Longitude
Example: **location.longitude:[15.0 TO 18.5]**

Operating Systems & Products

Operating System – Search by operating system type
Examples: **metadata.os:Windows**

Product (Web Service) – Search by known product name
Example: **443.https.get.metadata.product:nginx**

Manufacturer – Search for known manufacturers
Example: **metadata.manufacturer:"Huawei"**

Microsoft SMBv1 – Search for instances of SMBv1
Example: **445.smb.banner.smbv1_support:true**

Dates & Ranges

Date: After – Search for findings that appear after a date
Example: **updated_at:[2018-12-15 TO *]**

Date: Before – Search for findings that appear before a date
Example: **updated_at:[* TO 2018-12-31]**

Date : Range – Search for findings that appear within a range
Example: **updated_at:[2018-12-15 TO 2018-12-31]**

Web Apps

Page's Title – Search for text in page's title
Example: **443.https.get.title:"Index of /ftp"**

Page's HTML Body – Search body of webpage for text string
Example: **443.https.get.body:"XML-RPC server accepts"**

Web Technologies – Search for specific web technologies
Example: **443.https.get.metadata.product: php**

TLS Version – Determine most recent version supported
Example: **443.https.tls.version:TLSv1.2**

SSLv3 – Find instances of SSLv3
Example: **443.https.ssl_3.support:true**

Expired Certificates – Search for expired HTTPS certs
Example:
443.https.tls.certificate.parsed.validity.end:[2018-12-31 TO *]

Self-Signed Certificates – Search for expired HTTPS certs
Example:
443.https.tls.certificate.parsed.signature.self_signed:true

Invalid Cert Signatures – Find invalid cert signatures
Example:
443.https.tls.certificate.parsed.signature.valid:false

Trusted Certs – Determine trusted certs by browsers
Example: **443.https.tls.validation.browser_trusted**

Heartbleed – Find potential instances of Heartbleed vuln
Example:
443.https.heartbleed.heartbleed_vulnerable:true

Tags

A list of common tags that I've found useful:
bacnet, database, DSL/cable modem, embedded, Heartbleed, industrial control system, known-private-key, modbus, mssql, mysql, network, oracle, postgres, printer, rdp, remote_display, raspberry pi, scada, smb, vnc

A special thank you to the Censys team (@censysio) and the University of Michigan (@UMich) !!!

By Michael Holcomb (@mdholcomb)

Figura 160. Cheatsheet con consultas frecuentes de *Censys*

DNS: Herramientas específicas para servidores y dominios

MX ToolBox (https://mxtoolbox.com/)

Si nuestro objetivo tiene un dominio a su nombre, esta web puede ofrecernos mucha información pública acerca del mismo que podemos usar para seguir buscando información sobre él. También sirve para diagnóstico de dominios, no solo búsquedas.

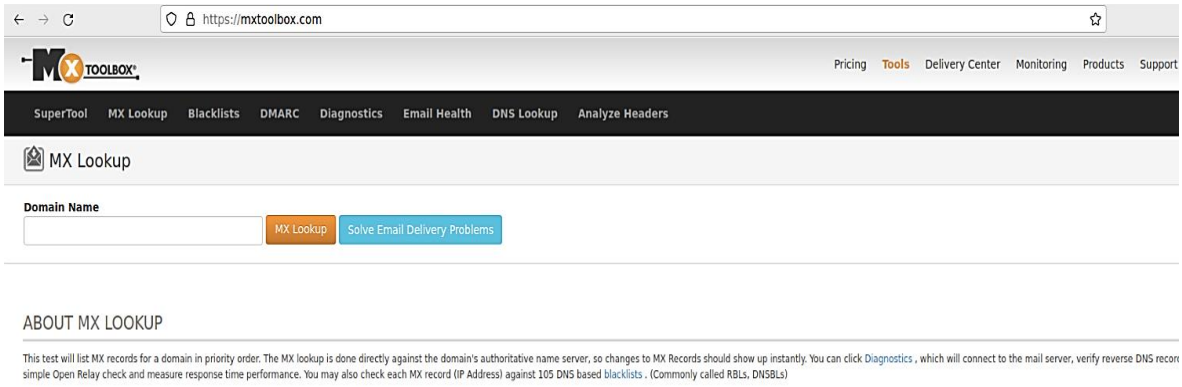


Figura 161. *MX Toolbox*

Domain investigation Toolbox (https://cipher387.github.io/domain_investigation_toolbox/ip.html)

Esta página nos ofrece información sobre IPs y DNS agregando la información ofrecida por varias herramientas de terceros, como se ve en la imagen. Es algo así como una especie de agregador de información que permite saber una gran cantidad de cosas de un objetivo en una sola operación.

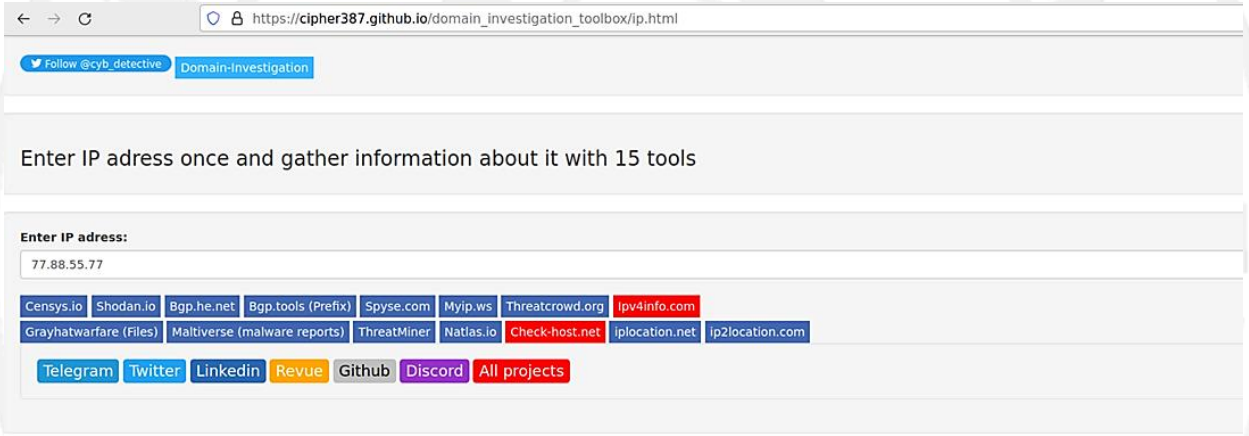


Figura 162. *Domain Investigation Toolbox*

Herramienta web con operaciones individuales: View DNS (https://viewdns.info/)

A veces lo que necesitamos ya no es tanto una cantidad de información masiva sobre un nombre DNS, sino más bien una serie de datos específicos. Si bien todo esto se puede hacer mediante herramientas en línea de comandos, esta web permite hacerlo online, separando cada posible operación de las demás como se ve en la imagen.

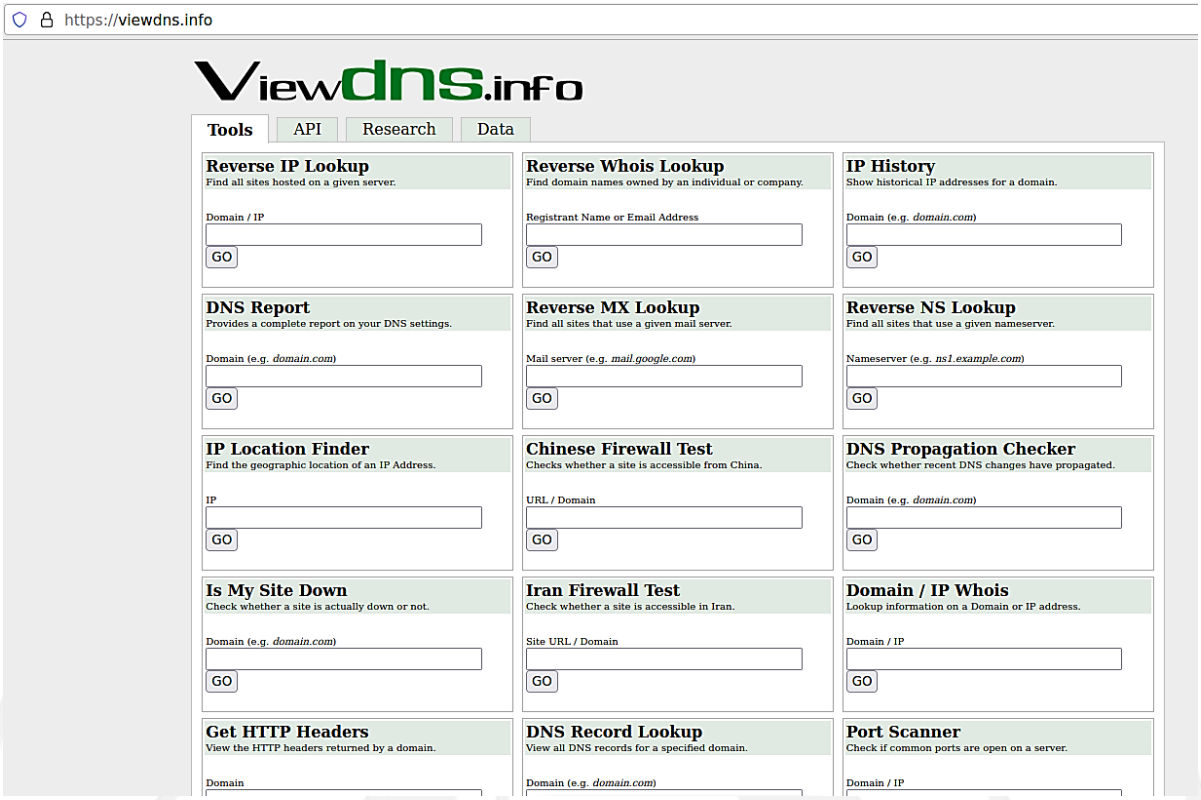


Figura 163. View DNS con una colección de operaciones DNS individuales concretas

FOFA (https://en.fofa.info/)

Un motor de búsqueda para *Threat Intelligence*, que además permite gestionar la superficie de ataque de una empresa más fácilmente. Puedes encontrar una **guía muy útil de consultas** con este motor aquí: https://github.com/FofaInfo/Awesome-FOFA/blob/main/Get%20Started%20with%20FOFA/A%20Beginner%E2%80%98s%20Guide.md

Figura 164. *FOFA*

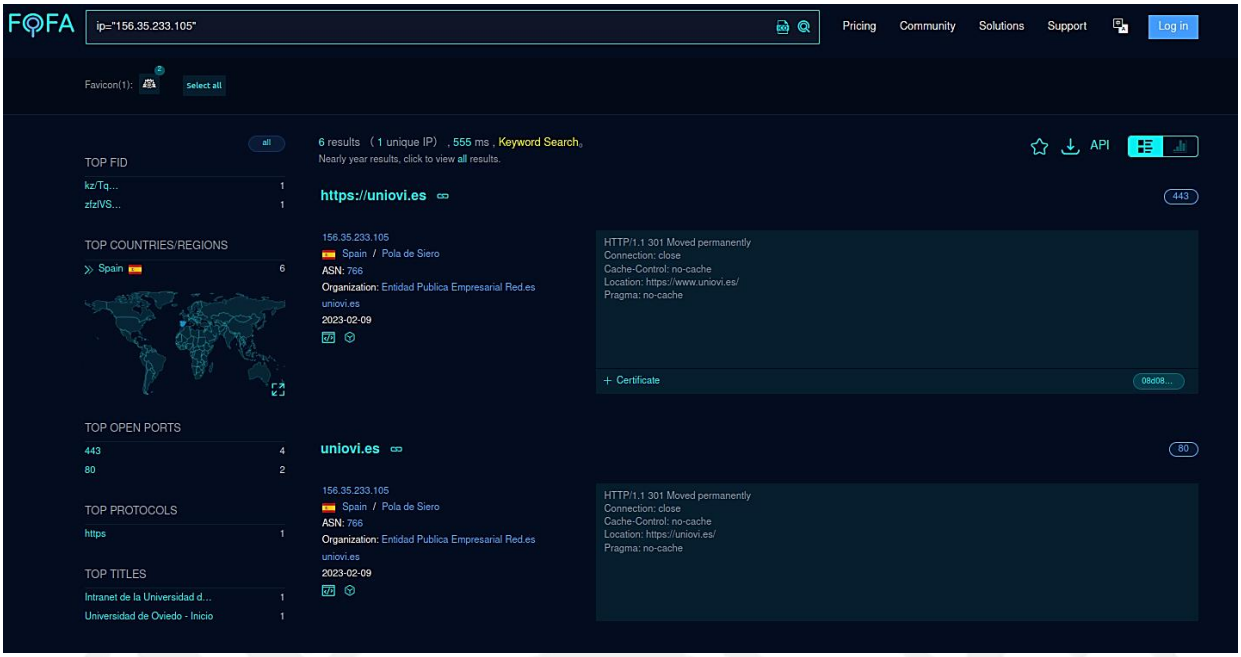


Figura 165. Ejemplo de resultado de búsqueda en *FOFA* sobre Uniovi

FullHunt (https://fullhunt.io/)

Es una base de datos de superficies de ataque de todo Internet. Permite descubrir superficies de ataque, monitorizarlas para detectar la exposición y escanearlas continuamente en busca de las últimas vulnerabilidades de seguridad.

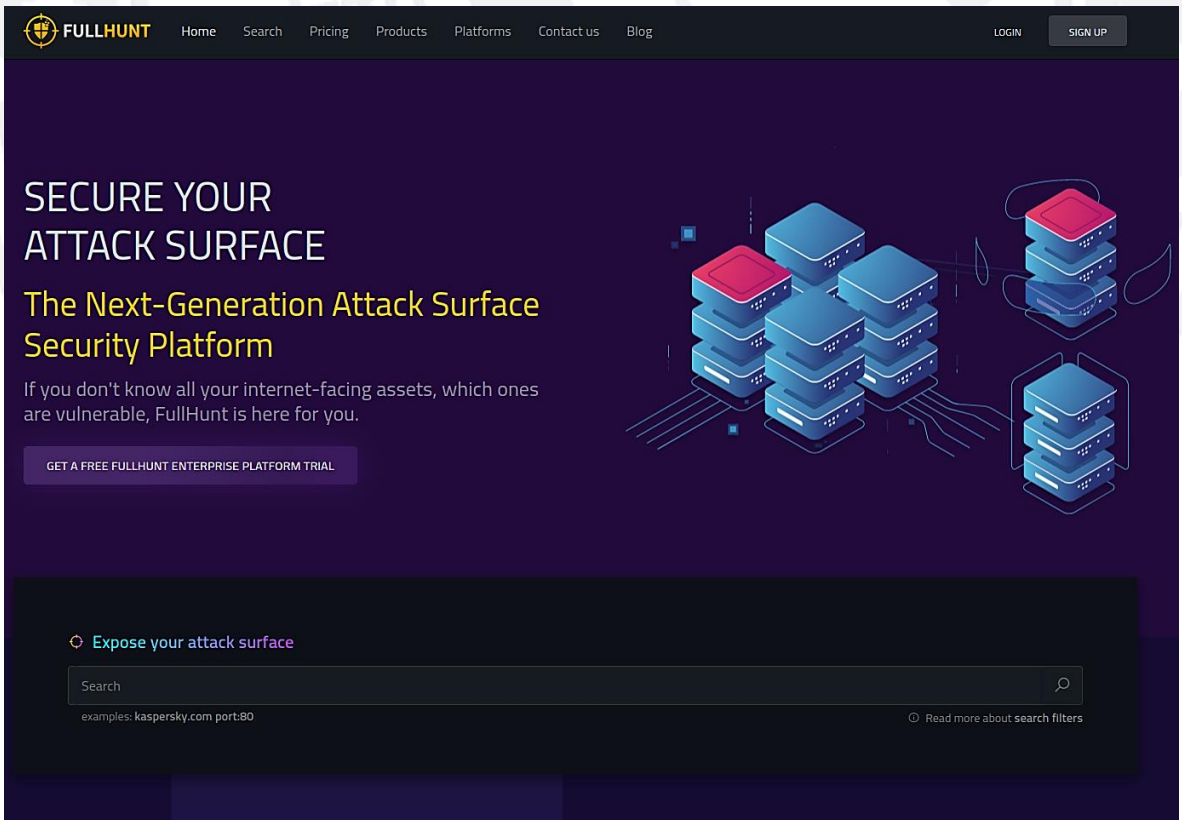


Figura 166. *FullHunt*

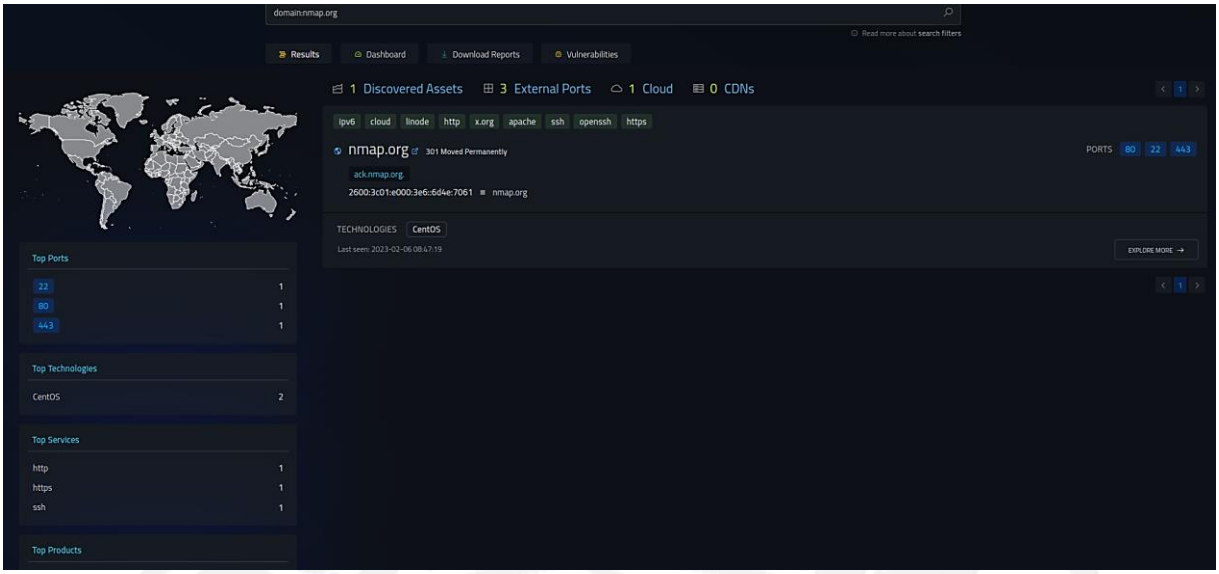


Figura 167. Ejemplo de resultado de búsqueda en *FullHunt* sobre el dominio *nmap.org*

GreyNoise (https://viz.greynoise.io/)

Busca dispositivos conectados a Internet al igual que otros motores de esta sección, trabajando con IPs, identificadores de vulnerabilidades CVE (qué dispositivos tienen ciertas vulnerabilidades) y otros. Su principal utilidad es averiguar quién te ha escaneado alguna de tus máquinas, y saber sus intenciones: si es un ataque dirigido o simplemente una de las muchas arañas web que pululan por Internet de buscadores y otras aplicaciones del estilo. Puedes encontrar un listado de posible información que puede encontrar aquí: https://viz.greynoise.io/cheat-sheet, https://www.netmux.com/blog/greynoise-cheat-sheet

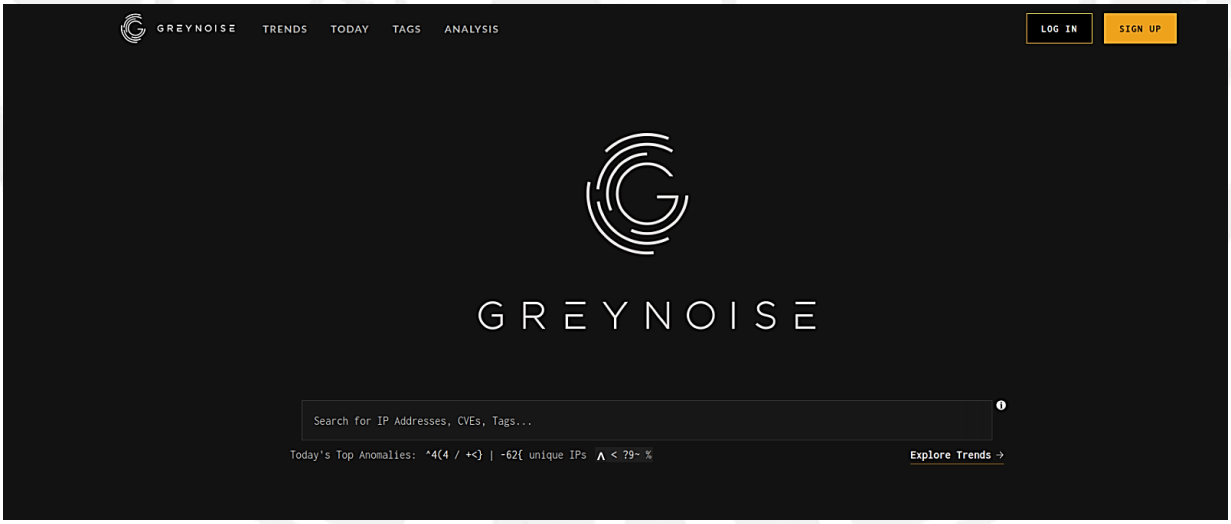


Figura 168. *GreyNoise*

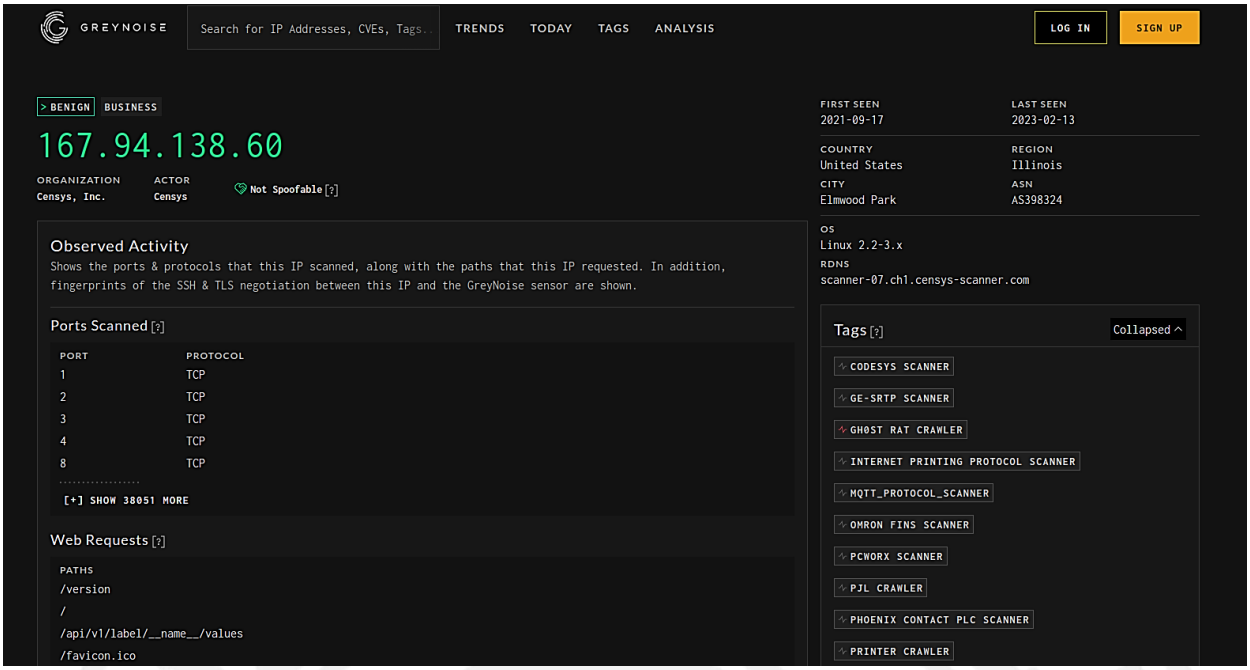


Figura 169. *GreyNoise* detectando un escaneo de *Censys* por una IP que tenía en el log de un *firewall*

IVRE (https://ivre.rocks/)

Framework open-source para hacer reconocimiento de red que se basa en herramientas de este tipo muy conocidas como *Nmap*, *Masscan*, *ZGrab2*, *ZDNS* o *Zeek* para obtener datos (inteligencia de redes). Puede instalarse localmente y guarda la información en una base de datos configurable, incorporando además herramientas para analizarlos.

Figura 170. *IVRE*

Netlas (https://app.netlas.io/responses/)

Busca y supervisa activos conectados a Internet escaneando direcciones IPv4 y rastreando todos los sitios web y aplicaciones web conocidos utilizando protocolos como HTTP, FTP, SMTP, POP3, IMAP, SMB / CIFS, SSH, Telnet, SQL y otros. En https://twitter.com/Netlas_io/status/1719625320565813394 puedes ver muchos de sus **filtros comunes**:

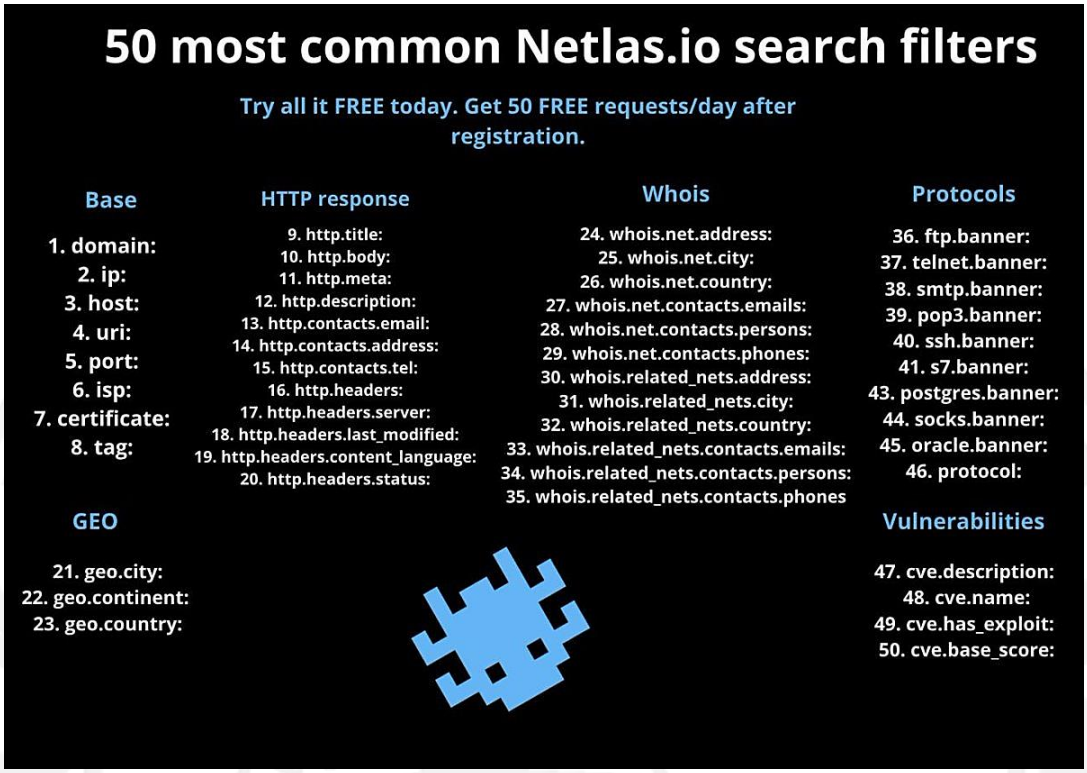


Figura 171. Filtros comunes de Netlas

Además, en esta dirección tenemos una **referencia completa de sus consultas**: https://github.com/netlas-io/netlas-cookbook

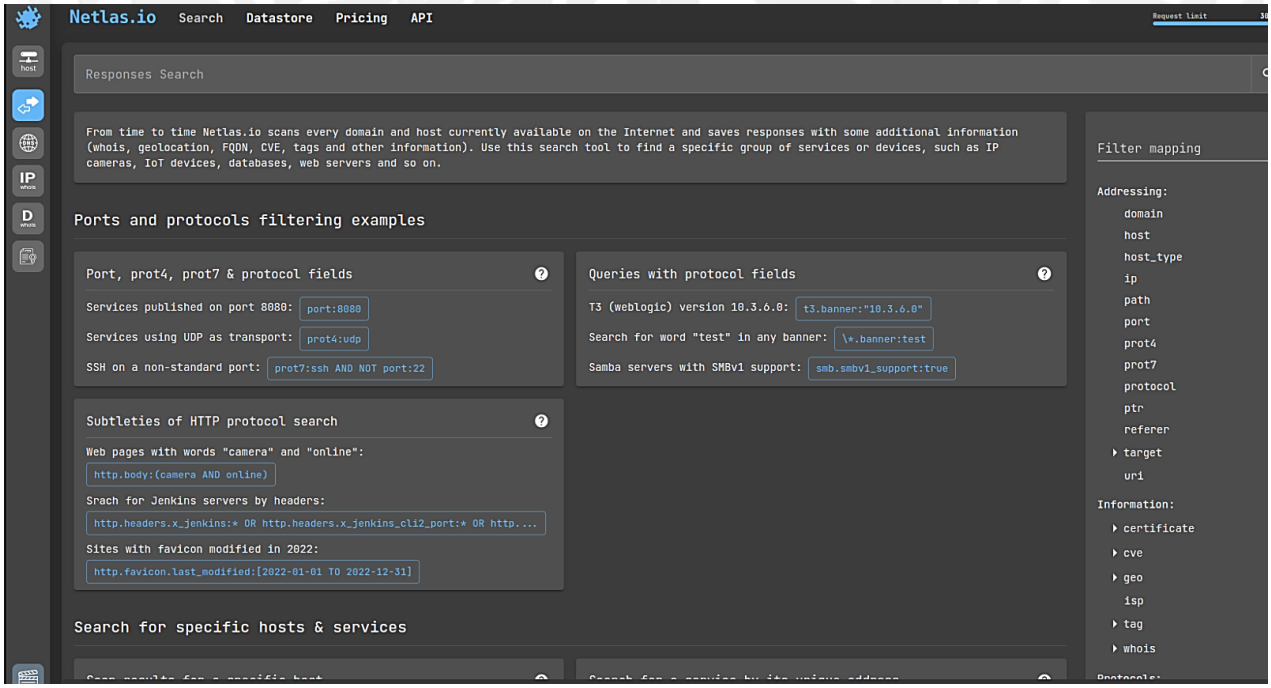


Figura 172. *Netlas*

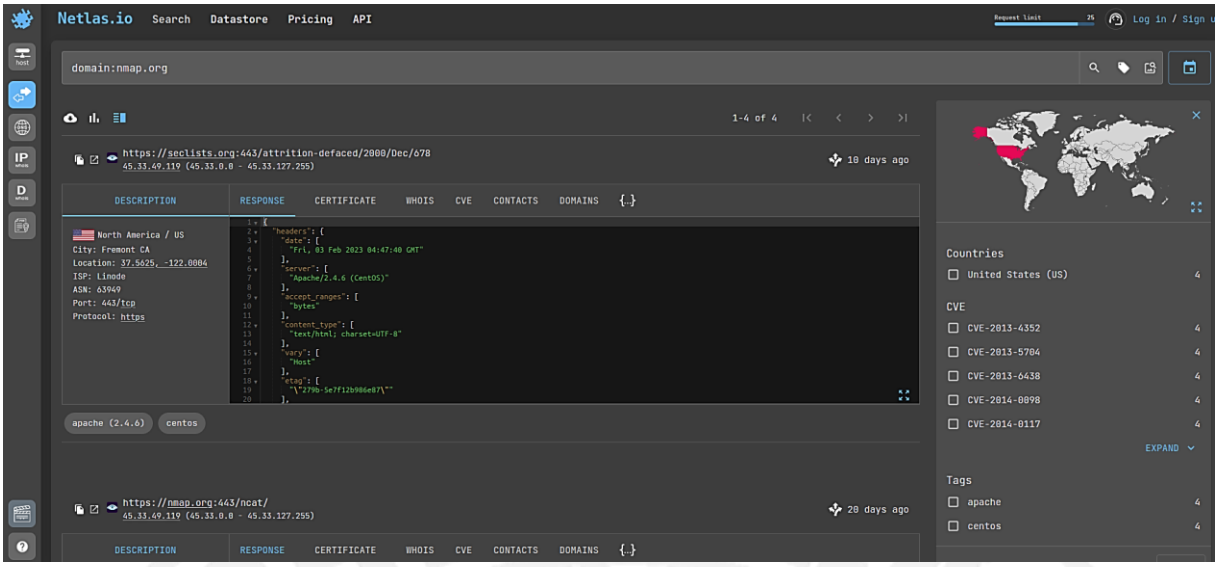


Figura 173. Ejemplo de uso de *Netlas* contra el dominio *nmap.org*

OnHype (https://www.onyphe.io/)

Motor de búsqueda orientado a la defensa cibernética para datos de inteligencia de amenazas variados recopilados mediante el rastreo de varias fuentes disponibles en Internet. En este enlace puedes ver una serie de **consultas típicas** para sacarle partido al motor: https://www.onyphe.io/files/onyphecheatsheetfreeview.pdf

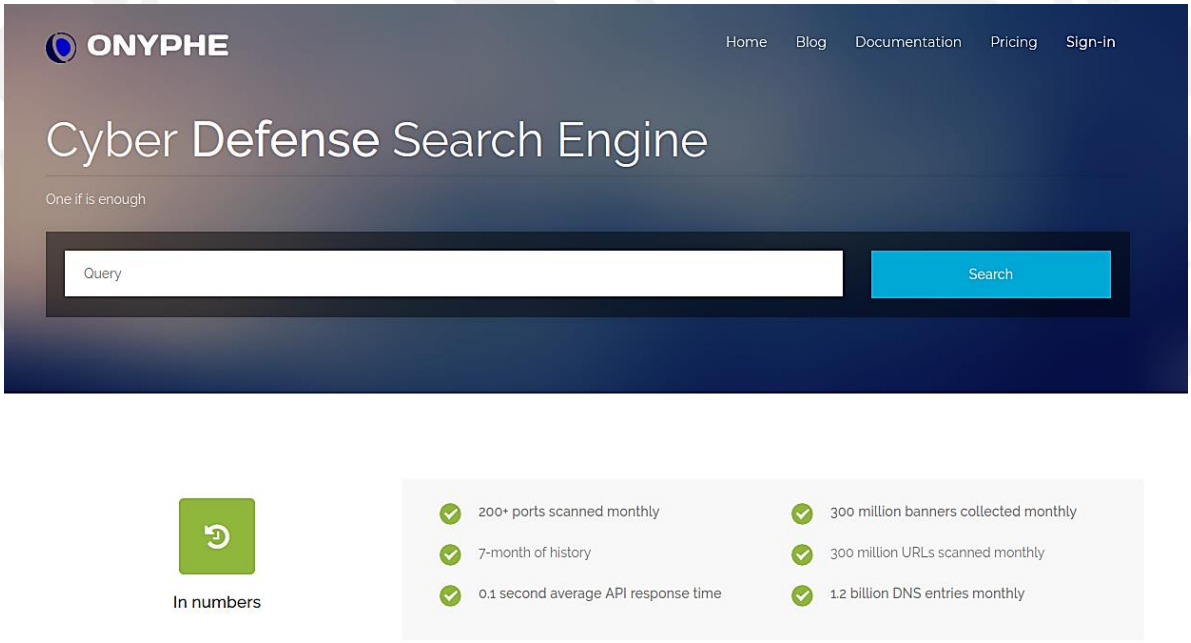


Figura 174. *OnHype*

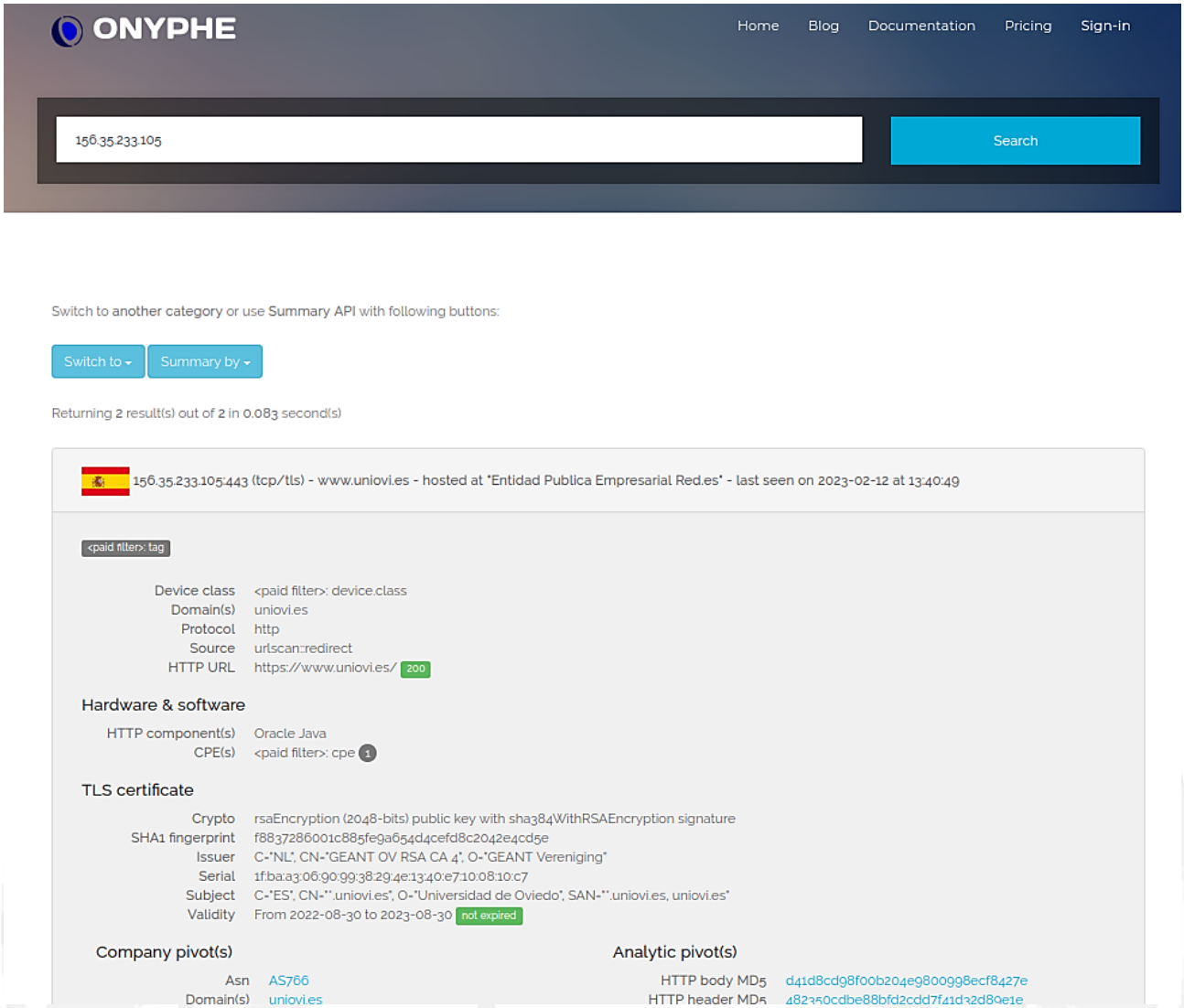


Figura 175. Ejemplo de búsqueda de *OnHype* contra uniovi.es

PulseDive (https://pulsedive.com/)

Ofrece inteligencia de amenazas de calidad buscando e investigando millones de IP, dominios, URL y amenazas, todo de forma gratuita, indicando si suponen o no un riesgo de seguridad entre otras cosas.

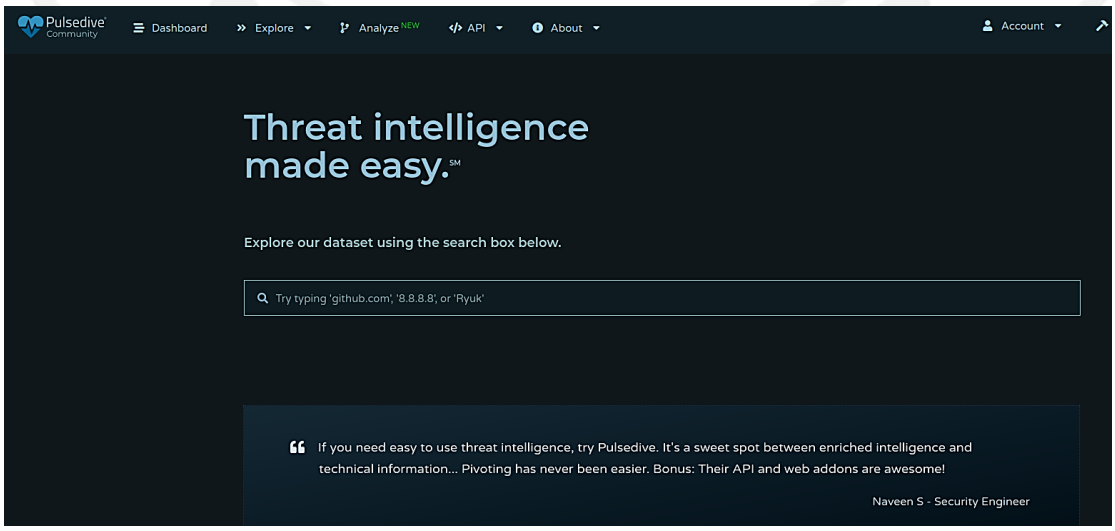


Figura 176. *PulseDive*

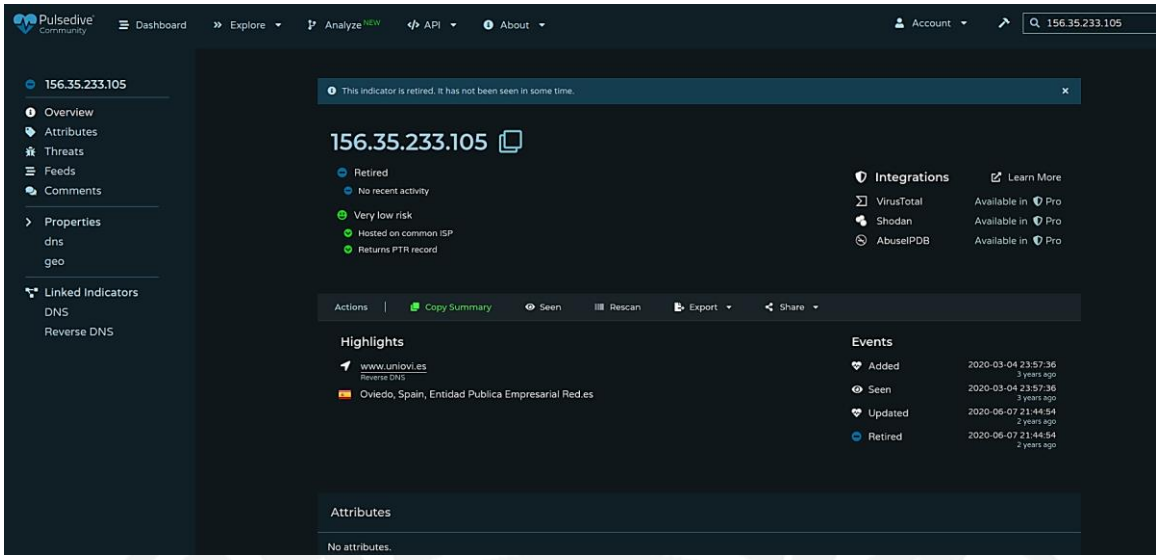


Figura 177. Ejemplo de uso de *PulseDive* contra *uniovi.es*

Shodan (https://www.shodan.io/)

El popular motor de investigación para realizar investigaciones de seguridad en dispositivos IoT conectados a Internet que vimos en el **Laboratorio 2** y en el **Seminario 1**. Esta herramienta es utilizada por miles de expertos en seguridad, investigadores, CERT, grandes organizaciones y otros en todo el mundo.

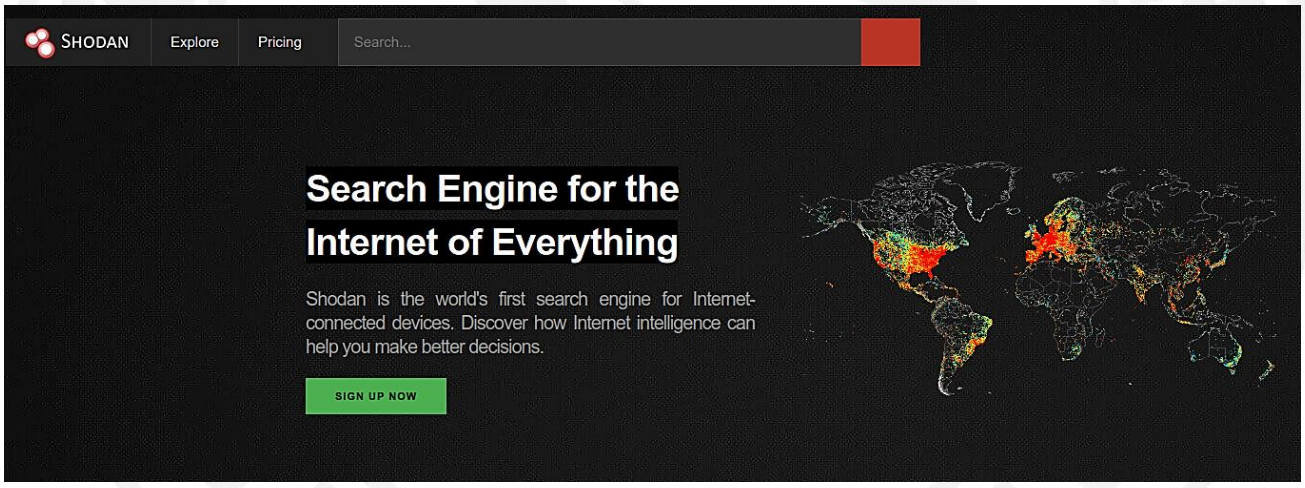


Figura 178. *Shodan*

El siguiente *cheatsheet* resume las **consultas más frecuentes** que puedes hacerle:

Cheatography

Shodan Cheat Sheet

by sir_slammington via cheatography.com/147122/cs/31955/

Common General Search Filters

ip:	Filter results by specific IP address.
asn:	Filter results by specific ASN ID.
hostname:	Filter results by specific hostname.
port:	Filter results by specific port number of service.
net:	Filter results from specified CIDR block.
isp:	Filter results by devices assigned a particular address (space) from a specified ISP.
city:	Filter results by specific city.
country:	Filter results by specific two-digit country code.
os:	Filter results by particular OS.
product:	Filter results by particular software.
version:	Filter results by specified version of software.

Common Premium API Search Filters

vuln:	Filter results by particular vulnerability ID (commonly CVE).
tag:	Filter results by tags on device.

HTTP Filters

http.component:	Filter results by a particular web technology.
http.status:	Filter results by specific status code.
http.html:	Filter results by strings found in HTML of files served.
http.title:	Filter results by string found in title of web pages served.

Common CLI Commands

count	Returns the number of results for a search.
domain	View all available information for a domain.
download	Download search results and save them in a compressed JSON file.
honeyscore	Check whether the IP is a honeypot or not.
host	View all available information for an IP address.
parse	Extract information out of compressed JSON files.
scan	Scan an IP/ netblock using Shodan.
search	Search the Shodan database.

Common CLI Search Fields

ip_str
port
org
hostnames
os
country
city

These will display their values upon a search, but won't provide statistics.

Common CLI Stats Facets

asn
city
country
cloud.provider
cloud.service
device
domain
ip
org
os

Common CLI Stats Facets (cont)

version

These will return statistical information about a given series of devices found on the public facing Internet. For example, it could be used to return the most common version found among devices running MariaDB in a particular ASN.

Use Case Examples

host: 8.8.8.8	shodan host 8.8.8.8	Display information about a Google's public DNS.
asn:15169 product:mysql	shodan stats asn:15169 product:mysql	Show information about devices within Google's ASN that run MySQL.
microsoft iis 6.0	shodan search --fields ip_str,port,org,hostnames microsoft iis 6.0	Detect IIS servers running on 6.0.
Navigate to https://honeyscore.shodan.io/ and enter target IP.	shodan honeyscore [TARGET]	Detect if given target is a honeypot or not.

Column one is the search you would perform in the Web UI. Column two is the search you would perform using the CLI utility, and the third column is an explanation of the search.

By sir_slammington
cheatography.com/sir-slammington/

Published 2nd May, 2022.
Last updated 2nd May, 2022.
Page 1 of 1.

Sponsored by Readable.com
Measure your website readability!
https://readable.com

Figura 179. Consultas más frecuentes que puedes usar en *Shodan*

Otro algo más sencillo de usar es este:

Figura 180. Más consultas frecuentes que puedes usar en *Shodan*

Wigle (https://wigle.net/)

WiGLE (o *Wireless Geographic Logging Engine*) es un sitio web para **recopilar información sobre los diferentes puntos de acceso inalámbricos en todo el mundo**. Los usuarios pueden registrarse en el sitio web y cargar datos de puntos de acceso como coordenadas GPS, SSID, dirección MAC y el tipo de cifrado utilizado en los puntos de acceso descubiertos. Es como el *Warchalking* (**Seminario 2**), pero a lo bestia 😊

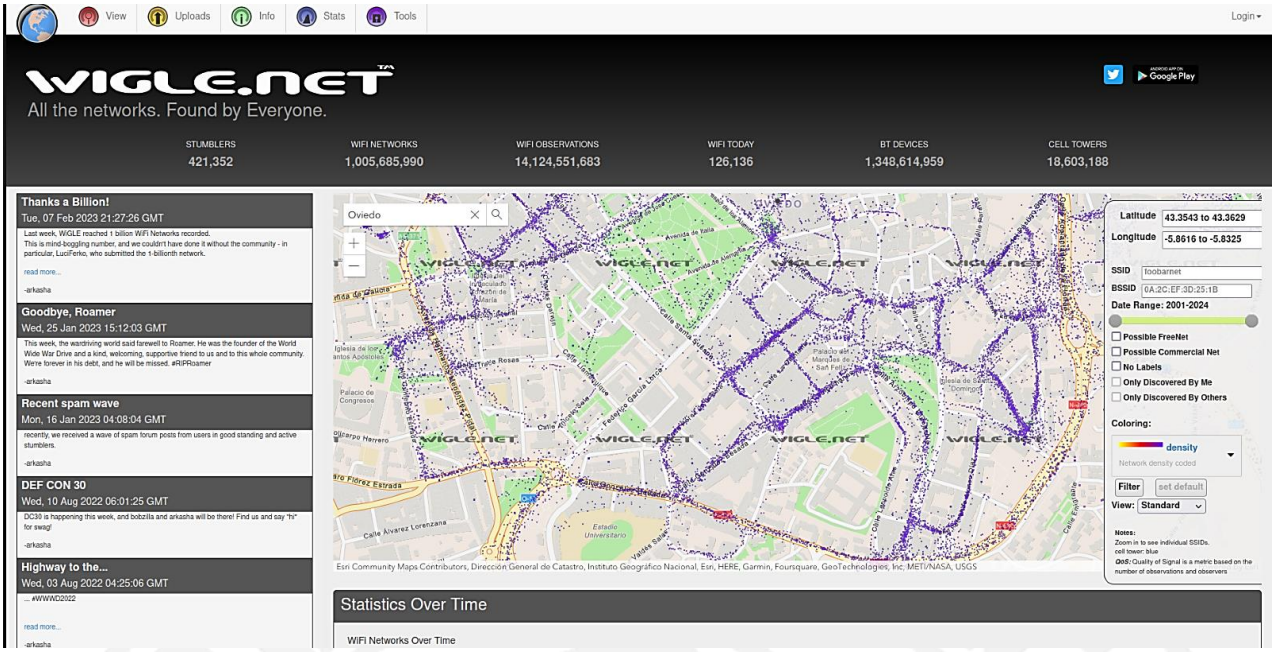


Figura 181. *Wigle* (mostrando Wi-Fis de Oviedo)

ZoomEye (https://www.zoomeye.org/)

Hace un mapeo de máquinas y su topografía a nivel local o global basándose en IPv6, IPv4 y bases de datos de nombres de dominio. Encuentra protocolos de servicio y puertos activos en un instante determinado de tiempo usando un lenguaje de consultas sencillo. Trabaja por ejemplo sobre IPs y ficheros y **requiere registro gratuito**. Este fichero contiene una serie de **consultas típicas útiles**: https://github.com/interference-security/zoomeye-data/blob/master/usage.txt

Figura 182. *ZoomEye*

Motores de búsqueda para sitios web

Advance Online Utilities (https://centralops.net/co/)

En esta dirección podemos encontrar diversas herramientas de exploración de sitios web

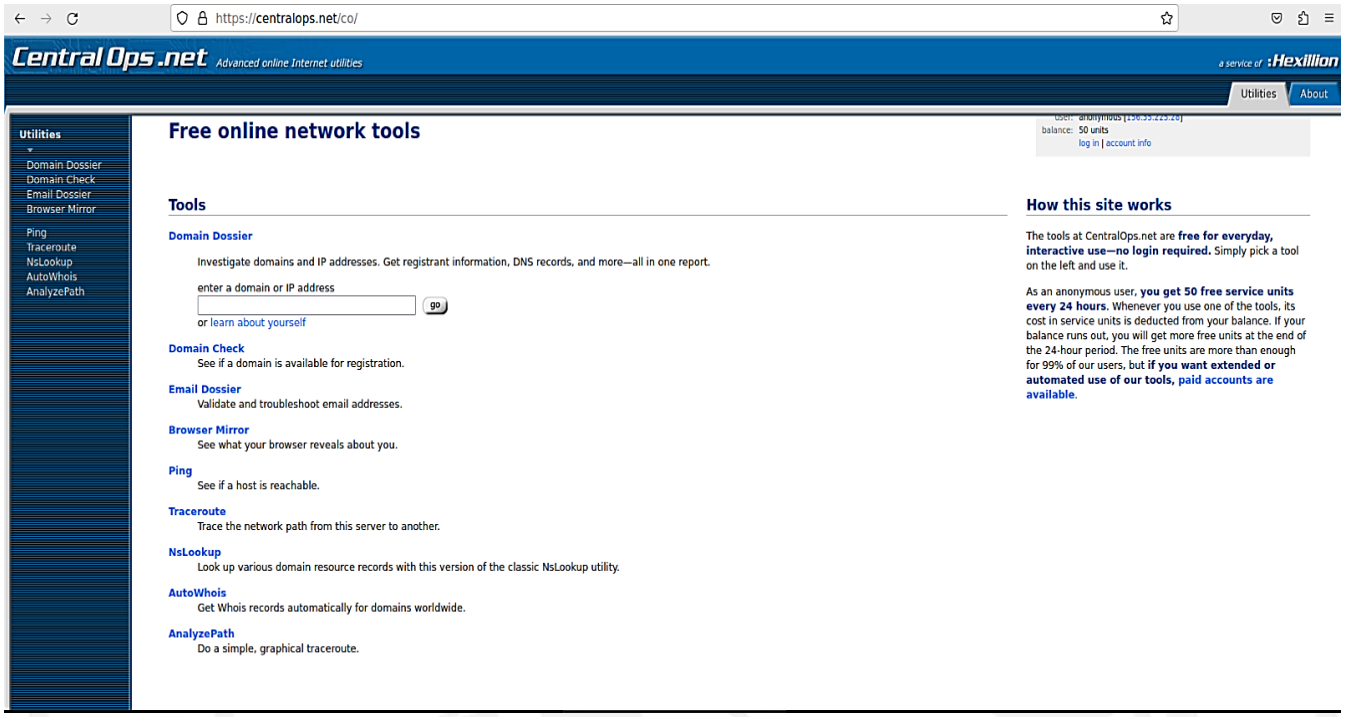


Figura 183. *Advance Online Utilities*

Cisco Reputation Check (https://www.talosintelligence.com/)

Web que comprueba la reputación de una URL o IP que le pasemos

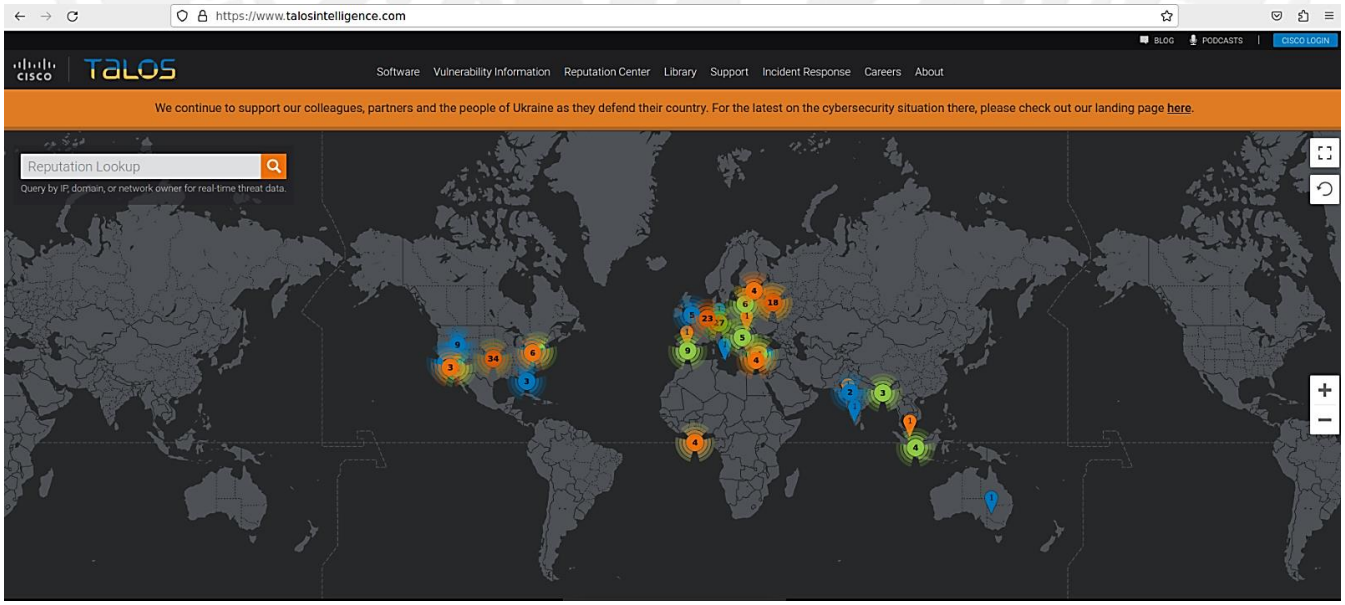


Figura 184. *CISCO Reputation check*

Cloudflare Radar URL Scanner (https://radar.cloudflare.com/scan)

Motor de búsqueda que a partir de una URL te hace un informe con muchos detalles técnicos de la web que se encuentra tras las misma: si tiene indicios de ser phishing, datos de su certificado TLS, datos de peticiones y respuestas, registros DNS, si trata con seguridad las cookies, las tecnologías y librerías usadas...

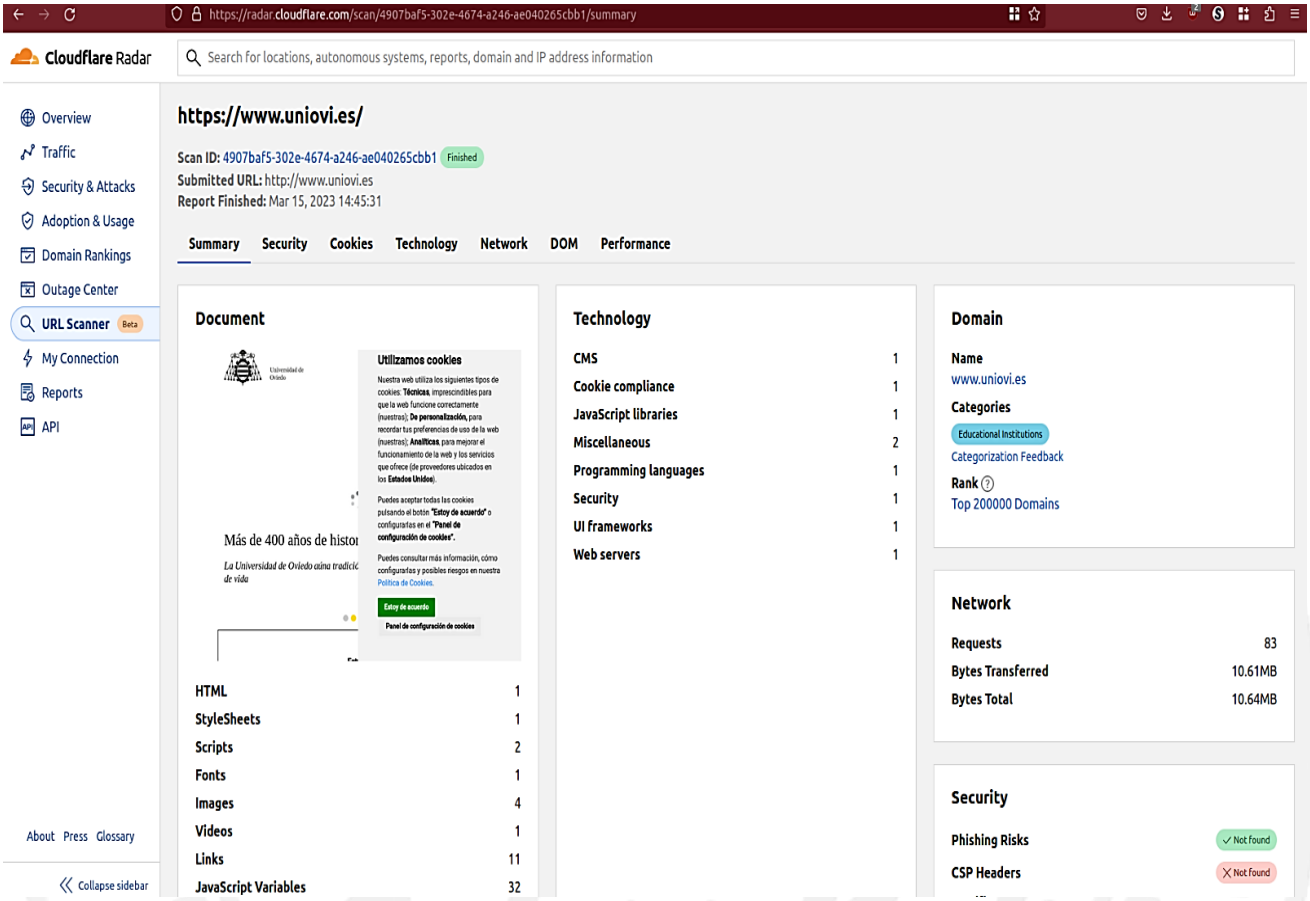


Figura 185. *Cloudflare Radar URL Scanner*

CyberGordon (https://cybergordon.com/)

Web que comprueba la reputación online de una IP o web para saber si puede ser o no peligrosa.

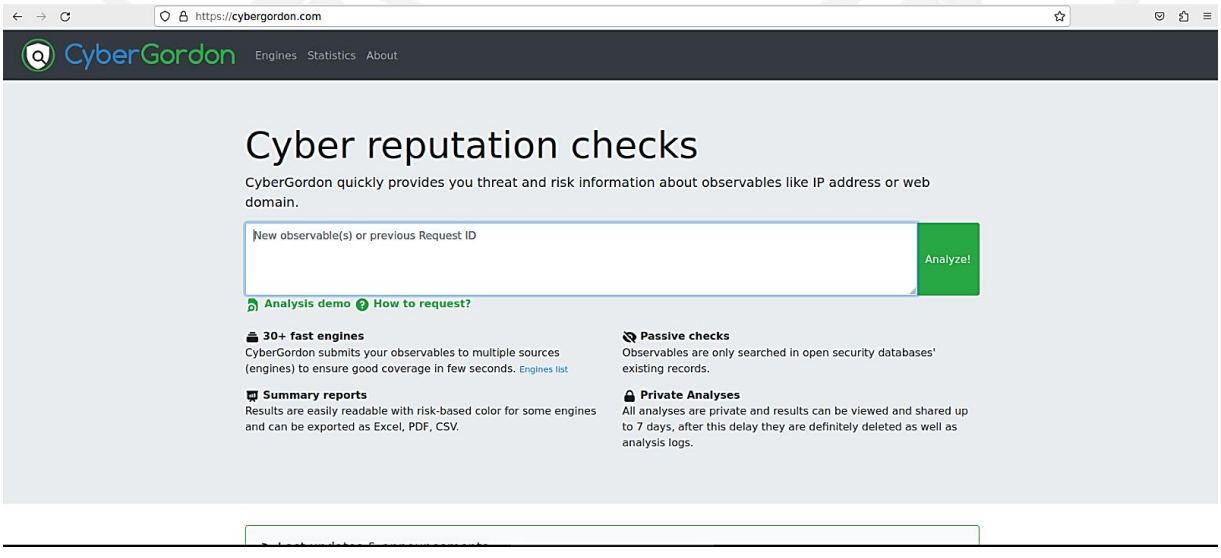


Figura 186. *CyberGordon*

IBM Reputation Check (https://exchange.xforce.ibmcloud.com/)

Otra web que comprueba la reputación de otras, si bien esta es de pago pero muy popular en entornos empresariales (por si necesitas una web de esta clase en tu trabajo futuro)

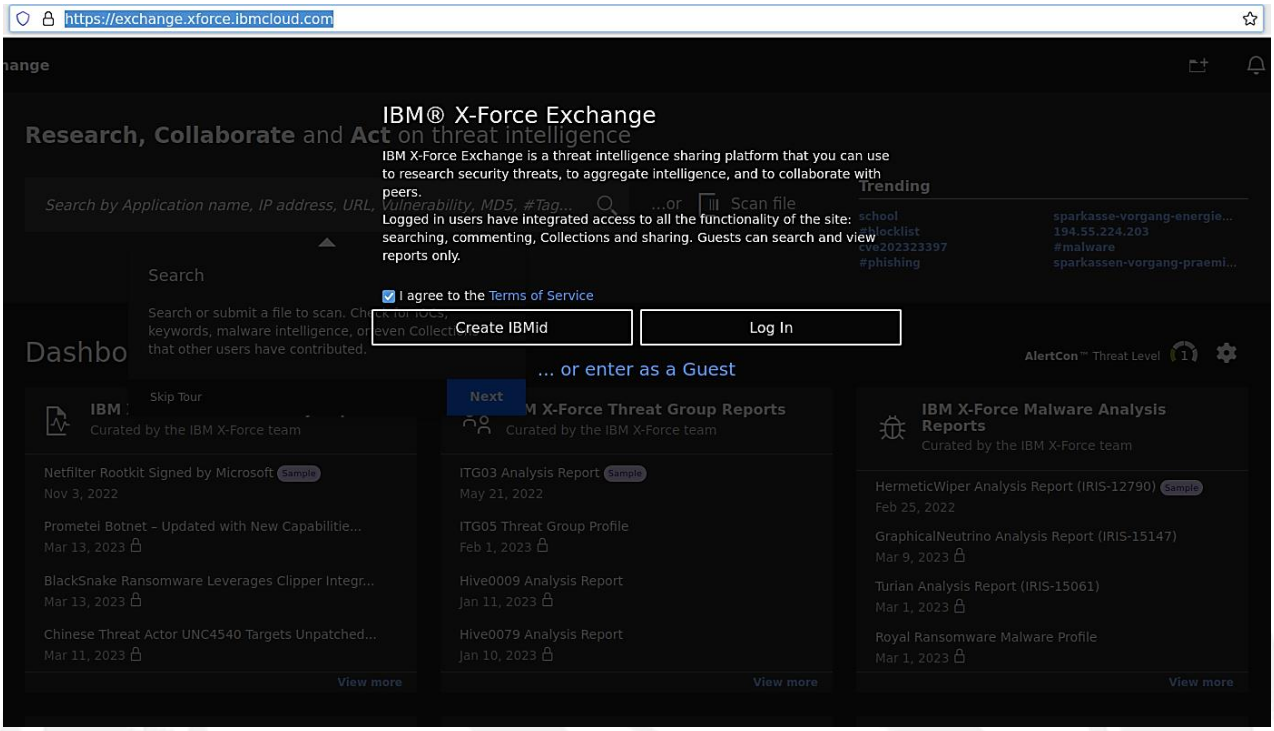


Figura 187. IBM *X-Force Exchange*

IP Reputation Check (https://www.abuseipdb.com/)

Otra web que permite saber si una IP se ha identificado como origen de *spam*, ataques y otras formas de abuso online.

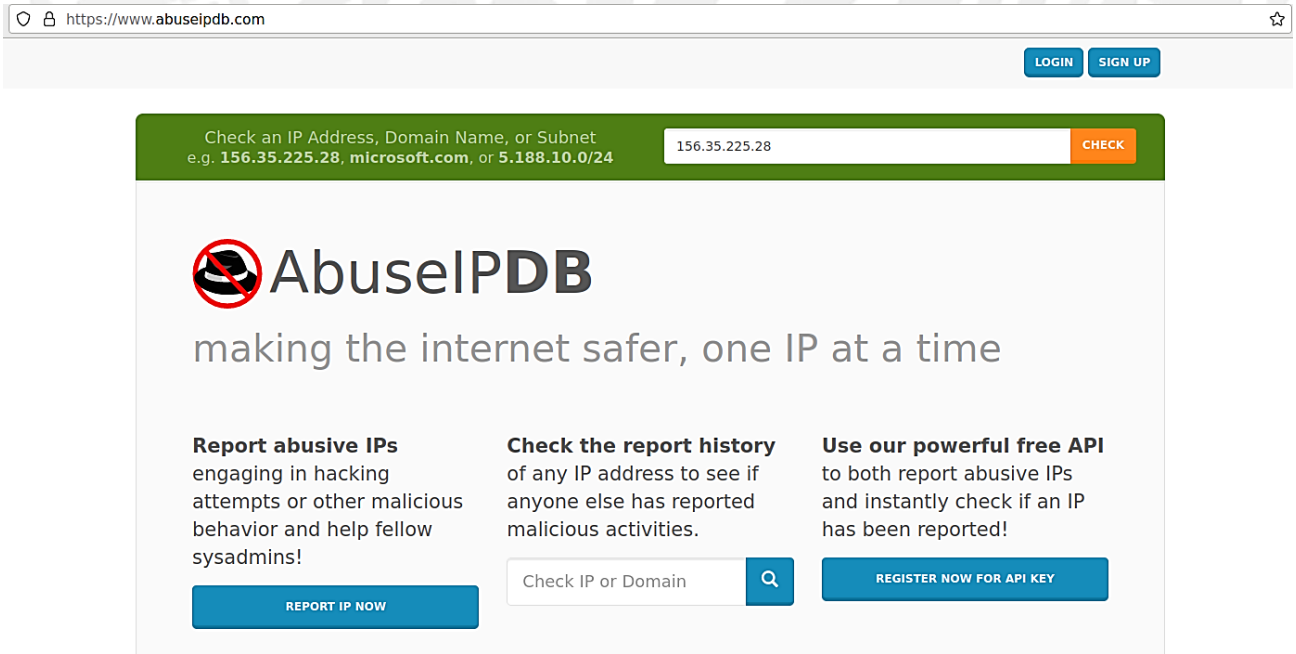


Figura 188. *Abuse IP DB*

PublicWWW (https://publicwww.com/)

Es una herramienta de marketing digital e investigación de marketing de afiliados que encuentra cualquier fragmento alfanumérico, firma o palabra clave en el código HTML, JS y CSS de las páginas web. Es una herramienta útil para el marketing digital y la investigación de marketing de afiliados, pero también para localizar determinados trozos de código claves en distintas web.

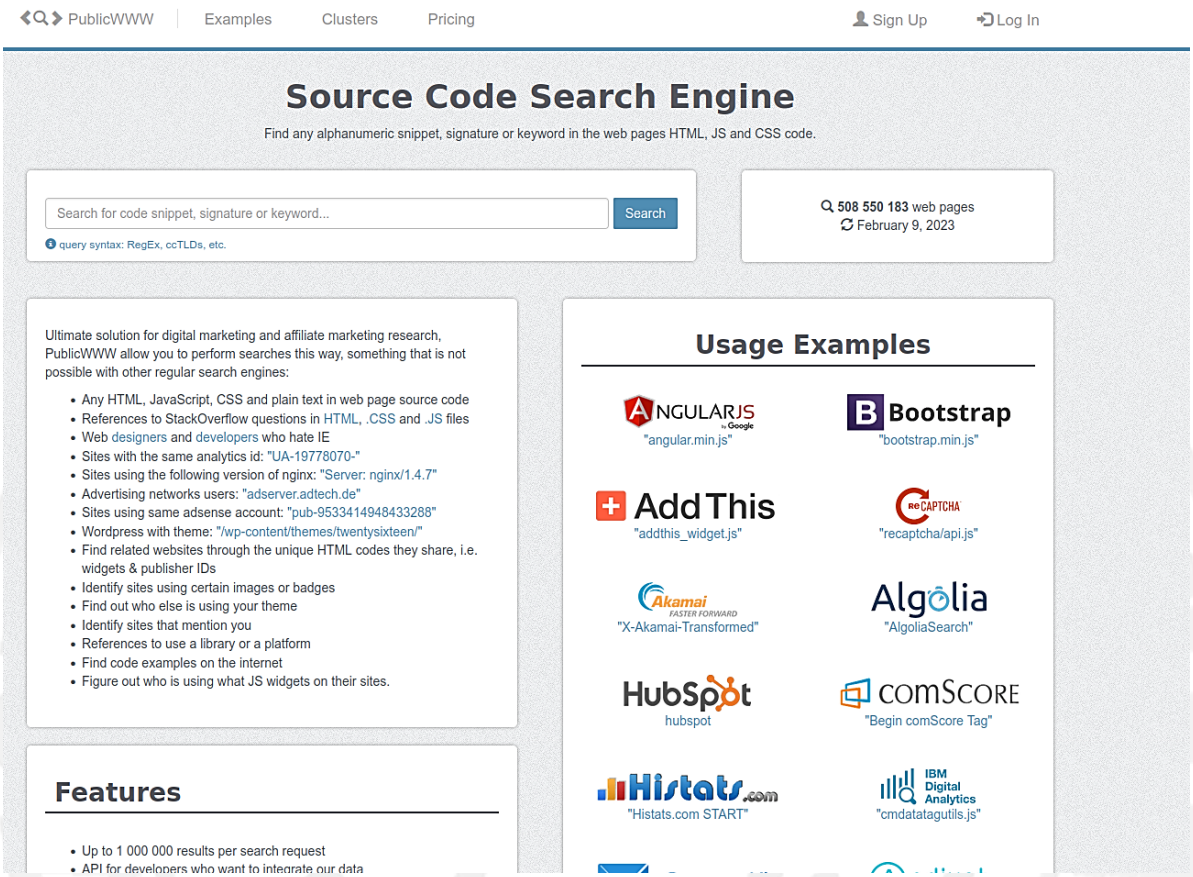


Figura 189. *PublicWWW*

URL Category Finder (https://sitereview.bluecoat.com/#/)

Esta web permite averiguar qué tipo de web es la que se sirve desde una URL que le demos, para saber qué nos podremos encontrar si la visitamos.

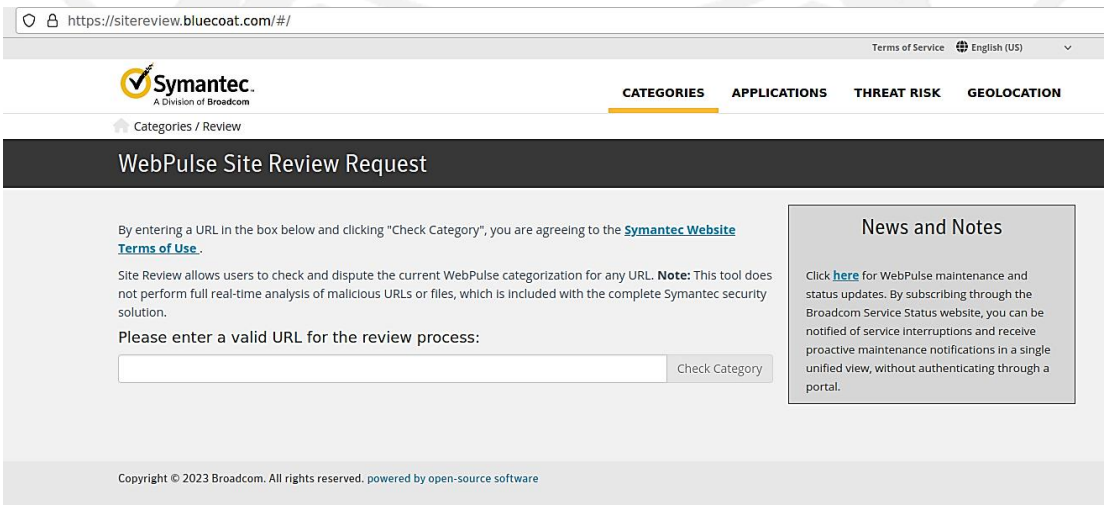


Figura 190. *URL Category Finder*

URLScan (https://urlscan.io)

Servicio gratuito para escanear y analizar sitios web. Tiene procesos automatizados que navegarán a las URL suministradas como un usuario normal y registrará la actividad que crea esta navegación de página, ofreciendo información variada pero de una forma que nuestra máquina queda aislada de la navegación.

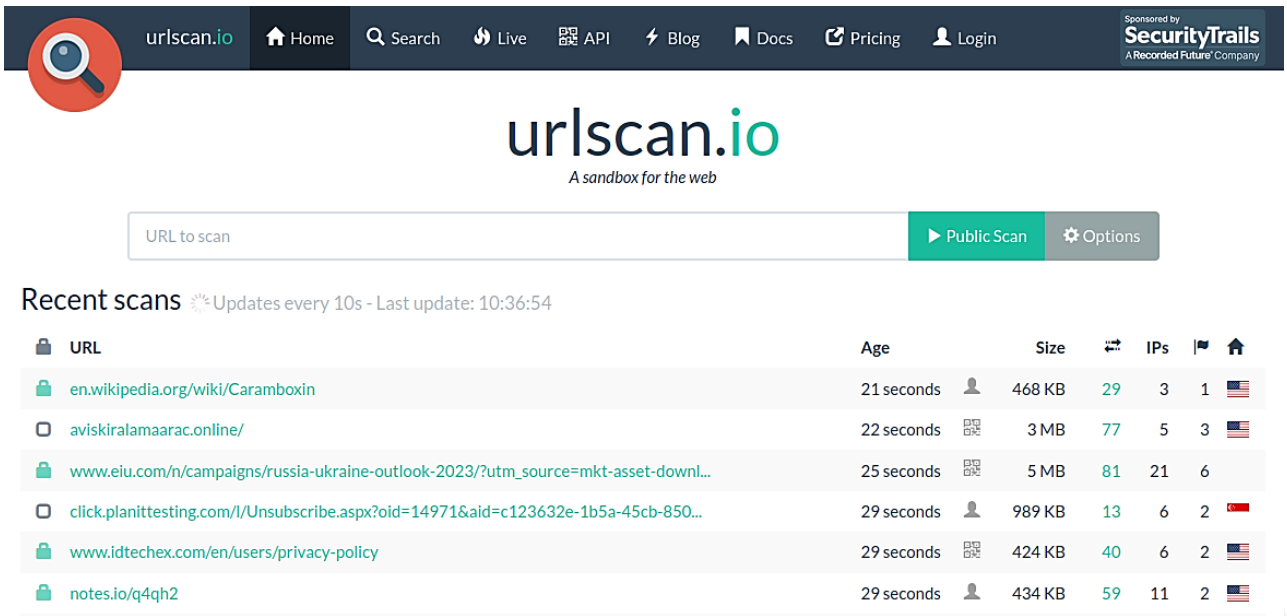


Figura 191. *URLScan*

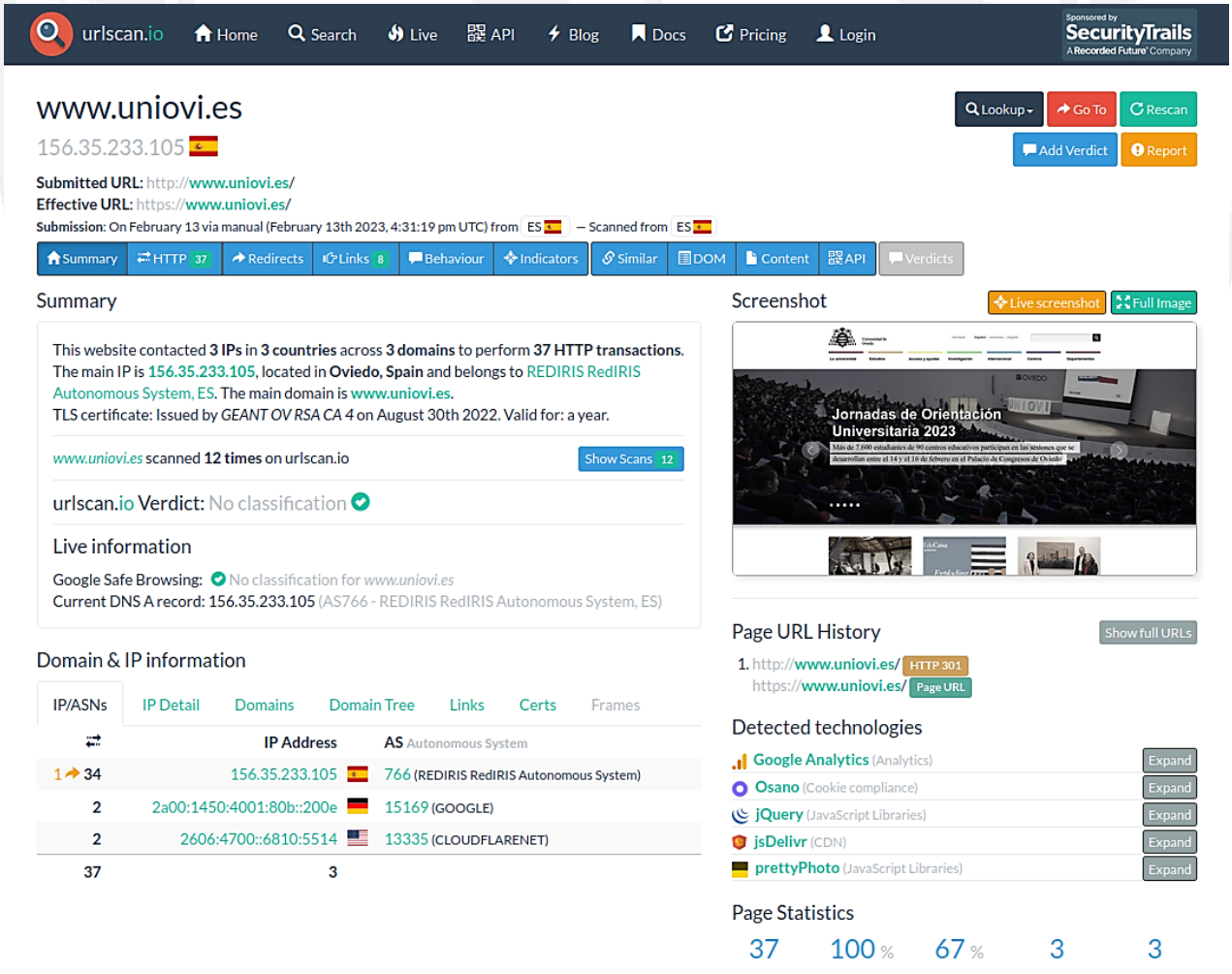


Figura 192. Escaneo de *URLScan* contra *uniovi.es*

Website Profiler Tool (https://builtwith.com/)

Esta web permite identificar las tecnologías que se han usado para construir las web cuya URL le consultemos, lo que puede usarse luego para continuar con nuestra investigación del propietario.

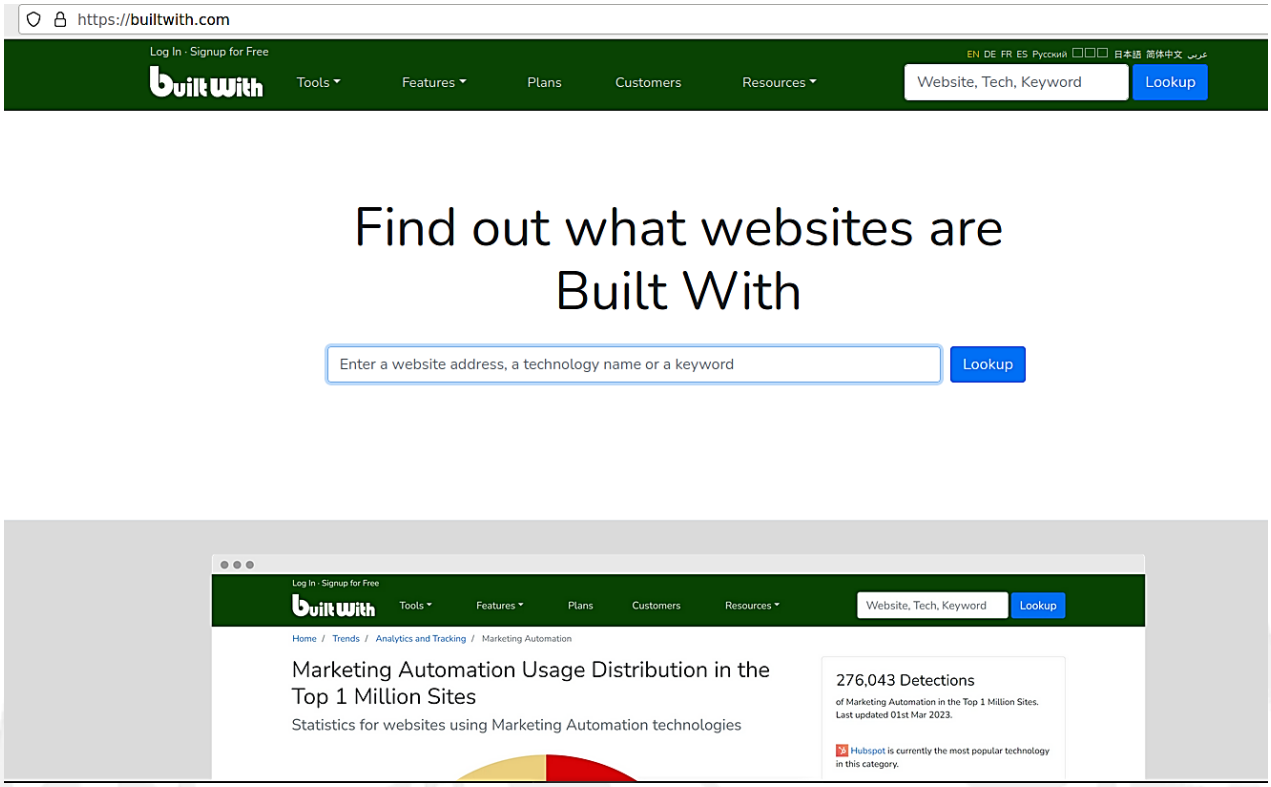


Figura 193. ***Website Profiler Tool***

Motores de búsqueda de vulnerabilidades

CyberChef (https://gchq.github.io/CyberChef/)

Esta es una herramienta para manipular datos de muchas formas distintas: codificación, criptografía...**no es en sí un motor de búsqueda de vulnerabilidades**, pero puede usarse para descubrir vulnerabilidades a partir de procesar cosas encontradas por otros motores, al revelar realmente qué son y por tanto poder suponer un peligro para la web que las tiene.

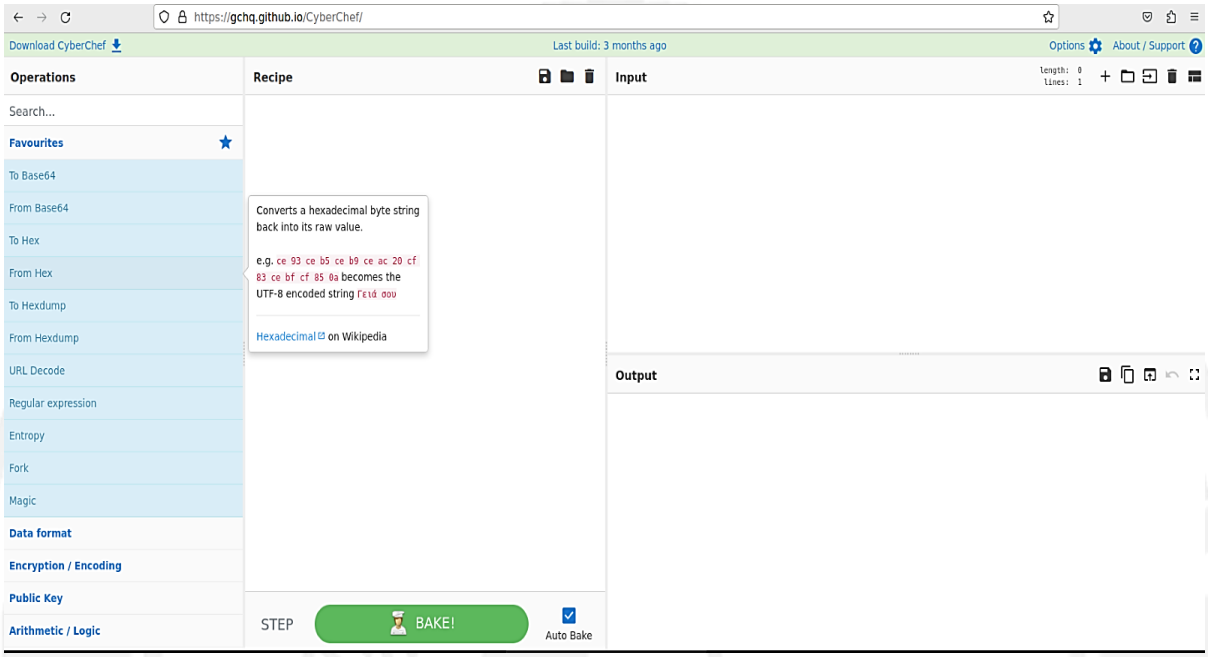


Figura 194. *CyberChef*

Packet Storm Security (https://packetstormsecurity.com/)

Sitio web que ofrece herramientas de seguridad informática actuales e históricas, *exploits* y avisos de seguridad. Operado por un grupo de entusiastas de la seguridad que publican nueva información de seguridad y ofrecen herramientas con fines educativos y de prueba.

Figura 195. Packet Storm Security

Vulners (https://vulners.com/)

Base de datos de seguridad que contiene descripciones de una gran cantidad de vulnerabilidades de software en un formato legible por una máquina y que se mantiene muy al día de las últimas amenazas de seguridad, permitiendo además, previo registro, hacer escaneo de vulnerabilidades a las IPs que especifiquemos.

Figura 196. *Vulners*

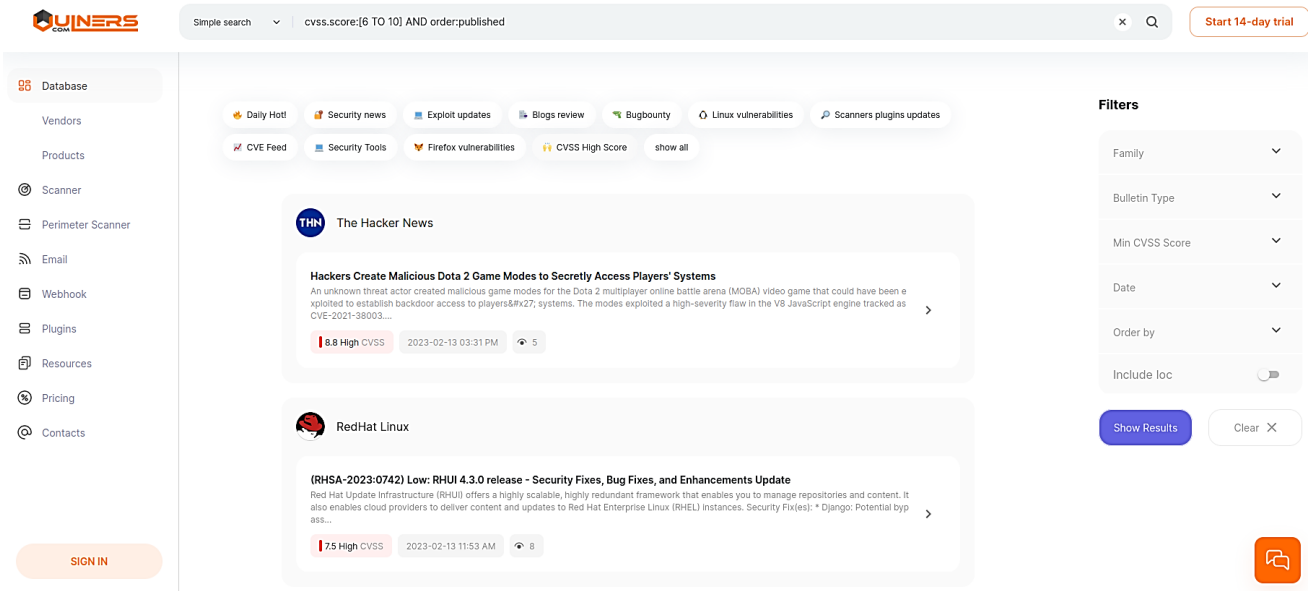


Figura 197. Ejemplo de búsqueda de vulnerabilidades conocidas con *Vulners*

Motores de búsqueda para filtraciones de datos generales

CRT.sh (https://ui.ctsearch.entrust.com/ui/ctsearchui)

Buscador de certificados que hayan sido registrados por *Certificate Transparency* (CT). Es un sitio donde se pueden encontrar todos los certificados SSL o TLS del dominio objetivo en particular y por tanto **consultar los datos que se usaron al registrarlos**, que pueden tener nombres y otra información del objetivo que se puede usar en otras operaciones OSINT. Es también una forma efectiva de descubrir dominios del objetivo que puedan no ser visibles de otra forma.

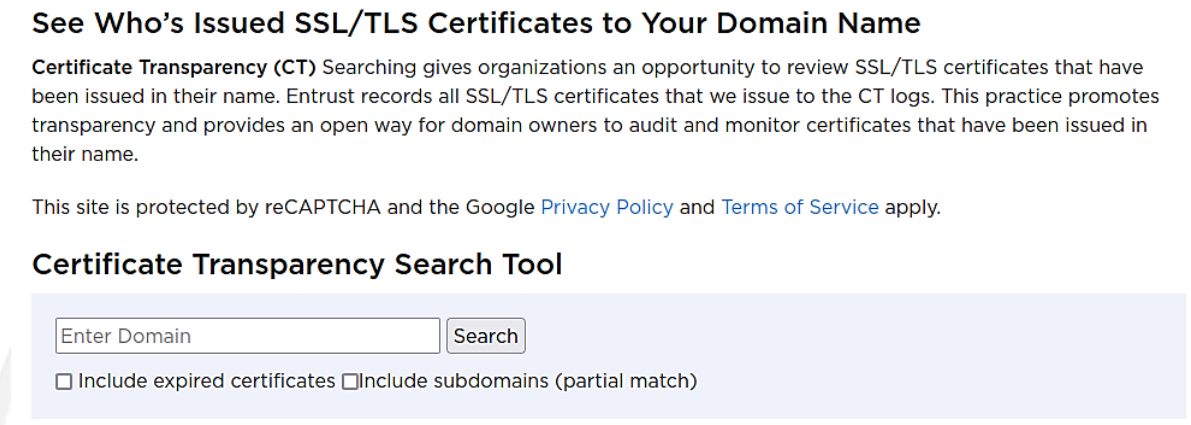


Figura 198. *CRT.sh*

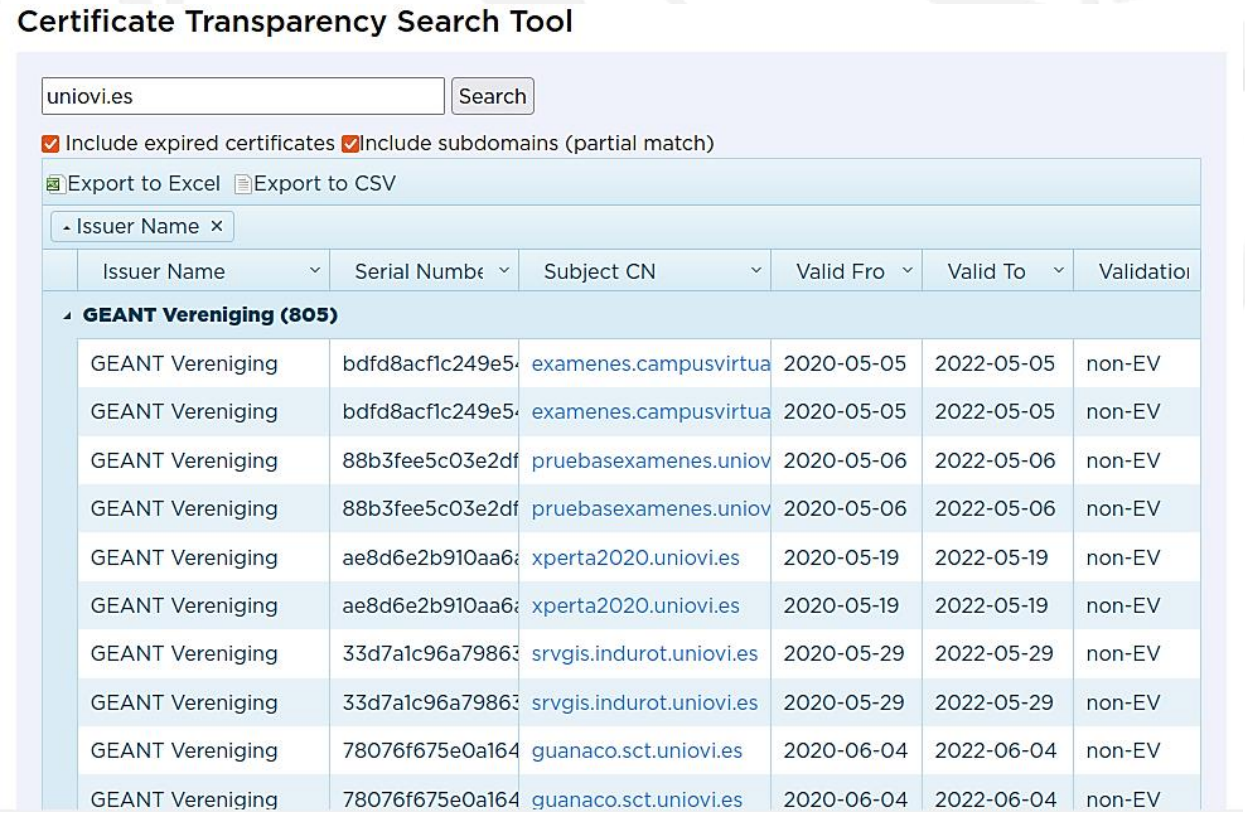


Figura 199. Búsqueda de certificados asociados a *uniovi.es* con *CRT.sh*

FilePursuit (https://filepursuit.com/)

Un motor de búsqueda para todo tipo de archivos en Internet, que obviamente pueden corresponder a filtraciones de un objetivo que estemos estudiando

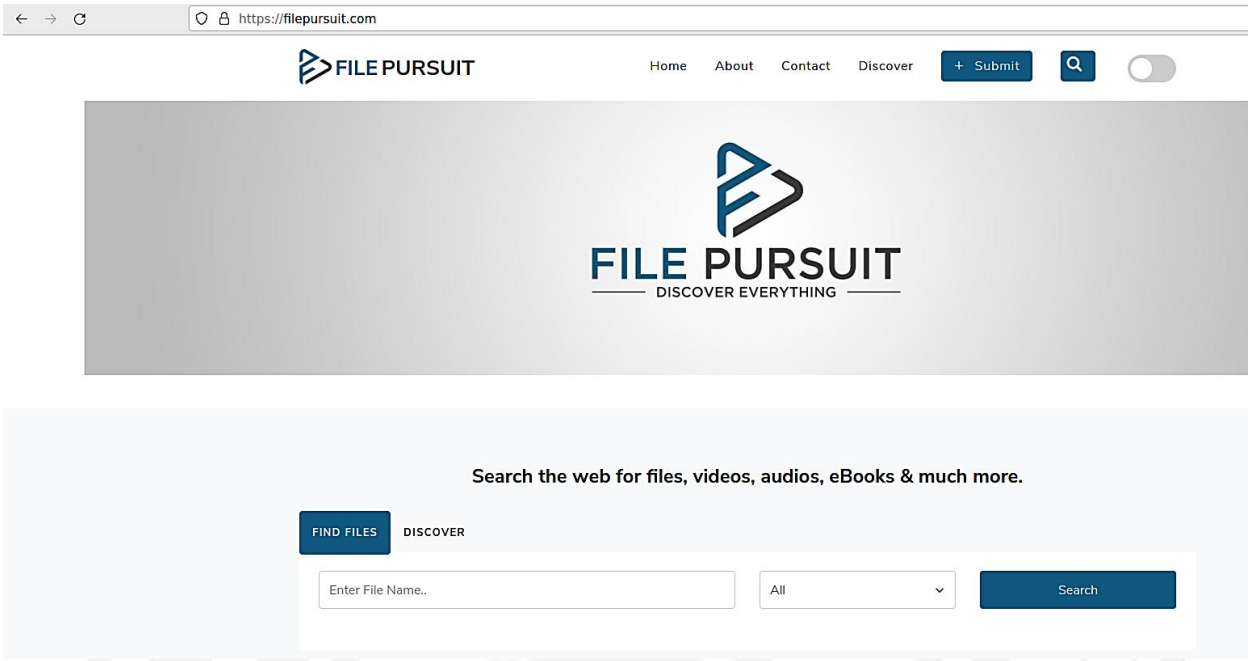


Figura 200. *FilePursuit*

Gray Hat Warfare (https://buckets.grayhatwarfare.com/)

Esta herramienta web permite encontrar fácilmente *buckets* abiertos en proveedores de nube (Ej. *Amazon S3*) con potencial información de los propietarios. También permite encontrar otros documentos rápidamente buscando varios entre tipos de archivos.

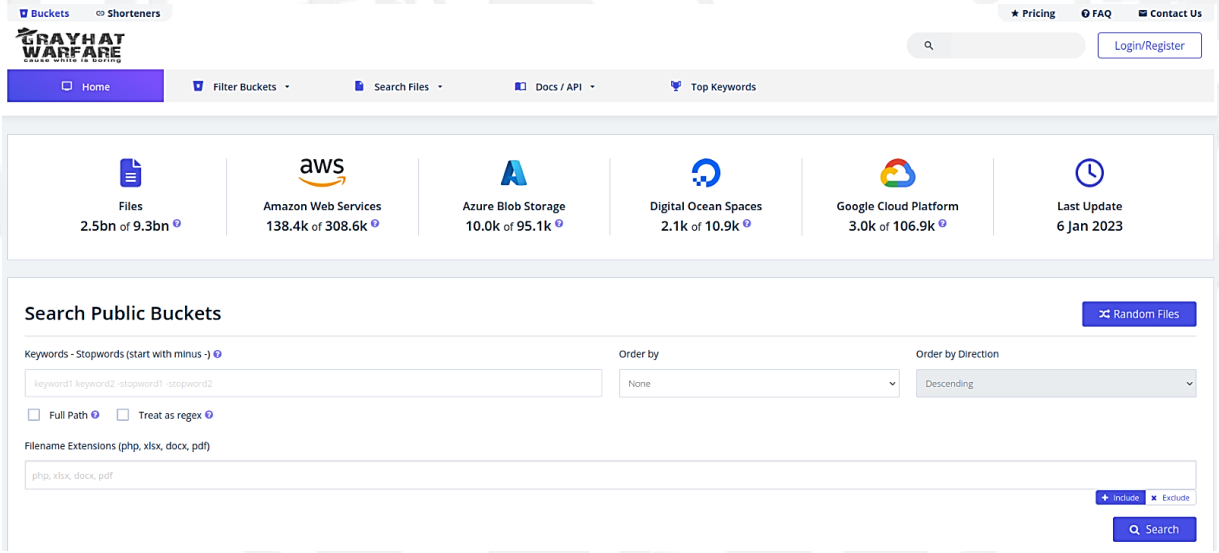


Figura 201. *Gray Hat Warfare*

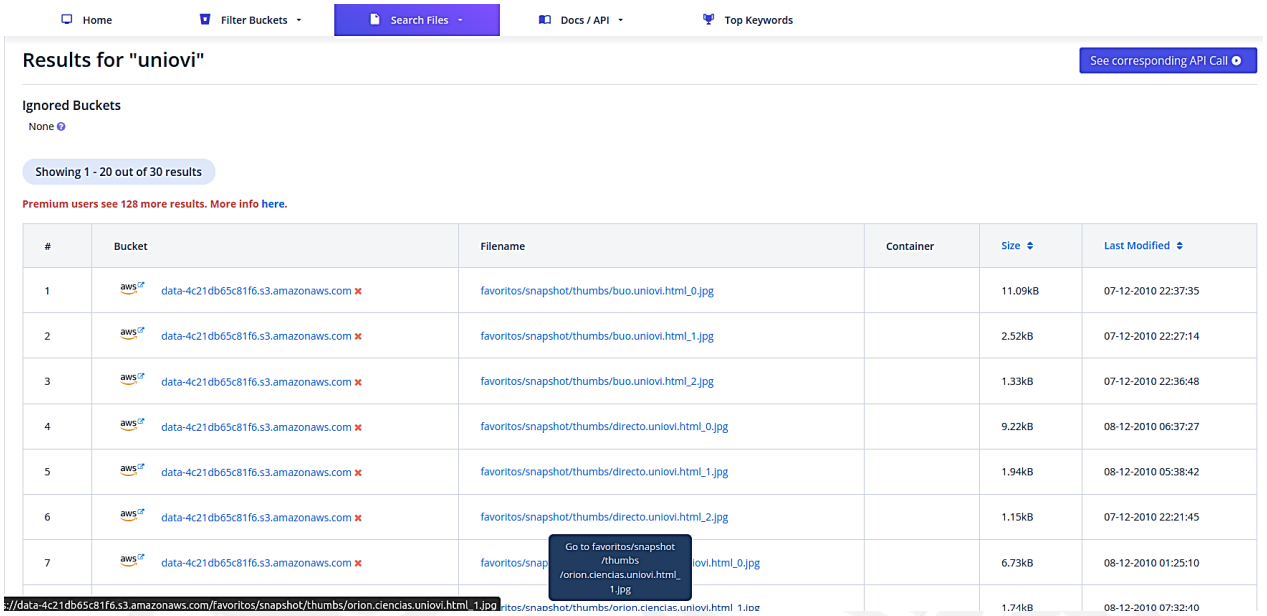


Figura 202. Búsqueda de *buckets* asociados a Uniovi con *Gray Hat Warfare*

IntelligenceX (https://intelx.io/tools)

Es una colección de herramientas para buscar en sitios de tercero y extraer inteligencia de redes sociales como *Facebook*, *Twitter*, *LinkedIn*, *Telegram* y *YouTube*. También busca *Tor*, *I2P*, fugas de datos, dominios y correos electrónicos. En otras palabras, funciona como un agregador de muchos servicios de búsqueda útiles (algunos incluidos en este documento y en la asignatura y otros no, ya que hay muchos) de forma que se pueda coordinar su uso.

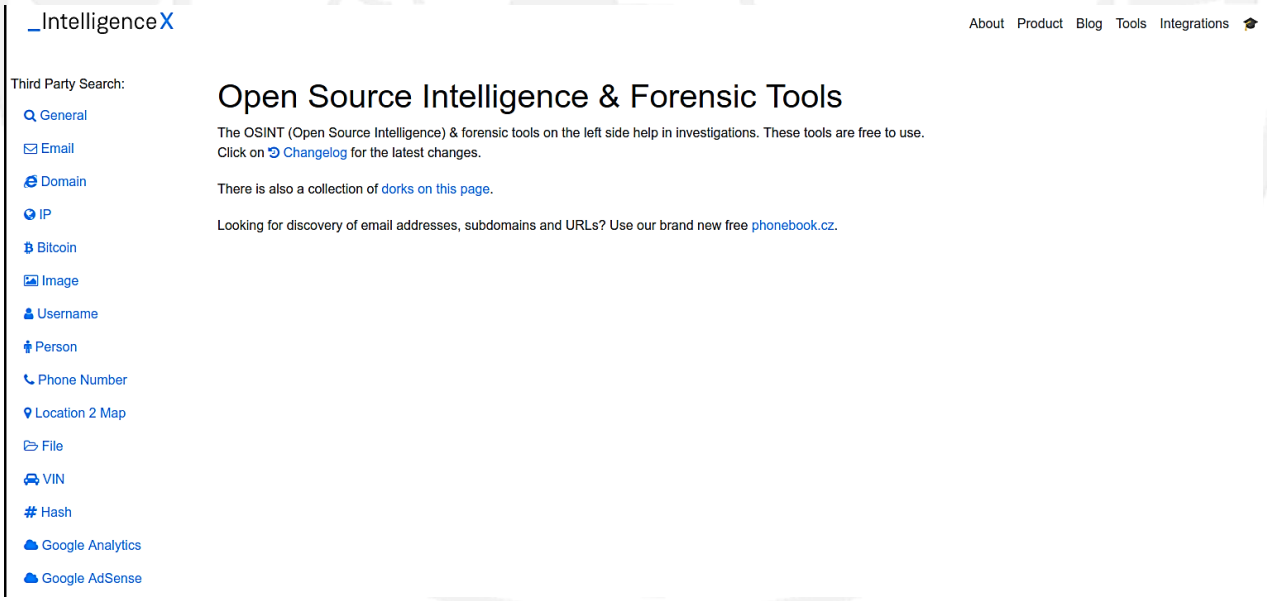


Figura 203. *IntelligenceX*

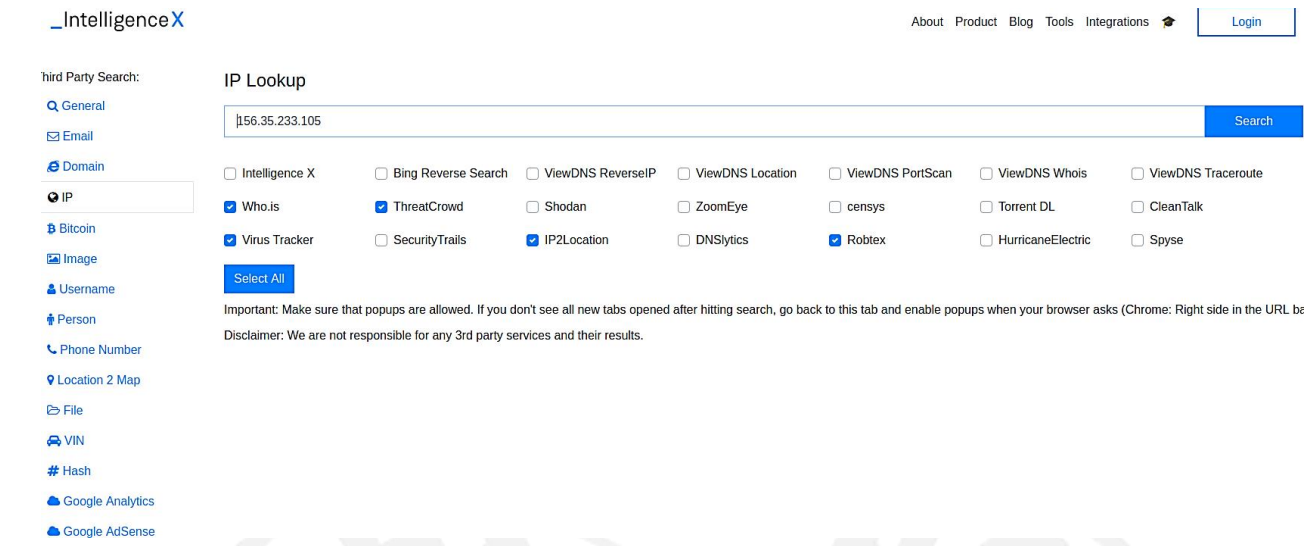


Figura 204. *IntelligenceX* y su catálogo de servicios sobre una determinada IP (Uniovi)

LeakIX (https://leakix.net/)

Plataforma que combina un motor de búsqueda que indexa información pública y una plataforma de informes abierta vinculada a los resultados obtenidos por dichas búsquedas.

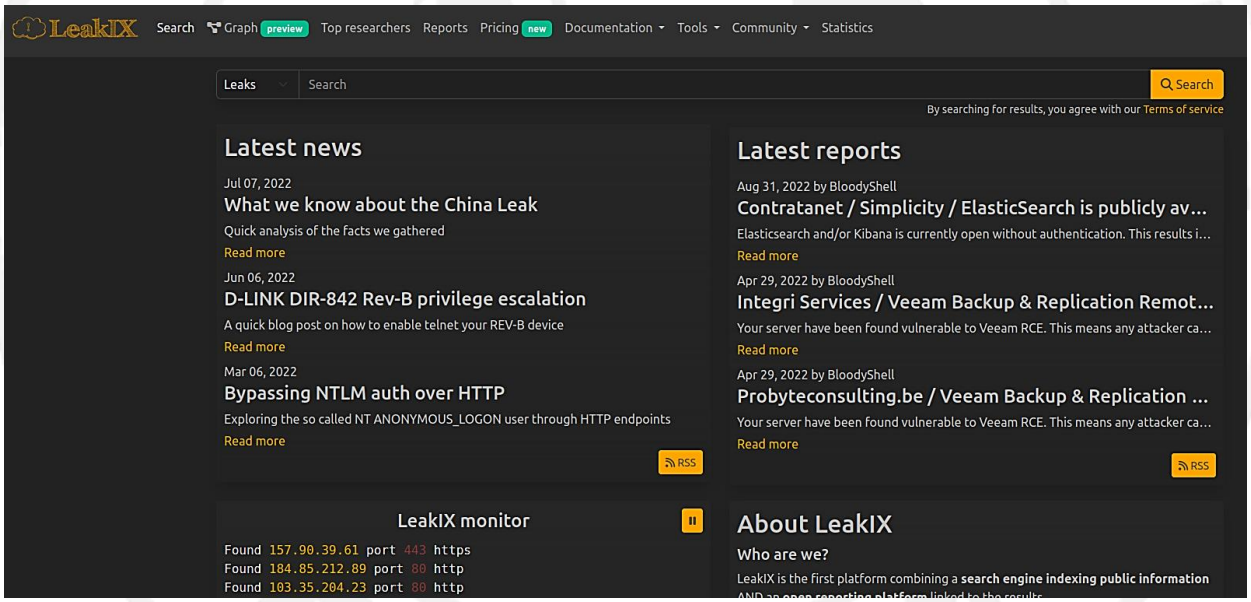


Figura 205. *LeakIX*

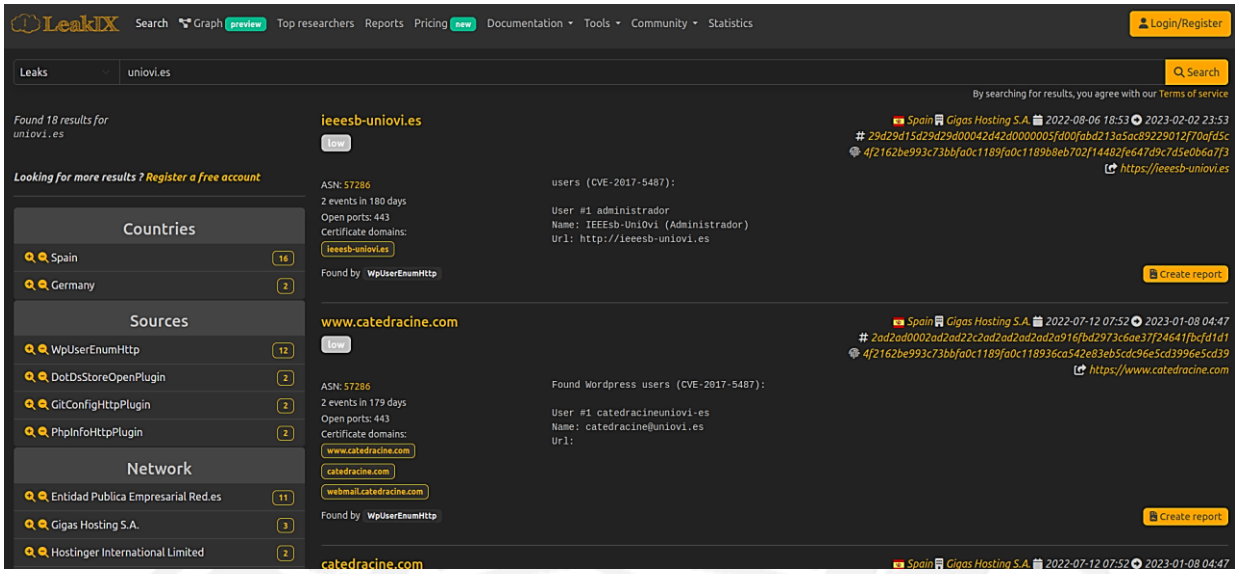


Figura 206. Filtraciones de datos asociadas a *uniovi.es* localizadas por *LeakIX*

Para cada filtración de datos, *LeakIX* además ofrece la razón y los datos que se han filtrado (en este caso usuarios válidos en una web hecha con *WordPress*).

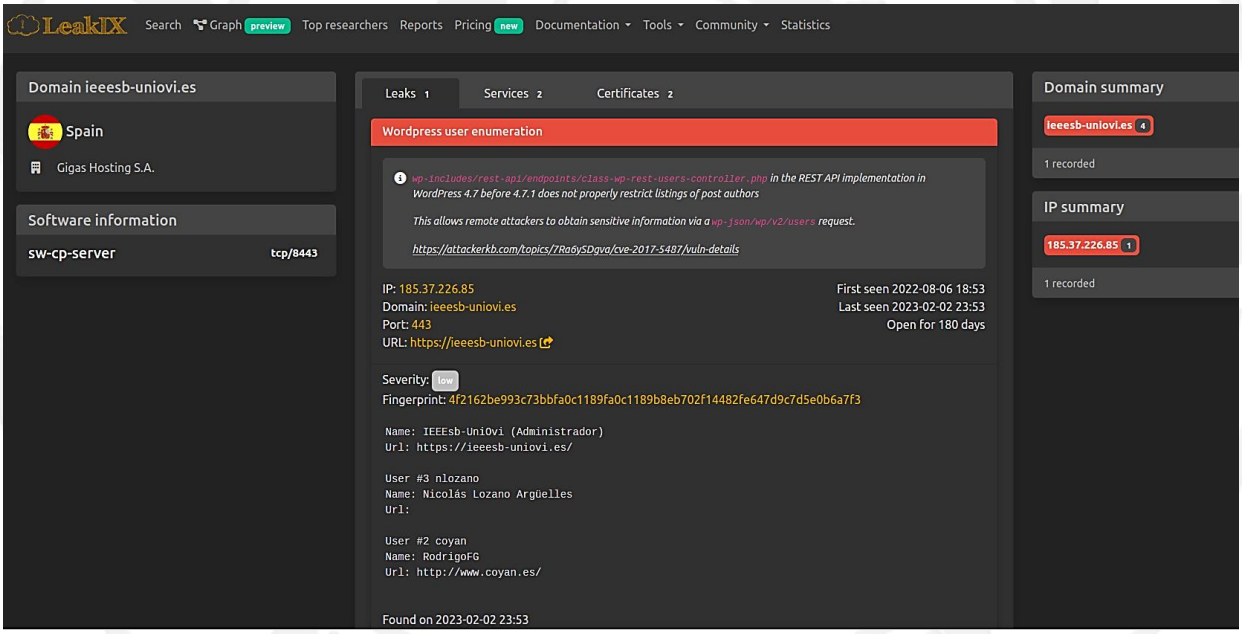


Figura 207. Descripción de un *leak* encontrado con *LeakIX*

TinEye Reverse Image search (https://tineye.com)

Un servicio alternativo a la famosa búsqueda inversa de imágenes de *Google*, con la misma funcionalidad de búsqueda por imágenes.

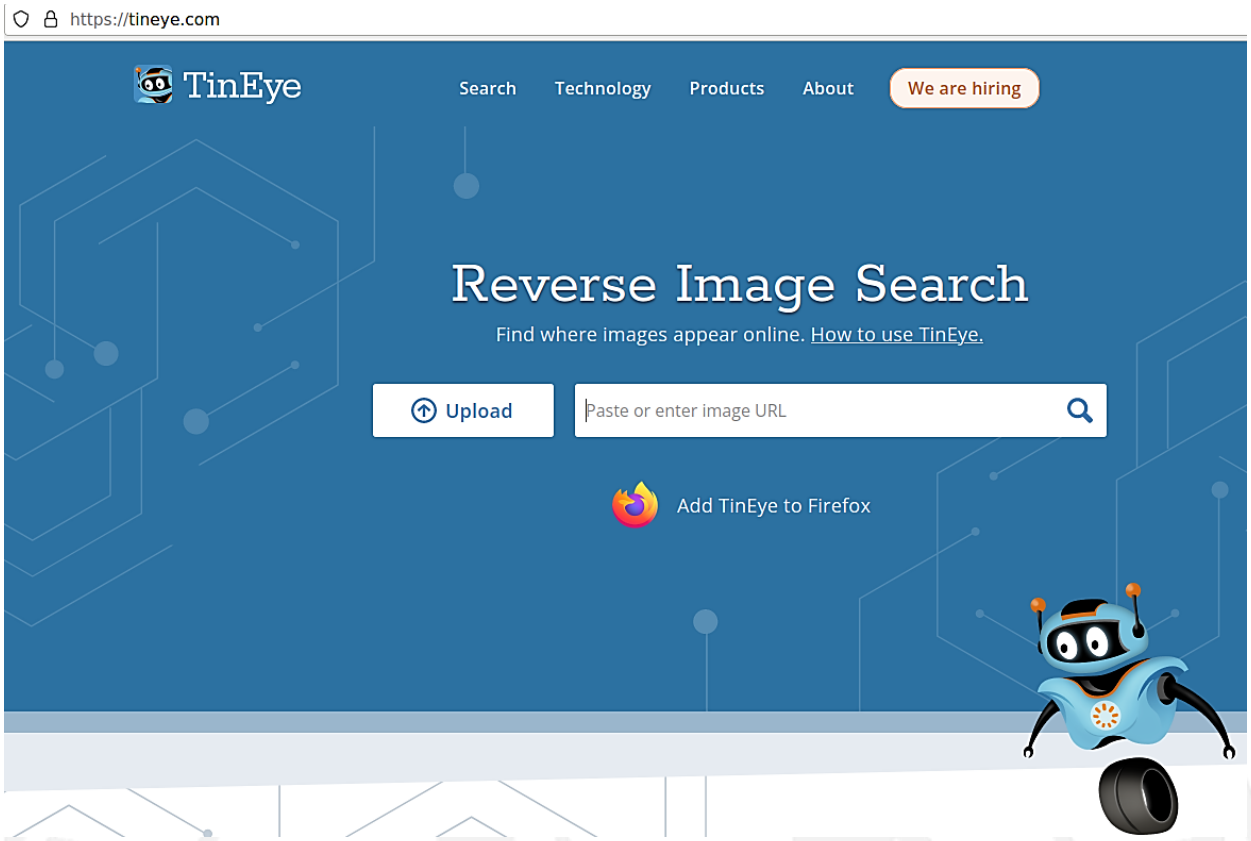


Figura 208. TinEye

Motores de búsqueda específicos para filtraciones relativas a cuentas de usuario

Hunter.io (https://hunter.io/)

Mencionado en el **Seminario 1**, busca direcciones de correo electrónico que pertenezcan a un sitio web, a partir de su nombre de dominio. También tiene una versión como plugin de *Firefox*. Una cuenta gratuita ofrece 25 búsquedas al mes sin coste, ideal para pequeñas búsquedas como las de esta asignatura.

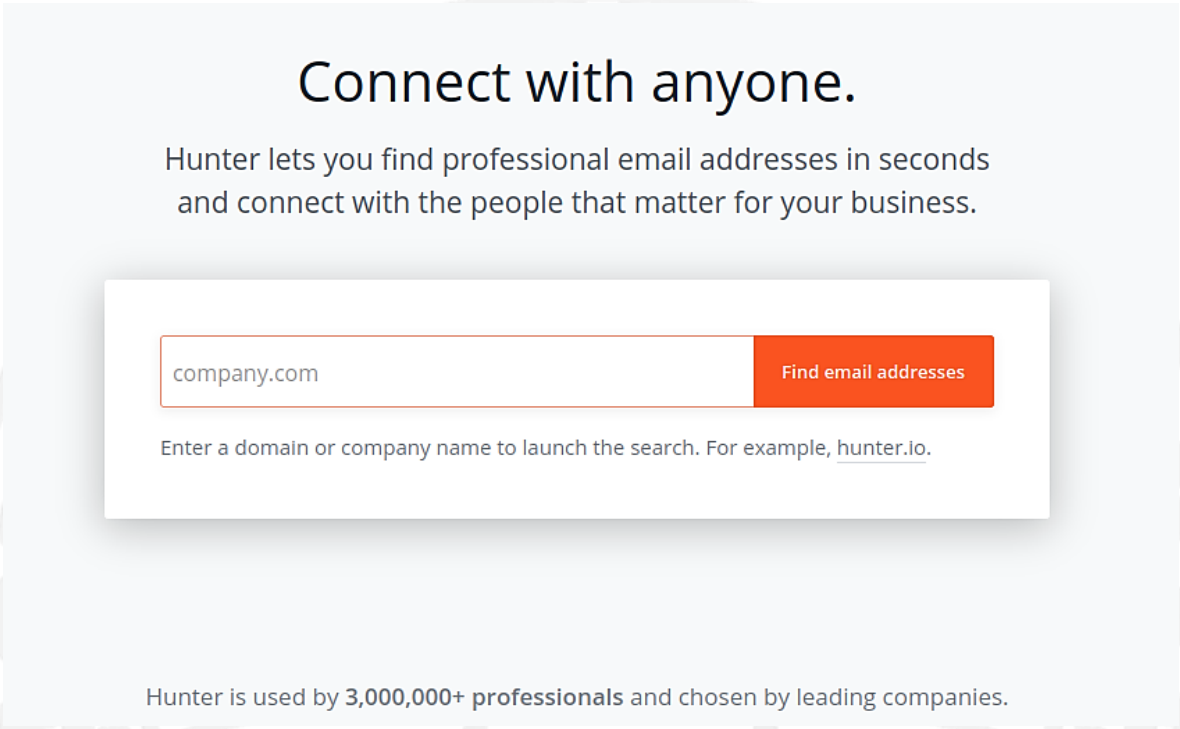


Figura 209. Hunter.io

LeakCheck (https://leakcheck.io/) y Have I been pwnd? (https://haveibeenpwned.com/)

Ambos son motores de búsqueda para localizar claves filtradas asociadas a emails en servicios variopintos. El uso gratuito de *LeakCheck* limita el detalle de los resultados y el nº de búsquedas mensuales, pero es plenamente operativo.

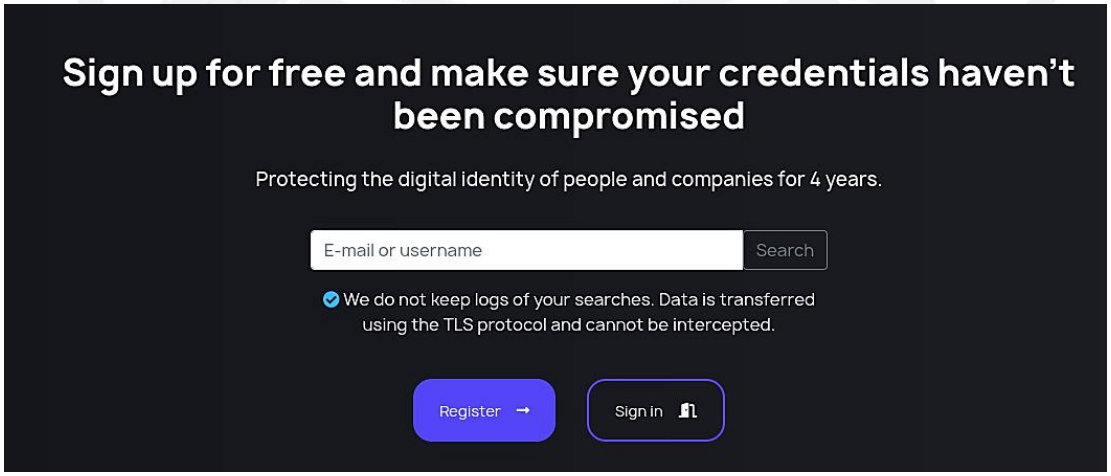


Figura 210. *LeakCheck*

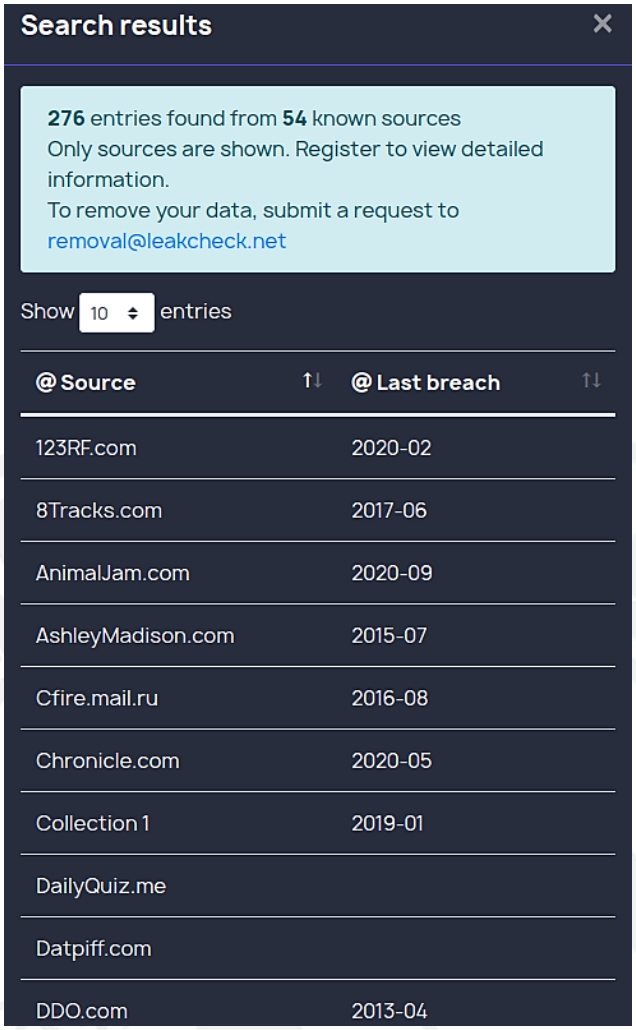


Figura 211. Ejemplo de fuentes de filtraciones localizadas por *LeakCheck*

`Haveibeenpwnd` no solo permite eso como vimos en el **Laboratorio 2**, sino que también contempla listar sitios que han sido comprometidos, saber si una password concreta ha sido filtrada, si un dominio ha sido comprometido (si verificas que es tuyo primero) y notificarte vía email si tu dirección aparece en alguna de los volcados que la web actualiza frecuentemente.

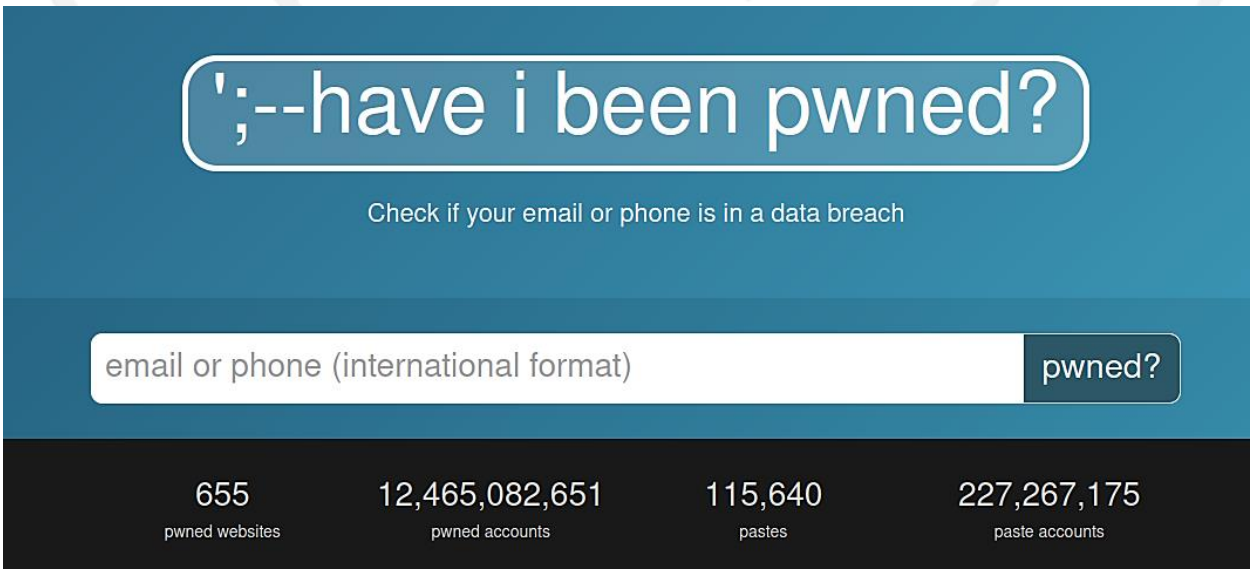


Figura 212. *Have I been pwnd?*

Motores de búsqueda para código fuente o ejecutable

Grep.App (https://grep.app/)

Busca en medio millón de repositorios *Git*, admitiendo expresiones regulares.

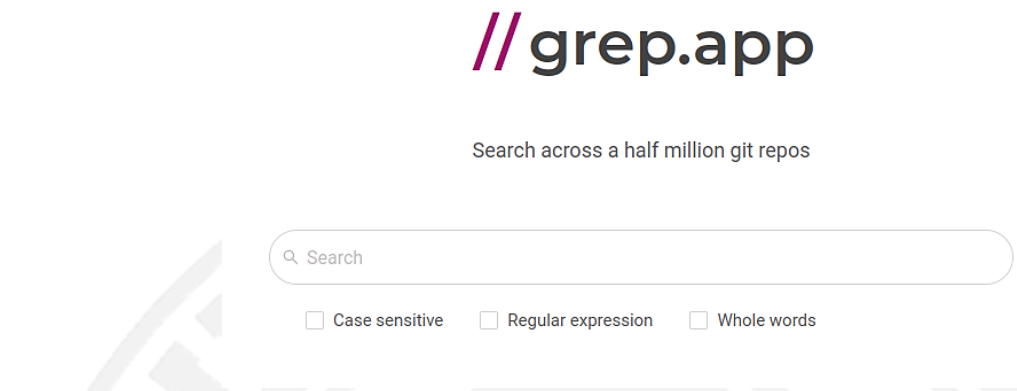


Figura 213. *Grep.app*

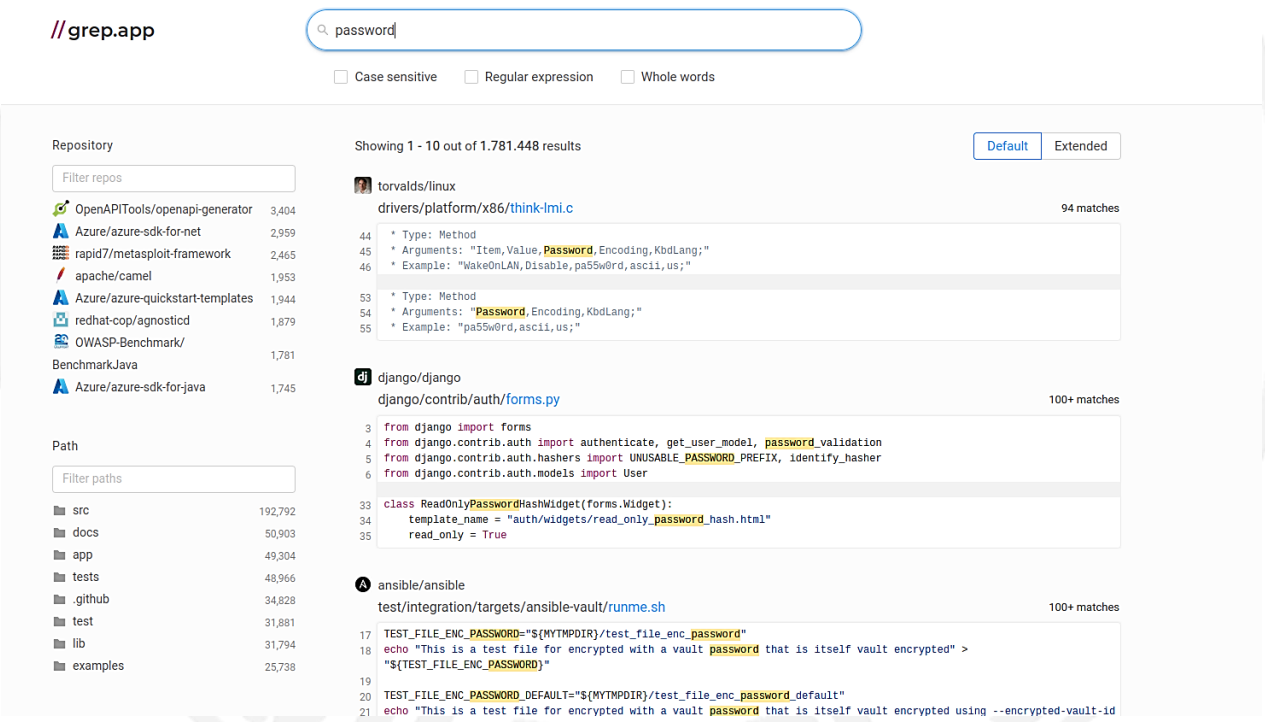


Figura 214. Sistema de búsqueda de código de *grep.app*

Joe SandBox (https://www.joesandbox.com/)

Ver esto aquí te puede parecer bastante extraño, puesto que no parece una herramienta que se pueda usar para averiguar nada de una víctima, pero no te dejes engañar por el propósito de una ***sandbox*** (ejecutar programas y ver si tienen comportamientos extraños que puedan indicar que son *malware*); aplicado sobre un ejecutable o documento que ofrezca un objetivo en sus webs, por ejemplo, parte del trabajo que deben hacer **puede revelar datos interesantes para nuestras investigaciones**:

- Si el archivo está **firmado**, puede revelar **datos de su autor** de los que "tirar" en una investigación.
- Si durante la ejecución **contacta con algún sitio externo**, puede ser una forma ideal de descubrir dominios pertenecientes al objetivo que de otra forma pueden pasar desapercibido. Sí, estoy

hablando de detectar lugares “escondidos” de un objetivo que encontramos porque algún programa que tiene el mismo se conecta a ellos.

- Si resulta que el objetivo está **ofreciendo un *malware* a sus clientes**, esto quiere decir dos cosas:
  - Tu objetivo **es un delincuente**
  - Tu objetivo **ha sido troyanizado** y tiene una infección en marcha, que ha llegado hasta los archivos que ofrece a los demás *(supply-chain attack*). Aparte de lo que esto implica a nivel de tu auditoría (probablemente tengas que lanzar unas cuantas alarmas...), el tipo de infección y el comportamiento del ejecutable **pueden revelar una puerta trasera en el objetivo**, es decir, un sitio por el que el atacante se “cuela” para manipular los sistemas del objetivo y...que ahora también puedes usar tú si sirve para algo en tu investigación. No obstante, estoy es “*engagement*” y no forma parte de la asignatura.

Joe Sandbox es probablemente uno de estos entornos de análisis más conocidos y, si bien tiene un uso más bien defensivo, acabamos de ver como se pueden usar sus capacidades de análisis profundo del comportamiento de un ejecutable para algo de carácter ofensivo 😉

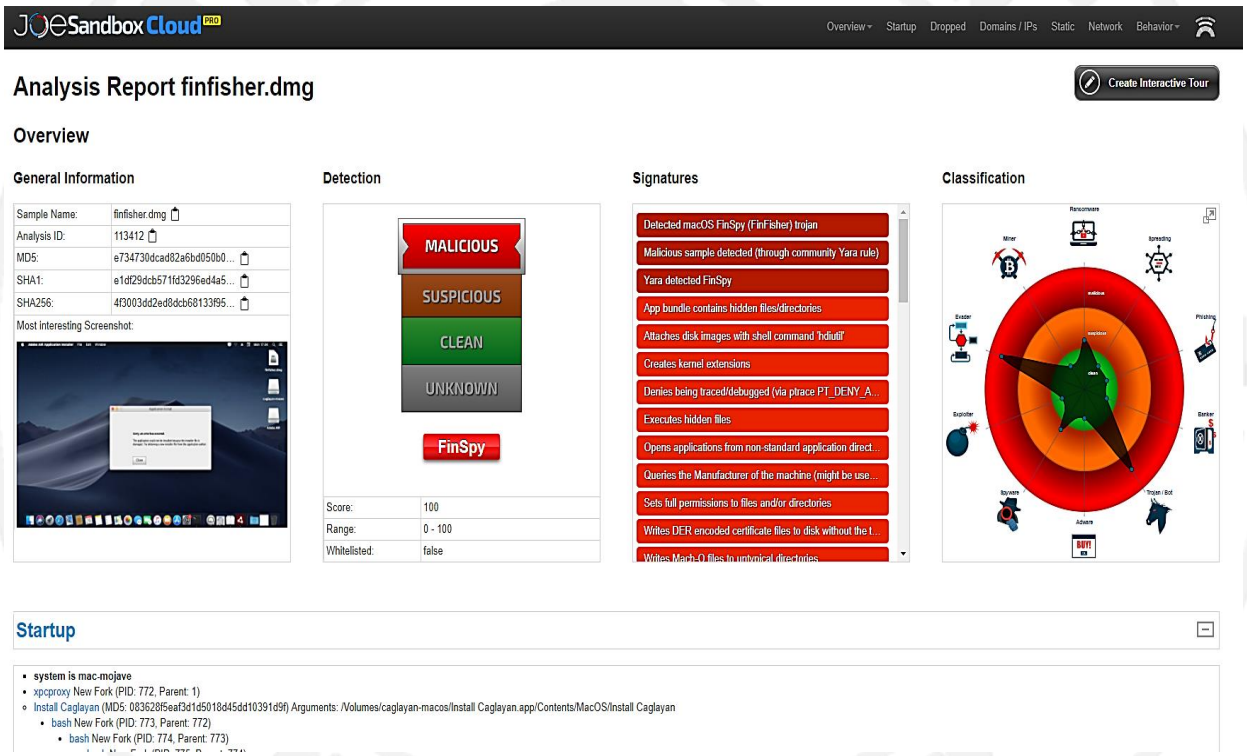


Figura 215. Joe "destripando” un *malware*. Fuente: https://www.joesecurity.org/blog/1896065926254527264

SearchCode (https://searchcode.com/)

Otro buscador de código dentro de repositorios, aunque en este caso con una base de datos mucho mayor que otras soluciones similares en el mercado.

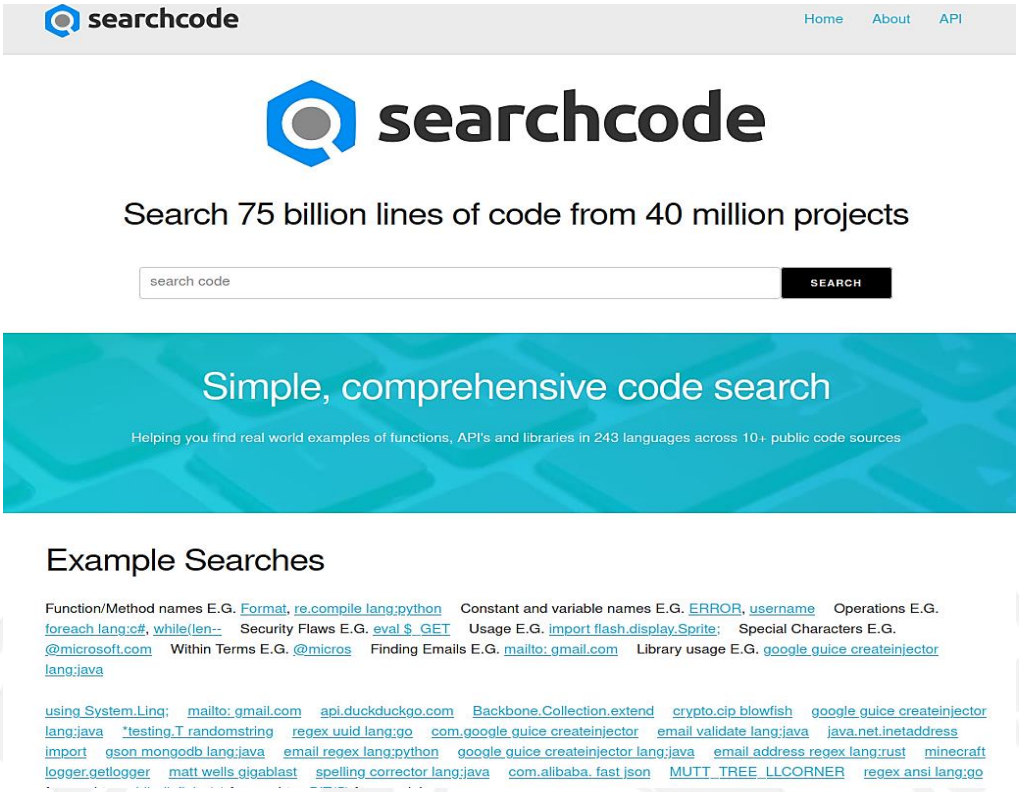


Figura 216. *SearchCode*

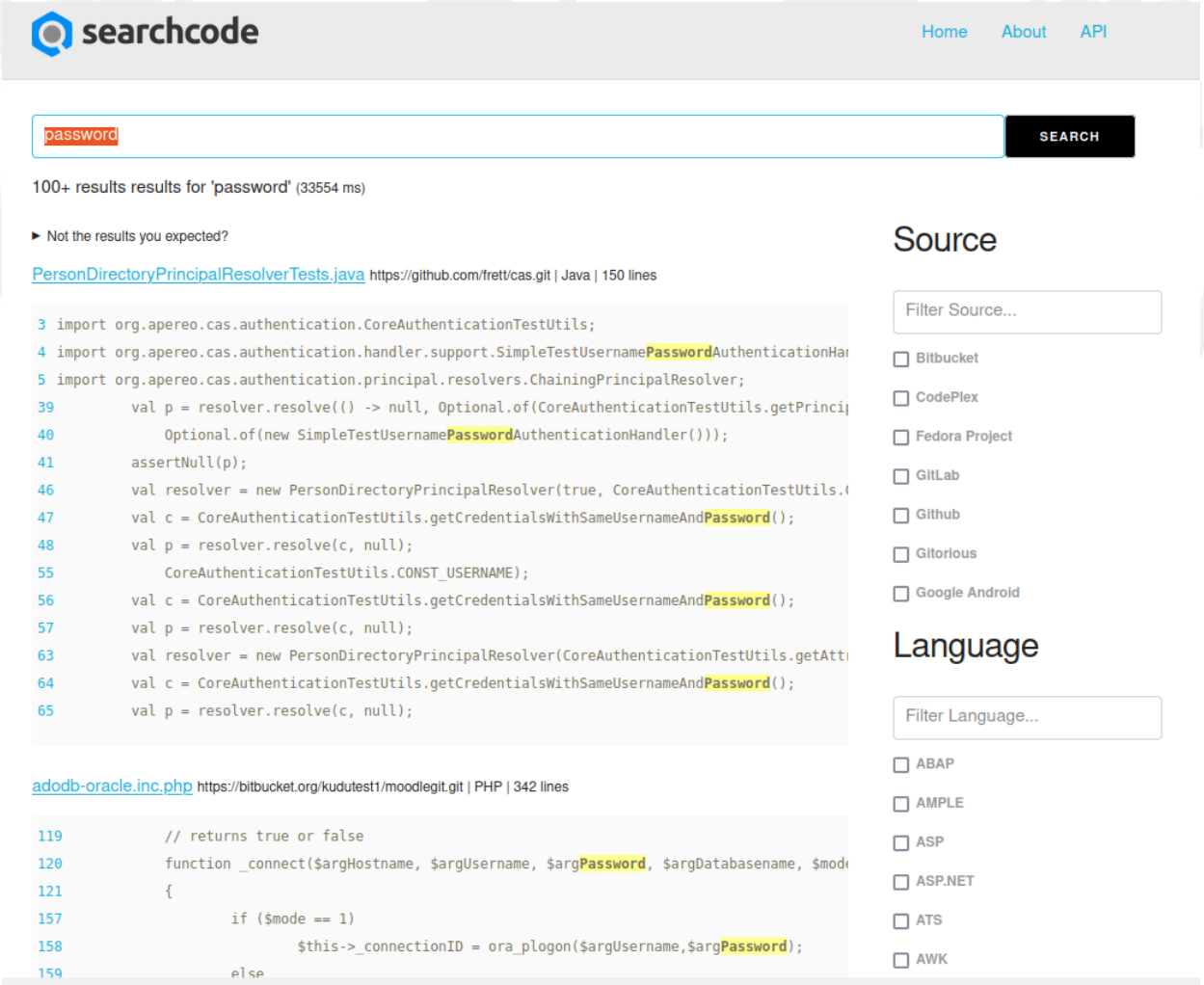
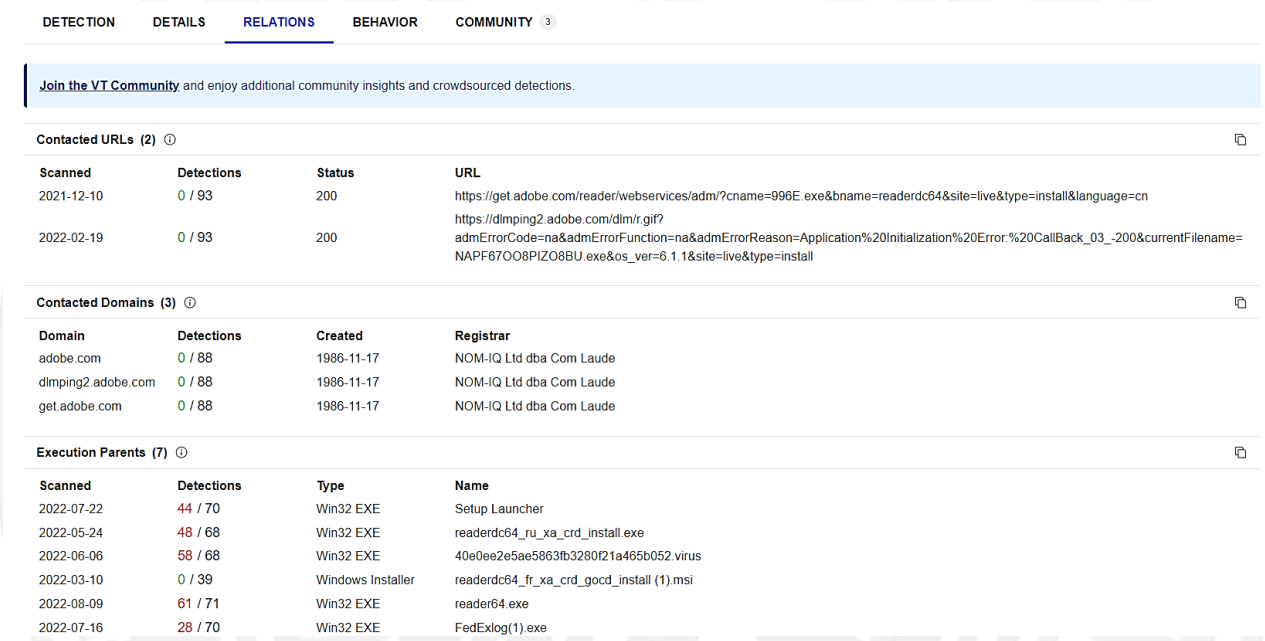


Figura 217. Ejemplo de búsqueda de *SearchCode*

Virustotal (https://www.virustotal.com/gui/home/upload) y VT Intelligence (https://www.virustotal.com/gui/intelligence-overview)

Virustotal se ha renovado recientemente y ahora **puede usarse para lo mismo que vimos en *Joe Sandbox***: revelar rutas y datos de un ejecutable sin que tengamos que ejecutarlo nosotros mismos: es la *sandbox* de *Virustotal* la que lo hará por nosotros, en sus propios servidores. Como ves, su uso va mucho más allá que el uso básico que vimos en el **Laboratorio 1**. Las dos imágenes siguientes muestran ejemplos claro de ello. Es además **más fácil de usar** que *Joe*: ¡no requiere registro alguno! 😊.

La plataforma ***VT Intelligence*** en cambio es de pago y promete ser algo así como un motor de búsqueda de inteligencia de amenazas a nivel mundial de la escala de Google, estableciendo relaciones entre los datos como una red social de personas. En otras palabras, es una plataforma que permite analizar amenazas mundiales y hacer foco en detalle solo en aquellas que nos interesen. Por desgracia, al ser de pago está fuera del objetivo de esta asignatura.



DETECTION DETAILS RELATIONS BEHAVIOR COMMUNITY 3

Join the VT Community and enjoy additional community insights and crowdsourced detections.

Contacted URLs (2)

Scanned	Detections	Status	URL
2021-12-10	0 / 93	200	https://get.adobe.com/reader/webservices/adm/?cname=996E.exe&bname=readerdc64&site=live&type=install&language=cn
2022-02-19	0 / 93	200	https://dlmping2.adobe.com/dlm/r.gif?admErrorCode=na&admErrorFunction=na&admErrorReason=Application%20Initialization%20Error:%20CallBack_03_-200&currentFilename=NAPF67OO8PIZO8BU.exe&os_ver=6.1.1&site=live&type=install

Contacted Domains (3)

Domain	Detections	Created	Registrar
adobe.com	0 / 88	1986-11-17	NOM-IQ Ltd dba Com Laude
dlmping2.adobe.com	0 / 88	1986-11-17	NOM-IQ Ltd dba Com Laude
get.adobe.com	0 / 88	1986-11-17	NOM-IQ Ltd dba Com Laude

Execution Parents (7)

Scanned	Detections	Type	Name
2022-07-22	44 / 70	Win32 EXE	Setup Launcher
2022-05-24	48 / 68	Win32 EXE	readerdc64_ru_xa_crd_install.exe
2022-06-06	58 / 68	Win32 EXE	40e0ee2e5ae5863fb3280f21a465b052.virus
2022-03-10	0 / 39	Windows Installer	readerdc64_fr_xa_crd_gocd_install (1).msi
2022-08-09	61 / 71	Win32 EXE	reader64.exe
2022-07-16	28 / 70	Win32 EXE	FedExlog(1).exe

Figura 218. *Virustotal* revelando sitios de terceros al conectarse un ejecutable con ellos

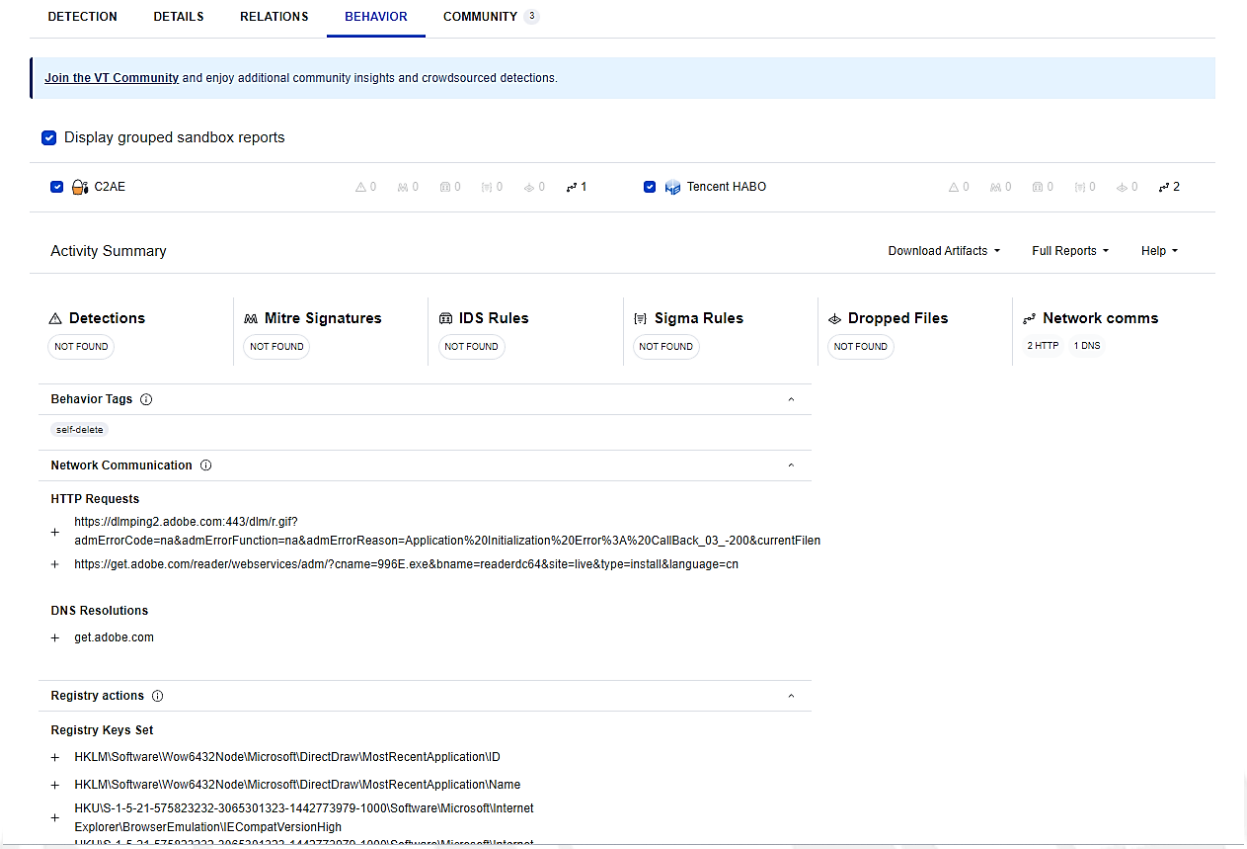


Figura 219. La inspección del comportamiento del ejecutable de *Virustotal* también puede revelar estas rutas externas, entre otras cosas

Cuando el fin justifica los medios

Phishing y *Spear phishing* para la obtención de información

Este capítulo te puede hacer pensar que me he vuelto loco. No, tranquilo 😊. Solo quiero que entiendas que el OSINT tiene una parte "malvada", es decir, que se puede usar para cometer delitos. Toda la información que se recopile mediante las técnicas vistas **se puede usar para distintos ataque que usan ingeniería social**. Y, como puedes ver en esta imagen, no son precisamente pocos:

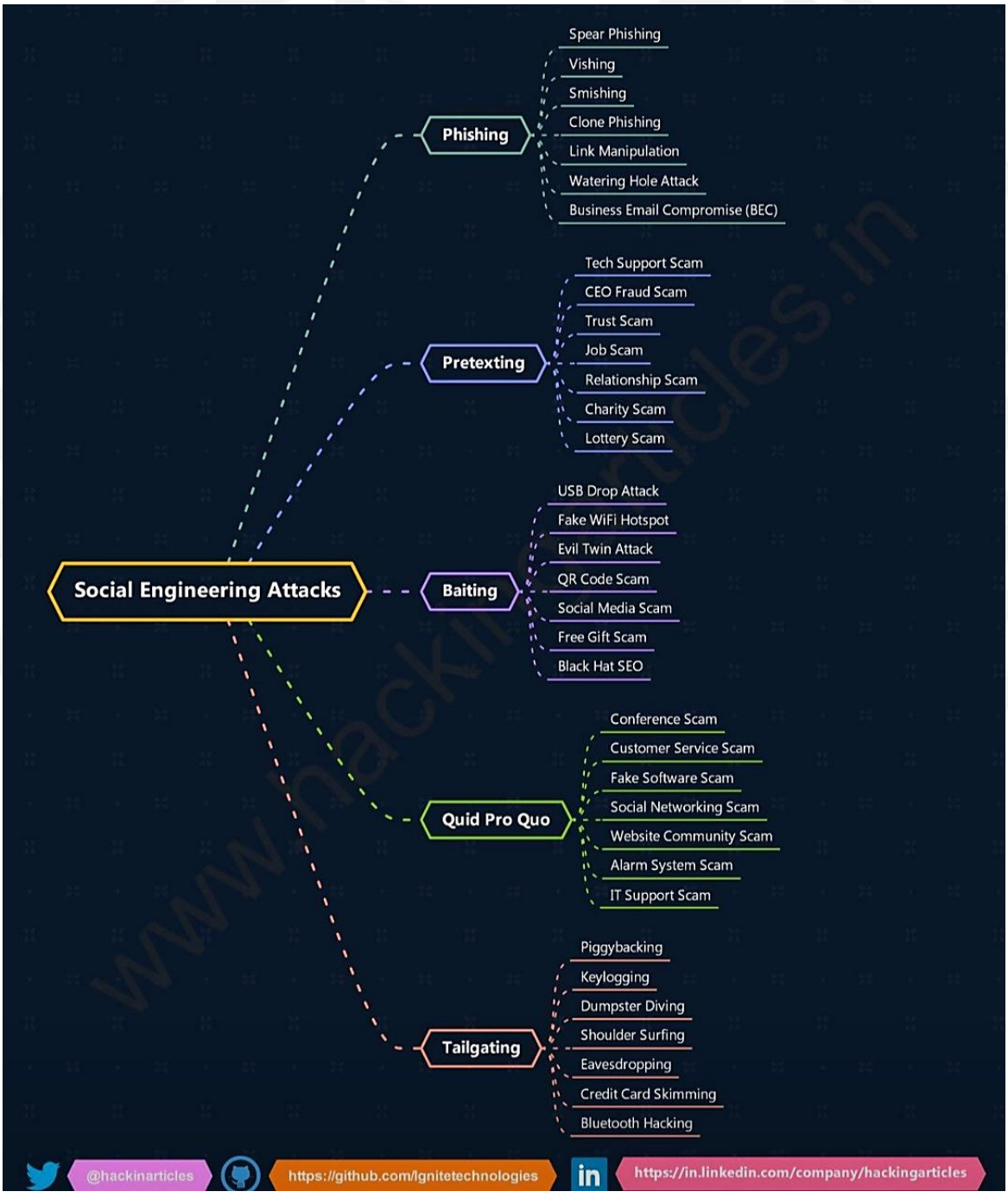


Figura 220. Mucha gente no entiende la gravedad e importancia de los ataques de ingeniería social

Pero en este apartado vamos a centrarnos en uno de los más importantes, **el *phishing* y las herramientas que lo facilitan**, que tienen un uso legal. Y este uso legal existe porque aunque te parezca increíble **muchas empresas hacen simulaciones de ataques de *phishing* y *spear phishing*** como entrenamiento para sus empleados. Dicho de otra forma, usan herramientas de *phishing* profesionales **para lanzar campañas simuladas** pero exactamente igual que lo harían los delincuentes, extraen la información que puedan con ellas y, con eso, hacen un informe de la vulnerabilidad de la empresa ante estos ataques[2]. Considerando lo que dijimos al principio de este documento acerca de la cantidad de ataques que vienen por un descuido humano, hacer este tipo de formaciones ya tiene más sentido, *¿verdad?* 😊. Y, ya que las hacemos, las hacemos bien 😉.

En este apartado vamos a hablar de tres herramientas para este fin, que son ***SET***, ***Gophish*** y ***Evilginx2***. Otra herramienta similar es ***SniperPhish*** (https://github.com/GemGeorge/SniperPhish). Por otro lado, aunque vamos a hablar de comunicación por email, hacer una estafa vía SMS también es viable con servicios como este: https://www.hackers-arise.com/post/social-engineering-attacks-creating-a-fake-sms-message

SET

Introducción

Social-Engineer Toolkit (SET) es una herramienta de pentesting *open source* destinada a hacer ataques de *phishing* a través de ingeniería social. Por tanto, los ataques que hagas pueden "alimentarse" de la información que hayas averiguado mediante OSINT sobre la víctima en todo lo que hemos descrito hasta ahora.

SET tiene un conjunto de vectores de ataque que se pueden usar para las fases de ***TA0001 Initial Access*** y ***TA0002 Execution*** del *MITRE ATT&CK*, y no es una herramienta que se deba tomar a broma: los ataques que genera **pueden ser perfectamente reales**. El éxito o no de los mismos dependerá de cómo de bueno o malo sea el mensaje que hagas llegar a la víctima, siguiendo las técnicas y pautas sicológicas que hemos visto. SET viene instalado por defecto en *Kali Linux*, por lo que simplemente tenemos que abrirlo.

[2] A lo mejor piensas: *¿Qué pasa si caigo en una? ¿me despiden?* A ver, todo es posible, pero yo hasta ahora solo vi como consecuencia hacer un curso de formación del tema. Tampoco es tan grave 😊

Figura 221. Abriendo SET en *Kali Linux*

Como decía anteriormente, la posibilidad de hacer verdadero daño con esta herramienta existe y no es desdeñable, por lo que usándola admites conocer los riegos de su uso. La herramienta se creó para comprobar la vulnerabilidad de una víctima a ataques de *phishing* e ingeniería social, pero como supondrás, hacer esto se puede convertir fácilmente en un delito si no sabemos cuándo parar (lo que nos dirá el contrato que hayamos firmado).

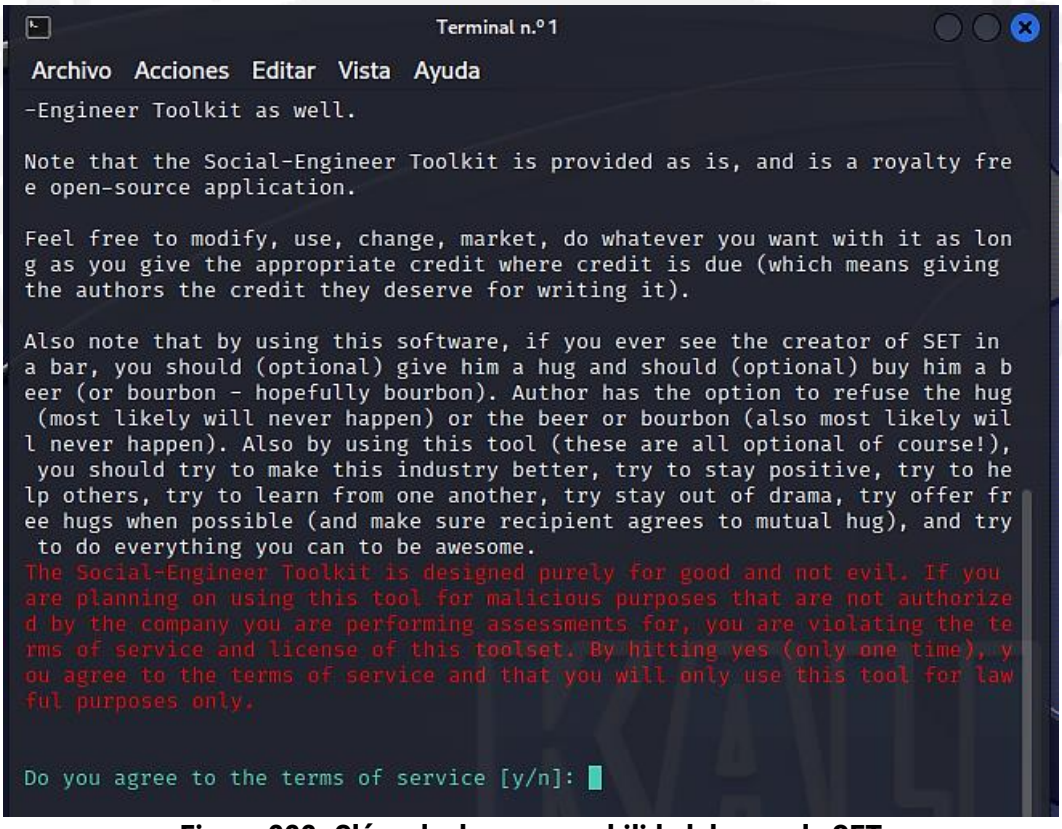


Figura 222. Cláusula de responsabilidad de uso de SET

SET es una herramienta en línea de comandos **basada en menús** que se seleccionan por número, al estilo de los programas que uno hace en primer año de carrera. Cada opción abre las opciones a hacer un tipo de ataque distinto, pero en este apartado nos interesa primero hacer uno **de ingeniería social**.

Figura 223. Menú principal de SET

Mails de phishing con SET

Dentro de la primera opción del menú tenemos la posibilidad de crear **mensajes de *spear phising***. Esto nos permite generar mensajes de este estilo y enviarlos a una víctima con *payloads* adjuntos. Por ejemplo, puedes enviar un PDF malicioso de manera que cuando la víctima lo abra comprometa el sistema. También permite falsificar la dirección de envío fácilmente si además tienes instalado el paquete `sendmail` (`apt-get install sendmail`) y cambiar en `config/set_config` la entrada `SENDMAIL=OFF` a `SENDMAIL=ON`. La imagen siguiente muestra el submenú que aparece cuando elegimos la opción de ataques de ingeniería social.

Figura 224. Ataques de ingeniería social disponibles en SET

Aunque hasta aquí la cosa parece muy fácil, hay que ser realista: Si envías tú mismo el mail a las víctimas, y estas tienen el típico proveedor de mail de gran empresa en nube (*Outlook*, *Gmail*...) las posibilidades de que el mail llegue o que llegue pero no lo haga a la carpeta de *spam* **son muy bajas**. Hoy en día lo servicios de correo son un oligopolio y las grandes empresas tienen en marcha protecciones en el mail (SPF, DKIM, DMARC) **que detectarán que el emisor no es fiable** o que directamente viene con *malware*. Lo mismo ocurre si copiamos y pegamos el mail e intentamos enviarlo a través de una cuenta de uno de estos proveedores, nuestra o "adquirida" (*TA0042 Resource Development*). Si tienes un *payload* malicioso lo más normal es que el proveedor lo detecte.

No obstante, esto no significa que vaya a fracasar siempre. Todavía hay victimas que tienen su propio sistema de correo (normalmente muy mal protegido), y el *payload* puede no detectarse como *malware* (especialmente si se genera con `msfvenom` **y se codifica** (**Tema 7**). También hay la opción de enviar un enlace en el correo, que apunte a una copia de una web real controlada por nosotros.

Hay dos formas de hacer un ataque de *spear phishing*:

1. Hacer un ataque masivo de mails.
2. Crear un *payload* y una plantilla para ingeniería social.

Para el primero **SET hace todo el trabajo por nosotros** (opción 1). En el caso del segundo hay que crear un *payload* y usarlo en un ataque que configuramos. Corresponden a las opciones 1 y 2 de este menú. En la siguiente secuencia de imágenes se ven las sucesivas opciones de menú a elegir para hacer un ataque de esta clase. Vemos que la secuencia de pasos a seguir se alinea con lo explicado en los **laboratorios 11 y 12** de la asignatura.

```
The Spearphishing module allows you to specially craft email messages and send
them to a large (or small) number of people with attached fileformat malicious
payloads. If you want to spoof your email address, be sure "Sendmail" is in-
stalled (apt-get install sendmail) and change the config/set_config SENDMAIL=OFF
flag to SENDMAIL=ON.

There are two options, one is getting your feet wet and letting SET do
everything for you (option 1), the second is to create your own FileFormat
payload and use it in your own attack. Either way, good luck and enjoy!

  1) Perform a Mass Email Attack
  2) Create a FileFormat Payload
  3) Create a Social-Engineering Template

 99) Return to Main Menu

set:phishing>
```

Figura 225. Tipos de ataque de *spear phishing*

```
set:phishing>1
/usr/share/metasploit-framework/

 Select the file format exploit you want.
 The default is the PDF embedded EXE.

           ********** PAYLOADS **********

   1) SET Custom Written DLL Hijacking Attack Vector (RAR, ZIP)
   2) SET Custom Written Document UNC LM SMB Capture Attack
   3) MS15-100 Microsoft Windows Media Center MCL Vulnerability
   4) MS14-017 Microsoft Word RTF Object Confusion (2014-04-01)
   5) Microsoft Windows CreateSizedDIBSECTION Stack Buffer Overflow
   6) Microsoft Word RTF pFragments Stack Buffer Overflow (MS10-087)
   7) Adobe Flash Player "Button" Remote Code Execution
   8) Adobe CoolType SING Table "uniqueName" Overflow
   9) Adobe Flash Player "newfunction" Invalid Pointer Use
  10) Adobe Collab.collectEmailInfo Buffer Overflow
  11) Adobe Collab.getIcon Buffer Overflow
  12) Adobe JBIG2Decode Memory Corruption Exploit
  13) Adobe PDF Embedded EXE Social Engineering
  14) Adobe util.printf() Buffer Overflow
  15) Custom EXE to VBA (sent via RAR) (RAR required)
  16) Adobe U3D CLODProgressiveMeshDeclaration Array Overrun
  17) Adobe PDF Embedded EXE Social Engineering (NOJS)
  18) Foxit PDF Reader v4.1.1 Title Stack Buffer Overflow
  19) Apple QuickTime PICT PnSize Buffer Overflow
  20) Nuance PDF Reader v6.0 Launch Stack Buffer Overflow
  21) Adobe Reader u3D Memory Corruption Vulnerability
  22) MSCOMCTL ActiveX Buffer Overflow (ms12-027)

set:payloads>
```

Figura 226. Posibles *payloads* del mail de *phishing* a crear

```
[-] Default payload creation selected. SET will generate a normal PDF with embedded
 EXE.

    1. Use your own PDF for attack
    2. Use built-in BLANK PDF for attack

set:payloads>
```

Figura 227. Creando un payload tipo PDF con un ejecutable dentro, bien a partir de un PDF existente o de uno en blanco

```
   1) Windows Reverse TCP Shell              Spawn a command shell on victim and se
nd back to attacker
   2) Windows Meterpreter Reverse_TCP        Spawn a meterpreter shell on victim an
d send back to attacker
   3) Windows Reverse VNC DLL                Spawn a VNC server on victim and send
back to attacker
   4) Windows Reverse TCP Shell (x64)        Windows X64 Command Shell, Reverse TCP
 Inline
   5) Windows Meterpreter Reverse_TCP (X64)  Connect back to the attacker (Windows
x64), Meterpreter
   6) Windows Shell Bind_TCP (X64)           Execute payload and create an acceptin
g port on remote system
   7) Windows Meterpreter Reverse HTTPS      Tunnel communication over HTTP using S
SL and use Meterpreter

set:payloads>
```

Figura 228. Efecto del *payload* a generar. Nótese como son todo *reverse shells* (Tema 7)

```
set:payloads>2
set> IP address or URL (www.ex.com) for the payload listener (LHOST) [10.0.2.15]:
set:payloads> Port to connect back on [443]:
[-] Defaulting to port 443...
[!] Metasploit has not been previously run on the system. This means that the msf d
irectories haven't been created yet. Running Metasploit for you.
[*] Waiting 10 seconds for the directories to be created...
```

Figura 229. Configurando la IP del servidor que escuchará los intentos de conexión de las víctimas que hayan abierto el PDF

```
[-] As an added bonus, use the file-format creator in SET to create your attachment
.

   Right now the attachment will be imported with filename of 'template.whatever'

   Do you want to rename the file?

   example Enter the new filename: moo.pdf

    1. Keep the filename, I don't care.
    2. Rename the file, I want to be cool.

set:phishing>
```

Figura 230. Finalizar el PDF con "bicho" creado

En este punto SET nos permite (al igual que con la opción de menú llamada "*Mass Mailer Attack*") **enviar emails a las víctimas** automáticamente con un mensaje personalizado que podemos elaborar a partir de la información de OSINT que hayamos recabado. Tiene dos opciones:

- *E-Mail Attack Single Email Address*
- *E-Mail Attack Mass Mailer*

La primera permite enviar un email a una sola dirección de email (típico ***spear phishing***), mientras que la segunda lo haría a una lista de receptores que importemos previamente, tan amplia como queramos. Con esta opción se puede hacer un **ataque dirigido contra un grupo de personas**, por ejemplo, ya que el mensaje se enviará a todos los miembros de esta lista.

```
   Social Engineer Toolkit Mass E-Mailer

   There are two options on the mass e-mailer, the first would
   be to send an email to one individual person. The second option
   will allow you to import a list and send it to as many people as
   you want within that list.

   What do you want to do:

   1.  E-Mail Attack Single Email Address
   2.  E-Mail Attack Mass Mailer

   99. Return to main menu.

set:phishing>
```

Figura 231. Configuración del tipo de envío de emails a realizar

```
   Do you want to use a predefined template or craft
   a one time email template.

   1. Pre-Defined Template
   2. One-Time Use Email Template

set:phishing>
```

Figura 232. Composición del correo de *phishing*

En este preciso momento **puedes componer tu propia plantilla de correo electrónico** o utilizar una predefinida. La segunda opción es solo para pruebas rápidas, ya que estas plantillas predeterminadas son conocidas y las posibilidades de ser identificadas por el software defensivo son muy altas. Este es el momento en el que **debes poner en juego toda la información que has recopilado del objetivo** para componer tu propio correo electrónico personalizado, aumentando la **credibilidad** (y por lo tanto, la **efectividad**) del ataque.

```
set:phishing>1
[-] Available templates:
1: New Update
2: How long has it been?
3: Have you seen this?
4: Status Report
5: Dan Brown's Angels & Demons
6: Strange internet usage from your computer
7: Computer Issue
8: Baby Pics
9: WOAAAA!!!!!!!!!! This is crazy...
10: Order Confirmation
set:phishing>10
set:phishing> Send email to:
```

Figura 233. Correos de *phishing* predefinidos

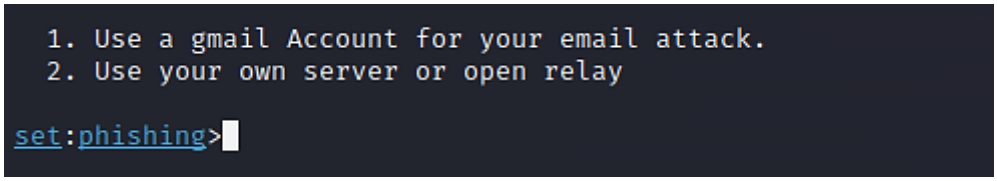


Figura 234. Elección del tipo de servidor a través del cual se enviarán los mails

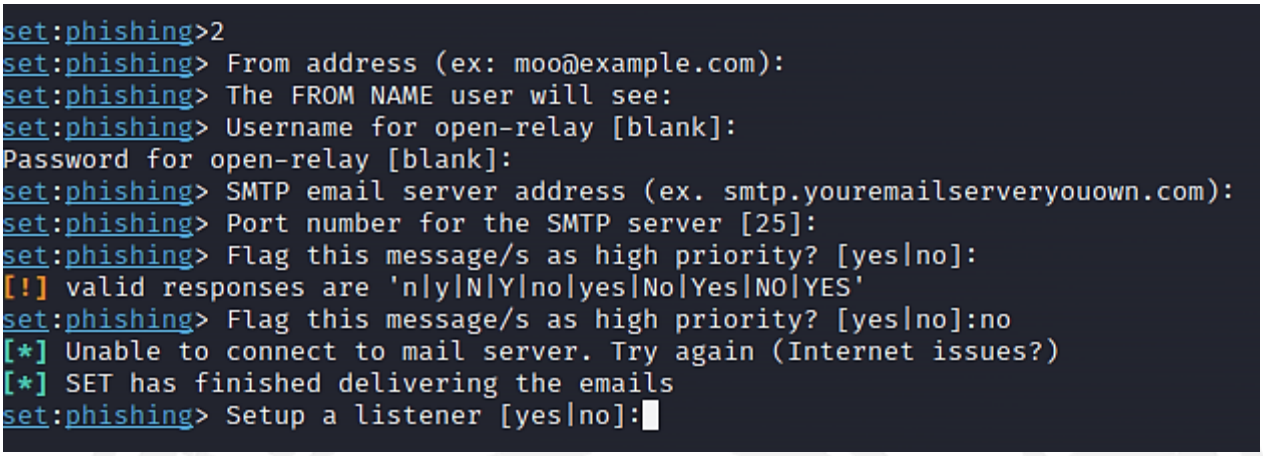


Figura 235. Proceso de envío de email (no funciona, porque no hemos configurado el servidor a propósito)

Ataques de phishing vía web con SET

Si vamos al menú anterior y seleccionamos "***The web attack vectors***", accederemos a una forma sencilla de usar ataques basados en web que puedan comprometer a víctima. Básicamente se usan para crear ataques de *phishing* que comprometan a todo aquel que haga clic en un enlace que genere una de estas opciones, como los que aparecen en esta imagen. Date cuenta de que cosas como "***Credential harverster***" son más adecuadas para lo que estamos describiendo en este documento.

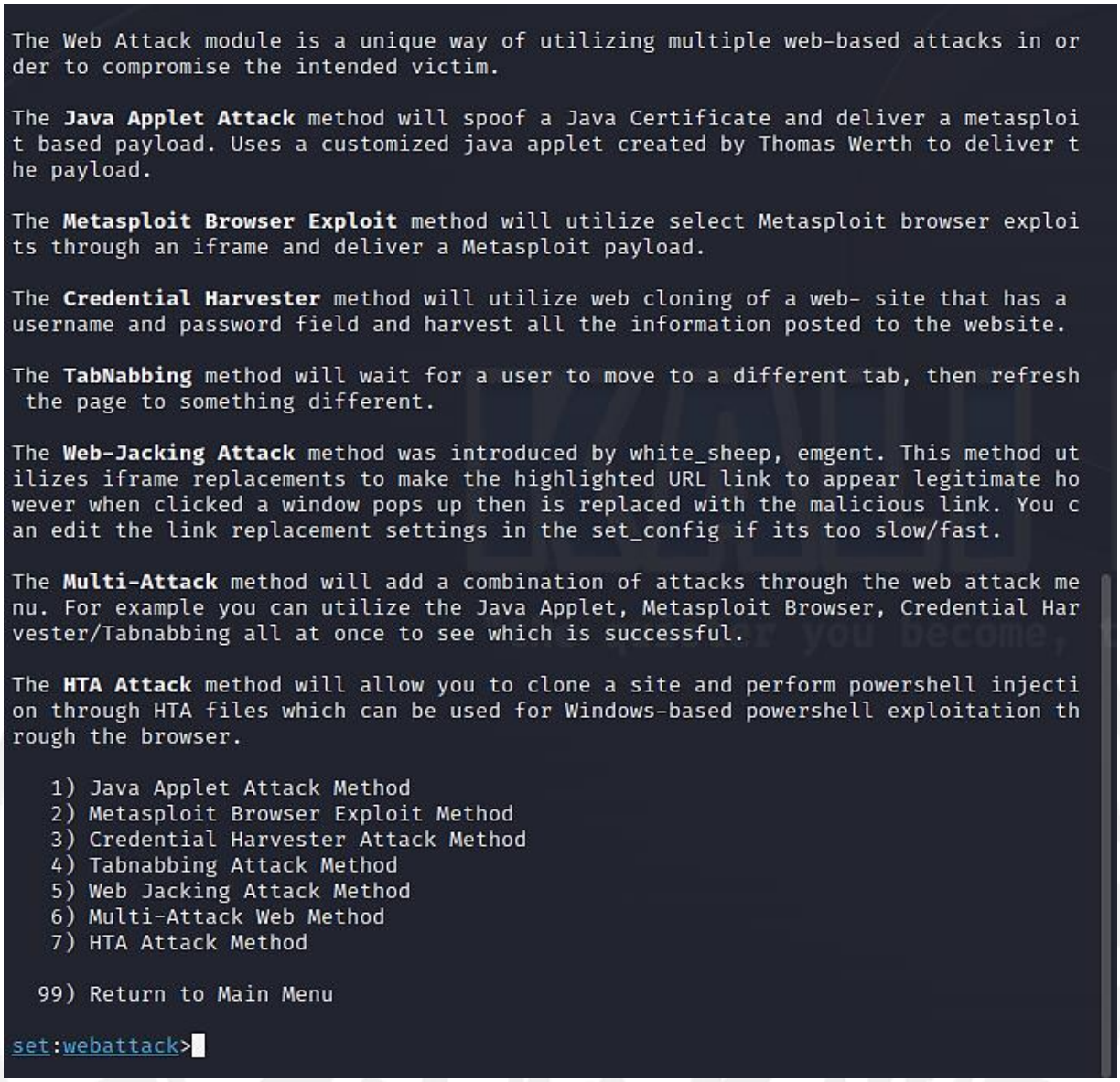


Figura 236. Tipos de ataques web (y su descripción) disponibles en SET

En el siguiente menú vemos claramente que tenemos la opción de usar una plantilla web predefinida, la nuestra propia o simplemente **clonar un sitio web real**, lo que puede ser muy efectivo con poco esfuerzo.

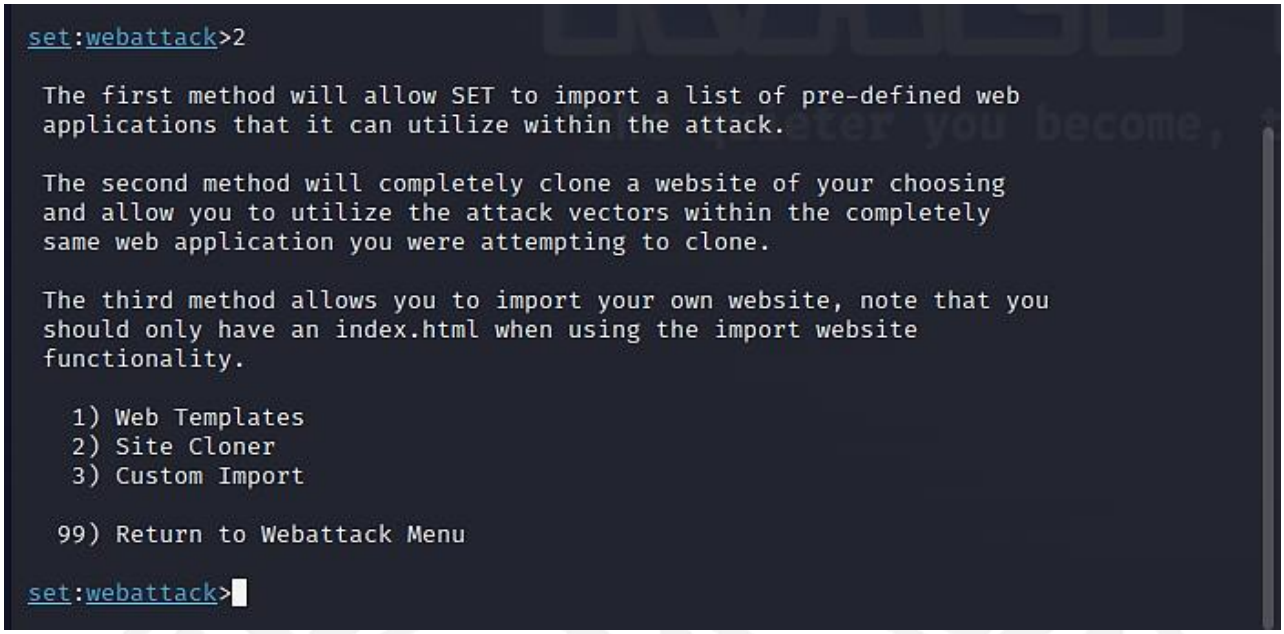


Figura 237. Obtención de la web falsa para desencadenar el ataque

Medios infectados con SET

En el menú principal también tenemos el módulo "***Infectious USB/CD/DVD***" que creará un fichero `autorun.inf` y un payload de `metasploit` (**Tema 7**, **Laboratorios 11 y 12**). Si copiamos ambas cosas a un DVD/USB y la victima inserta ese medio en su máquina, activará la opción de auto ejecución y se ejecutará dicho *payload*, comprometiendo el sistema. Para generar el payload se pueden usar estas dos opciones

- *File-Format Exploits*
- *Standard Metasploit Executable*

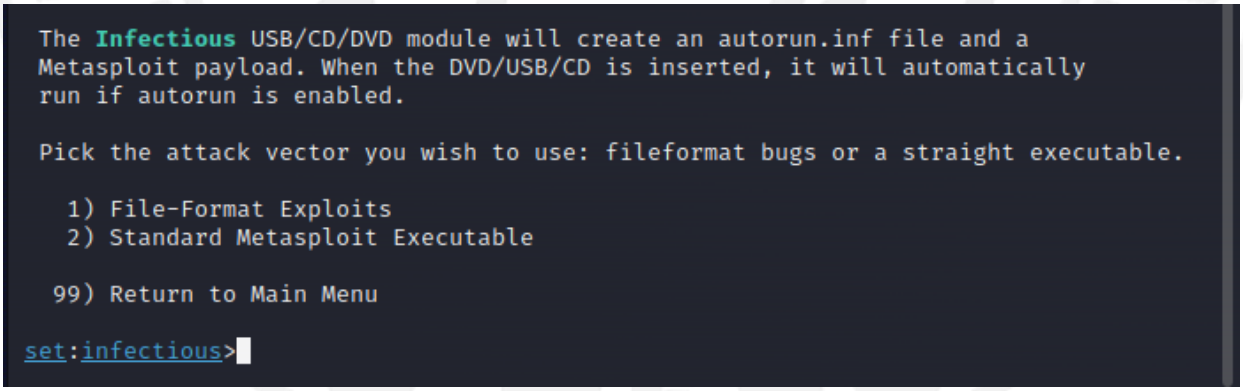


Figura 238. Tipos de payloads para medios infectados de SET

Payloads y listeners

La opción "***Create payload and listener***" del menú principal es una forma simple de crear un *payload Metasploit*, mucho más que hacerlo nosotros a mano (como se explica en el **Tema 7**). Esta opción exportará el ejecutable que tenemos que hacer llegar a la víctima (y convencerla para que lo ejecute) y un *listener* para él. Cualquier victima que lo ejecute enviará una petición a nuestro *listener*, y con ella podemos hacer un ***reverse shell***. En otras palabras, la misma secuencia de eventos que hemos visto en la parte de ejecución de una cadena de ataque, pero con esta opción tendremos el *payload* creado **para usar donde creamos oportuno**.

```
   1) Windows Shell Reverse_TCP               Spawn a command shell on victim and s
end back to attacker
   2) Windows Reverse_TCP Meterpreter         Spawn a meterpreter shell on victim a
nd send back to attacker
   3) Windows Reverse_TCP VNC DLL             Spawn a VNC server on victim and send
 back to attacker
   4) Windows Shell Reverse_TCP X64           Windows X64 Command Shell, Reverse TC
P Inline
   5) Windows Meterpreter Reverse_TCP X64     Connect back to the attacker (Windows
 x64), Meterpreter
   6) Windows Meterpreter Egress Buster       Spawn a meterpreter shell and find a
port home via multiple ports
   7) Windows Meterpreter Reverse HTTPS       Tunnel communication over HTTP using
SSL and use Meterpreter
   8) Windows Meterpreter Reverse DNS         Use a hostname instead of an IP addre
ss and use Reverse Meterpreter
   9) Download/Run your Own Executable        Downloads an executable and runs it

set:payloads>
```

Figura 239. *Payloads* (y sus *listeners*) creables con SET

Powershell

Finalmente, si volvemos al menú anterior y pulsamos "9" vamos a "***Powershell Attack Vector***" donde tendremos estas opciones.

```
set> 9

The Powershell Attack Vector module allows you to create PowerShell specific attack
s. These attacks will allow you to use PowerShell which is available by default in
all operating systems Windows Vista and above. PowerShell provides a fruitful  land
scape for deploying payloads and performing functions that  do not get triggered by
 preventative technologies.

   1) Powershell Alphanumeric Shellcode Injector
   2) Powershell Reverse Shell
   3) Powershell Bind Shell
   4) Powershell Dump SAM Database

  99) Return to Main Menu

set:powershell>
```

Figura 240. Ataques vía *PowerShell* disponibles con SET

Este módulo permite crear ataques concretos con *PowerShell*, que viene instalado por defecto en sistemas operativos Windows y en algunos Linux. Este vector de ataque tiene un gran nº de *payloads* con funciones que pueden fácilmente no ser detectados por tecnologías antivirus modernas, aunque se escapa del ámbito de esta asignatura.

- *PowerShell Alphanumeric Shellcode Injector*
- *PowerShell Reverse Shell*
- *PowerShell Bind Shell*
- *PowerShell Dump SAM Database*

Gophish

Gophish (https://getgophish.com/) es un ***framework* de *phishing*** que hace que **simular un ataque real de *phishing*** sea muy sencillo. La idea simplemente es que cualquiera pueda poner en marcha uno de estos ataques con el objeto de entrenar a sus empleados de forma rápida y sin problema. Es gratuito y muy usado en empresas que se dedican a la seguridad. El motivo es sencillo: permite lanzar campañas por separado, total configuración de los emails a enviar, elegir el tipo de *payload*, gestionar en tiempo real todas las "víctimas" de los emails de *phishing* lanzados, y otras herramientas de gestión que la hacen muy avanzada, todo vía web.

No obstante, para ponerlo en marcha **necesitaríamos una IP contactable por las máquinas de las posibles víctimas del *phishing*** (normalmente pública) y un servidor de correo o cuenta a través de la cual enviar los correos de *phishing* de manera automatizada (la herramienta lo haría a través del que le configuremos). Debido a la dificultad extra que esto supone, y que obtener una infraestructura de pruebas adecuada puede suponer un gasto para vosotros (servidores en nube), no vamos a incluirlo en la asignatura. No obstante, mencionarlo para que conozcáis su existencia es obligatorio debido al uso que tiene en la empresa privada, y que por tanto podríais tener la opción de orquestar estas campañas en vuestro trabajo. Si crees que tienes la infraestructura de pruebas necesaria, la instalación y puesta en marcha se hace de forma muy sencilla siguiendo los pasos de este enlace: https://docs.getgophish.com/user-guide/

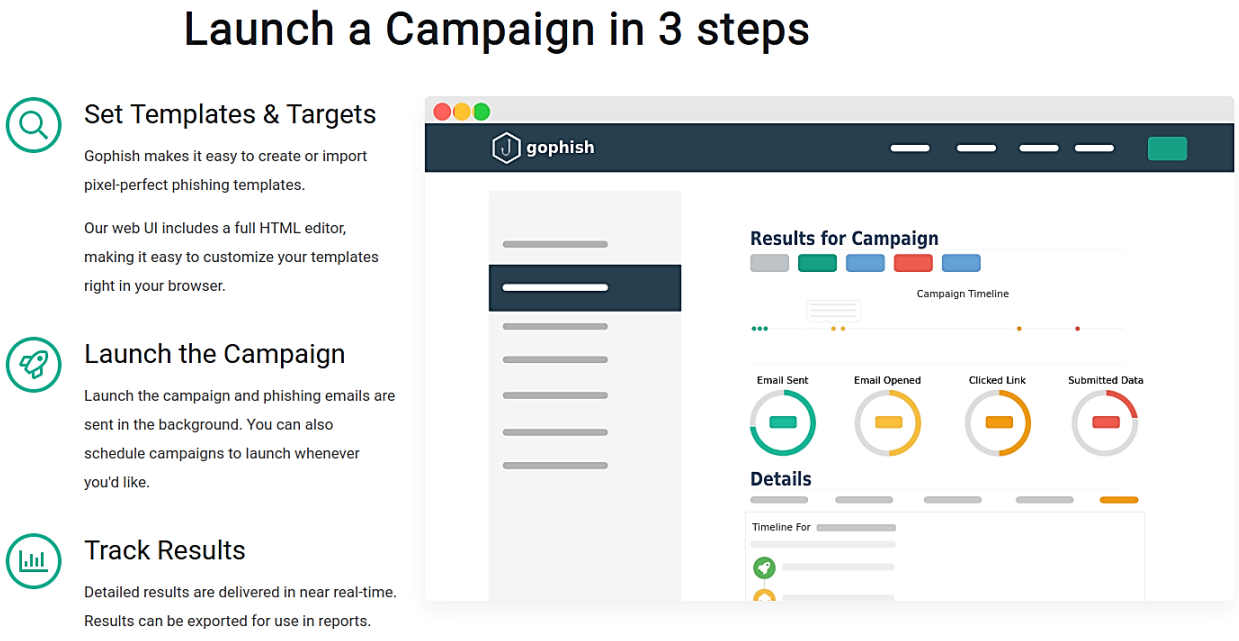


Figura 241. *GoPhish*

Evilgnx2

Evilgnx2 (https://github.com/kgretzky/evilginx2) es un *framework* para **crear ataques *Attacker-in-the-Middle* (AitM)** con el objetivo de obtener credenciales de **usuarios y sus *cookies* de sesión** mediante ataques de *phishing*. Obtener *cookies* tiene además la consecuencia de poder **saltarse medidas de seguridad avanzadas, como 2FA**. El uso de esta herramienta requiere lo mismo que *Gophish*: una IP contactable por las víctimas, normalmente en Internet. Es por ese motivo por lo que no lo incluiremos en esta asignatura tampoco, aunque si tienes posibilidad de tener esa infraestructura, puedes seguir este tutorial para ponerlo en marcha y "exprimirlo".

- https://thehackerway.com/2019/11/06/phishing-2-0-con-evilginx2-parte-1/
- https://thehackerway.com/2020/01/02/phishing-2-0-con-evilginx2-parte-2/

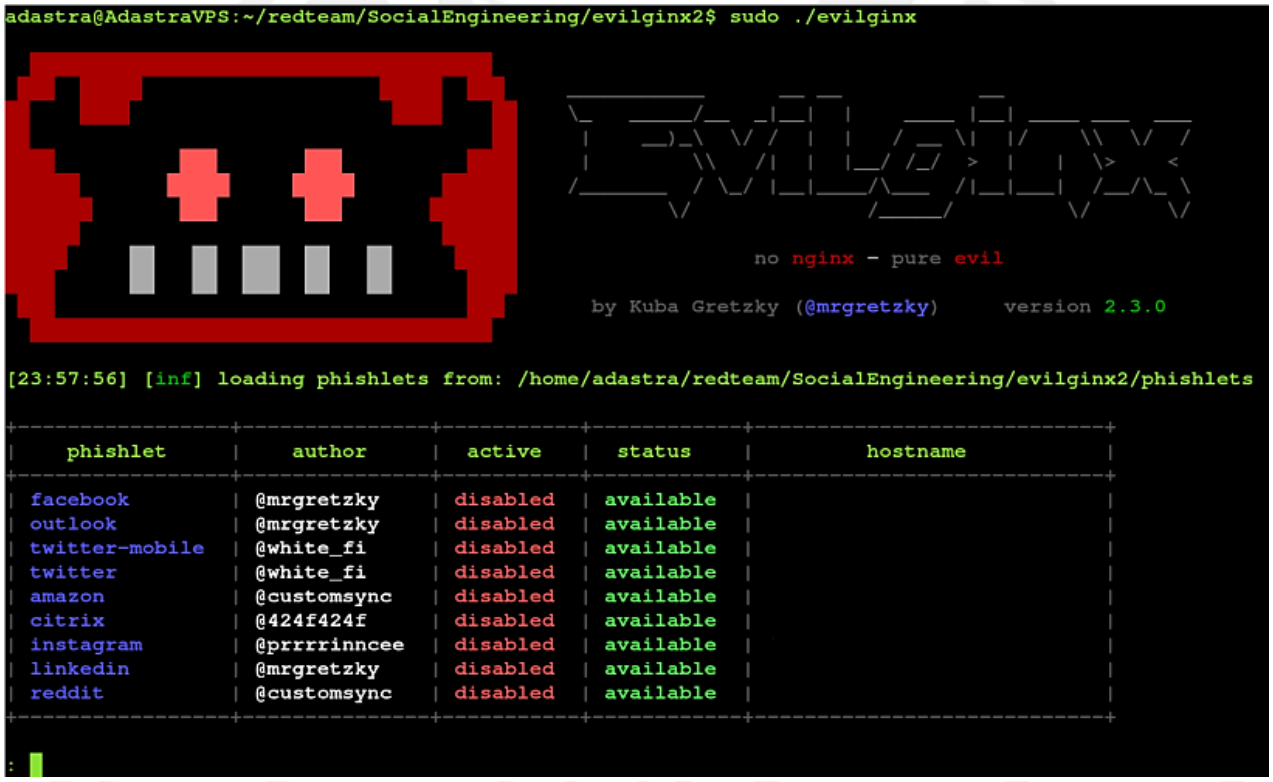


Figura 242. *EvilGinx2*

Investigaciones en la *Dark web*

Este apartado de la documentación soy consciente de que es polémico, porque puedes pensar (y realmente es así en su mayor parte) que en la *dark web* solo hay **delitos e información para cometer crímenes**. Esta información **no se da para que lo hagas como parte del trabajo**, sino para que te des cuenta de que realmente es importante para empresas y organizaciones monitorizar la *dark web* por varias razones:

- **Prevención de ciberataques**: La *dark web* es un lugar donde los ciberdelincuentes pueden compartir y vender información robada o filtrada, como contraseñas comprometidas, credenciales, propiedad intelectual y otros datos sensibles. Si una empresa u organización monitoriza la *dark web* **puede detectar si su información ha sido comprometida** y tomar medidas para responder al problema.

  En una investigación OSINT encontrar información privada comprometida puede dar "las llaves del reino" a un delincuente. Por todo ello, vigilar la *dark web* (o tener un servicio contratado que lo haga por ti si eres una empresa), es una necesidad, ya que si en lugar de hacerlo tú lo hace

un delincuente, puede entrar en tu infraestructura simplemente con datos que ha encontrado "por ahí".

- **Respuesta rápida a incidentes**: Sin una monitorización bien hecha de la *dark web*, una empresa u organización **puede pasar meses sin saber que ha ocurrido una filtración de datos**. Cuando se quieran dar cuenta en estos casos, ya será seguramente demasiado tarde.
- **Protección contra pérdidas financieras**: Las consecuencias financieras de una violación de datos pueden costar a una organización o empresa una gran cantidad de dinero, que puede disminuirse si se sabe qué datos han sido comprometidos exactamente y, en función de ellos, los sistemas afectados que hay que corregir.
- **Cumplimiento legal**: Dependiendo de la naturaleza de los datos que se hayan violado, las organizaciones pueden estar sujetas a acciones legales de los organismos reguladores, **incluyendo multas si se demuestra negligencia** (Ej.: la AEPD). Detectar la filtración pronto y poner remedio lo antes posible al problema puede ahorrar mucho dinero en multas.
- **Daño a la reputación**: La pérdida de confianza que causa un ataque puede llevar a la pérdida de clientes y también crear dificultades para ganar nuevos. Pero si se sabe el alcance de la filtración puede elaborarse una **política de comunicación adecuada** que intente minimizar estos impactos.

Monitorizar la web requiere experiencia y herramientas adecuadas. Para algo básico, puedes usar una conexión desde una máquina virtual con un ***ToR Browser*** en modo estricto (**Laboratorio 7**) para evitar el mayor nº de problemas posible al navegar . Solo necesitarías entonces los enlaces `.onion` de sitios de la *dark web* donde se vuelcan los datos filtrados, algo que no es muy difícil de encontrar 😉.

Estos enlaces te pueden servir de ayuda. No obstante, insisto en que aunque esto no sea *engagement* **tiene un riesgo no despreciable y solo se da para vuestra información**, no siendo necesario hacerlo como parte del trabajo:

- https://www.fortinet.com/resources/cyberglossary/dark-web-monitoring
- https://webz.io/blog/dark-web-monitoring/dark-web-monitoring-your-essential-guide/
- https://blog.elhacker.net/2023/04/investigaciones-osint-en-la-dark-web.html
- https://osintteam.blog/dark-web-how-to-get-started-658409184c06?gi=b37e78471105

😇 Anexo: Un amigo/a quiere saber más de estos temas (bueno, vale, el amigo/a soy yo): Sitios donde aprender más sobre OSINT

Como dijimos en la introducción este documento es una introducción a esta disciplina de la seguridad y aunque se ha tratado de cubrir una gran cantidad de casos de uso comunes, hay mucho más donde explorar. Por eso hay planes de **desarrollar esta documentación en una asignatura completa en el futuro**. Si quieres saber más de OSINT mientras eso ocurre, te puedo recomendar los siguientes enlaces.

Guías y artículos

- **Artículo introductorio "*A Guide To Open Source Intelligence (OSINT)*":** https://itsec.group/blog-post-osint-guide-part-1.html
- **Guía "*A Beginners Guide to OSINT*":** https://www.csnp.org/post/a-beginners-guide-to-osint
- ***Telegram OSINT: The Ultimate Guide to Unveiling Insights Beyond What Meets the Eye***: https://espysys.com/blog/telegram-osint-the-ultimate-guide-to-unveiling-insights-beyond-what-meets-the-eye/

Repositorios de documentación, guías y herramientas

- **Repositorio "*Awesome Social Engineering*"**, con un gran nº de materiales relacionados con cómo obtener información de distinta clase: https://github.com/giuliacassara/awesome-social-engineering
- **Repositorio "*Awesome OSINT*"**, con un gran nº de materiales relacionados con cómo obtener información de distinta clase**:** https://github.com/jivoi/awesome-osint
- **Repositorio "*OSINT Anonymous*"**, especializado en herramientas para crear identidades falsas complementarias a las vistas en el documento**:** https://github.com/CScorza/OSINTAnonymous
- **Repositorio "*Image Reseach OSINT*"**, especializado en herramientas para hacer investigaciones de varios tipos sobre imágenes: https://github.com/cqcore/Image-Research-OSINT
- **Repositorio del curso "*Python for OSINT in 21 days*"**, con los materiales de dicho curso introductorio al OSINT**:** https://github.com/cipher387/python-for-OSINT-21-days
- **Blog "*OH SHINT!*",** con una enorme cantidad de información acerca de cómo obtener datos públicos de una gran cantidad de temas distintos: https://ohshint.gitbook.io/oh-shint-its-a-blog/
- **Repositorio "*OSINT*"**, con un gran nº de materiales relacionados con cómo obtener información de distinta clase: https://github.com/sinwindie/OSINT
- ***OSINT4ALL*:** Una guía masiva de herramientas adaptadas a cualquier necesidad que podamos tener en una investigación OSINT: https://start.me/p/L1rEYQ/osint4all?locale=es