Trabajo Fin de Máster

Técnicas y herramientas de recolección de información mediante OSINT en RRSS

Máster Universitario en Dirección de Proyectos Informáticos

Presentado por:
D. Víctor Pablo Prado Sánchez

Dirigido por:
D. José Javier Martínez Herraiz

Alcalá de Henares, a 21 de Junio de 2023

UNIVERSIDAD DE ALCALÁ

Escuela Politécnica Superior

Máster en Dirección de Proyectos Informáticos

Trabajo Fin de Máster

Técnicas y herramientas de recolección de información mediante OSINT en RRSS

Autor: Víctor Pablo Prado Sánchez
Director: José Javier Martínez Herraiz

TRIBUNAL:

Presidente:

Vocal 1º:

Vocal 2º:

CALIFICACIÓN: _____ **FECHA:**

Índice

Índice de figuras

Resumen

Hoy en día tenemos millones de datos y de información publicada en internet y de acceso libre debido a la contante y continua publicación de nuevos contenidos mediante diversas formas, redes sociales, blogs... Esto ha favorecido que se almacene una desorbitada cantidad de datos online a partir de los cuales se puede obtener información de gran valor y utilidad mediante técnicas como OSINT.

En este trabajo estudiaremos e investigaremos diferentes herramientas y técnicas OSINT existentes, cómo se utilizan, su formación, sus objetivos y sus consecuencias en un enfoque centralizado en redes sociales.

Palabras clave: OSINT, ciberseguridad, ciberdelincuente, información, inteligencia, fuentes abiertas, redes sociales.

Abstract

Nowadays we have millions of data and information published on the Internet and freely accessible due to the constant and continuous publication of new content through various forms, social networks, blogs... This fact has provoked the storage of a huge amount of online data from wich we can obtain, through techniques and tools such as OSINT, useful and great value information.

In this paper we will study and investigate different OSINT tools and techniques, how they are used, their formation, their goals and their consequences in a social media-centric approach.

Keywords: OSINT, cybersecurity, cybercriminal, information, intelligence, open sources, social media.

1. Planteamientos y objetivos

1.1. Objetivos

El objetivo del proyecto será buscar, recolectar y explicar diferentes herramientas ya existentes, así como ver su origen, funcionalidad, estructura de control, métodos de explotación, estudios e informes hechos por grandes compañías de ciberseguridad.

Para cada tipo de herramienta que se va a utilizar, se va a estructurar cada parte de la siguiente manera:

- Taxonomía y procedimiento: clasificación y procesos a seguir para la consecución de diferentes propuestas de valor.
- Instrumentos empleados: donde se tratarán las diferentes herramientas, los distintos equipos y entornos usados.
- Puesta en escena: describiremos cómo y de qué forma se llevará a cabo el lanzamiento de las herramientas empleadas.
- Desarrollo de la propuesta: se desplegará y analizará en profundidad las herramientas a utilizar y las opciones disponibles en ellas. A su vez, se mostrarán diferentes capturas de pantalla donde se verá reflejado todo el trabajo realizado de forma secuencial.
- Conclusiones detalladas: nos encargaremos de analizar bien todo el conjunto de pruebas y resultados obtenidos, para de esta forma obtener información clara acerca del objetivo.

1.2. Planteamientos

En este proyecto se van a usar varias herramientas con las cuales se analizarán la recolección de información con el fin de poder mostrar diferentes métodos de recolección con inteligencia. Las herramientas que se van a usar son las siguientes:

- ✓ **Namechk**: herramienta en línea que se emplea para verificar la disponibilidad de un nombre en múltiples plataformas en línea.
- ✓ **Sherlock:** herramienta de código abierto desarrolla en Python que tiene como objetivo la búsqueda activa de un nombre de usuario en las principales redes sociales.

- ✓ **theHarvester:** herramienta que recopila cuentas de correo y dominios utilizando diferentes fuentes abiertas como Bing, Yahoo, VirusTotal...

- ✓ **Dorks:** técnica que, a través de operadores, ponen su puesta en escena mediante los buscadores con el fin de conseguir alcanzar una búsqueda avanzada.

- ✓ **SXDork**: herramienta que se encarga realizar Google Dorking (empleo de dorks en la búsqueda en Google) para buscar información específica en internet.

- ✓ **Accountanalysis:** herramienta que se encarga de analizar el comportamiento y evaluar las diferentes cuentas de Twitter que se encuentren públicas.

- ✓ **Cree.py:** herramienta de ingeniería social enfocada a redes sociales que nos permite realizar la recopilación de información por medio de la geolocalización a través de plataformas de redes sociales.

- ✓ **FOCA**: herramienta que nos permite obtener información oculta y metadatos a partir de documentos; estos se pueden subir a la misma, o bien, son encontrados en dominios.

- ✓ **OSINTGram:** herramienta OSINT enfocada a la red social de Instagram que se encarga de analizar el comportamiento y evaluar las diferentes cuentas de Instagram que se encuentren públicas.

2. OSINT: Open Source Intelligence

A la hora de plantear el concepto OSINT, debemos fijar su definición Open Source Intelligence, o lo que es lo mismo, Inteligencia de Fuentes Abiertas. En esto último existen diferentes cuestiones, ya que a qué llamamos fuentes abiertas; aquellas a las que se pueden acceder desde la surface web, es decir, desde la primera capa de Internet en la cual no existe ningún tipo de restricción.

En este trabajo enfocamos esas fuentes abiertas a las redes sociales, siempre dentro del marco de la surface web, mediante iremos realizando el procedimiento OSINT a través de diferentes técnicas y herramientas hasta poder obtener la inteligencia a través de datos e información previamente recolectada y convertida en conocimiento.

2.1. OSINT en la actualidad

OSINT es la Inteligencia de Fuentes Abiertas, es decir, obtener datos e información de internet con los que obtener inteligencia de forma posterior.

Hoy en día el **68% de toda la población (5.44/8.01 billones) tiene un teléfono móvil**, un **64.4% (5.16/8.01 billones) usa internet**, y el **59.4% (4.76/8.01 billones) tiene redes sociales.**

Con estos datos del informe realizado por *Meltwater*, podemos ver reflejado la importancia y punto de inflexión de las redes en la actualidad.

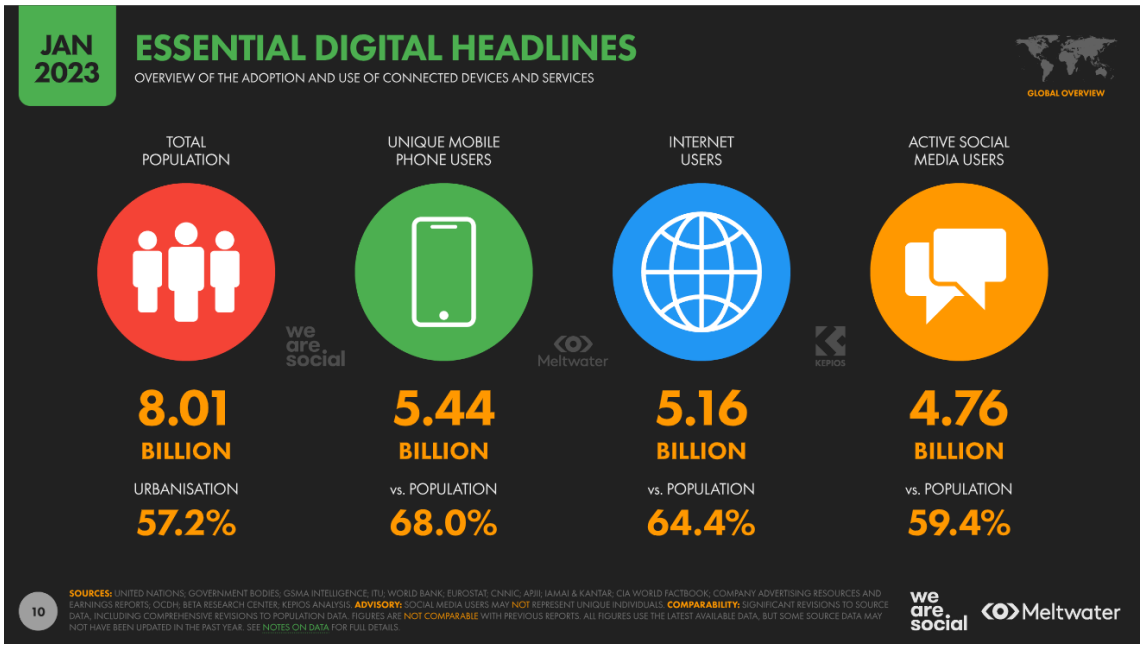


Ilustración 1. Visión de la adopción y uso de los servicios y dispositivos conectados, informe Digital 2023 Global Overview Report v01 de Meltwater

Y es que, una de las principales redes a la que el ser humano se ha vuelto adicto es internet. A través de internet podemos tener acceso a cualquier cosa que queramos hoy en día, desde compras, noticias, películas, reuniones…

Este hecho se puede ver reflejado en la gráfica realizado por *Meltwater*, donde apreciar el incremento de forma exponencial del número de usuarios que usa internet, desde su origen 1990, hasta 2023.

Incluso, un hecho notable y hay que apreciar, es el hecho que cada año, el número de usuarios que navegan por la red de internet es mayor.

- **En comparación con el año pasado, 2022**, los usuarios han **aumentado un 2%.**
- **Respecto a 2021, una subida del 4%.**
- Con un intervalo de cinco años, **desde 2019, su uso ha incrementado el 23%.**

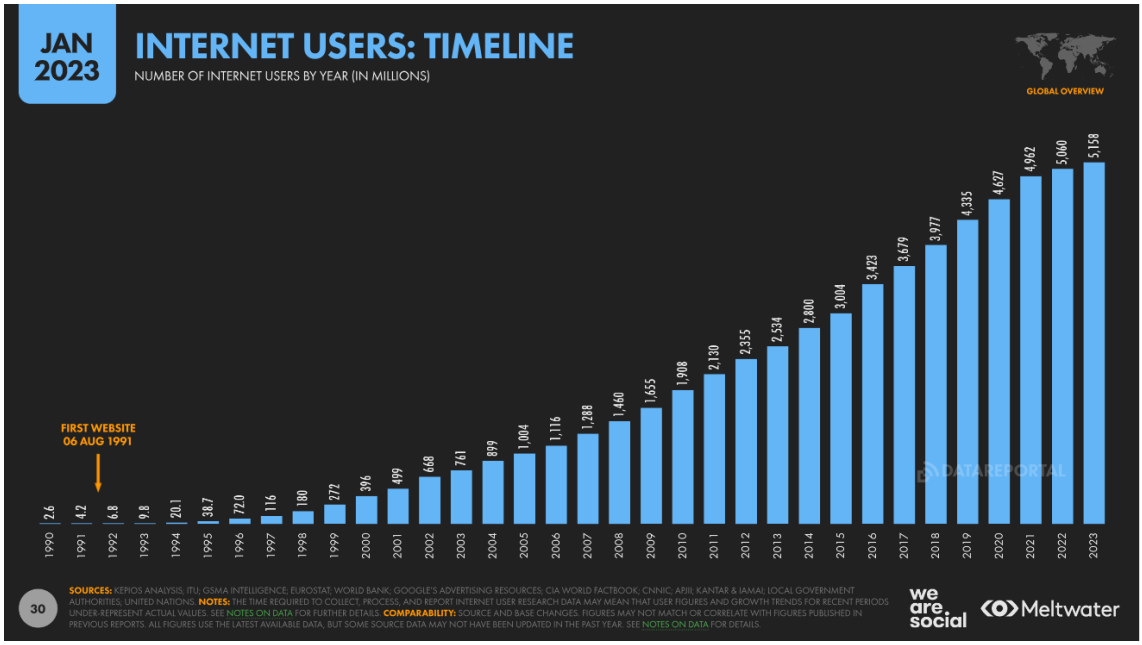


Ilustración 2. Gráfico de los usuarios en Internet a lo largo del tiempo, informe Digital 2023 Global Overview Report v01 de Meltwater

Como podemos apreciar, el uso de internet en las personas incrementa cada año y de forma constante, esto a su vez provoca que cada instante se recojan de forma constante diferentes datos e información de cada usuario, que dependiendo de dónde y cómo se almacenen puede provocar una cesión involuntaria o voluntaria de estos.

OSINT es una práctica ampliamente utilizada en diversos campos, desde el ámbito de seguridad y defensa hasta el empresarial. Según los objetivos establecidos,

gracias a ello podemos adquirir datos de entidades, individuos, corporaciones u otros tipos de fuentes, siempre y cuando la información esté disponible en línea.

2.2. OSINT en RRSS

Las redes sociales más usadas en la actualidad como podemos apreciar en la gráfica realizado por *Meltwater* entre la gente que comprende la edad de dentro del intervalo entre 16 y 64 años, son: Instagram, Facebook, TikTok y Douyin.

No obstante, hay diferentes aplicaciones de mensajería instantánea que son muy usadas, pero que no se consideran redes sociales, como son Whatsapp, Facebook Messenger, Telegram…

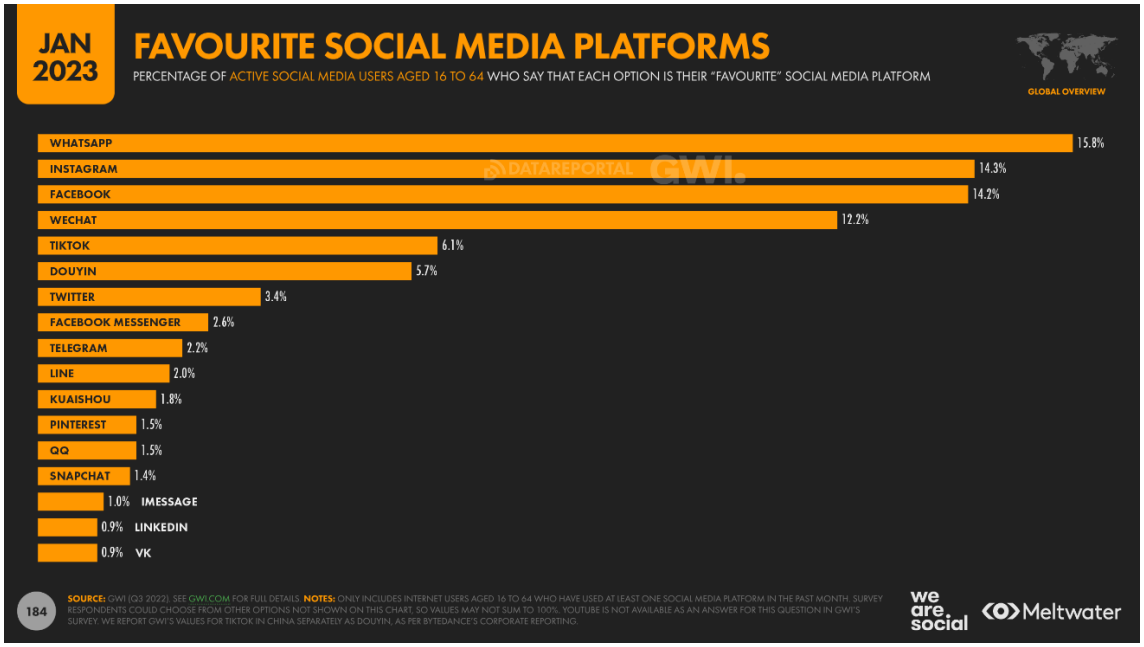


Ilustración 3. Redes sociales favoritas para la gente entre 16 y 64 años, informe Digital 2023 Global Overview Report v01 de Meltwater

Todos los usuarios, ya sean personas o empresas tienden a subir diferentes datos e información en redes sociales, blogs o en aplicaciones de mensajería instantánea.

Desde contenido multimedia, mensajes, conversaciones públicas… Es decir, de todo tipo de información con la cual se puede luego obtener inteligencia. Nos encontramos muy expuestos con internet, y más cuando nosotros mismos somos los que ampliamos la información existente añadiendo datos e información de valor de nuestro entorno.

Las técnicas y herramientas de OSINT pueden dar lugar a resultados muy extensos dependiendo del objetivo; es importante categorizar toda información correctamente, y más aún cuando lo enfocamos a las redes sociales.

2.3. Regulación OSINT

En España no se cuenta con un organismo que rija y controle OSINT, pero en otros países, como, por ejemplo, Estados Unidos, Australia, Francia… Sí cuentan con diferentes organismos que regulan la Inteligencia de Fuentes Abiertas.

En los Estados Unidos podemos apreciar dos organismos que se encargan de regular toda la inteligencia e información obtenida por medio OSINT: **OSE**, **Open Source Enterprise**, o lo que es lo mismo, Proyecto de Fuentes Abiertas y **NOSIC**, **National Open Source Intelligence Centre**, o lo que es lo mismo, Centro Nacional de Inteligencia de Fuentes Abiertas.

OSE, es una organización del Gobierno de Estados Unidos dedicada a la inteligencia de fuentes abiertas. Suministran material al Servicio Nacional de Información Técnica, National Technical Information Service (NTIS) y a otros funcionarios gubernamentales a través del servicio de noticias en línea World News Connection.

NOSIC, establecido como fuente de alerta en Australia, Francia, Países Bajos, Nueva Zelanda, Reino Unido y Estados Unidos. Su objetivo es proporcionar a las agencias estatales y federales una capacidad dedicada de monitoreo, investigación y apoyo analítico de fuentes abiertas; este se especializa en inteligencia de orden público, conciencia de amenazas transnacionales y apoyo a la inteligencia criminal.

2.4. Beneficios OSINT

Algunos de los beneficios clave del OSINT son:

- **Menos riesgo que otras formas de inteligencia**: OSINT nos permite recolectar diferentes datos e información pública que se encuentra en Internet dentro de la capa de Surface web. A diferencia de otro tipo de fuente información como es el HUMINT, donde la inteligencia es obtenida por medio de la información humana.

- **Rentabilidad y facilidad de acceso**: es rentable y accesible para empresas y particulares; ya que podemos realizar OSINT a través de herramientas o técnicas siempre que queramos y que por contexto tengamos acceso a internet para obtener los datos e información; o bien, a través de bases de datos obtenidos de la misma.

- **Ayuda a investigadores financieros a detectar evasores de impuestos**: con ello se pueden detectar evasores de impuestos tras un análisis e investigación exhaustiva de sus cuentas y perfiles públicos.

- **Permite combatir la falsificación en línea**: se pueden detectar productos y/o servicios fraudulentos y avisar de esta forma a los usuarios y clientes.

2.5. Grados de centralidad en RRSS

A la hora de exponer los grados de centralidad en redes sociales, tenemos que hablar de **la teoría de grafos**; esto se debe a que una red social, es un tipo de red que contiene diferentes relaciones e interacciones con un conjunto de entidades (personas, empresas…).

De esta forma todos los componentes de la teoría de grafos, los vemos reflejados en **los grados de centralidad: aristas, vértices, caminos…** Cabe destacar que ante mayor grado de centralidad se obtenga de una red, esto indica una ventaja a muchos rasgos; desde nivel estructural, a nivel de influencia, hasta a dominio.

Existen diferentes métricas para analizar una red social, estas son las que veremos a continuación.

2.5.1. Grado/Degree

Este caso, en **grado/degree**, tiene lugar en una red **cuando un nodo tiene el mayor número de aristas**. Esto provoca que exista mayor grado de centralidad. Es decir, las personas y empresas cuando mayor número de conexiones tengan, entonces estarán obteniendo mayor grado de información.

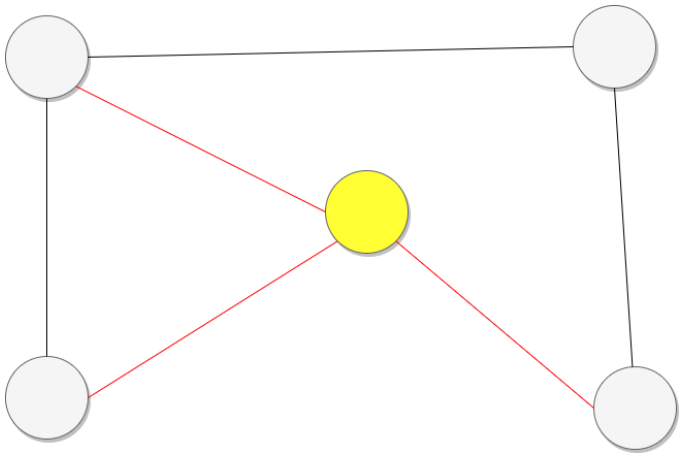


Ilustración 4. Red con grado de centralidad de Grado/Degree generado con draw.io.

2.5.2. Proximidad / Closeness

La **proximidad/closeness**, sucede cuando en una red **un nodo se encuentra ubicado en el centro de esta**. Esto provoca que exista mayor grado de centralidad. Es decir, las personas y empresas cuando más centradas en el centro se encuentren asignadas, entonces estarán mejor posicionadas, por lo que tendrán el mayor grado de información y con mayor rapidez.

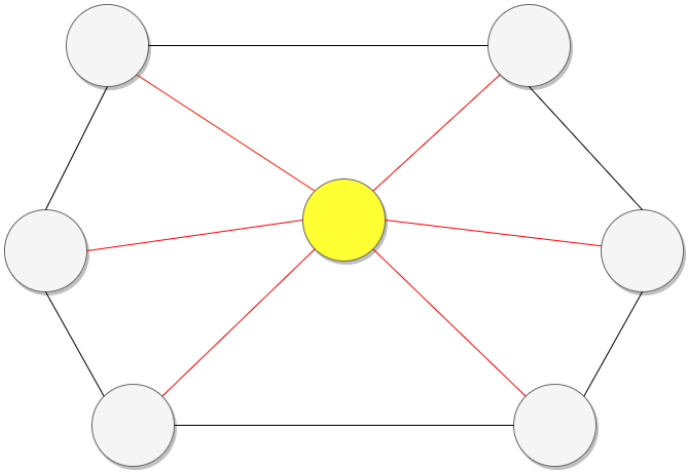


Ilustración 5. Red con grado de centralidad de Proximidad/Closeness generado con draw.io.

2.5.3. Intermediación / Betweenness

La **intermediación/betweenness** ocurre en la red cuando **un nodo actúa el máximo número de veces en la red de puente y camino más corto hacia otros nodos**. Esto provoca que exista mayor grado de centralidad. Es decir, las personas y empresas con una intermediación elevada, quiere decir, que muchas personas y/o empresas se relacionan a través de ellas.

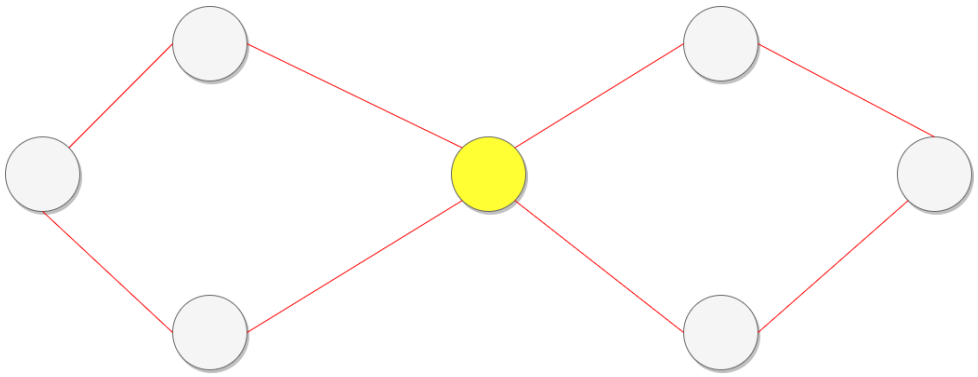


Ilustración 6. Red con grado de centralidad de Intermediación/Betweenness generado con draw.io

2.5.4. Vector propio / Eigenvector

Este caso en el **vector propio/eigenvector**, se da en **una red cuando se destaca la influencia del nodo en la misma**; es decir, nodos conectados a muchos nodos que se encuentran bien conectados. Esto provoca que exista mayor grado de centralidad. Es decir, las personas y empresas con esta métrica de grado de centralidad, son buenos candidatos para difundir información.

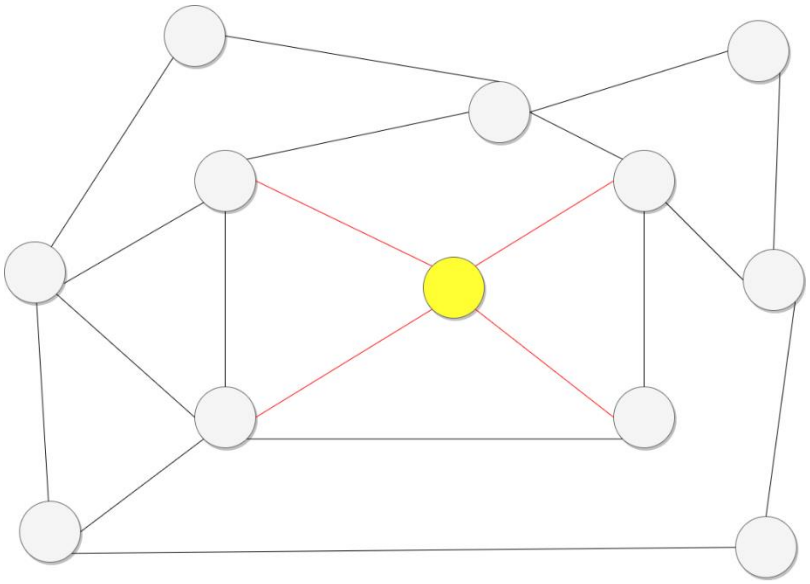


Ilustración 7. Red con grado de centralidad de Vector Propio/Eigenvector generado con draw.io

2.6. SOCMINT vs OSINT

SOCMINT (Social Media Intelligence) también conocido como **Inteligencia de Redes Sociales** se trata de la inteligencia que se obtiene tras recopilar, analizar y tratar grandes cantidades de datos de redes sociales, que después se convierten en información, que con ella obtenemos dicha inteligencia.

Sin embargo, a SOCMINT también le podemos dar el enfoque de la información que se puede llegar a obtener de los foros o plataformas de mensajería instantánea.

La Inteligencia de Redes Sociales anteriormente se conocía como SOCINT, pero este término se ha quedado centrado a la inteligencia que se obtiene de la sociedad y de la cultura.

Esta inteligencia obtenida de las redes sociales puede estar enfocada ya sea con un enfoque individual/colectivo, como puede ser una persona/s o empresa/s, o bien a nivel de acontecimientos.

Todo ello nos hace darnos cuenta de la importancia que tienen las redes sociales y lo expuestos que nos encontramos; son las propias personas y organizaciones las que por medio de estas vías se encuentran subiendo datos e información a internet: metadatos, información de perfiles, interacciones…

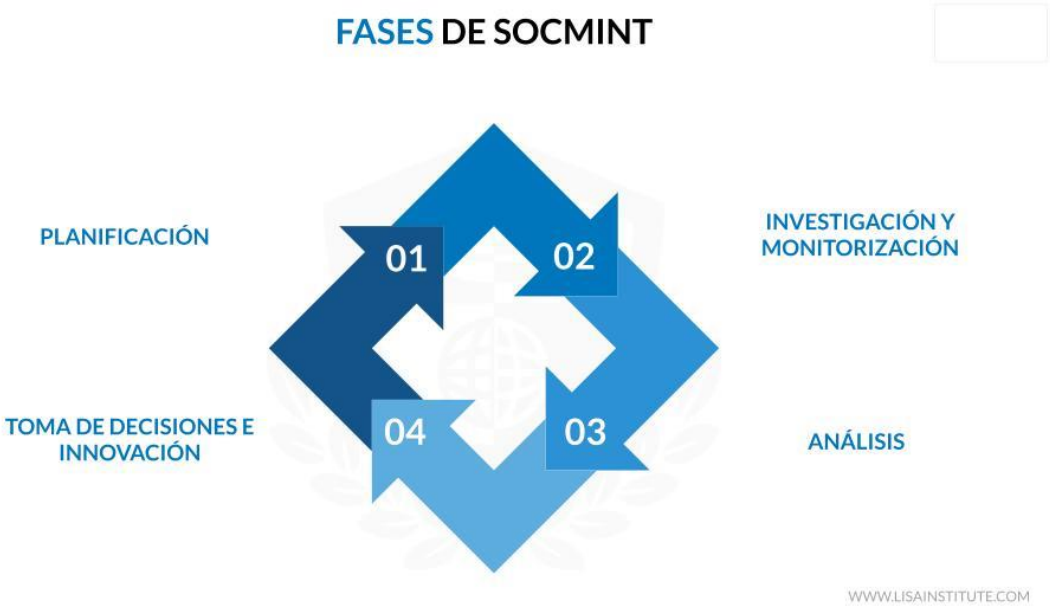


Ilustración 8. Ciclo de fases del SOCMINT vía LISA Institute.

El proceso con el cual obtenemos la Inteligencia de las Redes Sociales lo podemos ver reflejado con el proceso de **obtención de información y la taxonomía llevada a cabo en esta investigación.**

OSINT y SOCMINT tienen semejanzas, pero a su vez diferencias. Ambas técnicas de obtención de inteligencia la consiguen por medio de fuentes abiertas, OSINT desde un ámbito más amplio, en toda la surface web, y SOCMINT, únicamente la obtiene de redes sociales.

Por lo que podemos decir que SOCMINT tiene parte de OSINT, ya que es una técnica que deriva de ella, pero OSINT, no tiene parte de SOCMINT; esto se debe a que SOCMINT puede llegar a realizar técnicas y métodos con los cuales puede acceder a datos e información que no se encuentra de forma pública, algo “ilegal”, como pueden ser conversaciones en mensajes directos en redes sociales…

2.7. Taxonomía de OSINT en redes sociales

Llevaremos a cabo la clasificación y organización jerárquica para poder identificar las redes sociales, todo ello, lo llevaremos a cabo por medio del estudio e investigación que hemos llevado a cabo en este trabajo con el que hemos podido trabajar con diferentes redes sociales a la cuales las clasificaremos.

Por medio de esta taxonomía estableceremos una **estructura lógica y coherente para agrupar las redes sociales y diferenciarlas** de otras que puedan ser distintas. Además, de poder obtener mayor o menor nivel de información de estas.

La taxonomía creada contiene varias condicionantes, que, a la vez, se tratan de diferentes procesos para la clasificación de las redes sociales: **autenticación, API, multimedia y metadatos.**

Procedemos a desarrollar en profundidad los procesos fundamentales para la taxonomía de OSINT en redes sociales:

- **Autenticación**: proceso por el cual se verifica y confirma la identidad de un usuario que intenta acceder a una cuenta en las redes sociales. Se trata de un mecanismo de seguridad para proteger la privacidad de los usuarios y evitar el acceso no autorizado a sus cuentas.

  Existen diferentes métodos de autenticación, desde el más conocido de inicio de sesión con usuario y contraseña, hasta la autenticación por medio de tokens o aplicaciones…

- **API**: proceso por el cual permiten a las redes sociales actuar como un intermediario; de esta forma facilita la comunicación entre distintas aplicaciones o servicios, mejorando la interoperabilidad y la eficiencia en el desarrollo y uso de software.

  También hemos de tener en cuenta la política de privacidad de cada red social, ya que pueden permitir o no, recolectar información y datos de estas.

- **Multimedia**: proceso por el cual se presentan diferentes formas de presentación de datos e información; puede ser a través de texto, imágenes, audio, vídeo…

  Este proceso es fundamental, ya que todas las redes sociales hacen uso de la multimedia para subir contenido sobre ellas: publicaciones de fotos, tweets, vídeos en directos…

- **Metadatos**: se trata de la información y datos que se encuentran dentro de otras fuentes de información y datos. Existen metadatos en imágenes, vídeos, bases de datos, páginas webs… Que nos aportan información extra sobre los mismos.

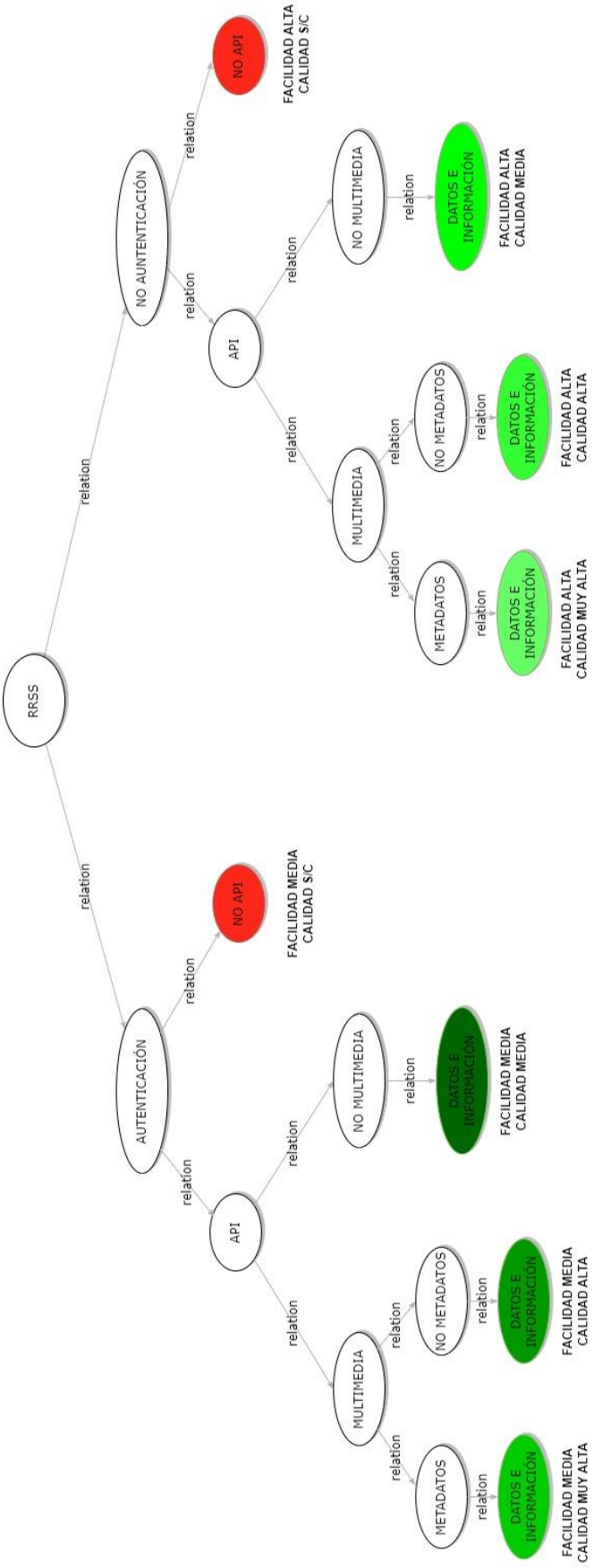


Ilustración 9. Taxonomía de redes sociales generado con draw.io.

Nuestra taxonomía cuenta con una estructura lógica y coherente para agrupar las redes sociales y diferenciarlas, es por ello, que podemos llegar a obtener un total de ocho caminos diferentes para medir el grado de la calidad de la información y datos y la dificultad para obtenerla.

El modo de obtener la información y datos de redes sociales depende del **nivel de facilidad** para su obtención, y el **nivel de la calidad** de esta.

Nivel	Camino	Facilidad	Calidad
Nivel 1	Autenticación No API	MEDIA	S/C
Nivel 2	Autenticación API No multimedia	MEDIA	MEDIA
Nivel 3	Autenticación API Multimedia No metadatos	MEDIA	ALTA
Nivel 4	Autenticación API Multimedia Metadatos	MEDIA	MUY ALTA
Nivel 5	No autenticación No API	ALTA	S/C
Nivel 6	No autenticación API No multimedia	ALTA	MEDIA
Nivel 7	No autenticación API Multimedia No metadatos	ALTA	ALTA
Nivel 8	No autenticación API Multimedia Metadatos	ALTA	MUY ALTA

Como podemos ver reflejado en la tabla, el **mejor nivel para obtener información y datos** es a través del **nivel 8**; ya que el nivel de facilidad es alto, y la calidad es muy alta: su camino pasa por qué no exista autenticación, que accedamos a su API sin problemas, que podamos hacer uso y obtener multimedia, y, por último, que este contenido multimedia contenga diferentes metadatos implícitos.

Por otro lado, el **peor nivel para obtener información y datos** es el **nivel 1;** el nivel de facilidad es medio y la calidad no se puede calificar al no poder obtener nada de información ni de datos; su camino pasa por que exista autenticación, y que no podamos acceder a su API.

2.8. Procedimiento OSINT en redes sociales

Llevamos a cabo un procedimiento sistemático de OSINT dentro de las redes sociales; con este procedimiento podremos ver los diferentes procesos a realizar para obtener un informe final basado en un caso de uso con inteligencia.

Para la elaboración del procedimiento, hemos identificado tres ciclos fundamentales: la recolección de **información y datos** de fuentes abiertas, que, tras procesarlos, conseguimos el **conocimiento**, con el cual, por último, obtenemos las conclusiones con base de **inteligencia**.

Ahora bien, desarrollaremos en profundidad los ciclos fundamentales para el procedimiento de obtención de inteligencia con OSINT en redes sociales:

- **Información y datos**: un dato es un valor o descripción que representa una entidad, objeto o evento específico; estos pueden ser números, texto, gráficos, audio... Los datos en sí no tienen un significado especial, es solo información sin procesar que no ha sido procesada ni interpretada.

  En cambio, la información es el resultado del procesamiento y análisis de datos; significa una comprensión más profunda y significativa de los datos y se puede utilizar para tomar decisiones, hacer predicciones o comprender situaciones complejas. La información generalmente se refiere a datos organizados, estructurados o contextualizados de manera que sean útiles y relevantes para la toma de decisiones.

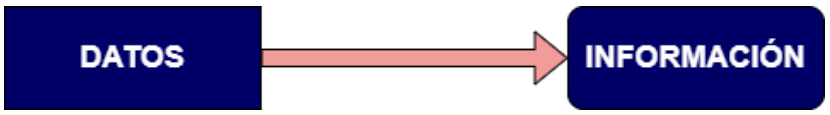


Ilustración 10. Proceso de transformación de datos en información generado con draw.io

  Como conclusión, los datos son la materia prima utilizada para crear información, y la información es el resultado del procesamiento y análisis de datos.

- **Conocimiento**: consiste en la comprensión de diferentes técnicas y principios con los que poder emplear para resolver determinadas situaciones.

  El conocimiento no se refiere a la información que se posee, sino a la capacidad de poder aplicarla de manera correcta y efectiva; es decir, el conocimiento es información procesada.

- **Inteligencia**: es el resultado de la recopilación, evaluación, análisis, integración e interpretación de toda la información disponible de importancia directa o potencialmente material para la planificación y las operaciones.

  No obstante, podemos definir la inteligencia como información procesada vista para una acción; o lo que es lo mismo, conocimiento que servirá para unas acciones posteriores.

La pirámide que se muestra en la imagen inferior muestra los tres ciclos del procedimiento de obtención de inteligencia con OSINT para redes sociales:

Ilustración 11. Pirámide del ciclo obtención de inteligencia en OSINT en redes sociales generado con draw.io

En la pirámide de los pilares del procedimiento de OSINT en redes sociales podemos apreciar el sentido unidireccional, es decir, una vez obtenida la información y datos, pasaremos a obtener conocimiento, y, por último, la inteligencia.

Ilustración 12. Procedimiento con procesos y fases de OSINT en redes sociales generado con draw.io

Información y datos:

1. **Búsqueda activa**: fase en la cual se fija un caso de uso como objetivo, y a partir de ello, comienza la recolección de información y datos de fuentes abiertas para de esta forma obtener la máximo cantidad de estos.

   Se debe localizar e identificar las diferentes fuentes de información relevantes y valorar qué volúmenes datos podemos encontrar; todo ello desde fuentes abiertas en las que todo usuario pude acceder a ellas.

2. **Verificación**: proceso por el cual se procede a validar y corroborar la legitimidad de la información obtenida en la fase de búsqueda activa.

   De esta forma se realiza un filtro de la información recogida que es válida; en cambio la información que es un falso positivo, tenderemos que prescindir de ella, ya que no nos ofrece una puesta de valor.

3. **Búsqueda centralizada**: una vez verificada toda la información obtenida, toca focalizar nuestro caso de uso y realizar una búsqueda centralizada con mayor profundidad con ayuda de la verificación previamente realizada.

   Con esta nueva búsqueda nos garantizamos que la información que recolectemos sea legítima y verdadera, para de forma posterior conseguir conocimiento e inteligencia de ella.

Conocimiento:

4. **Síntesis**: proceso por el cual se lleva a cabo la evaluación y síntesis de toda la información que se ha recolectado; esto se debe a que, si no han sido evaluada, entonces todavía no podremos pasar al ciclo de inteligencia.

   La información recolectada se analiza para dotarla de fiabilidad y credibilidad, utilizando distintas técnicas o métodos de análisis que conducen a la generación de inteligencia.

Inteligencia:

5. **Informe**: se trata de una herramienta que contribuye a la extracción de conclusiones y toma de decisiones, siendo el resultado final del proceso del ciclo de inteligencia y, por tanto, el final de la investigación.

3. Recolección de información mediante OSINT en RRSS, puesta en escena

En este apartado analizaremos y llevaremos a cabo el despliegue y análisis de diferentes herramientas y técnicas de recolección de información mediante OSINT en RRSS; por último, la realización de un informe final con la obtención de diferente información transformada en inteligencia de un objetivo establecido, la Universidad de Alcalá.

Este tipo de recolección y obtención de información se han llevado a cabo en entornos virtuales, para evitar cualquier tipo de percance en el equipo habitual, pero también en este último. En cada apartado se especificarán los instrumentos empleados (entornos, técnicas, máquinas y herramientas).

3.1. Namechk: búsqueda de dominios y usuarios en redes sociales

Namechk es una herramienta en línea que se emplea para verificar la disponibilidad de un nombre en múltiples plataformas en línea.

3.1.1. Instrumentos empleados

Entorno virtual: **VMWare**
Máquina atacante: **Kali Linux**
Herramientas:

- **Namechk: https://namechk.com/**

3.1.2. Puesta en escena de Namechk

Para el uso de Namechk, hemos optado por lanzarla en entorno virtual. Se ha empleado el navegador de forma online para hacer uso de la herramienta.

Namechk se encarga de verificar 36 posibilidades diferentes de nombres de dominio y más de 100 sitios web de redes sociales y plataformas en línea.

En esta puesta en escena se llevarán a cabo la **búsqueda a un objetivo.**

3.1.3. Namechk: el entorno

Abrimos la herramienta de Namechk en nuestro navegador, y vemos las opciones disponibles dentro de esta.

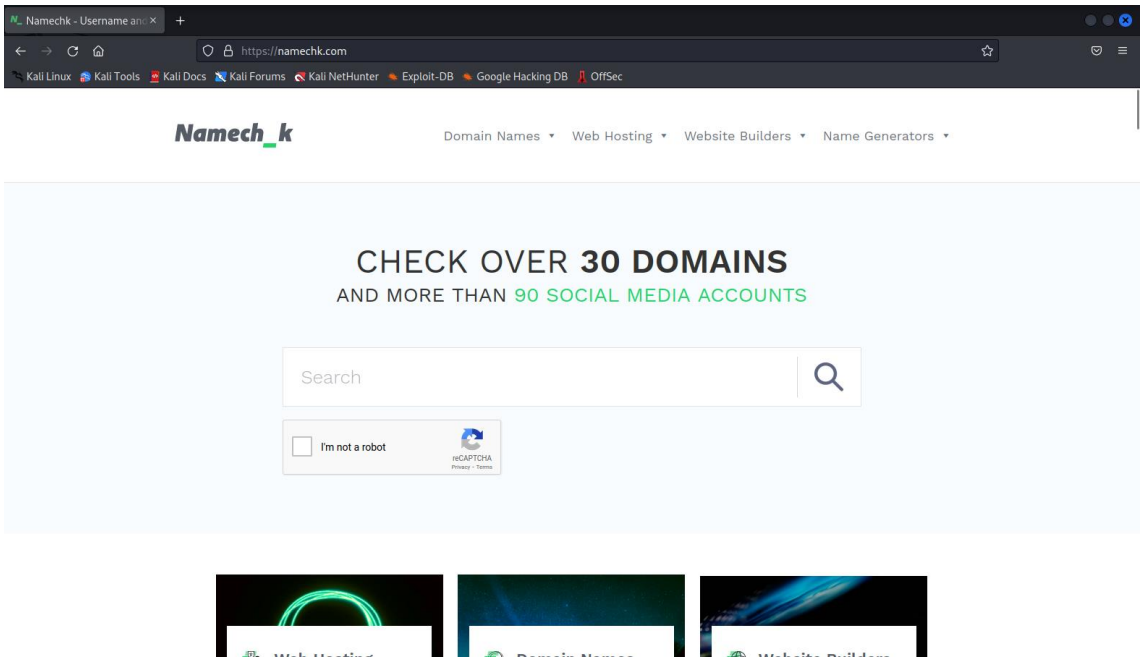


Ilustración 13. Interfaz de la página de inicio de Namechk.

Como podemos observar, hay un único buscador dentro de Nameck, y pondremos el objetivo que queremos investigar.

3.1.4. Búsqueda de un objetivo en dominios y redes sociales

Llevaremos a cabo la búsqueda del nombre de nuestra víctima por medio de la herramienta. Fijamos como objetivo a la Universidad de Alcalá, nuestro objetivo principal en la instigación. Por ello establecemos como nombre de usuario de búsqueda en Namechk UAH, y realizamos la búsqueda.

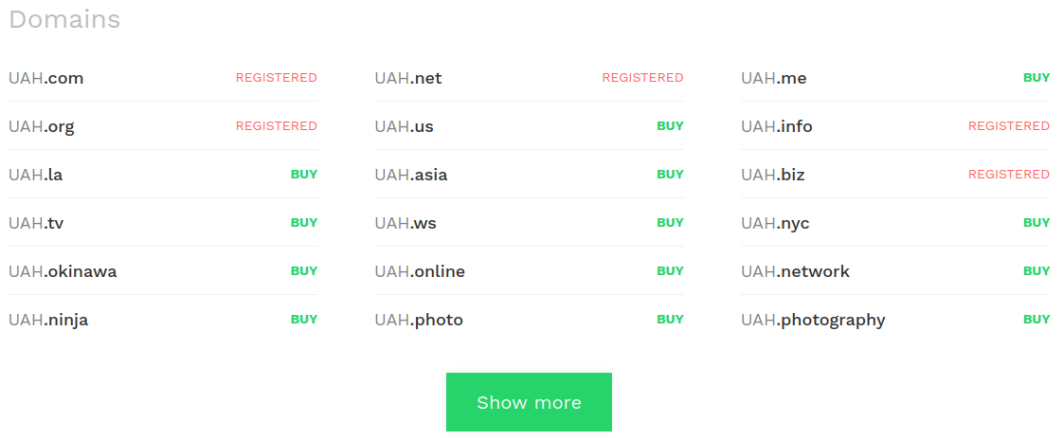


Ilustración 14. Dominios encontrados a partir de la búsqueda UAH en Namechk

Como podemos apreciar, la herramienta nos ofrece un amplio resumen de resultados dependiendo en este caso del tipo de dominio con el que se encuentra registrado "UAH".

En el caso de que se encuentre registrado, nos sale a la derecha del dominio, y si está disponible, nos aparece a su derecha la opción de buy, comprar.

Sabemos que la Universidad de Alcalá se trata de una corporación española, por lo que el filtro que realizamos a los dominios es el de dos casuísticas: .com (de forma global), y .es (de España).

UAH.mobi	REGISTERED	UAH.eu	BUY	UAH.be	BUY
UAH.am	BUY	UAH.pro	REGISTERED	UAH.org.pe	BUY
UAH.net.pe	BUY	UAH.nom.pe	BUY	UAH.com.pe	BUY
UAH.es	REGISTERED	UAH.com.es	BUY	UAH.nom.es	BUY
UAH.org.es	BUY	UAH.au	BUY	UAH.xxx	BUY
UAH.com.au	REGISTERED	UAH.org.au	BUY	UAH.net.au	BUY
UAH.ae.org	BUY	UAH.sg	BUY	UAH.ch	BUY
UAH.co	REGISTERED	UAH.com.co	BUY	UAH.nom.co	BUY
UAH.net.co	BUY				

Ilustración 15. Dominios de UAH en España .es, y de forma global .com

Realizamos la búsqueda de los diferentes dominios para comprobar cuál es el legítimo por parte de la Universidad de Alcalá.

- **UAH.com:** observamos que el dominio existe, y que el propietario no es la Universidad, sino la empresa dan.com quien se hace con diversos dominios para después alquilarlos o venderlos.

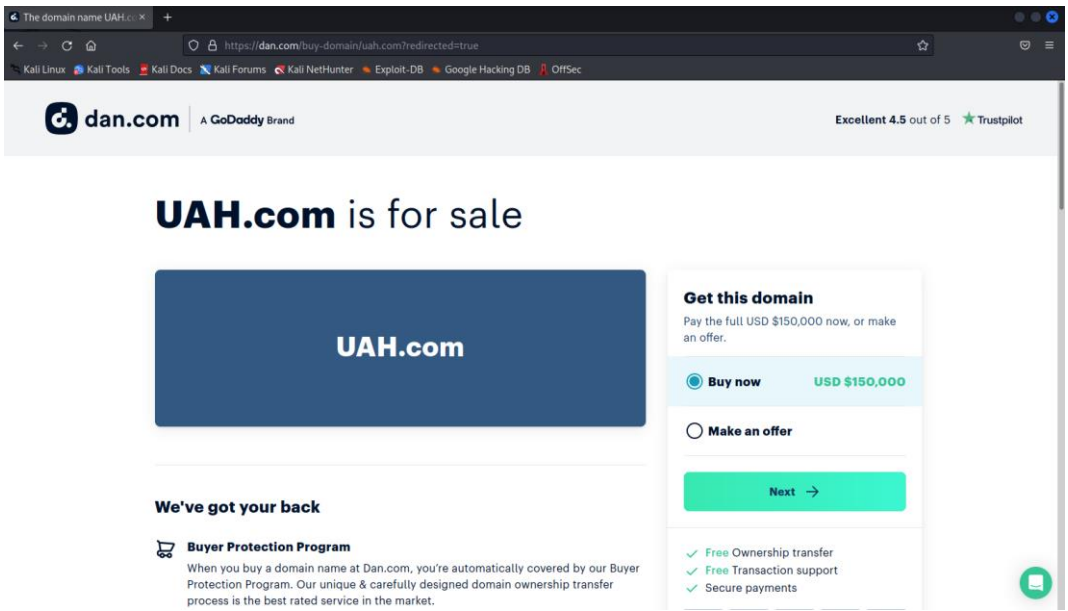


Ilustración 16. Dominio de UAH.com no legítimo, está en venta, no es de la Universidad de Alcalá.

- **UAH.es**: podemos comprobar que es una página legítima, y a su vez, es la página oficial de la Universidad de Alcalá.

Ilustración 17. Dominio UAH.es es legítimo ya que pertenece a la Universidad de Alcalá.

Ahora bien, tras llevar a cabo la búsqueda en Namechk y ver los dominios encontrados en los resultados de la búsqueda, también podemos apreciar los nombres en las redes sociales.

En rojo se encuentran los nombres que se encuentran registrados en cada red social correspondiente, y en verde, los que se encuentran disponibles.

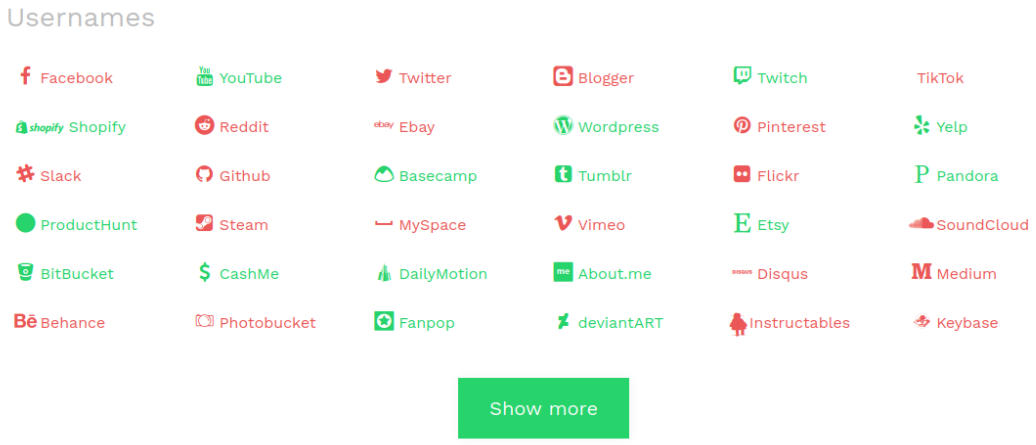


Ilustración 18. Redes sociales relacionadas con la UAH. Rojo nombres registrados, verdes, nombres disponibles.

Podemos apreciar que son varias las redes sociales que tienen ya creado un usuario con “UAH”, no obstante, podemos detectar un problema de esta herramienta en este apartado; se trata, que cuando intentas acceder a la red social requerida, se redirige a la página oficial de la misma y solicita le introducción de credenciales, es decir, no lleva al perfil del usuario buscado.

3.1.5. Conclusiones de Namechk

Tras realizar diferentes investigaciones con la herramienta de Namechk, hemos podido comprobar que tiene como fin encontrar tanto dominios como redes sociales registradas a partir del objetivo dado por el usuario.

No obstante, cabe destacar el fallo respecto a los resultados dentro de las redes sociales, ya que no te redirigen a las cuentas de estas. Por otro lado, en los resultados de los dominios, muestra todo de forma clara y redirige correctamente al resultado dado.

Por último, resaltar la sencillez y usabilidad que ofrece Namechk, que con únicamente el navegador y acceso a internet podemos llevar a cabo diferentes búsquedas sobre investigaciones de nuestros objetivos.

3.2. Sherlock: búsqueda de nombres de usuarios en redes sociales

Sherlock se trata de una herramienta de código abierto desarrolla en Python que tiene como objetivo la búsqueda activa de un nombre de usuario en las principales redes sociales. Además, esta aporta el hipervínculo pertinente por cada red social en la que se encuentra a la víctima.

3.2.1. Instrumentos empleados

Entorno virtual: **VMWare**
Máquina atacante: **Kali Linux**
Herramientas:

- **Sherlock: https://github.com/sherlock-project/sherlock**
- **Redes sociales:**
  - **About (https://about.me/).**
  - **Ask (https://ask.fm/).**
  - **Blogger (https://www.blogger.com/).**
  - **Coil (https://www.coil.com/).**
  - **Disqus (https://disqus.com/).**
  - **Docker (https://hub.docker.com/).**
  - **Flickr (https://www.flickr.com/).**
  - **G2G (https://www.g2g.com/).**
  - **Instagram (https://www.instagram.com/).**
  - **Periscope (https://www.periscope.tv/).**
  - **Roblox (https://www.roblox.com/).**
  - **Scribd (https://es.scribd.com/).**
  - **Slideshare (https://www.slideshare.net/).**
  - **Snapchat (https://www.snapchat.com/).**

- **Stream Community (https://steamcommunity.com/).**
- **TikTok (https://www.tiktok.com/).**
- **Twitch (https://www.twitch.tv/).**
- **Twitter (https://twitter.com/).**
- **Vimeo (https://vimeo.com/).**
- **YouTube (https://www.youtube.com/).**

3.2.2. Puesta en escena de Sherlock

Para el despliegue de la herramienta **Sherlock**, hemos optado por lanzarla en entorno virtual. También se han hecho uso de las redes sociales: *About, Ask, Blogger, Coil, Disqus, Docker, Flickr, G2G, Instagram, Periscope, Roblox, Scribd, Slideshare, Snapchat, Stream Community, TikTok, Twitch, Twitter, Vimeo y YouTube*; todo ello se encuentra de forma online en el navegador.

Cabe destacar que Sherlock tiene una característica que se trata de que puede lanzar uno o más nombres de usuario para consultar con las redes sociales, siendo esta forma una manera rápida y efectiva.

En esta puesta en escena se llevarán a cabo diferentes tipos de búsquedas: **búsqueda a un objetivo y búsqueda a objetivos anidados.**

3.2.3. Sherlock: el entorno

Nos descargamos la herramienta de Sherlock, la lanzamos y vemos en la shell información importante; donde nos muestra las opciones disponibles dentro de esta.

El comando de instalación es el siguiente: **sudo apt install sherlock**.

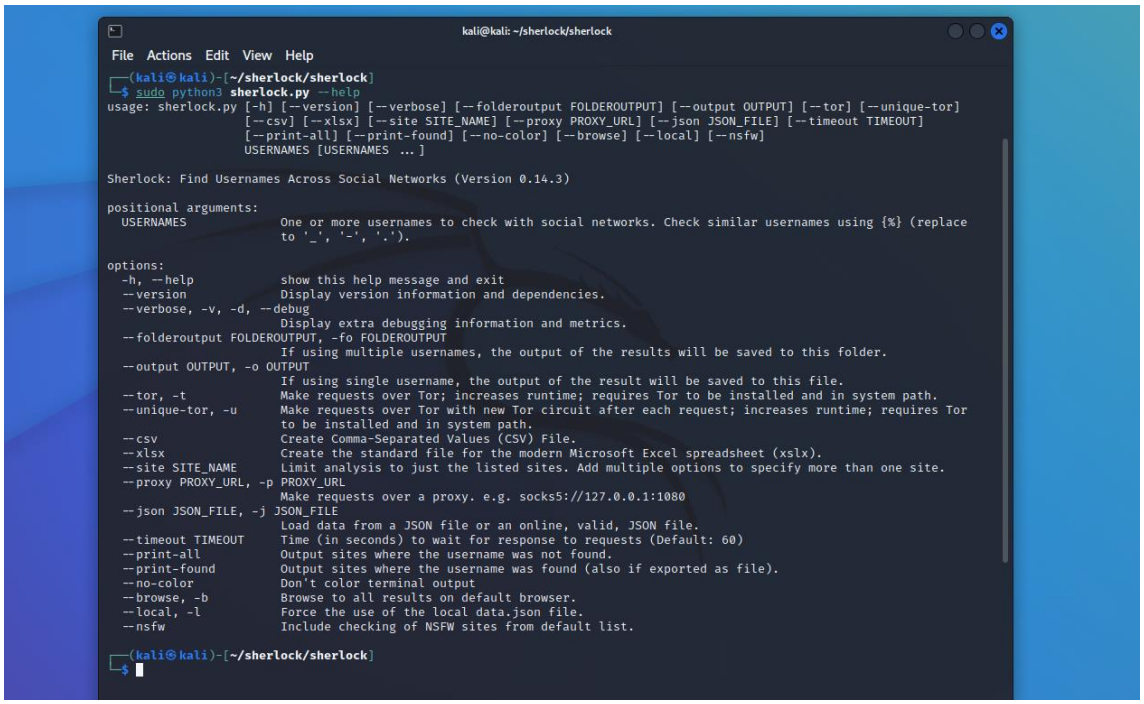


Ilustración 19. Opciones disponibles dentro de la herramienta Sherlock

Como podemos observar, hay diferentes opciones dentro de Sherlock, nosotros iremos al objetivo que queremos, encontrar las redes sociales del nombre de usuario determinado como objetivo.

3.2.4. Búsqueda de un objetivo en redes sociales

Llevaremos a cabo la búsqueda del nombre de nuestra víctima por medio de la herramienta. Fijamos como objetivo a la Universidad de Alcalá, nuestro objetivo principal en la instigación. Por ello establecemos como nombre de usuario de búsqueda en **Namechk: búsqueda de dominios y usuarios en redes sociales**, y realizamos la búsqueda.

Llevaremos a cabo la búsqueda de un nombre de usuario de nuestra víctima por medio de la herramienta. Fijamos como objetivo a la Universidad de Alcalá, nuestro objetivo principal en la investigación.

A su vez, como comprobamos anteriormente tras encontrar la web oficial de la Universidad, el nombre de usuario de esta en redes sociales es @uahes; es por ello por lo que establecemos como nombre de usuario de búsqueda en Sherlock uahes, y realizamos la búsqueda.

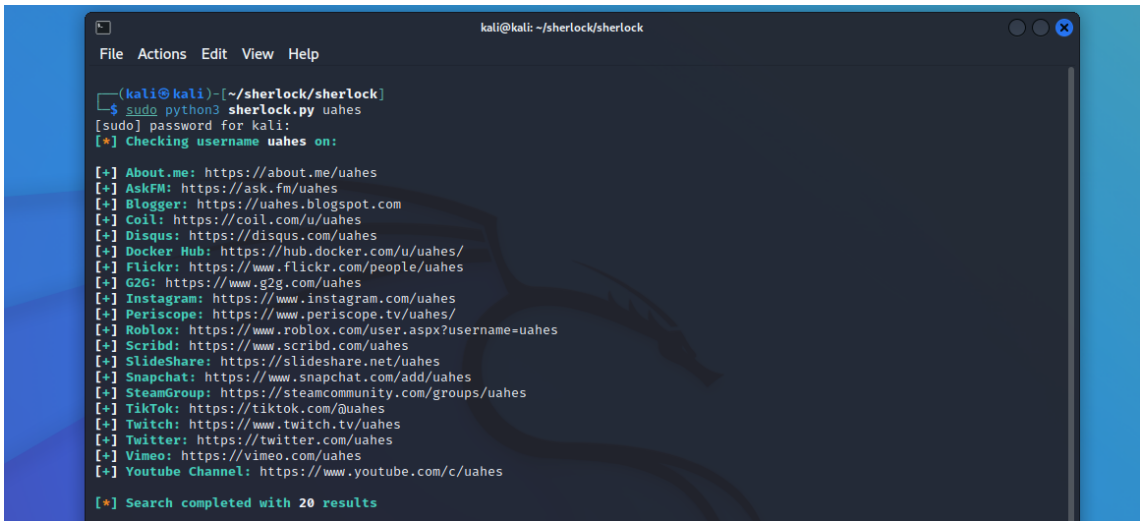


Ilustración 20. Búsqueda del usuario uahes en Sherlock, obtenemos un total de 20 resultados

La herramienta nos proporciona como resultado final un total de 20 resultados en redes sociales con la búsqueda del nombre de usuario de uahes.

Además, se genera por cada búsqueda un archivo .txt en el cual nos almacena los resultados obtenidos.

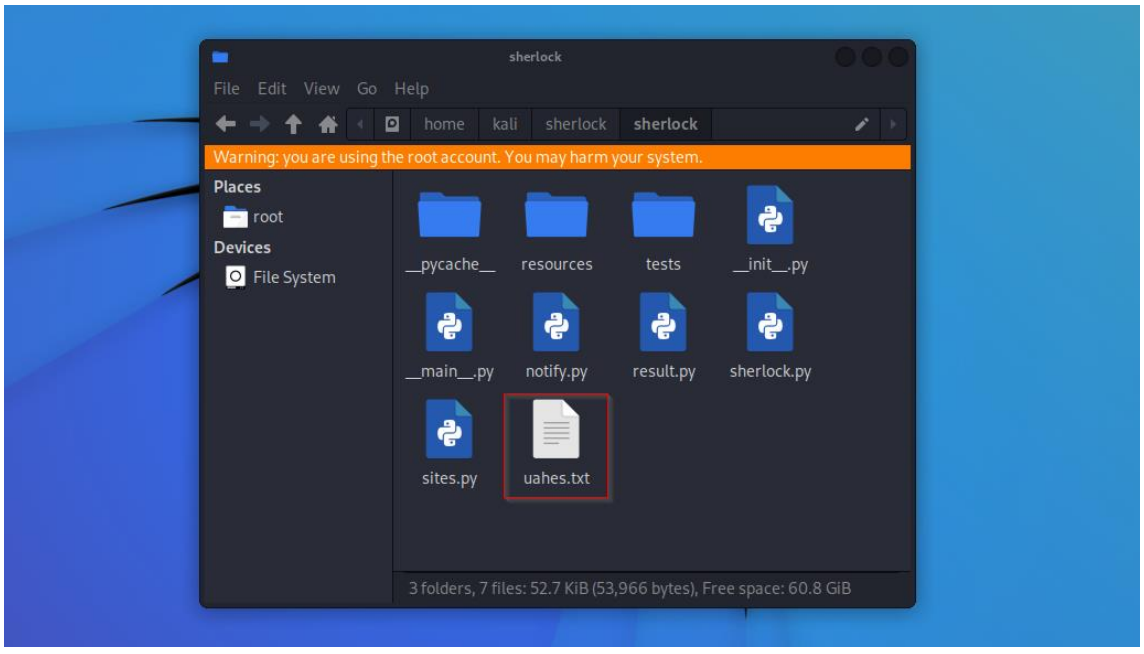


Ilustración 21. Fichero generado con los resultados obtenidos de la búsqueda del usuario uahes en Sherlock

El nombre del fichero varía en función del nombre de usuario de la víctima a averiguar.

Si abrimos el fichero .txt generado: uahes.txt, observamos como la información que este contiene, es la información que se mostró tras realizar la búsqueda por la shell.

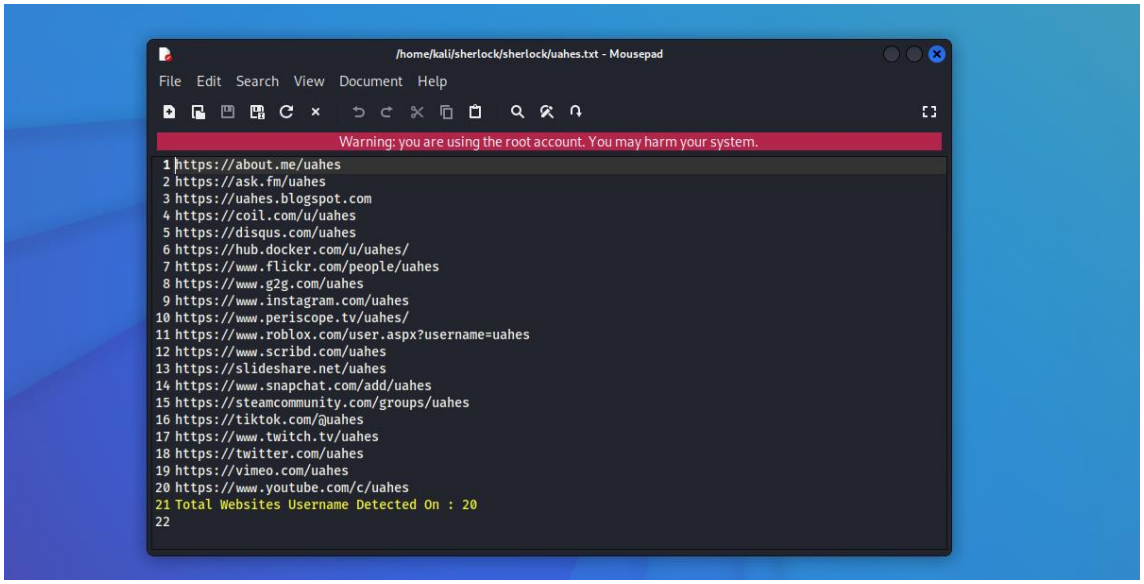


Ilustración 22. Contenido del fichero generado con los resultados obtenidos, un total de 20 redes sociales encontradas con el objetivo de uahes en Sherlock

3.2.5. Búsqueda de varios objetivos en redes sociales

Ahora bien, se realizará la búsqueda de varios nombres de usuarios. Seguimos con el objetivo de nuestra víctima, la Universidad de Alcalá. Por ello realizaremos la búsqueda de dos nombres de usuario de forma simultánea: uahes y uah.

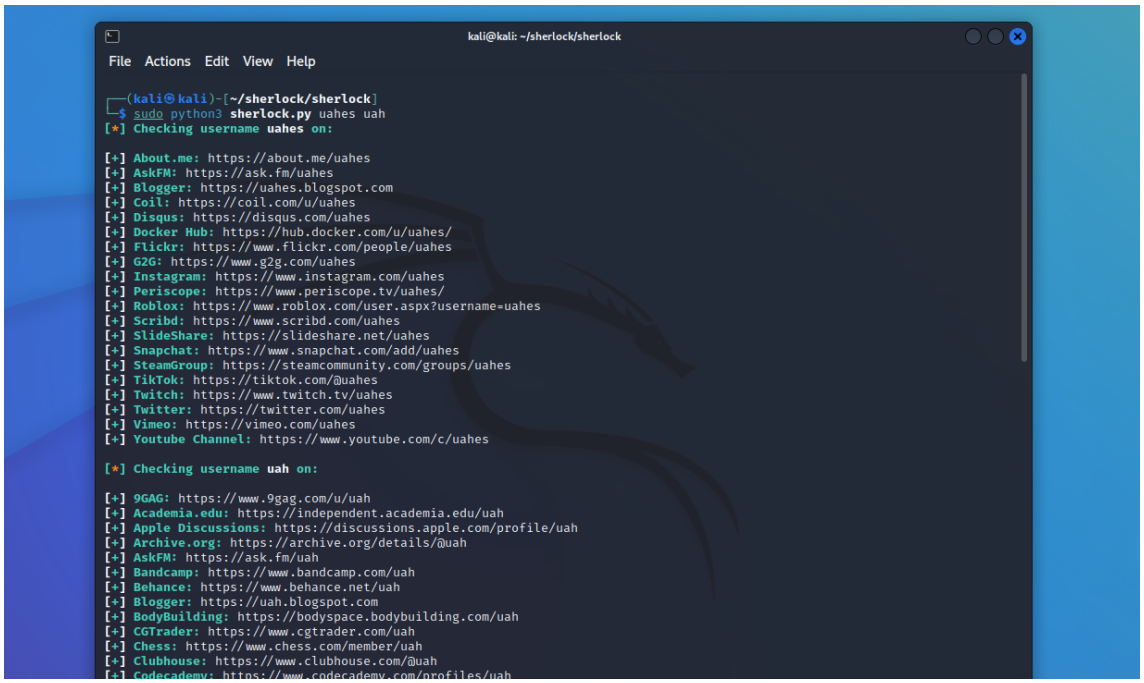


Ilustración 23. Búsqueda de los usuarios uahes y uah en Sherlock, de forma síncrona

Como podemos apreciar, la búsqueda de objetivos de forma simultánea se realiza de forma similar y pareja a la del objetivo individual, únicamente hay que añadir a continuación los otros nombres de usuarios a analizar.

En este caso, se generan un total de dos archivos .txt, el número de estos varía en función de los objetivos establecidos en la investigación; es decir, si tenemos dos nombres se usuarios, se generarán dos archivos, si tenemos tres, entonces tres.

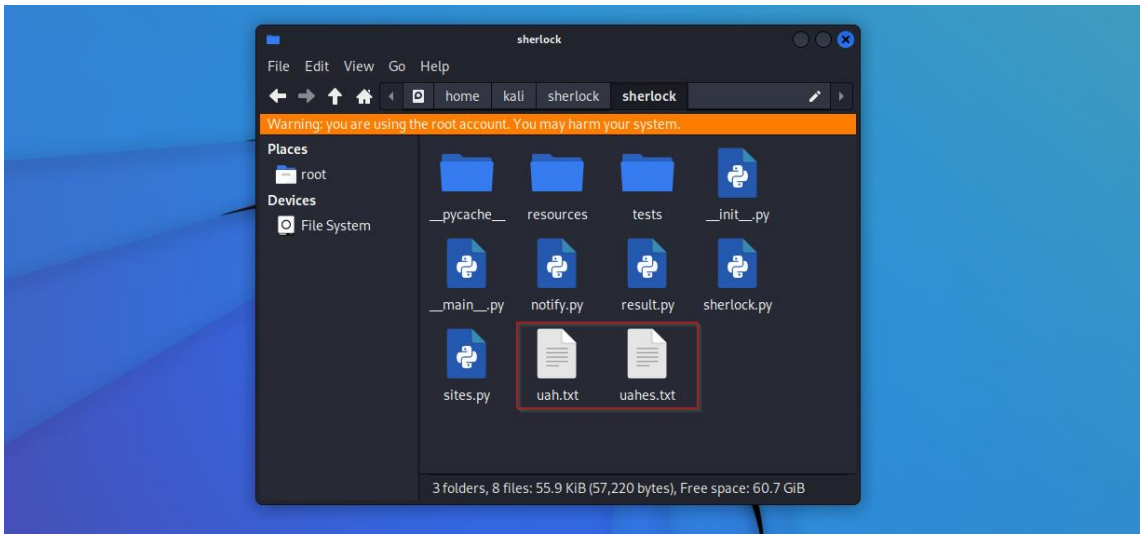


Ilustración 24. Ficheros generados con los resultados obtenidos de la búsqueda del usuario uahes y uah en Sherlock por separado

Si abrimos el fichero .txt generado: uahes.txt (es el mismo que en el de búsqueda de un nombre de usuario) y uah.txt, observamos como la información que este contiene, es la información que se mostró tras realizar la búsqueda por la shell.

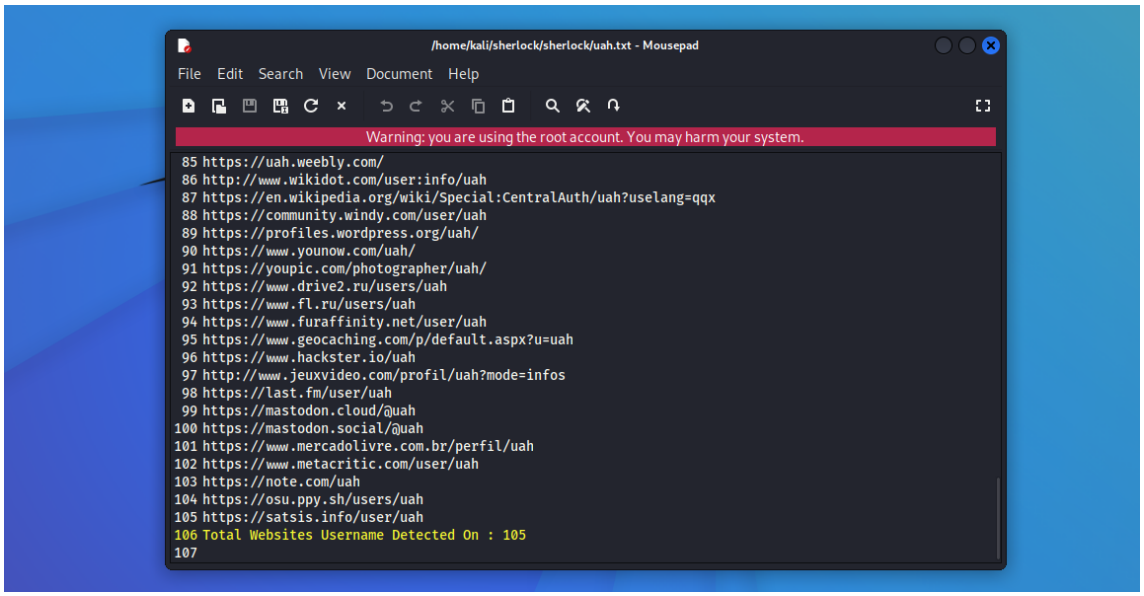


Ilustración 25. Contenido del fichero generado con los resultados obtenidos, un total de 105 redes sociales encontradas con el objetivo de uah en Sherlock

3.2.6. Análisis resultados obtenidos sobre uahes

Procedemos a validar los resultados obtenidos por medio de Sherlock en las redes sociales, de esta forma podremos identificar si estas son legítimas por parte de la Universidad de Alcalá.

Se han encontrado 20 resultados con el nombre de usuario uahes, de los cuales corresponden a diferentes redes sociales: *About, Ask, Blogger, Coil, Disqus, Docker, Flickr, G2G, Instagram, Periscope, Roblox, Scribd, Slideshare, Snapchat, Stream Community, TikTok, Twitch, Twitter, Vimeo y YouTube.*

Procedemos a llevar a cabo el análisis:

1. **About**, https://about.me/uahes: red social legítima por parte de la Universidad de Alcalá. Redirige al Twitter y Flickr oficial de la Universidad. Encontramos una novedad, el nombre de usuario en Flickr: univ_alcala.

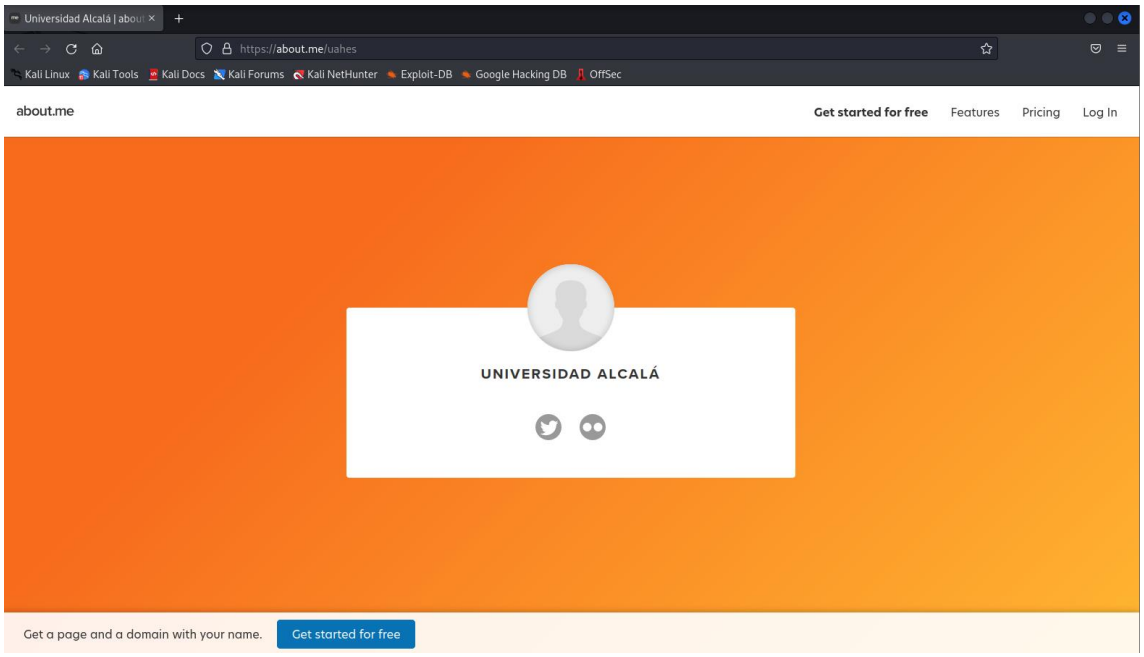


Ilustración 26. Validación del perfil de uahes en About

2. **Ask**, https://ask.fm/uahes: red social no legítima por parte de la Universidad de Alcalá. Se observa que el perfil no es corporativo.

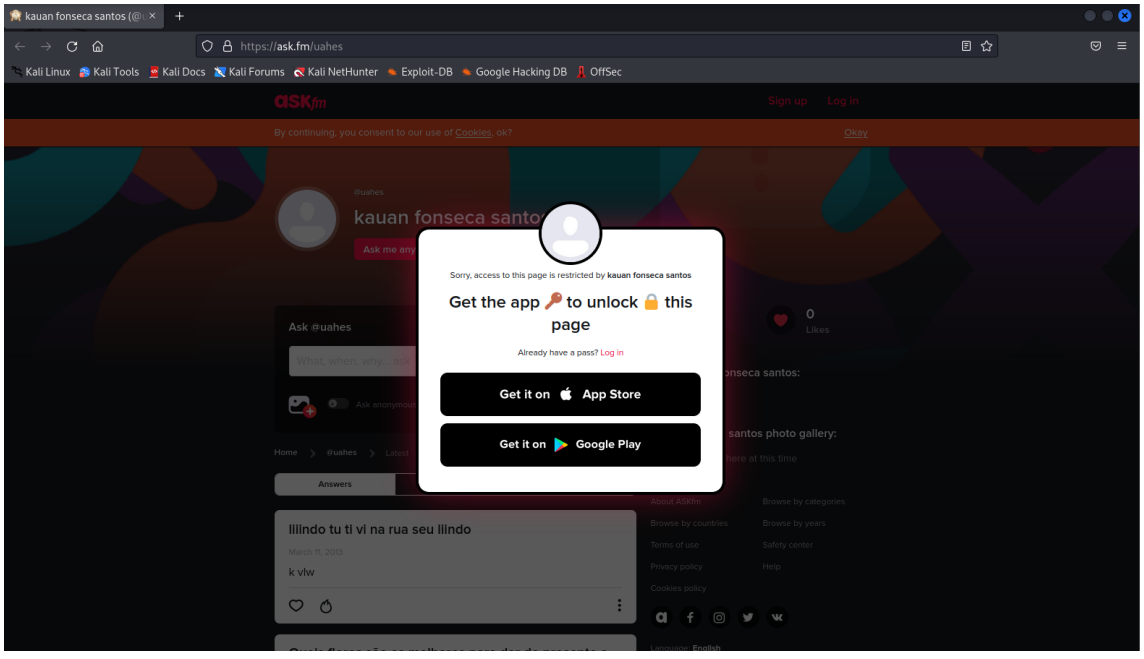


Ilustración 27. Validación del perfil de uahes en Ask

3. **Blogger**, https://uahes.blogspot.com/: red social sin clasificar por parte de la Universidad de Alcalá.

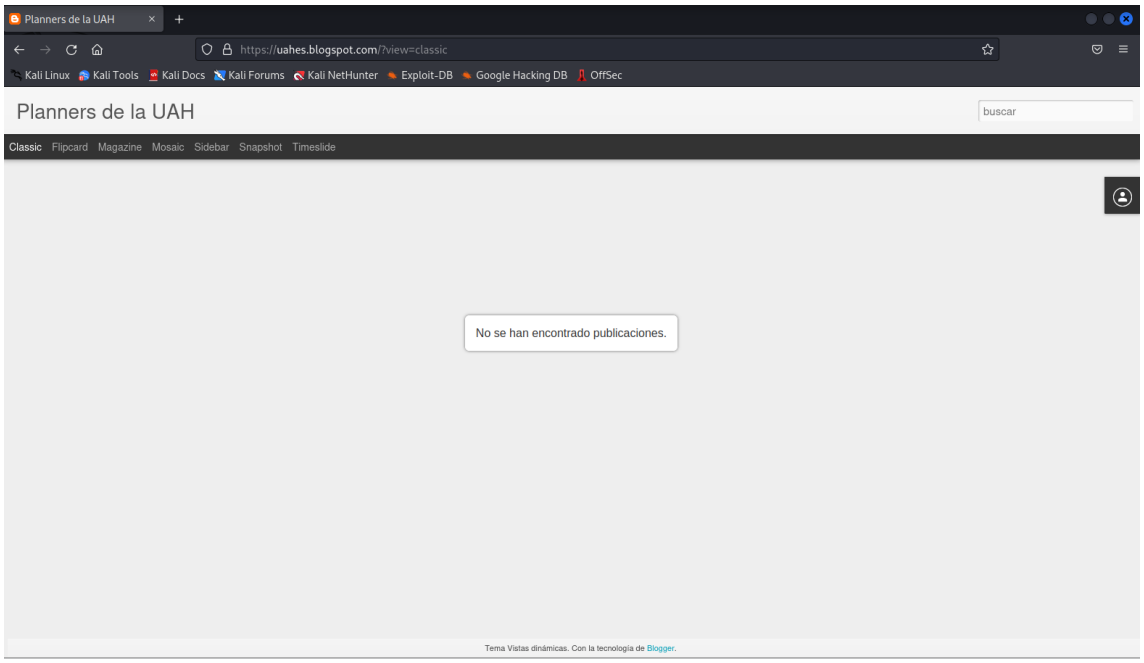


Ilustración 28. Validación del perfil de uahes en Blogger

4. **Coil**, https://coil.com/u/uahes: red social sin clasificar por parte de la Universidad de Alcalá. No existe.

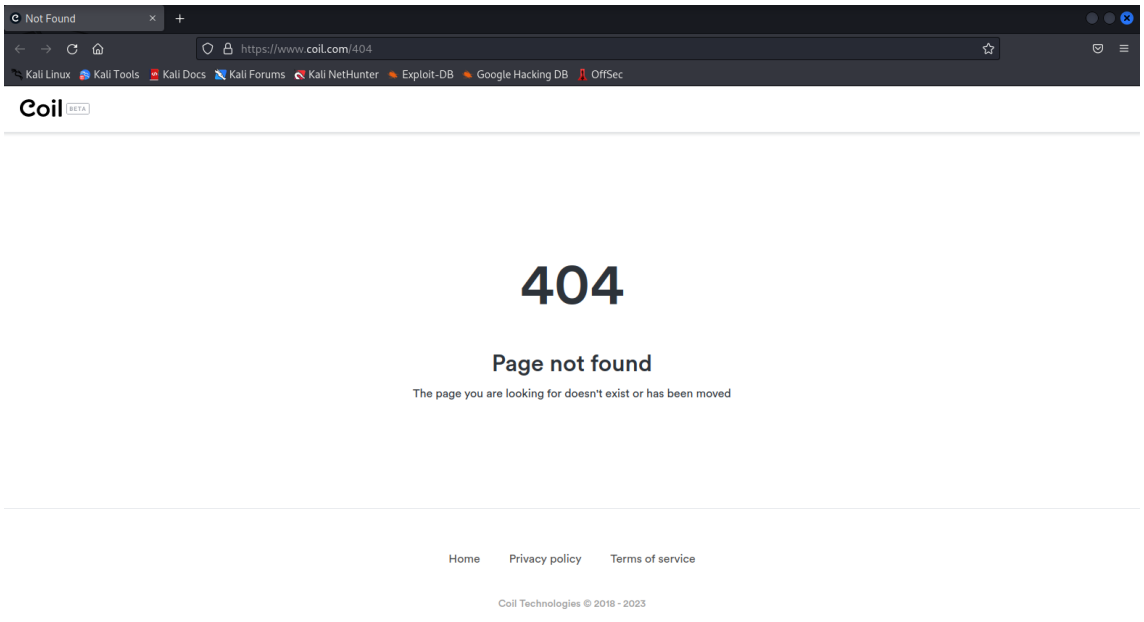


Ilustración 29. Validación del perfil de uahes en Coil

5. **Disqus**, https://disqus.com/by/UAHes/: red social sin clasificar por parte de la Universidad de Alcalá.

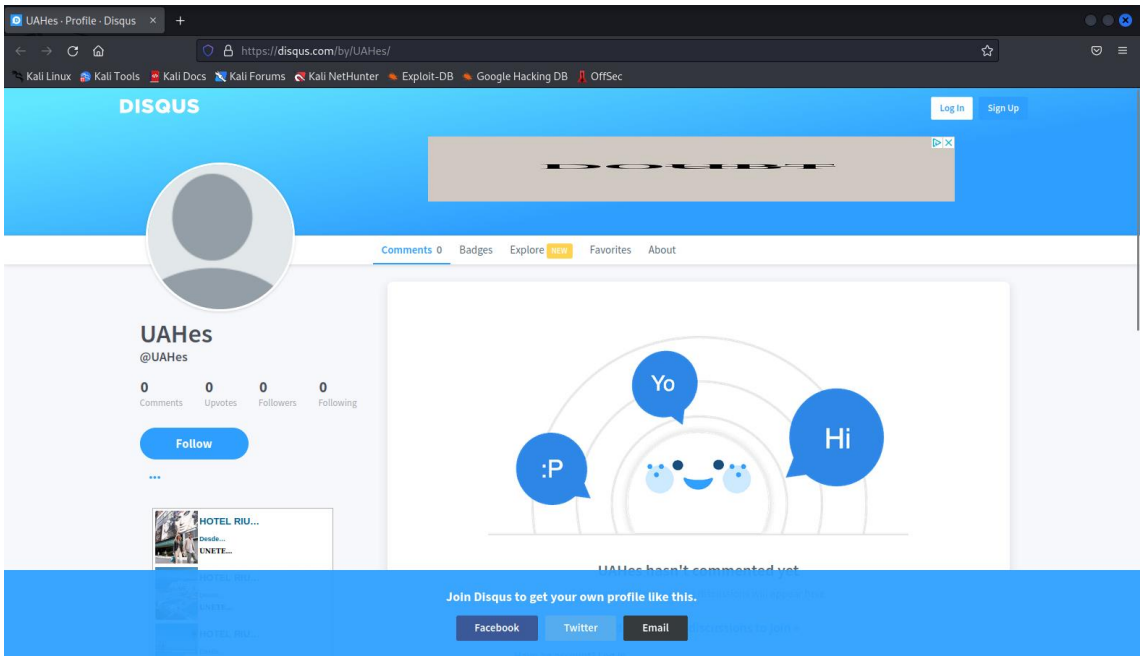


Ilustración 30. Validación del perfil de uahes en Disgus

6. **Docker**, https://hub.docker.com/u/uahes/: red social sin clasificar por parte de la Universidad de Alcalá.

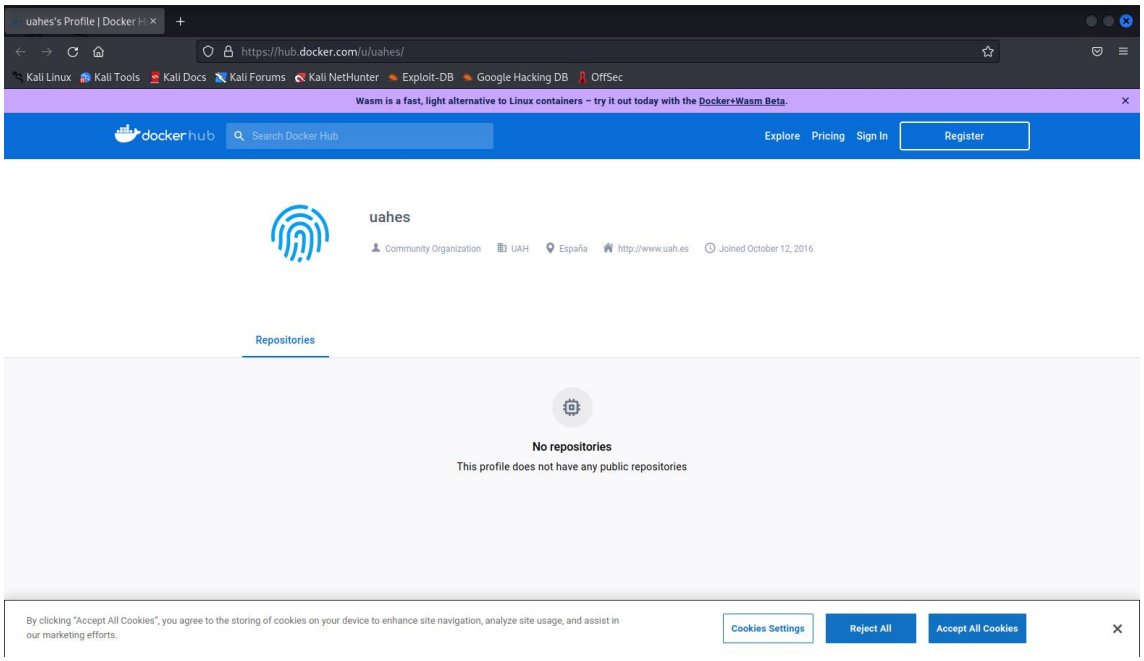


Ilustración 31. Validación del perfil de uahes en Docker

7. **Flickr**, https://www.flickr.com/people/uahes: red social legítima por parte de la Universidad de Alcalá. Actualmente usada por la Universidad.

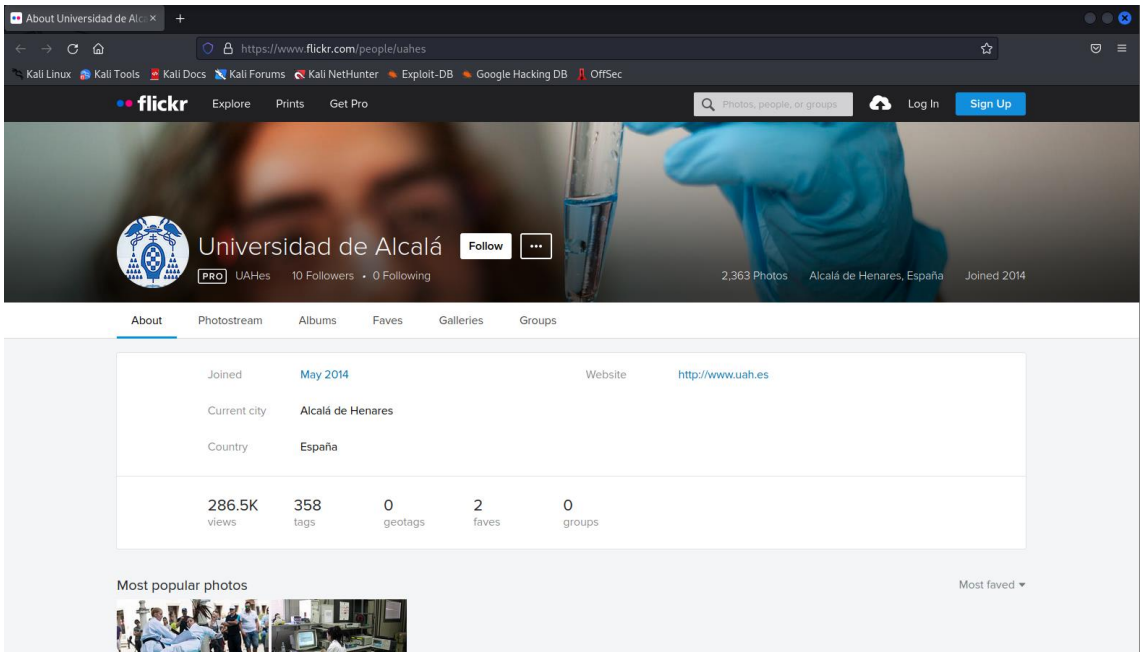


Ilustración 32. Validación del perfil de uahes en Flickr

8. **G2G**, https://www.g2g.com/uahes: red social sin clasificar por parte de la Universidad de Alcalá.

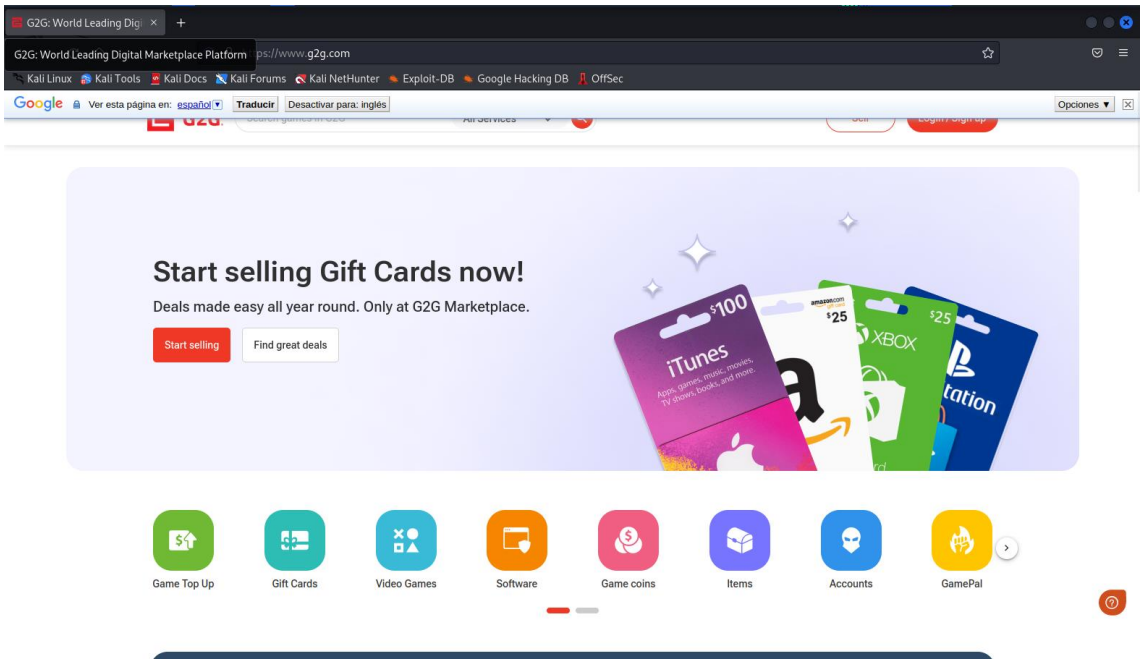


Ilustración 33. Validación del perfil de uahes en G2G

9. **Instagram**, https://www.instagram.com/uahes/: red social legítima por parte de la Universidad de Alcalá. Actualmente usada por la Universidad.

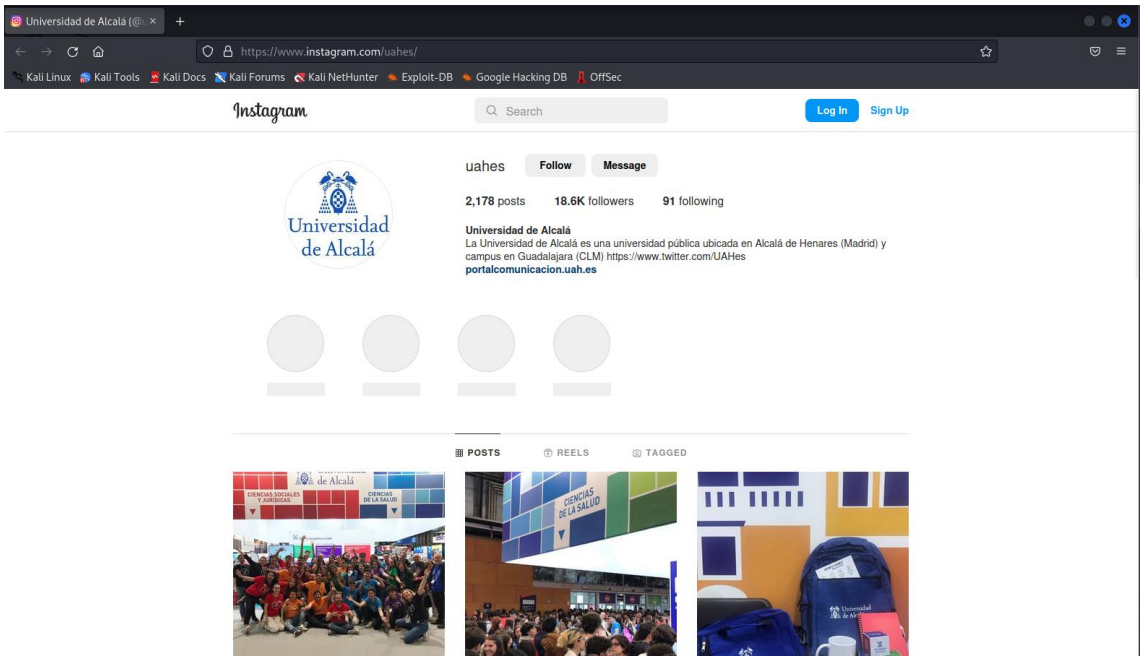


Ilustración 34. Validación del perfil de uahes en Instagram

10. **Periscope**, https://www.periscope.tv/uahes/: red social legítima por parte de la Universidad de Alcalá. Último uso por la Universidad hace tres años.

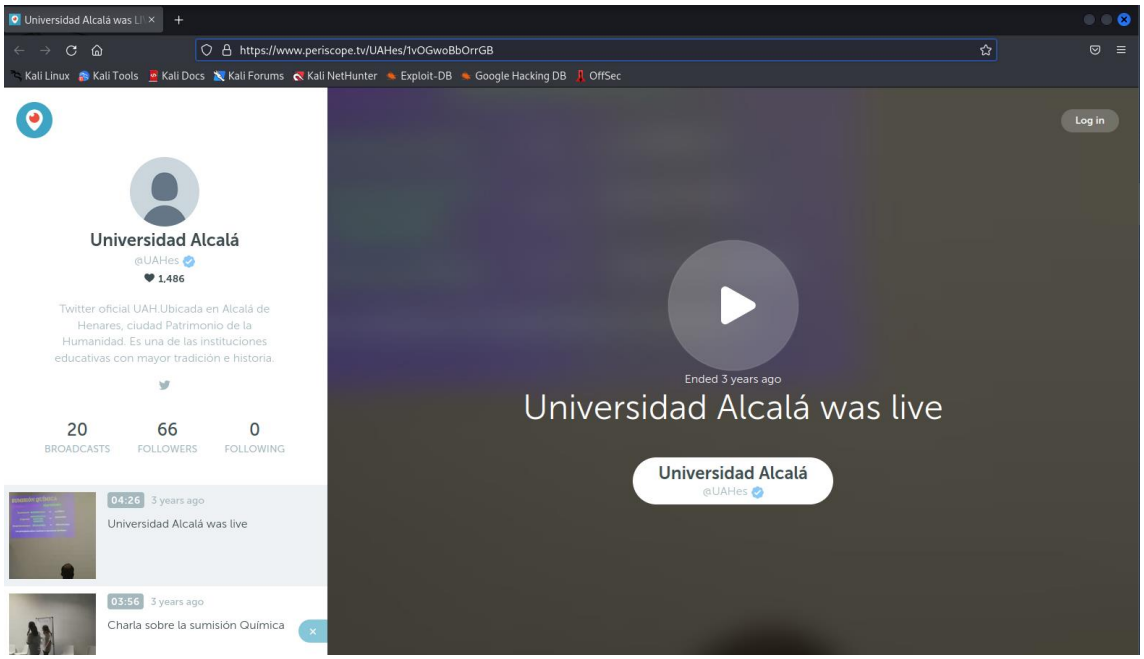


Ilustración 35. Validación del perfil de uahes en Periscope

11. **Roblox**, https://www.roblox.com/user.aspx?username=uahes: red social no legítima por parte de la Universidad de Alcalá. Se observa que el perfil no es corporativo.

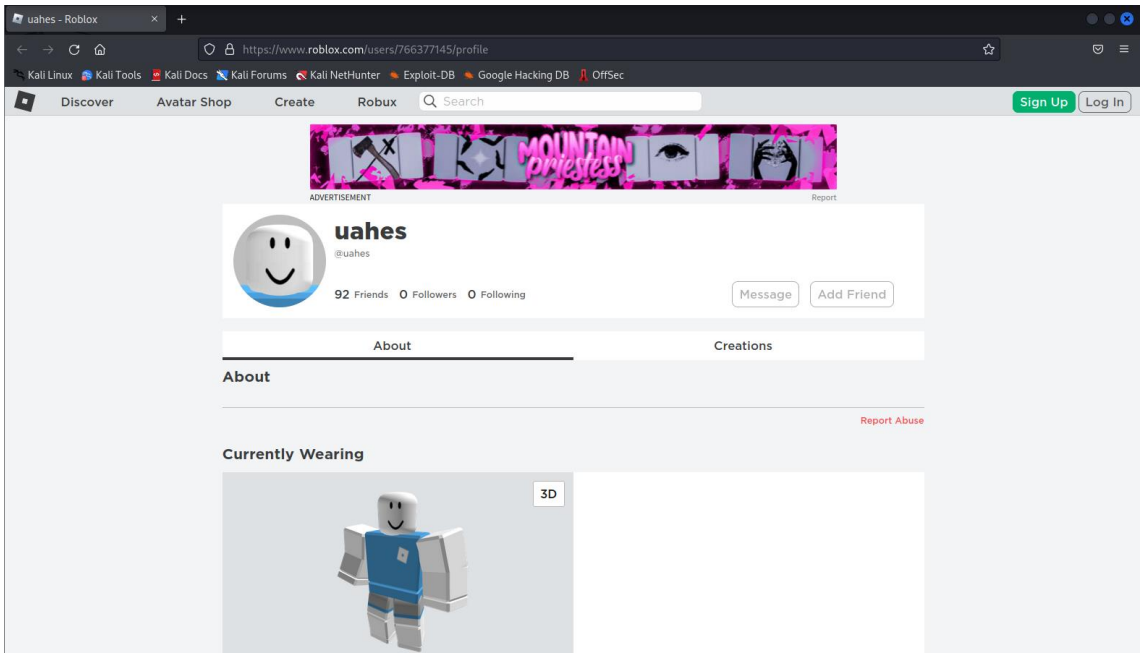


Ilustración 36. Validación del perfil de uahes en Roblox

12. **Scribd**, https://www.scribd.com/uahes: red social legítima por parte de la Universidad de Alcalá.

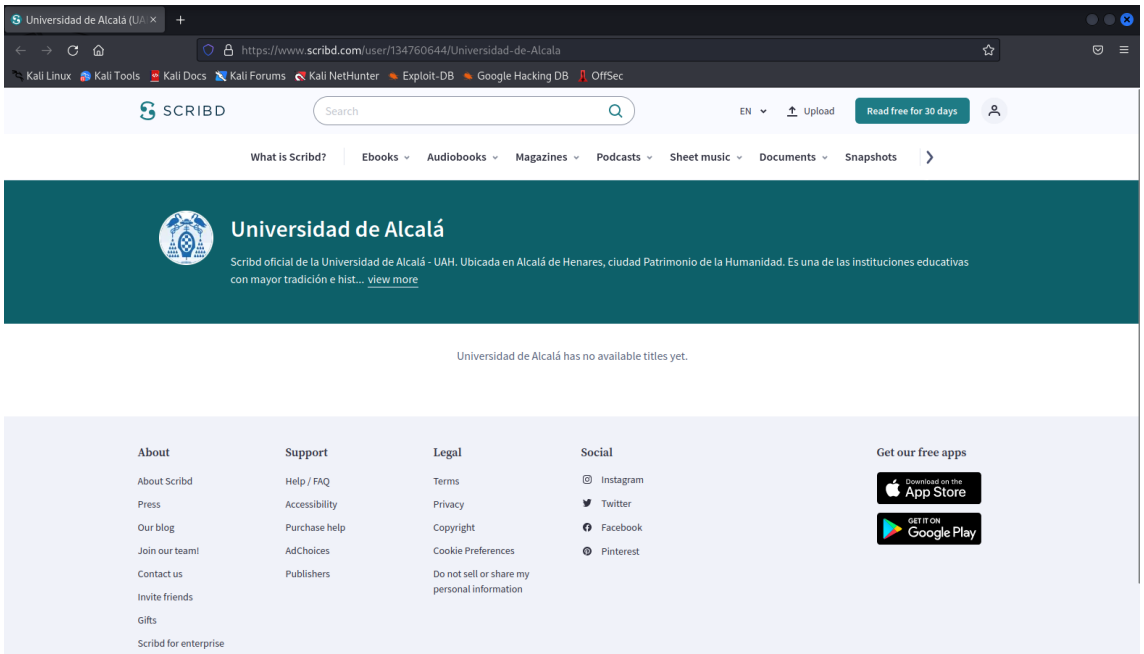


Ilustración 37. Validación del perfil de uahes en Scribd

13. **Slideshare**, https://www.slideshare.net/uahes: red social legítima por parte de la Universidad de Alcalá. Último uso por la Universidad hace diez años.

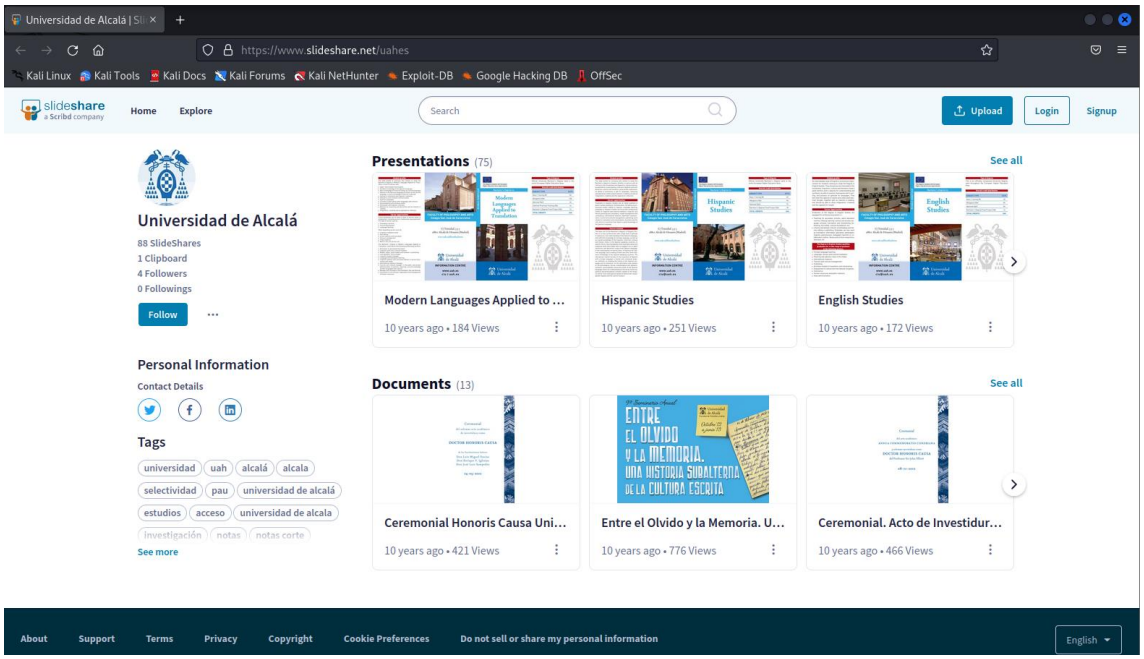


Ilustración 38. Validación del perfil de uahes en Slideshare

14. **Snapchat**, https://www.snapchat.com/add/uahes: red social no legítima por parte de la Universidad de Alcalá. Se observa que el perfil no es corporativo.

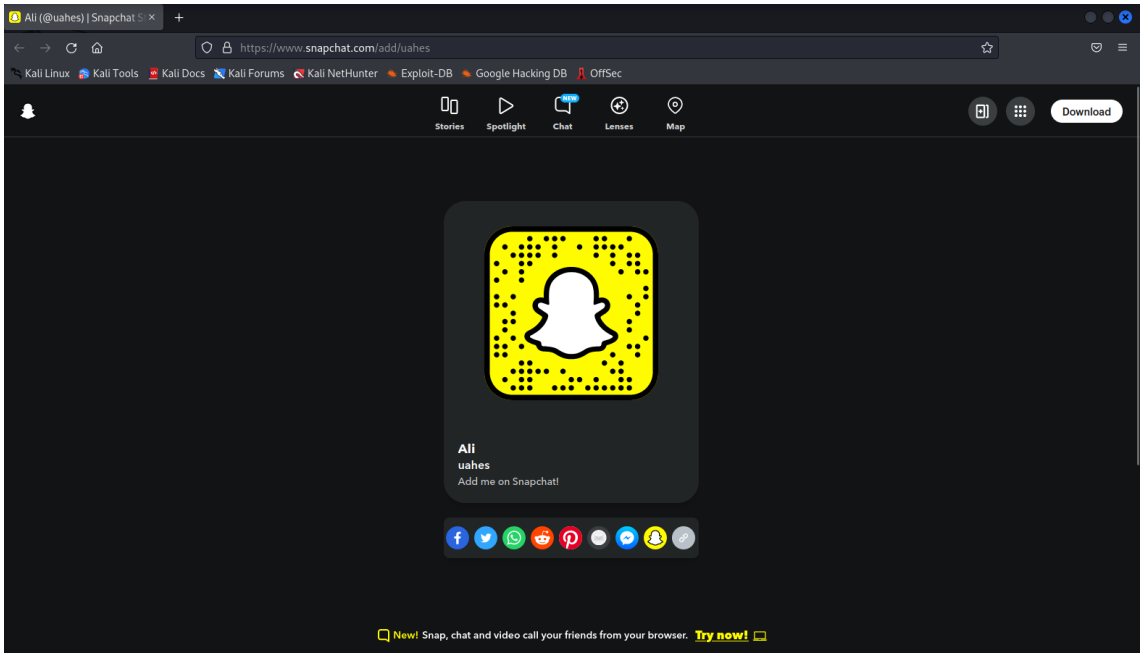


Ilustración 39. Validación del perfil de uahes en Snapchat

15. **Steam Community**, https://steamcommunity.com/groups/uahes: red social no legítima por parte de la Universidad de Alcalá. Se observa que el perfil no es corporativo.

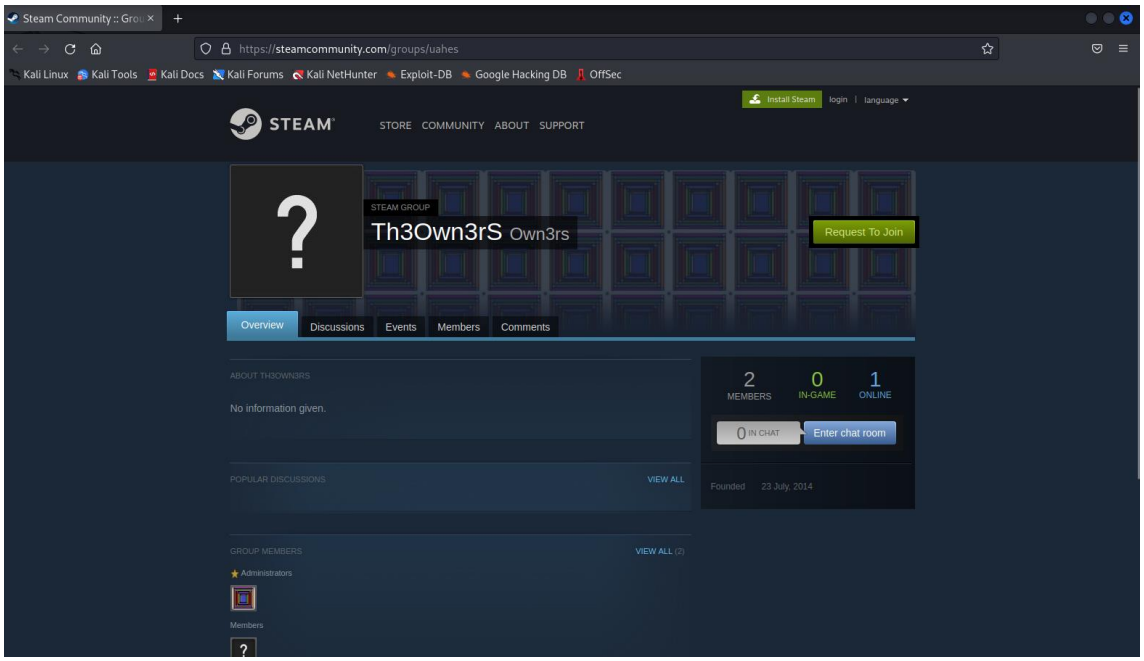


Ilustración 40. Validación del perfil de uahes en Steam Community

16. **TikTok**, https://www.tiktok.com/@uahes: red social no legítima por parte de la Universidad de Alcalá. Se observa que el perfil no es corporativo.

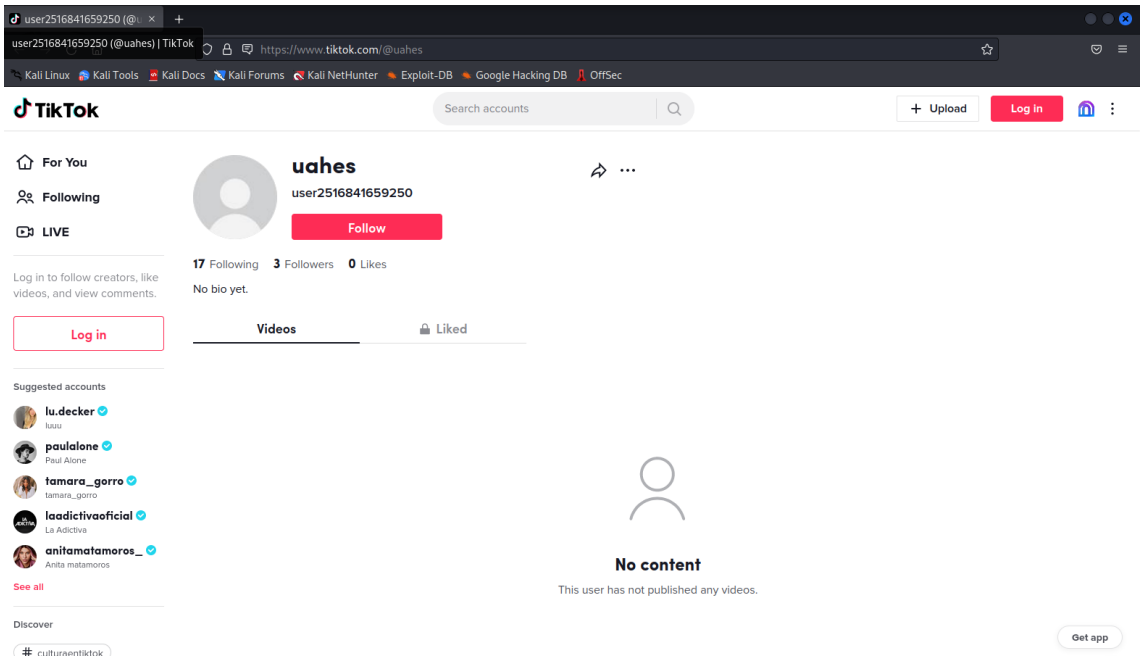


Ilustración 41. Validación del perfil de uahes en TikTok

17. **Twitch**, https://www.twitch.tv/uahes: red social sin clasificar por parte de la Universidad de Alcalá.

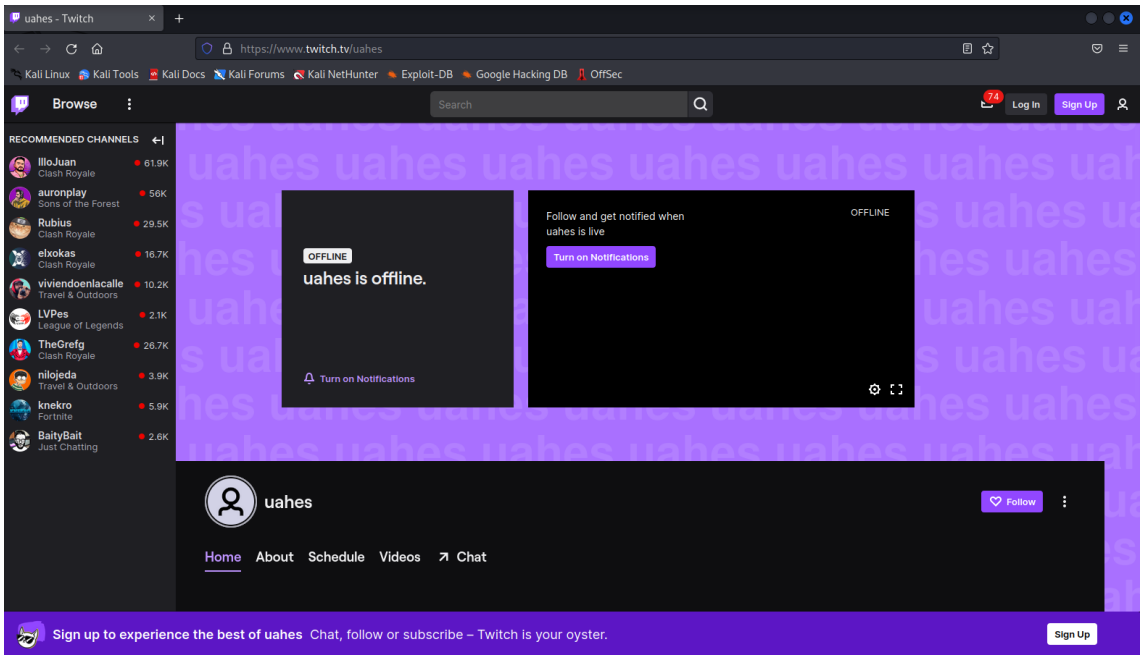


Ilustración 42. Validación del perfil de uahes en Twitch

18. **Twitter**, https://twitter.com/uahes: red social legítima por parte de la Universidad de Alcalá. Actualmente usada por la Universidad.

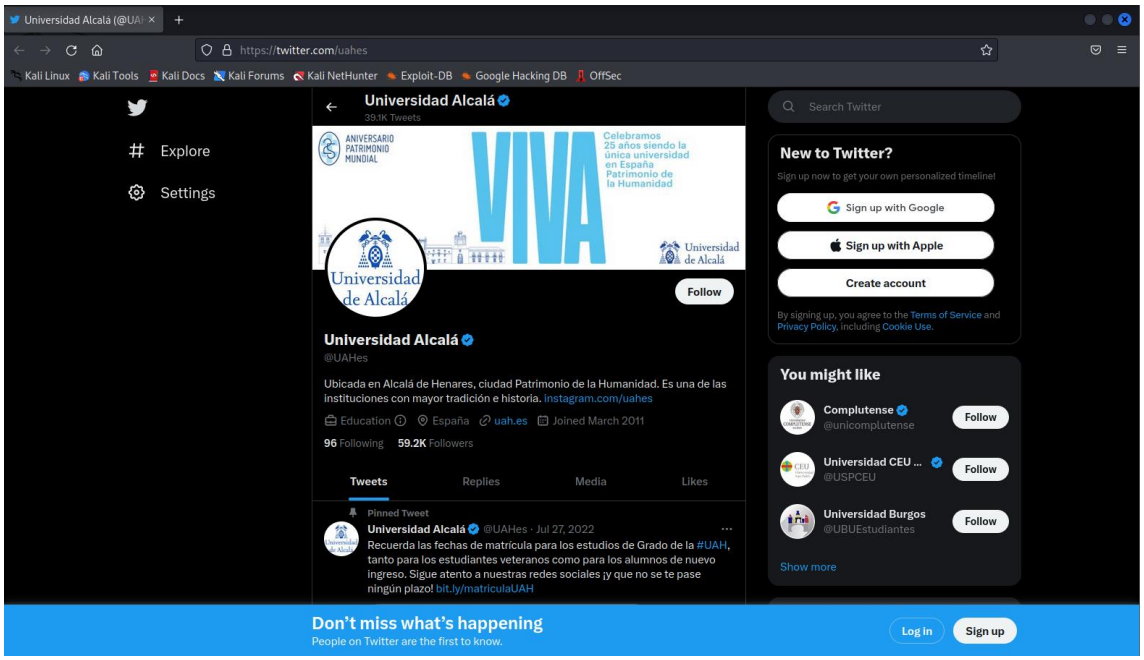


Ilustración 43. Validación del perfil de uahes en Twitter

19. **Vimeo**, https://vimeo.com/uahes: red social no legítima por parte de la Universidad de Alcalá. Se trata de la Universidad de Alabama en EEUU.

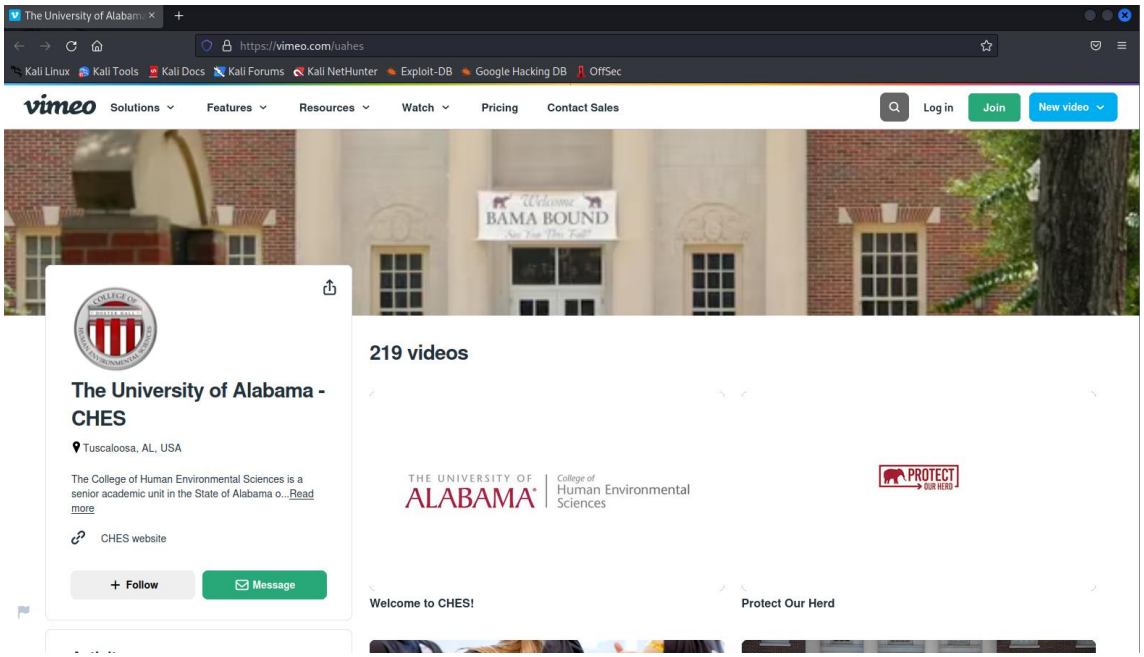


Ilustración 44. Validación del perfil de uahes en Vimeo

20. **YouTube**, https://www.youtube.com/c/uahes: red social legítima por parte de la Universidad de Alcalá. Actualmente usada por la Universidad.

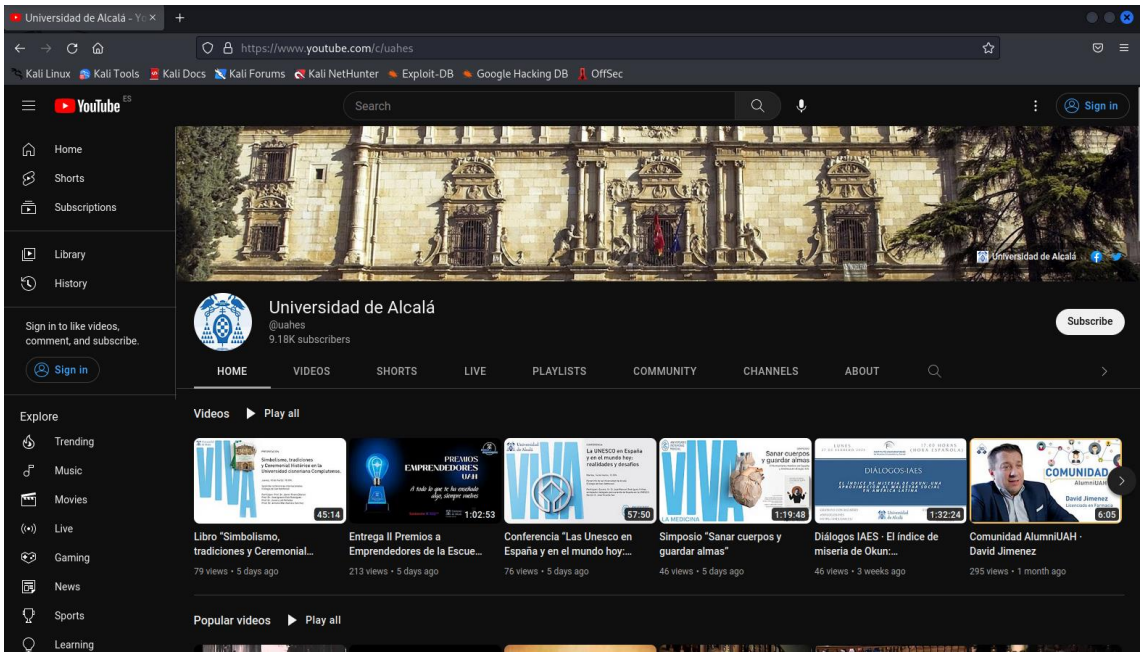


Ilustración 45. Validación del perfil de uahes en YouTube

3.2.6.1. Tabla comparativa de redes sociales detectadas

Red social	Legitimidad	Hipervínculo

(uahes)		
About	Legítimo	https://about.me/uahes
Ask	No legítimo	https://ask.fm/uahes
Blogger	Sin clasificar	https://uahes.blogspot.com/
Coil	None	https://coil.com/u/uahes
Disqus	Sin clasificar	https://disqus.com/by/UAHes/
Docker	Sin clasificar	https://hub.docker.com/u/uahes/
Flickr	Legítimo	https://www.flickr.com/people/uahes
G2G	Sin clasificar	https://www.g2g.com/uahes
Instagram	Legítimo	https://www.instagram.com/uahes/
Periscope	Legítimo	https://www.periscope.tv/uahes/
Roblox	No legítimo	https://www.roblox.com/user.aspx?username=uahes
Scribd	Legítimo	https://www.scribd.com/uahes
Slideshare	Legítimo	https://www.slideshare.net/uahes
Snapchat	No legítimo	https://www.snapchat.com/add/uahes
Steam Community	No legítimo	https://steamcommunity.com/groups/uahes
TikTok	No legítimo	https://www.tiktok.com/@uahes
Twitch	Sin clasificar	https://www.twitch.tv/uahes
Twitter	Legítimo	https://twitter.com/uahes
Vimeo	No legítimo	https://vimeo.com/uahes
YouTube	Legítimo	https://www.youtube.com/c/uahes

De las 20 redes sociales con resultados con el objetivo de uahes, **únicamente se encuentran legítimas por la Universidad de Alcalá 8 de ellas**, **5 no se pueden clasificar, 1 es un falso positivo, y 6 son no legítimas.**

3.2.7. Conclusiones de Sherlock

Tras realizar diferentes investigaciones con la herramienta de Sherlock, nos hemos dado cuenta de lo expuestos que estamos los usuarios en la red; con únicamente saber el nombre de usuario del objetivo que queremos buscar, podemos encontrar todas sus redes sociales.

A su vez, en el caso de tener posibles dudas de cómo se llama nuestra víctima en redes, podemos hacer una búsqueda anidada añadiendo diferentes posibilidades de nombres de usuario y Sherlock, realizará la búsqueda de resultados en las redes.

Por último, cabe destacar el fichero .txt de soporte que nos ofrece la herramienta por cada nombre de usuario investigado, de esta forma podemos almacenar todas las posibilidades de forma más sencilla.

3.3. theHarvester: recopilación de dominios y cuentas de correo

theHarvester se trata de una herramienta que recopila cuentas de correo y dominios utilizando diferentes fuentes abiertas como Bing, Yahoo, VirusTotal... Puede recopilar información sobre una organización o un individuo.

3.3.1. Instrumentos empleados

Entorno virtual: **VMWare**
Máquina atacante: **Kali Linux**
Herramientas:

- **theHarvester: https://github.com/laramies/theHarvester**

3.3.2. Puesta en escena de theHarvester

Para el despliegue de la herramienta **theHarvester**, hemos optado por lanzarla en entorno virtual.

Cabe destacar que theHarvester tiene una característica que se trata de que puede lanzar la búsqueda a través de diferentes fuentes: *anubis, baidu, bevigil, binaryedge, bing, bingapi, bufferoverun, censys, certspotter, crtsh, dnsdumpster, duckduckgo, fullhunt, github-code, hackertarget, hunter, intelx, omnisint, otx, pentesttools, projectdiscovery, qwant, rapiddns, rocketreach, securityTrails, sublist3r, threatcrowd, threatminer, urlscan, virustotal, yahoo, zoomeye*.

En esta puesta en escena se llevará a cabo la **búsqueda a un objetivo.**

3.3.3. theHarvester: el entorno

Nos descargamos la herramienta de theHarvester, la lanzamos y vemos en la shell información importante; donde nos muestra las opciones disponibles dentro de esta.

El comando de instalación es el siguiente: **sudo apt install theharvester**

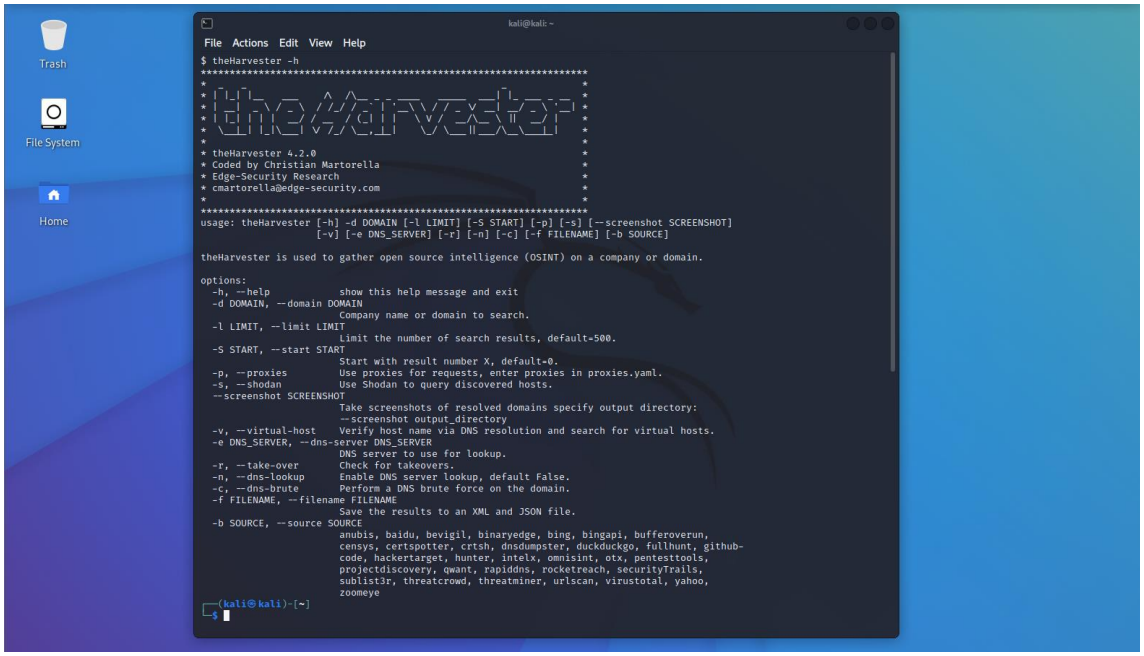


Ilustración 46. Opciones disponibles dentro de la herramienta theHarvester

Como podemos observar, hay diferentes opciones dentro de theHarvester, nosotros iremos al objetivo que queremos, y generaremos un fichero XML y JSON con los resultados de la búsqueda.

3.3.4. Búsqueda de correos electrónicos y dominios de un objetivo

Llevaremos a cabo la búsqueda de nuestro objetivo por medio de la herramienta. Fijamos como objetivo a la Universidad de Alcalá, nuestro objetivo principal en la instigación.

A su vez, como comprobamos con **Sherlock: búsqueda de nombres de usuarios en redes sociales,** sabemos que el dominio de la Universidad de Alcalá es uah.es; es por ello por lo que establecemos como dominio de búsqueda en theHarvester uah.es.

Resaltar también que limitamos la búsqueda a 500 resultados, que es el máximo por defecto; fijamos de buscador Bing, que al ser uno de los buscadores más empleados es donde más información vamos a encontrar; añadimos el nombre del fichero que queremos que se nos genere; y realizamos la búsqueda.

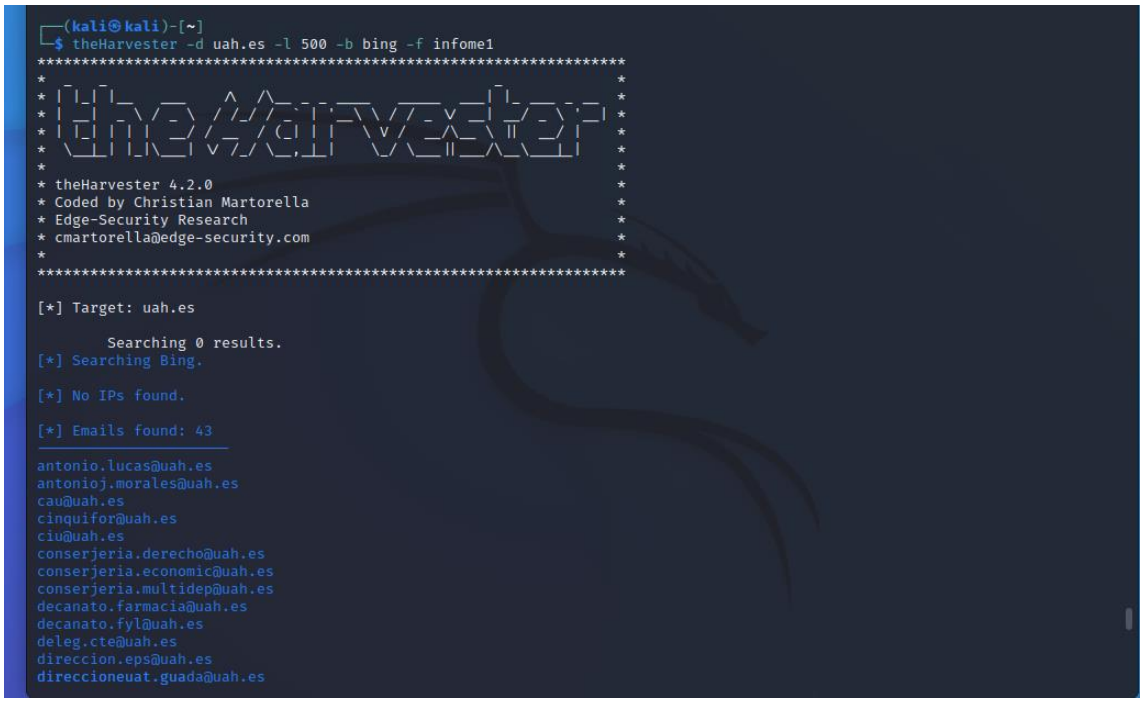


Ilustración 47. Búsqueda del dominio uah.es en theHarvester, obtenemos un total de 43 resultados

La herramienta nos proporciona como resultado final múltiples correos electrónicos y hosts encontrados como el resultado de la búsqueda del uah.es. Además, se genera por la búsqueda un archivo tanto XML como JSON, con el nombre de informe1, en el cual nos almacena los resultados obtenidos.

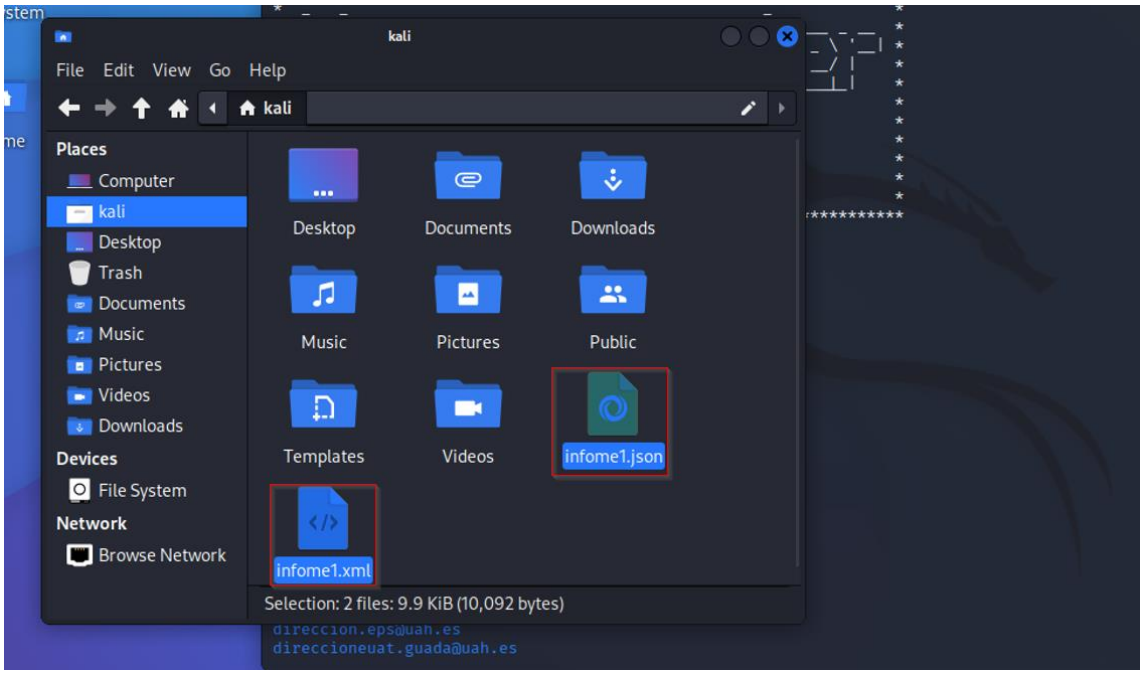


Ilustración 48. Ficheros generados en .json y .xml con los resultados obtenidos de la búsqueda del dominio uah.es en theHarvester

Si abrimos uno de ambos ficheros generados, obtendremos los siguientes resultados:

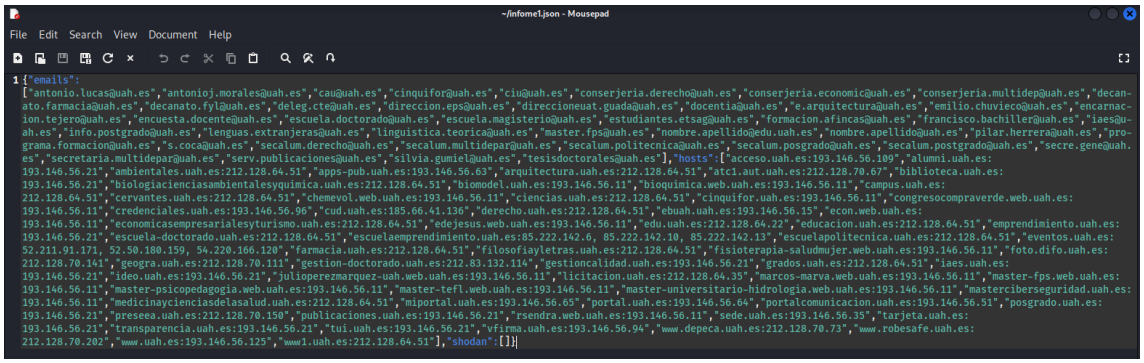


Ilustración 49. Resultados obtenidos de la búsqueda del dominio uah.es en theHarvester: emails y hosts de la Universidad de Alcalá

Vamos a desglosar los resultados obtenidos.

Correos electrónicos encontrados:

["antonio.lucas@uah.es","antonioj.morales@uah.es","cau@uah.es","cinquifor@uah.es","ciu@uah.es","conserjeria.derecho@uah.es","conserjeria.economic@uah.es","conserjeria.multidep@uah.es","decanato.farmacia@uah.es","decanato.fyl@uah.es","deleg.cte@uah.es","direccion.eps@uah.es","direccioneuat.guada@uah.es","docentia@uah.es","e.arquitectura@uah.es","emilio.chuvieco@uah.es","encarnacion.tejero@uah.es","encuesta.docente@uah.es","escuela.doctorado@uah.es","escuela.magisterio@uah.es","estudiantes.etsag@uah.es","formacion.afincas@uah.es","francisco.bachiller@uah.es","iaes@uah.es","info.postgrado@uah.es","lenguas.extranjeras@uah.es","linguistica.teorica@uah.es","master.fps@uah.es","nombre.apellido@edu.uah.es","nombre.apellido@uah.es","pilar.herrera@uah.es","programa.formacion@uah.es","s.coca@uah.es","secalum.derecho@uah.es","secalum.multidepar@uah.es","secalum.politecnica@uah.es","secalum.posgrado@uah.es","secalum.postgrado@uah.es","secre.gene@uah.es","secretaria.multidepar@uah.es","serv.publicaciones@uah.es","silvia.gumiel@uah.es","tesisdoctorales@uah.es"].

Hosts encontrados:

["acceso.uah.es:193.146.56.109","alumni.uah.es:193.146.56.21","ambientales.uah.es:212.128.64.51","apps-pub.uah.es:193.146.56.63","arquitectura.uah.es:212.128.64.51","atc1.aut.uah.es:212.128.70.67","biblioteca.uah.es:193.146.56.21","biologiacienciasambientalesyquimica.uah.es:212.128.64.51","biomodel.uah.es:193.146.56.11","bioquimica.web.uah.es:193.146.56.11","campus.uah.es:212.128.64.51","cervantes.uah.es:212.128.64.51","chemevol.web.uah.es:193.146.56.11","ciencias.uah.es:212.128.64.51","cinquifor.uah.es:193.146.56.11","congresocompraverde.web.uah.es:193.146.56.11","credenciales.uah.es:193.146.56.96","cud.uah.es:185.66.41.136","derecho.uah.es:212.128.64.51","ebuah.uah.es:193.146.56.15","econ.web.uah.es:193.146.56.11","economicasempresarialesyturismo.uah.es:212.128.64.51","edejesus.web.uah.es:193.146.56.11","edu.uah.es:212.128.64.22","educacion.uah.es:212.128.64.51","emprendimiento.uah.es:193.146.56.21","escuela-doctorado.uah.es:212.128.64.51","escuelaemprendimiento.uah.es:85.222.142.6,

85.222.142.10,
85.222.142.13","escuelapolitecnica.uah.es:212.128.64.51","eventos.uah.es:52.211.91.171, 52.50.180.159, 54.220.166.120","farmacia.uah.es:212.128.64.51","filosofiayletras.uah.es:212.128.64.51","fisioterapia-saludmujer.web.uah.es:193.146.56.11","foto.difo.uah.es:212.128.70.141","geogra.uah.es:212.128.70.111","gestion-doctorado.uah.es:212.83.132.114","gestioncalidad.uah.es:193.146.56.21","grados.uah.es:212.128.64.51","iaes.uah.es:193.146.56.21","ideo.uah.es:193.146.56.21","julioperezmarquez-uah.web.uah.es:193.146.56.11","licitacion.uah.es:212.128.64.35","marcos-marva.web.uah.es:193.146.56.11","master-fps.web.uah.es:193.146.56.11","master-psicopedagogia.web.uah.es:193.146.56.11","master-tefl.web.uah.es:193.146.56.11","master-universitario-hidrologia.web.uah.es:193.146.56.11","masterciberseguridad.uah.es:193.146.56.11","medicinaycienciasdelasalud.uah.es:212.128.64.51","miportal.uah.es:193.146.56.65","portal.uah.es:193.146.56.64","portalcomunicacion.uah.es:193.146.56.51","posgrado.uah.es:193.146.56.21","preseea.uah.es:212.128.70.150","publicaciones.uah.es:193.146.56.21","rsendra.web.uah.es:193.146.56.11","sede.uah.es:193.146.56.35","tarjeta.uah.es:193.146.56.21","transparencia.uah.es:193.146.56.21","tui.uah.es:193.146.56.21","vfirma.uah.es:193.146.56.94","www.depeca.uah.es:212.128.70.73","www.robesafe.uah.es:212.128.70.202","www.uah.es:193.146.56.125","www1.uah.es:212.128.64.51"].

3.3.5. Conclusiones de theHarvester

Después de probar la herramienta theHarvester, hemos podido comprobar toda la información perteneciente al dominio establecido como objetivo en el análisis respecto a los dominios y correos electrónicos.

Se trata de información que se encuentra expuesta en internet de forma pública, que, a lo mejor, no debería estar, como es el caso de correos electrónicos corporativos de gente que trabaja en la Universidad; o de diferentes hosts que no deberían ser indexados.

Por último, cabe destacar los ficheros de soporte que nos ofrece la herramienta por objetivo investigado, de esta forma podemos almacenar todas las posibilidades de forma más sencilla.

3.4. Dorks: uso de operadores en buscadores

Los **Dorks** se tratan de operadores que ponen su puesta en escena mediante los buscadores con el fin de conseguir alcanzar una búsqueda avanzada.

La herramienta **SXDork** se encarga realizar Google Dorking (empleo de dorks en la búsqueda en Google) para buscar información específica en internet; esta nos ofrece diferentes tipos de dorks dentro de ella: *login dork, wpadmin dork, SQL dork, archivo de configuración dorks, logfile dorks, dashboard dork, id_rsa dorks, ftp dorks, archivo de copia de seguridad dorks, archivo de correo dorks, contraseña dorks, DCIM fotos dork, y CCTV dorks.*

3.4.1. Instrumentos empleados

Entorno virtual: **VMWare**
Máquina atacante: **Kali Linux**
Herramientas:

- **Navegador.**
  - **Buscadores: Google.**
- **SXDork: https://github.com/samhaxr/SXDork**

3.4.2. Puesta en escena de los Dorks

Para el uso de los Dorks, hemos optado por lanzarla en entorno virtual. Se ha empleado el navegador de forma online y el uso de una herramienta para hacer uso de los diferentes operadores.

Los Dorks tienen como fin llevar a cabo una búsqueda con mayor profundidad con el uso de operadores. Nosotros haremos uso de Google Dorks y de SXDork.

En esta puesta en escena se llevarán a cabo la **búsqueda a un objetivo.**

3.4.3. Google Dorks: el entorno

Abrimos nuestro navegador habitual y fijaremos el buscador con el que trabajaremos: Google.

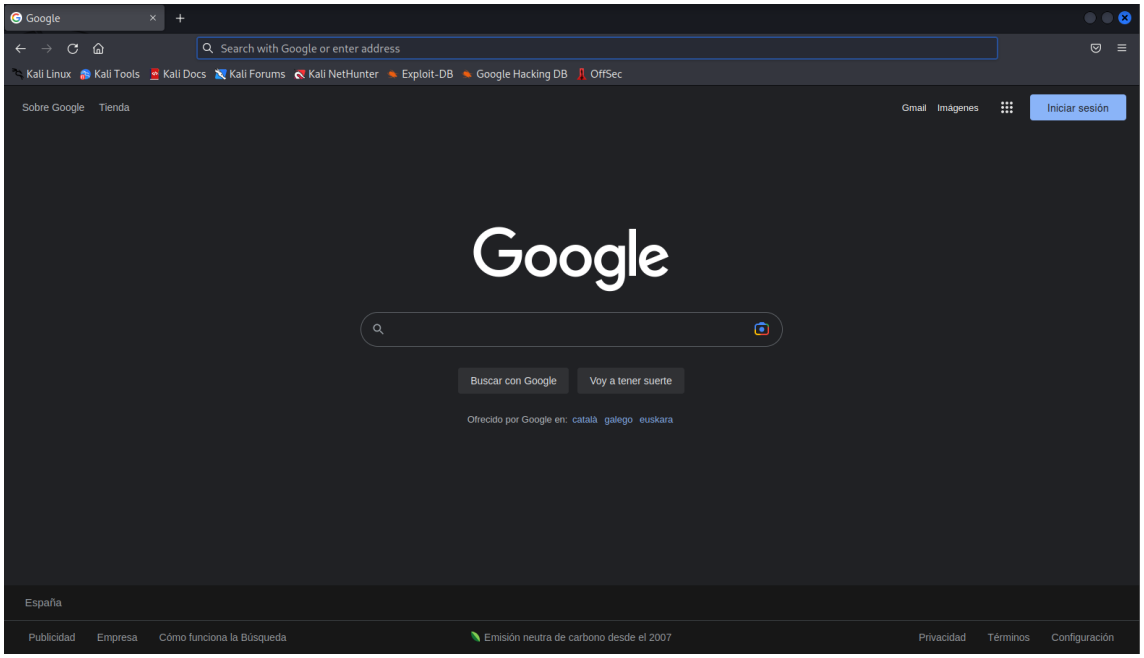


Ilustración 50. Captura de pantalla del inicio de Google Chrome con Google de buscador establecido

Cabe destacar que toda esta información que podamos extraer gracias a la ayuda de los dorks en Google viene ligada a la indexación de todos los contenidos de páginas webs que se recogen en el mismo, sino no sería posible.

3.4.4. Búsqueda de un objetivo en Google Dorks

Llevaremos a cabo la búsqueda del nombre de nuestra víctima por medio de la herramienta. Fijamos como objetivo a la Universidad de Alcalá, nuestro objetivo principal en la investigación.

Sabemos de antemano que la página oficial de la Universidad de Alcalá es uah.es por medio **de Sherlock: búsqueda de nombres de usuarios en redes sociales**, por ellos lo usaremos con dorks para obtener más información.

Related:objetivo; obtenemos toda la información relacionada con el objetivo que queremos encontrar. En nuestro caso al tratarse de una Universidad española nos salen otras Universidades.

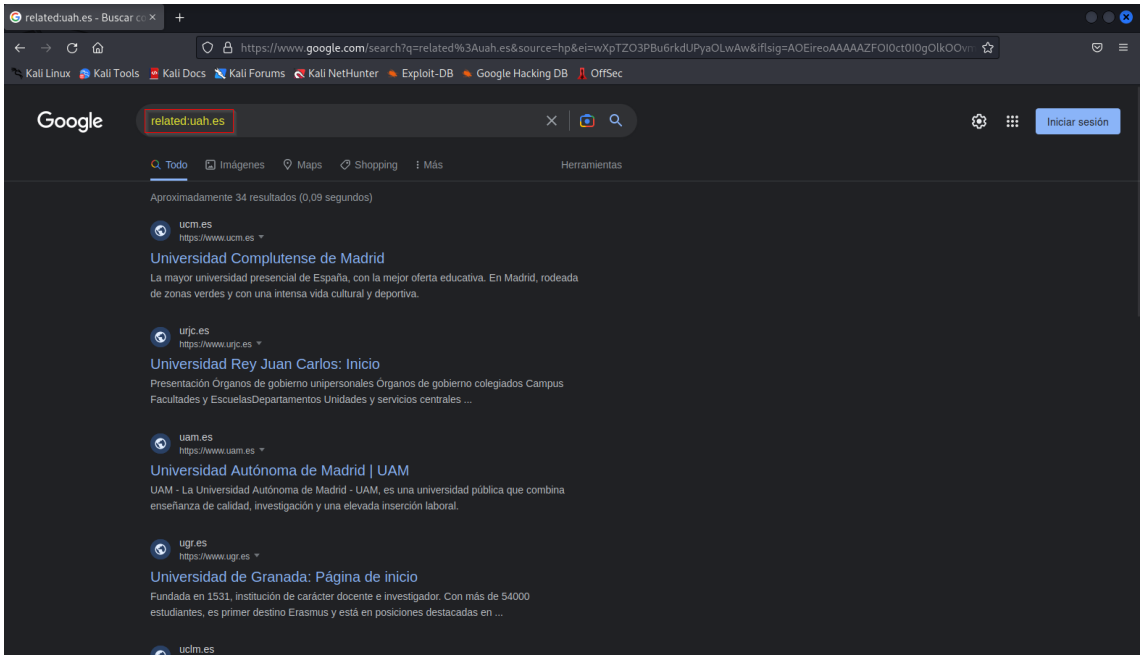


Ilustración 51. Hacemos uso de related con el dominio de la Universidad de Alcalá

Rss site:objetivo; sirve para encontrar diferentes RSS del objetivo que queremos seleccionar. La Universidad de Alcalá nos muestra diferentes eventos, licitaciones, newsletters…

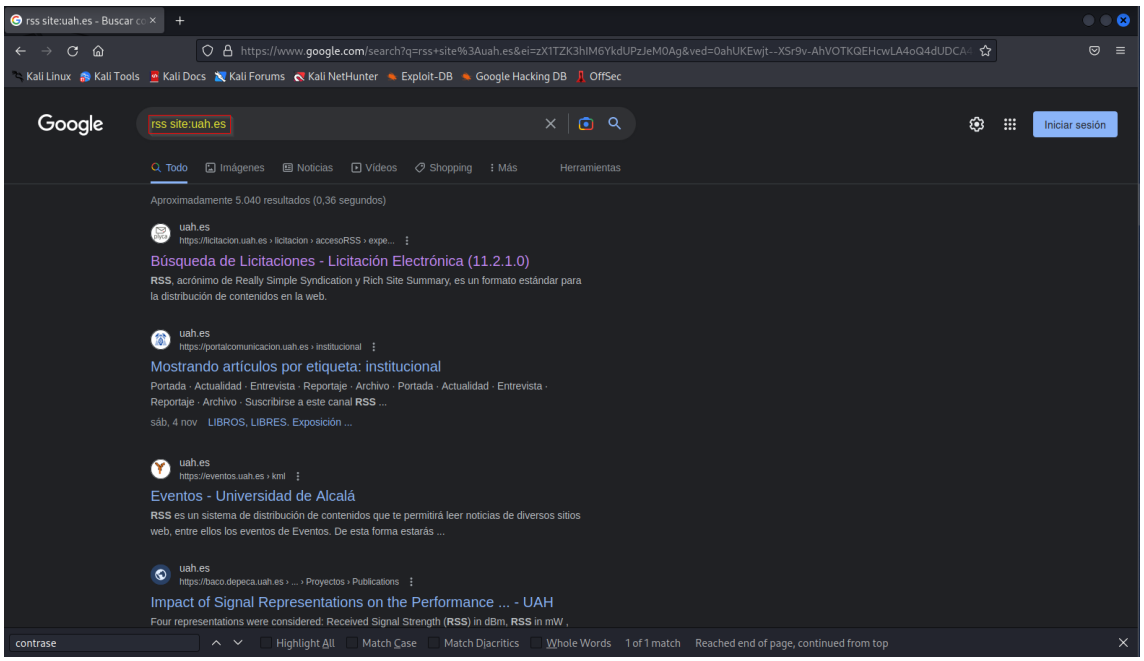


Ilustración 52. Hacemos uso de rss site con el dominio de la Universidad de Alcalá

Objetivo @redsocial: sirve para buscar en cualquier red social el objetivo que queremos desarrollar, por ello escribimos delante del @ el objetivo y seguido del @ la red social. Nosotros podemos apreciar el twitter oficial de la Universidad de Alcalá.

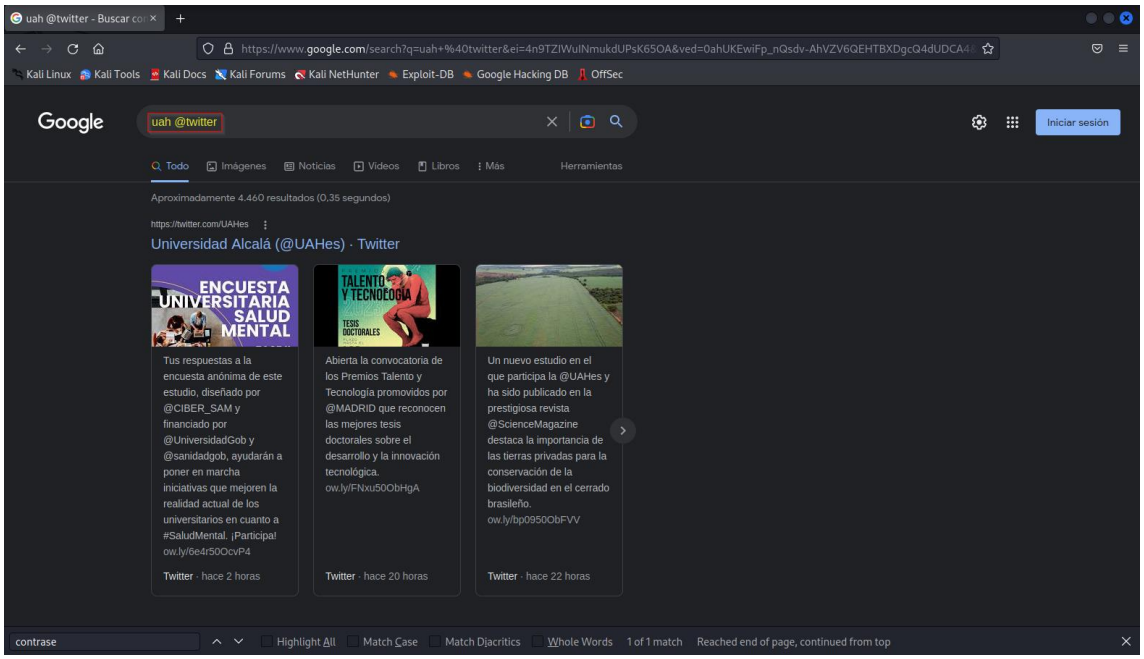


Ilustración 53. Hacemos uso de @redsocial con el dominio de la Universidad de Alcalá

3.4.5. SXDork: el entorno

Nos descargamos la herramienta de SXDork, la lanzamos y vemos en la shell información importante; donde nos muestra las opciones disponibles dentro de esta.

El comando de instalación es el siguiente:

sudo git clone https://github.com/samhaxr/SXDork.git

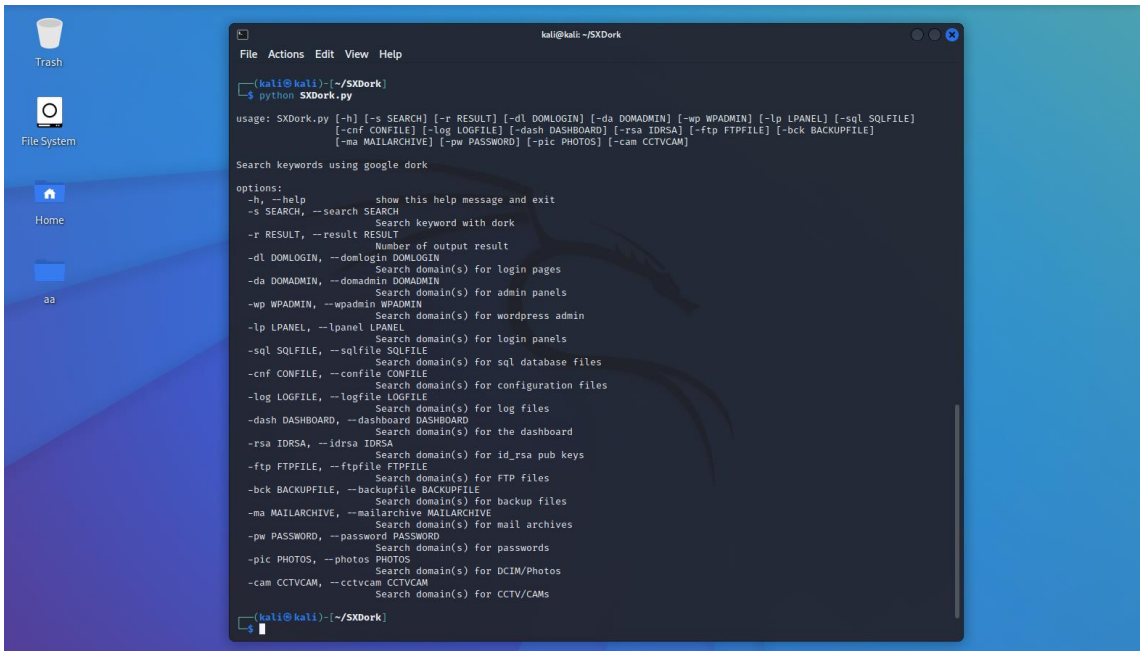


Ilustración 54. Opciones disponibles dentro de la herramienta SXDork

SXDork se encarga de realizar búsquedas con diferentes tipos de dorks, dominios y hasta filtrar a través de los resultados.

Por defecto nos da 10 resultados (se puede aumentar o reducir, -r), las búsquedas de dominios los realiza en pastebin.com y controlc.com, pero se puede añadir más dominios (dorks.py).

Como podemos observar, hay diferentes opciones dentro de SXDork, nosotros fijaremos los dorks que queremos emplear para encontrar toda la información acerca de un determinado como objetivo.

3.4.6. Búsqueda de un objetivo en SXDork

Llevaremos a cabo la búsqueda de un nombre de usuario de nuestra víctima por medio de la herramienta. Fijamos como objetivo a la Universidad de Alcalá, nuestro objetivo principal en la instigación.

Usaremos el dominio de la Universidad de Alcalá para lanzar las diferentes búsquedas dentro de la herramienta: uah.es.

Haremos la búsqueda por medio de la herramienta de las diferentes páginas de inicio de sesión con dominio uah.es.

Para ello haremos uso del siguiente comando: **python3 SXDork.py -dl "uah.es"**

Ilustración 55. Búsqueda del dominio uah.es en SXDork, obtenemos un total de 10 resultados

Como podemos apreciar nos salen un total de diez enlaces (diez resultados por defecto) que nos redirigen a páginas de login de la Universidad de Alcalá.

Un ejemplo de estas se trata del portal de Gestión de Calidad de la Universidad: https://gestioncalidad.uah.es/login-editar/

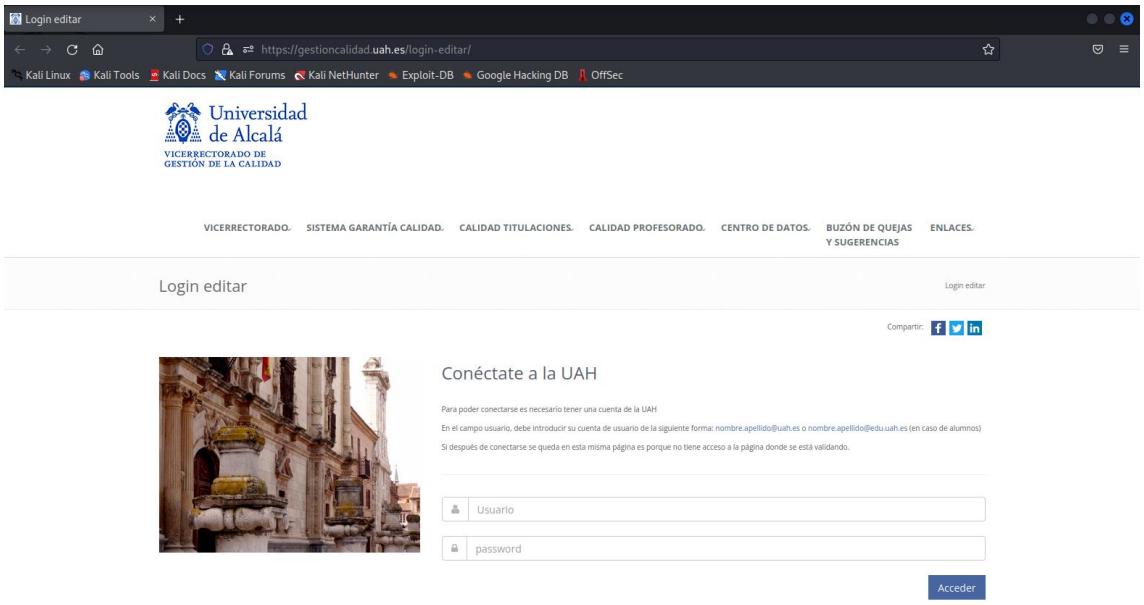


Ilustración 56. Página oficial de Gestión de Calidad de la Universidad de Alcalá

Ahora bien, iniciaremos la búsqueda de las diferentes páginas de dashboard con dominio uah.es.

Para ello haremos uso del siguiente comando: **python3 SXDork.py -dash "uah.es"**

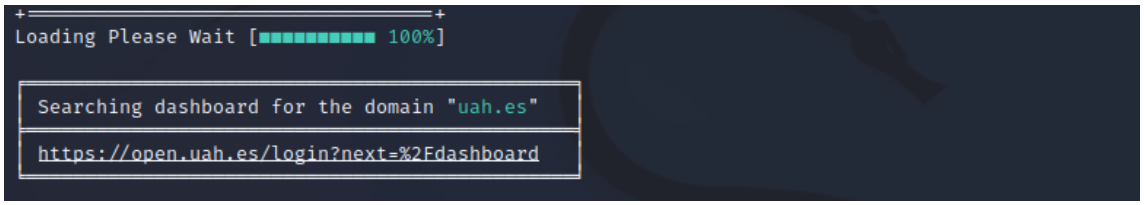


Ilustración 57. Búsqueda del dominio uah.es en SXDork sobre dashboard, obtenemos un total de 1 resultado

En este caso únicamente SDXDork puede encontrar un enlace de una página que nos debería llevar a la dashboard de la Universidad de Alcalá.

Esta se trata de la siguiente:
https://apps.open.uah.es/authn/login?next=%2Fdashboard

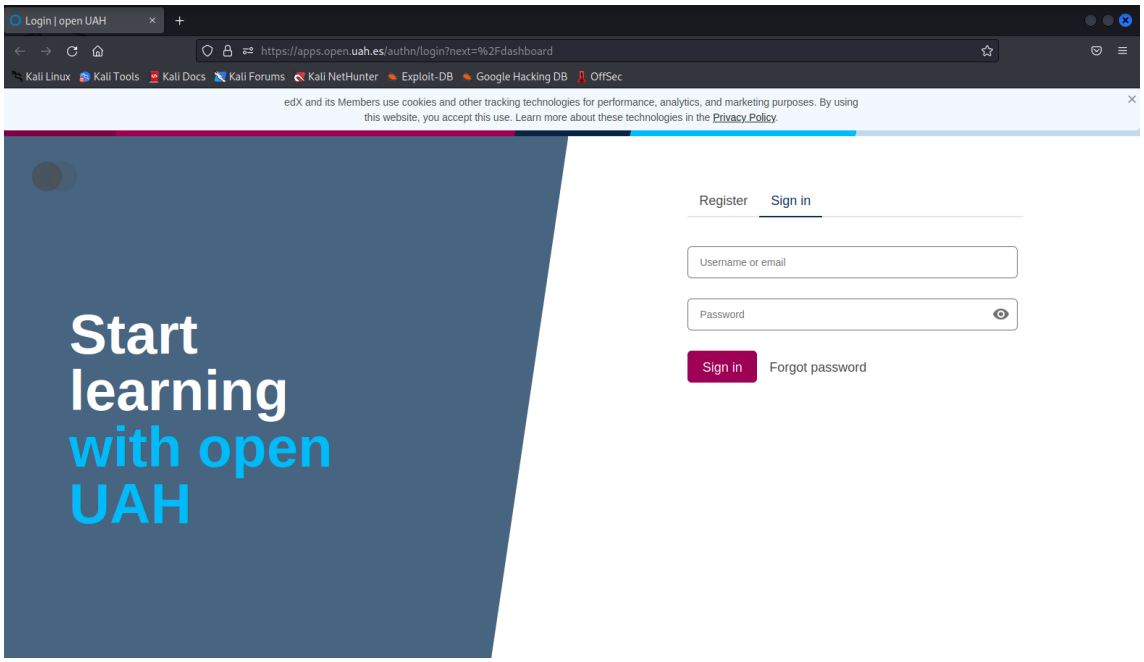


Ilustración 58. Página oficial de Start learning de la Universidad de Alcalá

En este caso nos solicitan un login previo para poder acceder a la dashboard. No obstante, si introdujéramos las credenciales correspondientes accederíamos a la dashboard correspondiente.

3.4.7. Conclusiones de los Dorks

Tras realizar diferentes investigaciones por medio de los dorks, operadores que realizan que las búsquedas se encuentren más avanzadas, hemos podido observar por medio de Google Dorks y por la herramienta de SXDork, la capacidad de estos operadores para obtener una información que en los buscadores de forma normal no podemos encontrar.

Con los Google Dorks hemos podido comprobar cómo de forma sencilla y desde el propio buscador de nuestro navegador, en este caso con Google como buscador establecido, podemos obtener información de mayor valor, esta se encuentra indexada, pero con los dorks podemos acceder a ella de forma más eficaz y sencilla.

Por último, hay que destacar la alta utilidad y facilidad de uso que ofrece SXDork, con esta herramienta hemos podido conseguir diferentes dominios con los que se podrían realizar diferentes pruebas de penetración para obtener información sensible del objetivo establecido.

3.5. accountanalysis: análisis y evaluación de cuentas de Twitter

accountanalysis se trata de una herramienta que se encarga de analizar el comportamiento y evaluar las diferentes cuentas de Twitter que se encuentren públicas. Desde los días de publicación, hasta las redirecciones a páginas webs.

3.5.1. Instrumentos empleados

Entorno virtual: **VMWare**
Máquina atacante: **Kali Linux**
Herramientas:

- **accountanalysis: https://accountanalysis.app/**

3.5.2. Puesta en escena de accountanalysis

Para el despliegue de la herramienta **accountanalysis**, hemos optado por lanzarla en entorno virtual.

Se ha empleado el navegador de forma online para hacer uso de la herramienta; esta nos permite observar el comportamiento en Twitter de las diferentes cuentas que se encuentren públicas.

El comportamiento se puede evaluar desde los seguidores y seguidos, frecuencia de tweets, cantidad de interacciones, horarios de mayor actividad y hashtags más utilizados…

En esta puesta en escena se llevará a cabo la **búsqueda a un objetivo.**

3.5.3. accountanalysis: el entorno

Abrimos la herramienta de accountanalysis en nuestro navegador, y vemos las opciones disponibles dentro de esta.

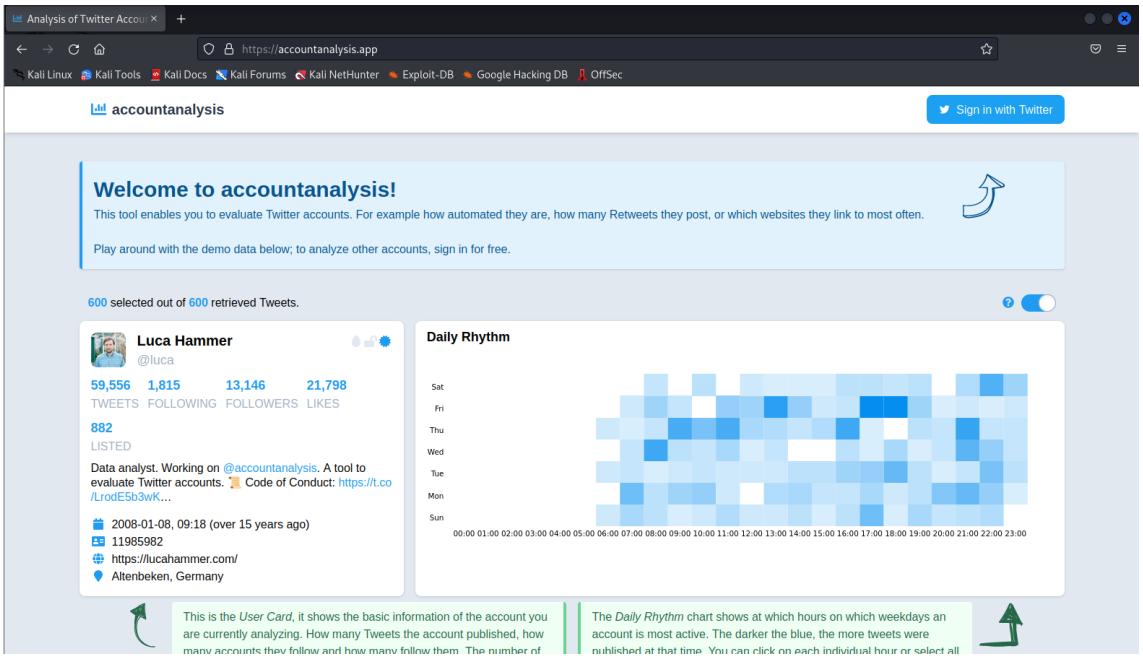


Ilustración 59. Interfaz de la página de inicio de accountanalysis.

Como podemos observar, para poder llevar a cabo la investigación dentro de la herramienta, es necesario vincular una cuenta de Twitter con la que

accountanalysis pueda acceder a la API y desde ahí llevar a cabo el análisis oportuno.

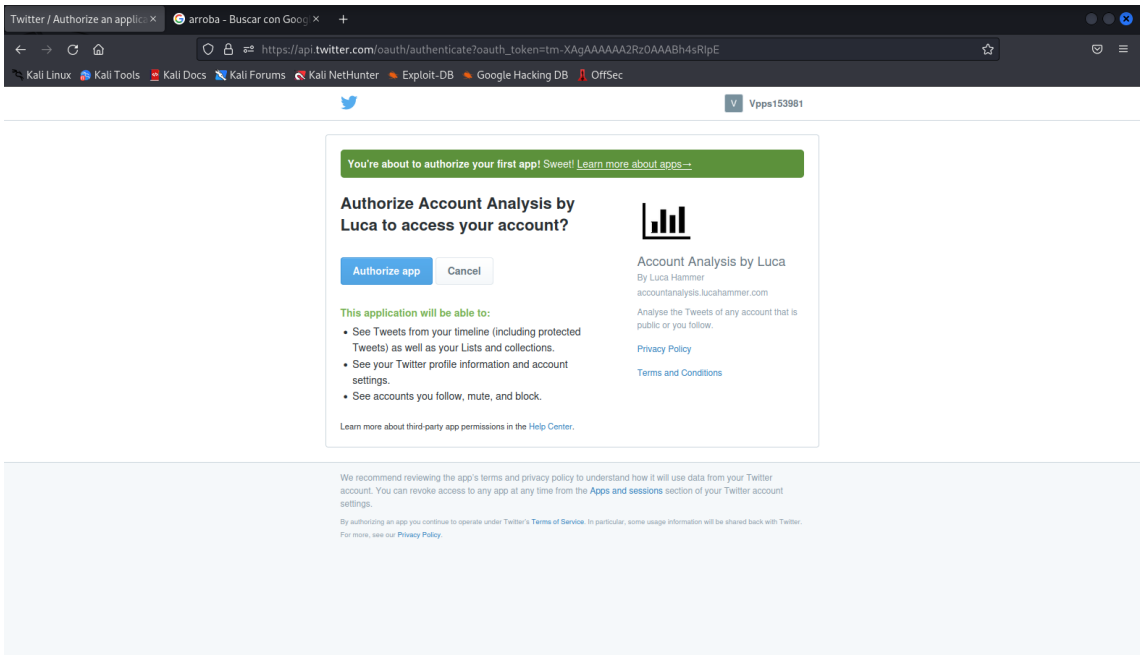


Ilustración 60. Autorización e inicio de sesión en Twitter dentro de accountanalysis.

3.5.4. Búsqueda de un objetivo y análisis de resultados

Llevaremos a cabo la búsqueda de nuestro objetivo por medio de la herramienta. Fijamos como objetivo a la Universidad de Alcalá, nuestro objetivo principal en la instigación.

Por ello establecemos el nombre de usuario de Twitter, estableceremos el de la UAH, tras corroborar anteriormente la oficialidad y legitimidad de esta: @UAHes.

Fijamos la investigación a 1600 tweets (es el máximo de tweets que plan gratito de la herramienta ofrece).

Con todo ello, realizamos la búsqueda.

Ilustración 61. Número de tweets de la Universidad de Alcalá para investigar y analizar.

Tras comenzar la búsqueda, en unos instantes accountanalysis nos mostrará por pantalla diferentes campos que muestran los resultados obtenidos tras el análisis del comportamiento del Twitter de la Universidad de Alcalá.

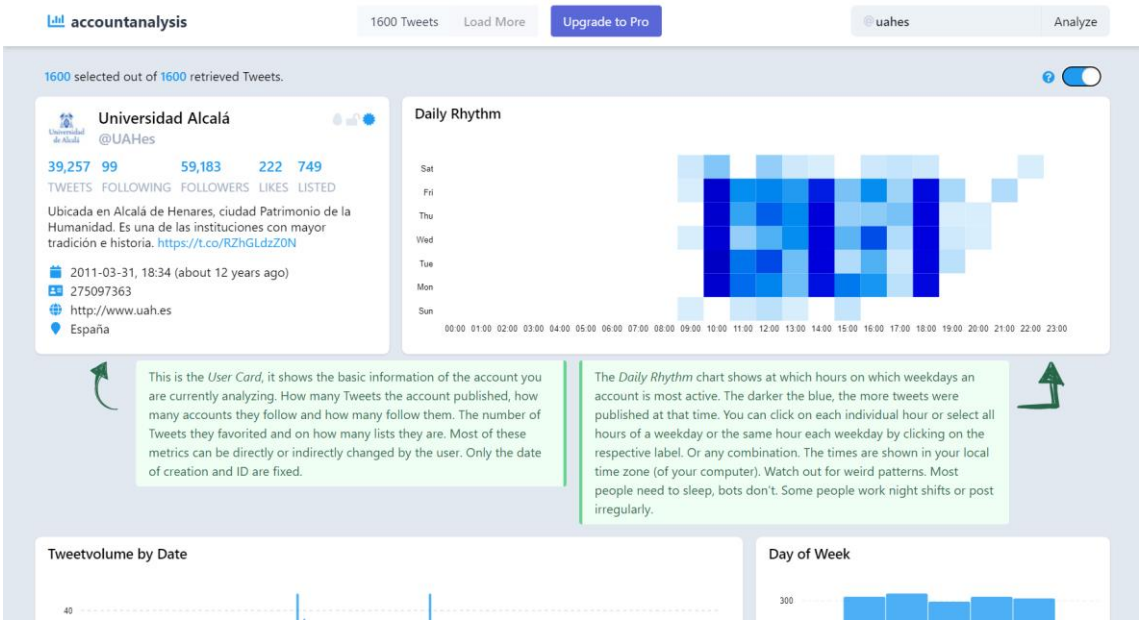


Ilustración 62. Interfaz de accountanalysis tras la búsqueda realizada sobre la Universidad de Alcalá en Twitter.

3.5.4.1. **User Card / Tarjeta de Usuario**

La Tarjeta de Usuario, muestra la información básica de la cuenta que está analizando actualmente. Cuántos tweets publicó la cuenta, cuántas cuentas siguen y cuántas las siguen. La cantidad de Tweets que marcaron como favoritos y en cuántas listas están. La mayoría de estas métricas pueden ser modificadas directa o indirectamente por el usuario. Solo se fija la fecha de creación y el ID.

Como vemos en este análisis del User Card / Tarjeta de Usuario, la Universidad de Alcalá cuenta con:

- 59.183 seguidores (16/04/2023).
- 99 seguidos (16/04/2023).
- 39.257 tweets publicados (16/04/2023).
- 749 listas (16/04/2023).

La descripción de la UAH en Twitter es: Ubicada en Alcalá de Henares, ciudad Patrimonio de la Humanidad. Es una de las instituciones con mayor tradición e historia. El enlace que tiene vinculado redirige a la cuenta oficial de la Universidad en Instagram: https://www.instagram.com/uahes/

Se unió a esta red social hace 12 años, el 31 de marzo de 2011; su identificador es el 275097363; la página web oficial: http://www.uah.es/; se encuentra ubicada en España.

Ilustración 63. User Card / Tarjeta de Usuario de la Universidad de Alcalá en Twitter dentro de accountanalisys.

3.5.4.2. **Daily Rhythm / Gráfico de ritmo diario**

El gráfico de ritmo diario muestra a qué horas y qué días de la semana una cuenta está más activa. Cuanto más oscuro era el azul, más tuits se publicaban en ese momento. Puede hacer clic en cada hora individual o seleccionar todas las horas de un día de la semana o la misma hora cada día de la semana haciendo clic en la etiqueta respectiva. O cualquier combinación. Los tiempos se muestran en su zona horaria local (de su computadora). Cuidado con los patrones extraños. La mayoría de la gente necesita dormir, los bots no. Algunas personas trabajan en turnos de noche o publican de manera irregular.

Como vemos en este análisis del Daily Rhythm / Gráfico de ritmo diario, la Universidad de Alcalá cuenta con:

- Hora de mayor nivel de publicación: 10:00h, 14:00h y 18:00h (16/04/2023).
- Hora de nivel medio de publicación: 11:00h, 12:00h y 13:00h (16/04/2023).

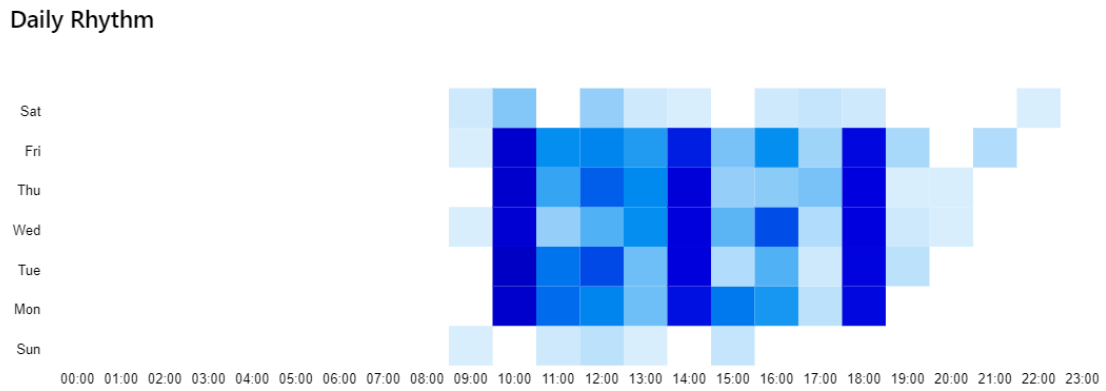


Ilustración 64. Daily Rhythm / Gráfico de ritmo diario de la Universidad de Alcalá en Twitter dentro de accountanalisys.

3.5.4.3. **Tweet volume by Date / Volumen de tweets por fecha**

Tweet volume by Date muestra cuántos Tweets publicó la cuenta en una fecha específica. Puede hacer clic y arrastrar para ver solo los Tweets que se publicaron en las fechas resaltadas. Solo unas pocas personas twittean la misma cantidad de tweets cada día. Los días individuales con alto volumen pueden indicar que ocurrió un evento. El gráfico Hashtag puede ayudar a identificar qué tipo de evento.

Como vemos en este análisis del Tweet volume by Date / Volumen de tweets por fecha, la Universidad de Alcalá cuenta con:

- Mayor nivel de tweets en enero de 2023 y diciembre de 2022 (16/04/2023).

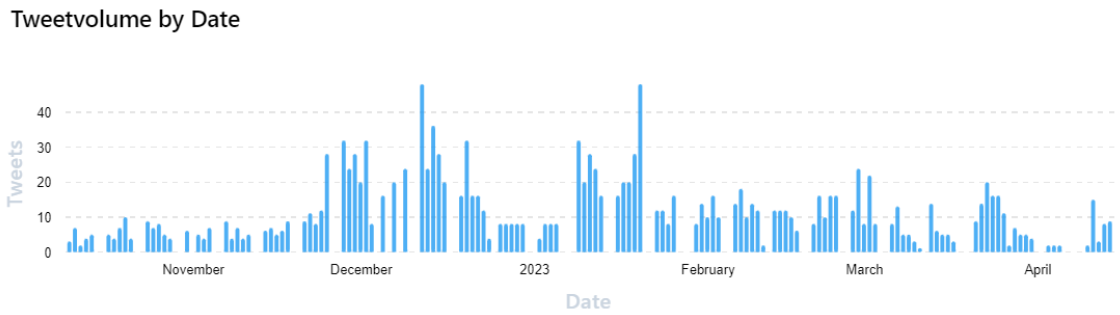


Ilustración 65. Tweet volume by Date / Volumen de tweets por fecha de la Universidad de Alcalá en Twitter dentro de accountanalisys.

3.5.4.4. **Day of Week / Día de la semana**

El gráfico Día de la semana muestra en qué días de la semana una cuenta está más activa. Puede hacer clic en cada día de la semana para seleccionar Tweets que se publicaron en ese día de la semana. El gráfico muestra la cantidad agregada de Tweets para ese día de la semana de todos los Tweets que se recuperaron de la cuenta. La mayoría de la gente tuitea más los días laborales o los fines de semana. Algunos tienen horarios diferentes.

Como vemos en este análisis del Day of Week / Día de la semana, la Universidad de Alcalá cuenta con:

- Mayor nivel de actividad en la cuenta: días laborables (lunes, martes, miércoles, jueves y viernes) (16/04/2023).

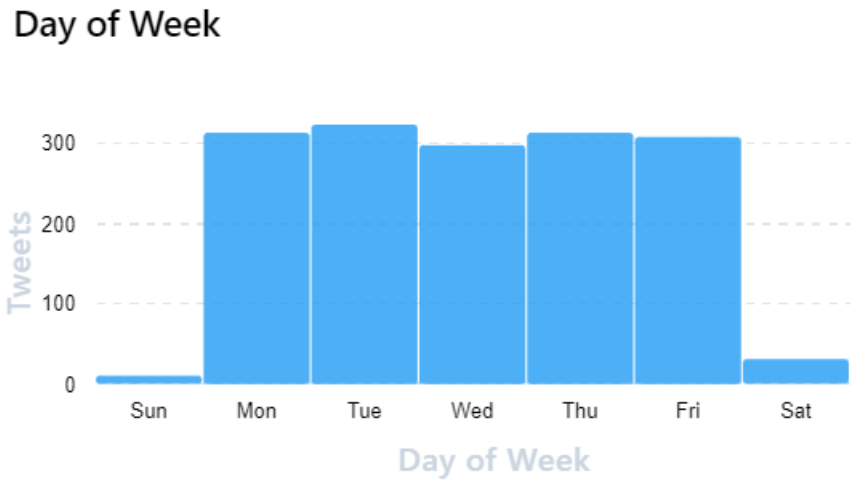


Ilustración 66. Day of Week / Día de la semana de la Universidad de Alcalá en Twitter dentro de accountanalisys.

3.5.4.5. **Tweet Type / Tipo de tweet**

Hay cinco tipos diferentes de Tweets.

- Un "Retweet" es un Retweet nativo de una cuenta diferente o de uno de sus propios Tweets. Solo se necesitan uno o dos toques para retuitear un Tweet. Por lo tanto, es fácil generar un gran número de Retweets.

- Una "Cita" es un Retweet con un comentario encima. A veces se denomina Quote-Tweet o Quote-Retweet.

- Una "Respuesta" es un Tweet que comienza con una @ y un nombre de usuario. En la mayoría de los casos, se adjunta a un Tweet de otra persona. Solo los usuarios que siguen ambas cuentas ven las respuestas en su línea de tiempo cronológica.

- Las "Auto-Respuestas" son Respuestas que se adjuntan a los Tweets de la misma cuenta. A menudo denominados subprocesos. Un hilo consta de un Tweet y una o más respuestas automáticas.

- El "Tweet" es un Tweet que no es un Retweet, una cita o una respuesta. Cada cuenta tiene un estilo de tweet diferente, pero la mayoría usa múltiples, si no todos, tipos de tweets con diferente intensidad.

Como vemos en este análisis del Tweet Tipe / Tipo de tweet, la Universidad de Alcalá cuenta con:

- Tweet: 954 (16/04/2023).
- Retweet: 651 (16/04/2023).
- Auto-Respuestas: 21 (16/04/2023).

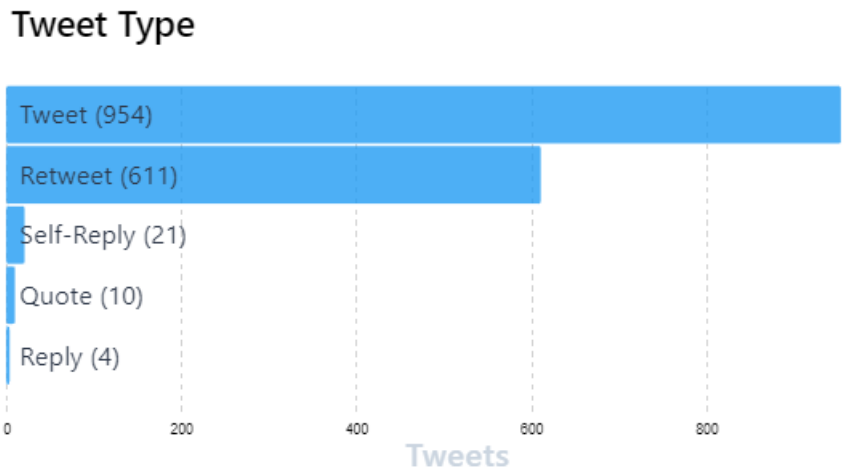


Ilustración 67. Tweet Type / Tipo de tweet de la Universidad de Alcalá en Twitter dentro de accountanalisys.

3.5.4.6. **Language of Tweets / Idioma de los tweets**

Twitter intenta determinar en qué idioma está escrito cada Tweet. El gráfico Idioma de los Tweets muestra esos idiomas y cuántos Tweets ha publicado la cuenta en ellos. El algoritmo de categorización no es perfecto. Los Tweets especialmente cortos o los Tweets con préstamos dan como resultado categorizaciones falsas. Los tuits con un alto grado de incertidumbre (p. ej., tuits solo multimedia) se clasifican como "Desconocidos". Pocas personas publican muchos Tweets en más de tres idiomas.

Como vemos en este análisis del Language of Tweets / Idioma de los tweets, la Universidad de Alcalá cuenta con:

- Español 95%, portugués 2%, inglés 2% y otros 1%.

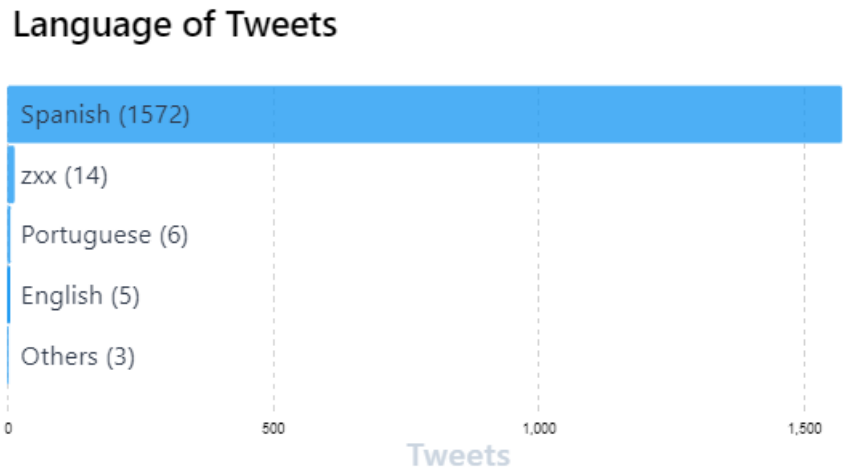


Ilustración 68. Language of Tweets / Idioma de los tweets de la Universidad de Alcalá en Twitter dentro de accountanalisys.

3.5.4.7. **Used Interface / Interfaz de uso**

El gráfico de Interfaz utilizada muestra qué aplicaciones usó la cuenta para publicar los Tweets. La mayoría de los usuarios se quedan con las oficiales (Twitter Web App, Twitter para Android, Twitter para iPhone, Twitter para iPad, TweetDeck), pero también hay algunas aplicaciones populares de terceros con todas las funciones (Tweetbot, Fenix, ...). Otras aplicaciones de terceros se utilizan principalmente para atención al cliente (swat.io, hootsuite, ...) o automatización (Buffer, IFTTT, ...). Los bots avanzados usan las aplicaciones oficiales para parecer menos sospechosos.

Como vemos en este análisis del Used Interface / Interfaz de uso, la Universidad de Alcalá cuenta con:

- Hootsuite 60%, Twitter Web App 25%, Twitter for iPhone 14% y Twitter for Android 1%.

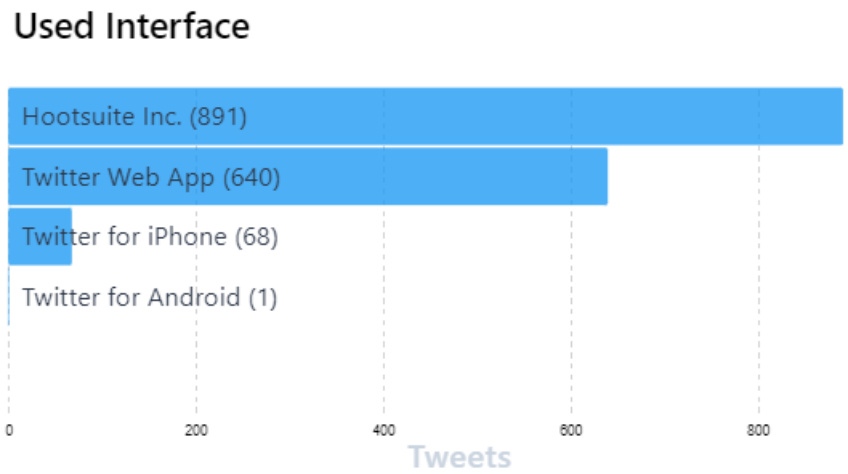


Ilustración 69. Used Interface / Interfaz de uso de la Universidad de Alcalá en Twitter dentro de accountanalisys.

3.5.4.8. **Used Hashtags / Hashtags usados**

Cada Tweet puede contener uno o más Hashtags. Los Hashtags usados los muestra todos y le permite filtrar los Tweets que contienen los Hashtags que le interesan. Los Hashtags tienen múltiples usos. La mayoría de las cuentas los utilizan para categorizar sus Tweets, participar en eventos y/o aumentar la visibilidad de los Tweets. Mirar los Hashtags más utilizados ayuda a comprender los temas principales sobre los que tuitea una cuenta.

Como vemos en este análisis del Used Hashtags / Hashtags usados, la Universidad de Alcalá cuenta con:

- #UAH
- #AlumniUAH
- #EmpleabilidadUAH
- #VoyASerUAH
- #EquipoMédulaUAH

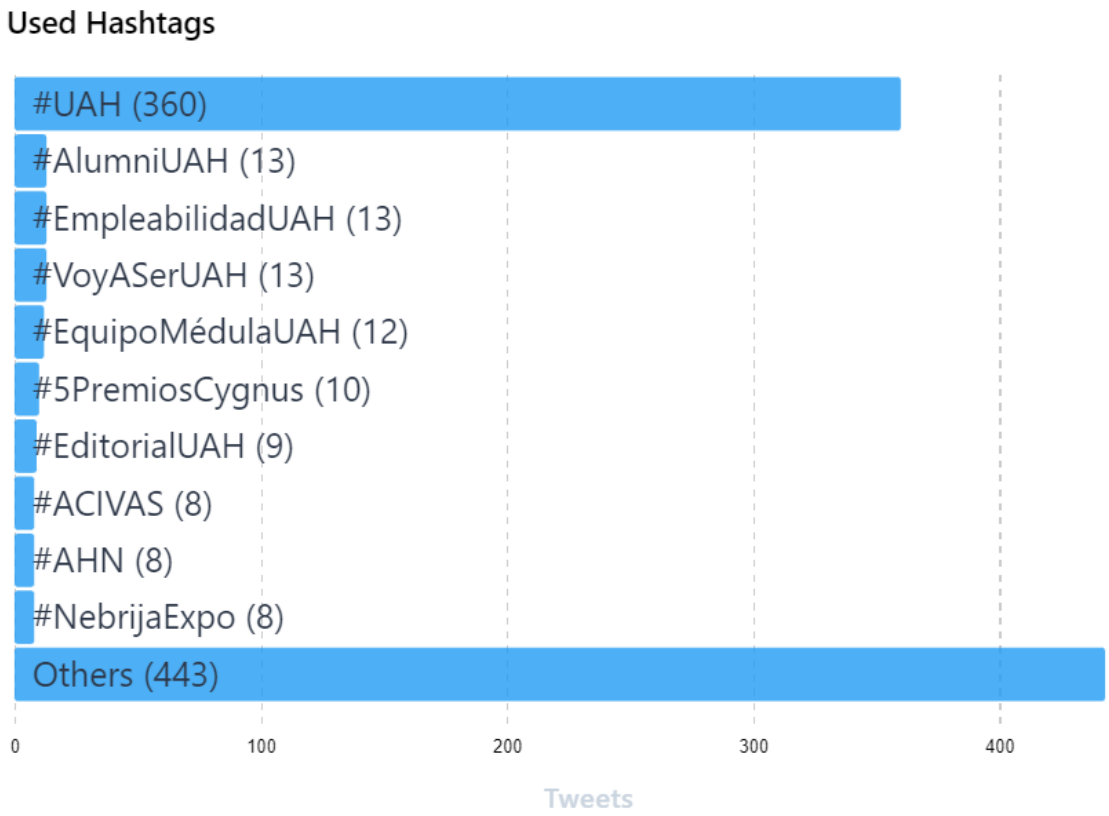


Ilustración 70. Used Hashtags / Hashtags usados de la Universidad de Alcalá en Twitter dentro de accountanalisys.

3.5.4.9. **Hostnames of URLs**

Cada Tweet puede contener una o varias URL. El gráfico Nombres de host de URL muestra los nombres de host de esas URL. El nombre de host contiene el subdominio, pero no la ruta, consulta o fragmento de la URL. "twitter.com" y "mobile.twitter.com" están excluidos ya que representan Tweets citados y cargas de medios.

Como vemos en este análisis del Hostnames of URLs, la Universidad de Alcalá cuenta con:

- YouTube
- Bit.ly
- Encuestas.uv.es
- Ow.ly
- Facebook
- Ifema

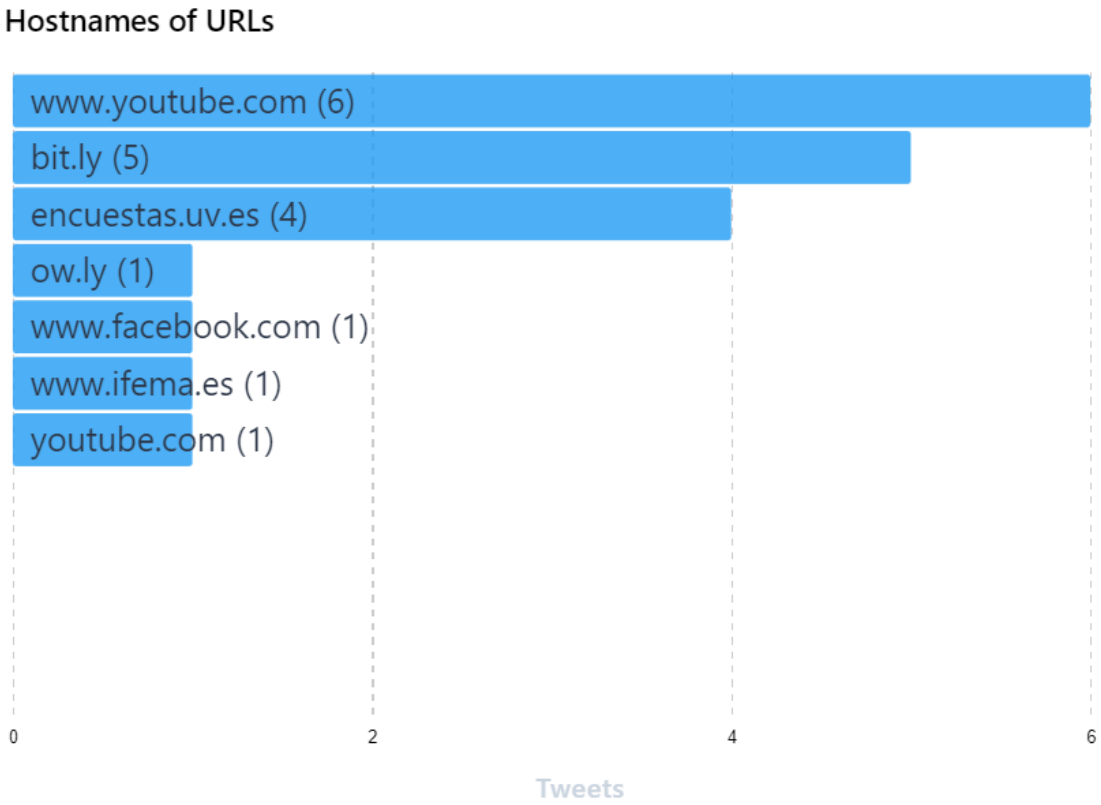


Ilustración 71. Hostnames of URLs de la Universidad de Alcalá en Twitter dentro de accountanalisys.

3.5.4.10. **Replied Users / Usuarios Respondidos**

El gráfico Usuarios respondidos muestra a qué cuentas escribe la cuenta analizada la mayoría de las respuestas. Solo se cuenta el usuario al que la cuenta respondió directamente. Si el usuario A publica un Tweet, el usuario B responde y el usuario C responde a esa respuesta, solo el usuario B aparece en el gráfico cuando se analiza al usuario C. Las respuestas automáticas están excluidas.

Como vemos en este análisis del Replied Users / Usuarios Respondidos, la Universidad de Alcalá cuenta con:

- _ele88
- alieth9409305
- cniostopcancer
- mrunowen_

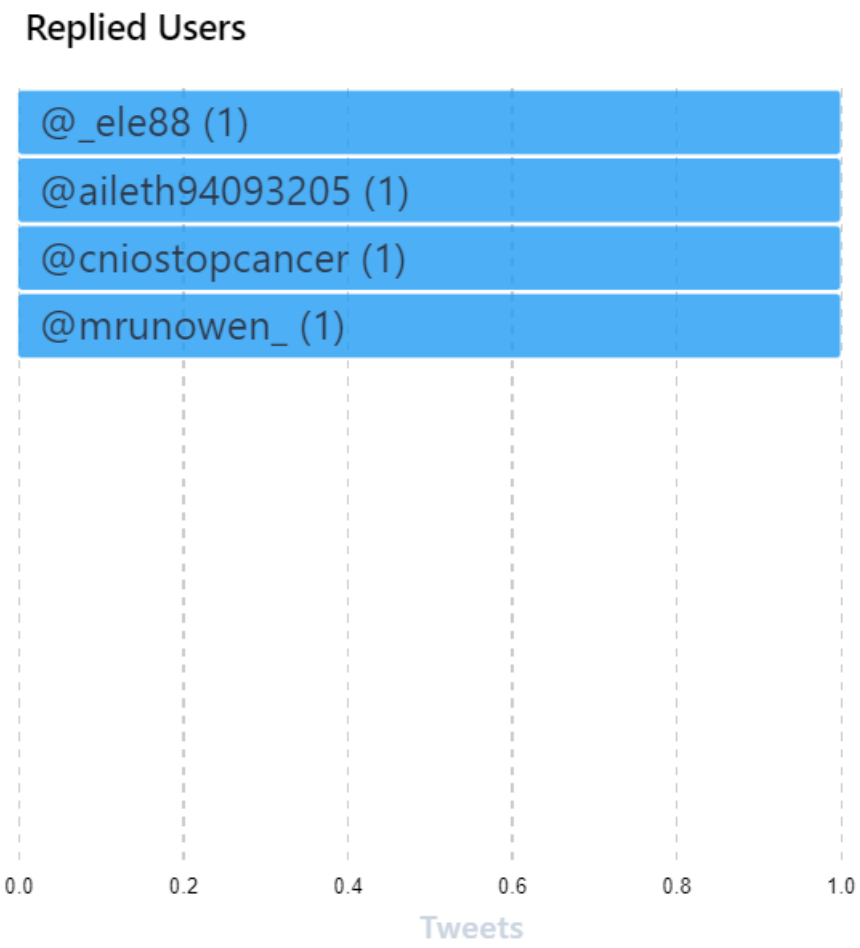


Ilustración 72. Replied Users / Usuarios Respondidos de la Universidad de Alcalá en Twitter dentro de accountanalisys.

3.5.4.11. **Retweeted Users / Usuarios retuiteados**

El gráfico de usuarios retuiteados muestra qué cuentas retuitea con más frecuencia la cuenta analizada.

Como vemos en este análisis del Retweeted Users / Usuarios retuiteados, la Universidad de Alcalá cuenta con:

- @otriuah (53)
- @culturauah (37)
- @madrimasd (20)

- @aytoalcalah (19)
- @equipomedula (19)
- @ib_franklin (18)
- @deportes_uah (15)
- @iq_humor (15)
- @instcervantes (14)
- @educacmadrid (12)

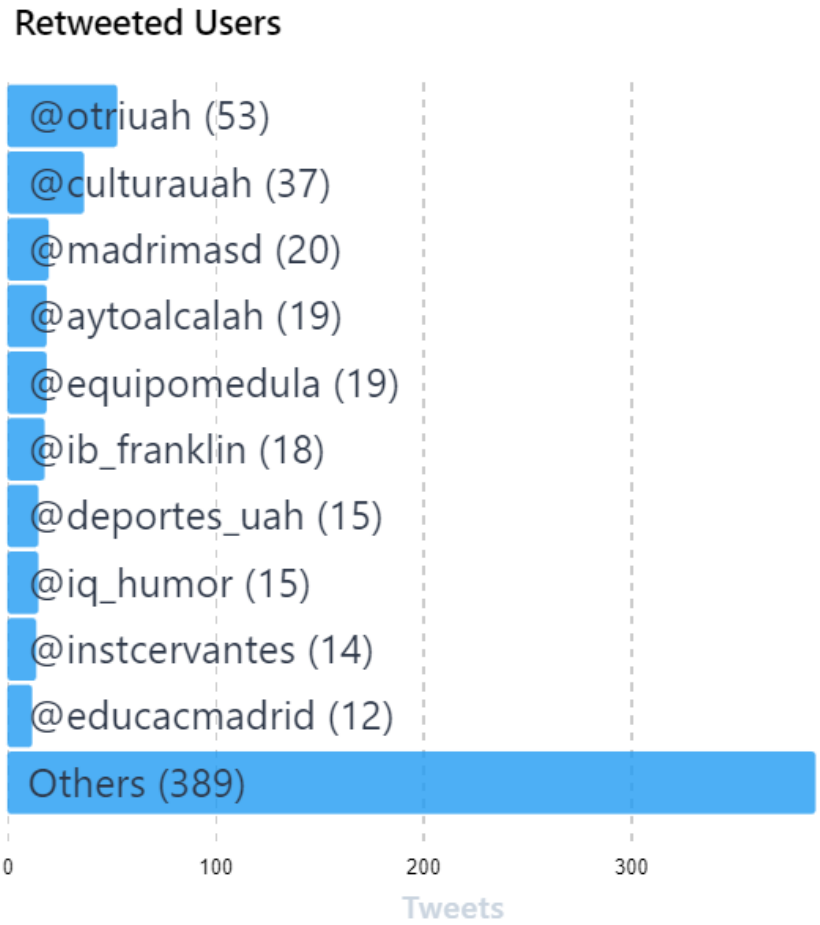


Ilustración 73. Retweeted Users / Usuarios retuiteados de la Universidad de Alcalá en Twitter dentro de accountanalisys.

3.5.4.12. **Quoted Users / Usuarios citados**

El gráfico de Usuarios citados muestra qué cuentas cita la cuenta analizada con más frecuencia.

Como vemos en este análisis del Quoted Users / Usuarios citados, la Universidad de Alcalá cuenta con:

- @culturauah (2)
- @educajccm (2)
- @foropostgrado (2)

- @jccmguadalajara (2)
- @faceducacionuah (1)

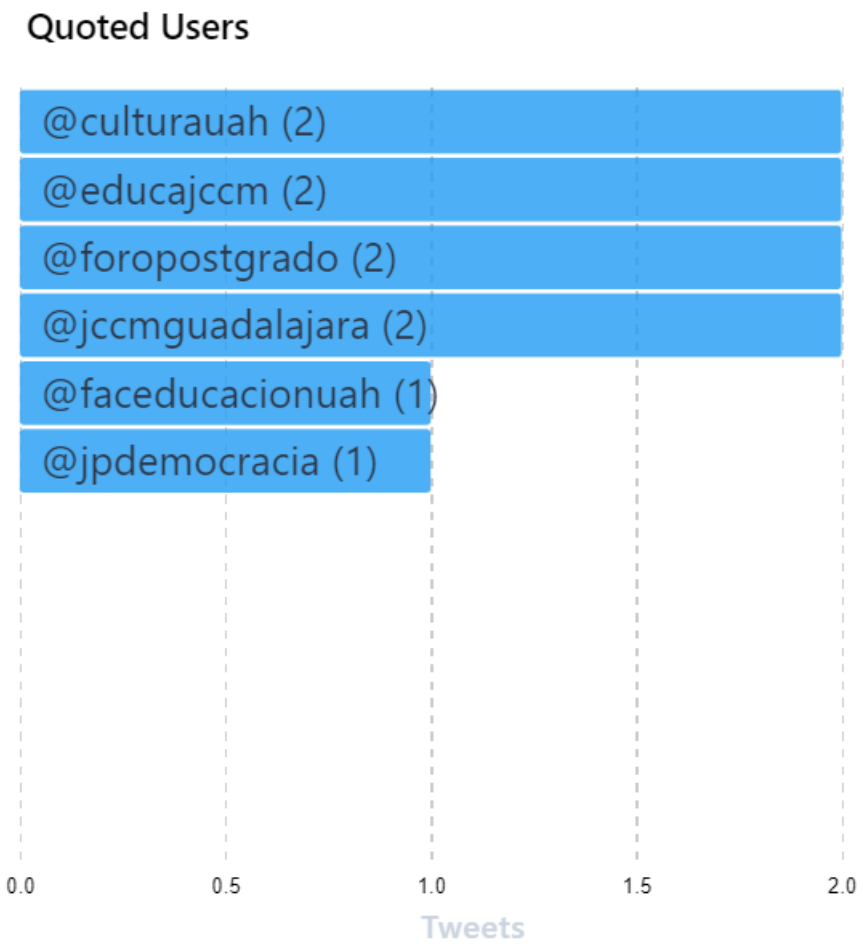


Ilustración 74. Quoted Users / Usuarios citados de la Universidad de Alcalá en Twitter dentro de accountanalisys.

3.5.5. Conclusiones de accountanalysis

Después de probar la herramienta accountanalysis, hemos podido comprobar toda el comportamiento e información que se puede extraer de una cuenta en la red social de Twitter.

A partir del informe final que nos ofrece, somos capaces de detectar de qué forma actúa la corporación o persona dentro de Twitter.

Por último, cabe destacar la facilidad y usabilidad de la herramienta a la hora de iniciar la búsqueda del objetivo y, tras ello, la forma de plasmar los resultados obtenidos de investigación.

3.6. Cree.py: análisis y evaluación de redes sociales

Cree.py es una herramienta de ingeniería social enfocada a redes sociales que nos permite realizar la recopilación de información por medio de la geolocalización a través de plataformas de redes sociales.

3.6.1. Instrumentos empleados

Entorno virtual: **VMWare**
Máquina atacante: **Kali Linux**
Herramientas:

- **Cree.py: http://www.geocreepy.com/**

3.6.2. Puesta en escena de Cree.py

Para el despliegue de la herramienta **Cree.py**, hemos optado por lanzarla en entorno virtual.

También se han hecho uso de las redes sociales: *Twitter e Instagram.* Donde se obtiene información correspondiente a la geolocalización de las cuentas que queremos investigar por medio de las redes sociales. A su vez, la herramienta nos ofrece un análisis de cada red social con la que investigamos.

En esta puesta en escena se llevará a cabo la **búsqueda a un objetivo.**

3.6.3. Cree.py: el entorno

Nos descargamos la herramienta de Cree,py, la lanzamos y vemos en la shell información importante; donde nos muestra las opciones disponibles dentro de esta.

El comando de instalación es el siguiente:

sudo pip install -U pytz python-qt flickrapi python-instagram yapsy tweepy google-api-python-client python-dateutil configobj dominate

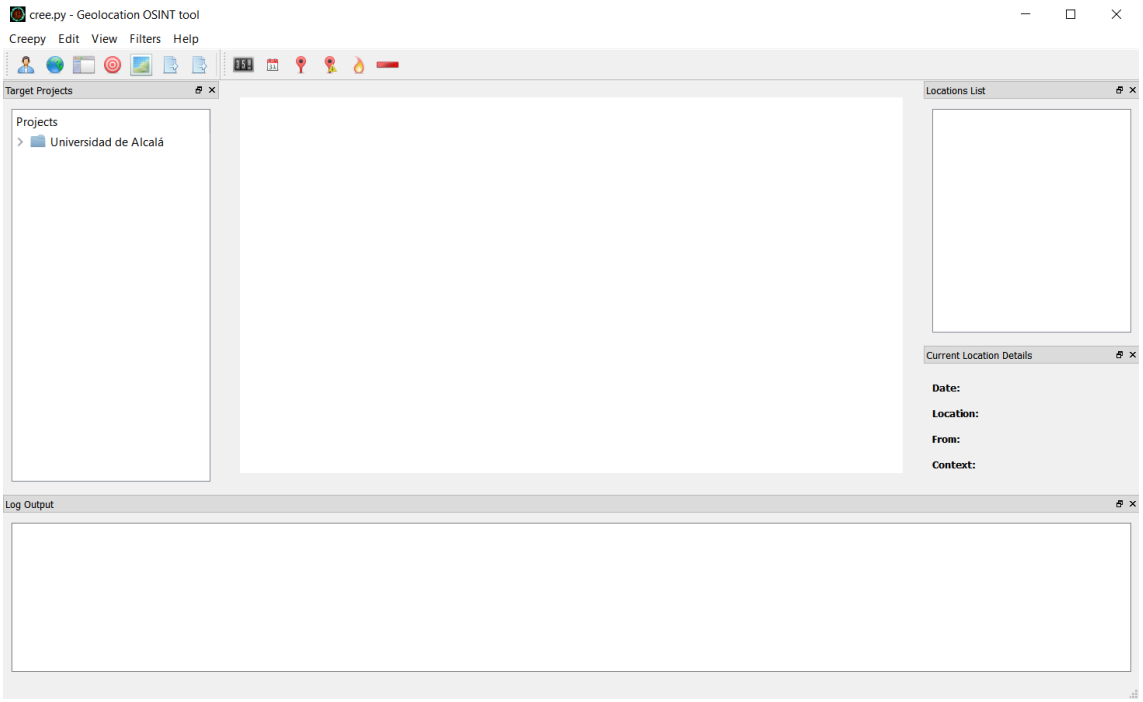
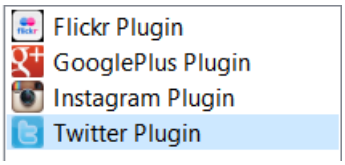


Ilustración 75. Interfaz de la herramienta de Cree.py.

Como podemos observar, hay diferentes opciones dentro de la herramienta, nosotros iremos al objetivo que queremos, encontrar las redes sociales del nombre de usuario determinado como objetivo.

3.6.4. Configuración de redes sociales

En primer lugar, hemos de configurar los diferentes plugin de redes sociales de los que vamos a hacer uso dentro de Cree.py.



La herramienta nos ofrece realizar investigaciones y análisis de personas/lugares en cuatro redes sociales: Flickr, GooglePlus, Instagram y Twitter.

Nosotros realizaremos la investigación de la Universidad de Alcalá en Twitter e Instagram, por ello, deberemos seguir los siguientes pasos de configuración previos a la búsqueda de información en estas redes sociales.

3.6.4.1. **Configuración de Twitter**

Seleccionaremos la opción de Twitter y tendremos que hacer un registro en la red social con una cuenta para obtener el PIN con el cual la herramienta pueda acceder a la API de Twitter y obtener diferente información sobre las ubicaciones del usuario seleccionado.

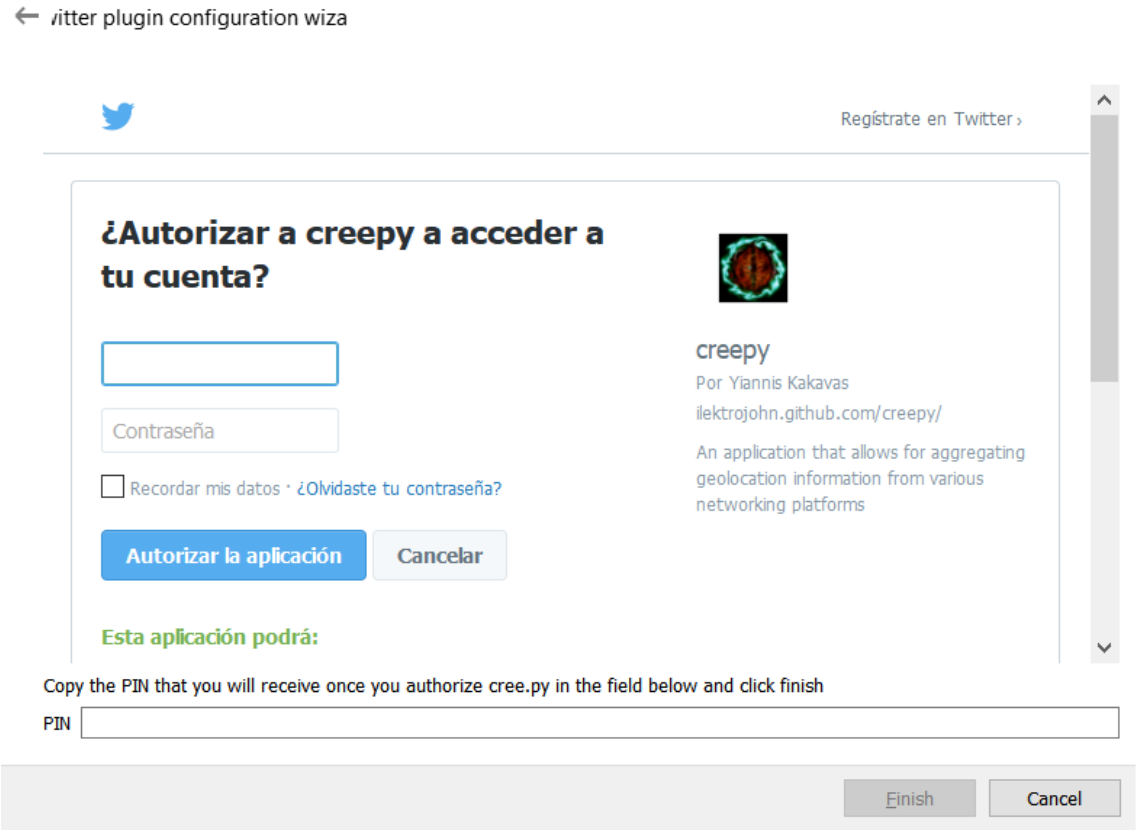


Ilustración 76. Configuración de Twitter dentro de Cree.py: autorización y PIN para conexión API.

Tras obtener el PIN y establecer una correcta conexión con la API de Twitter, tendremos toda la configuración de la herramienta en esta red social establecida.

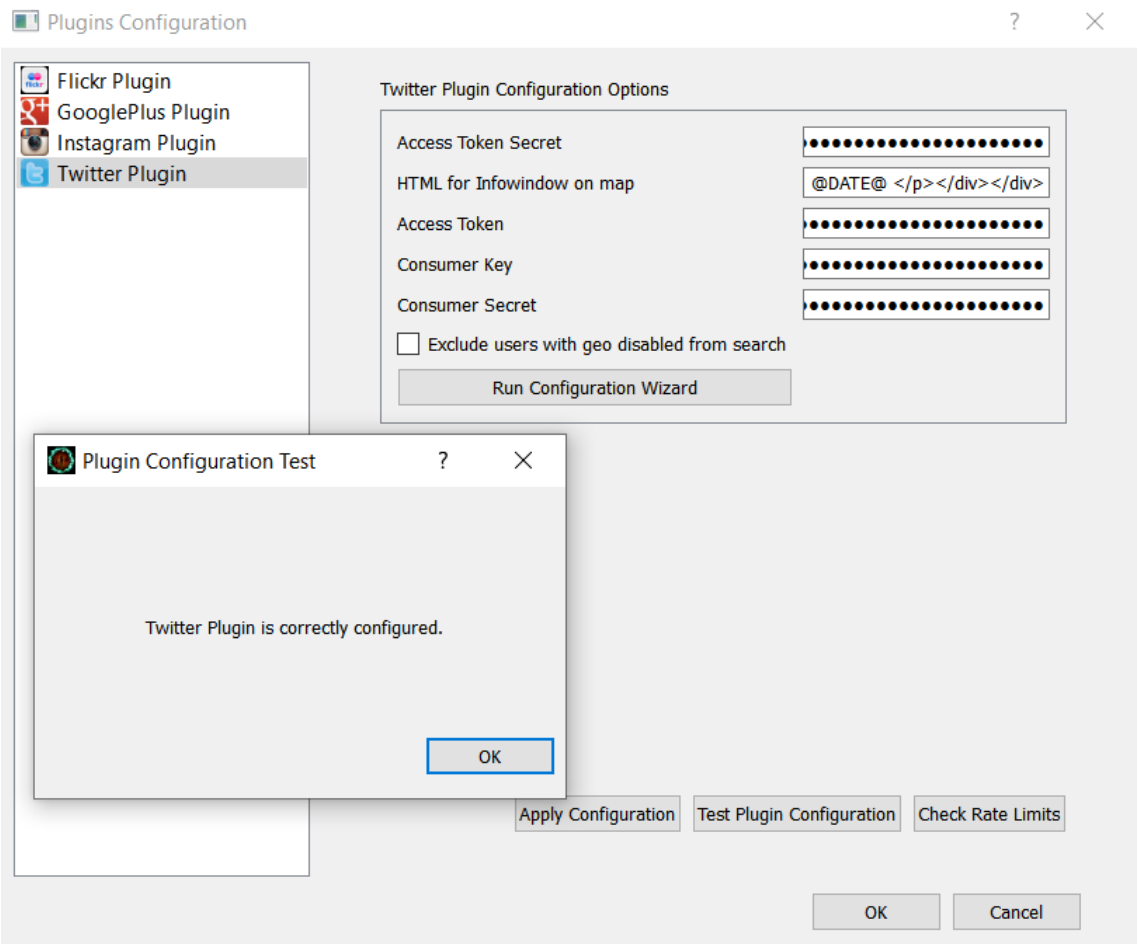


Ilustración 77. Configuración de Twitter y conexión API realizada correctamente.

3.6.4.2. **Configuración de Instagram**

Seleccionaremos la opción de Instagram y tendremos que hacer un registro en la red social con una cuenta para obtener el PIN con el cual la herramienta pueda

acceder a la API de Instagram y obtener diferente información sobre las ubicaciones del usuario seleccionado.

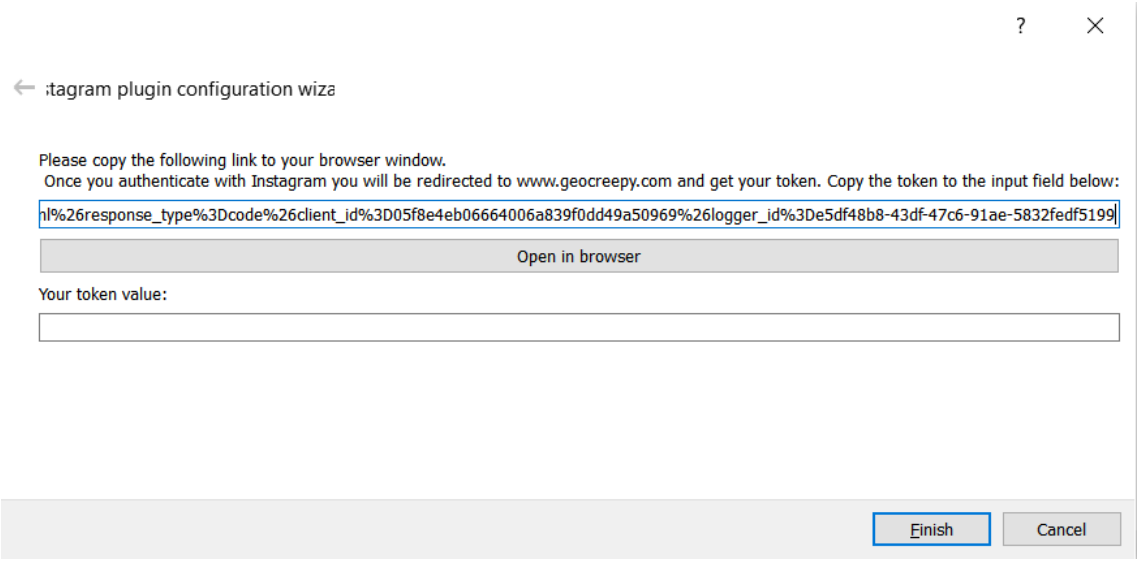


Ilustración 78. Configuración de Instagram dentro de Cree.py para conexión API.

En este caso tras iniciar sesión en Instagram con la cuenta que queremos usar para realizar la investigación, nos retorna error de conexión al realizar la conexión con la API de esta red social, por lo que no es posible obtener el PIN con el cual la herramienta pueda acceder.

{"error_type": "OAuthException", "code": 400, "error_message": "You must include a valid client_id, response_type, and redirect_uri parameters"}

Ilustración 79. Error en la configuración de Instagram y conexión API, conexión rechazada.

Este error 400 que nos retorna Instagram nos quiere decir que la conexión ha sido rechazada y de esta forma, que se produce un error de Bad Request 400.

3.6.5. Búsqueda de un objetivo en redes sociales

Llevaremos a cabo la búsqueda de un nombre de usuario de nuestra víctima por medio de la herramienta. Fijamos como objetivo a la Universidad de Alcalá, nuestro objetivo principal en la investigación.

A su vez, como comprobamos anteriormente tras encontrar la web oficial de la Universidad, el nombre de usuario de esta en redes sociales es @uahes; es por ello por lo que establecemos como nombre de usuario de búsqueda en Cree.py uahes, y realizamos la búsqueda.

Para ello, en primer lugar, crearemos un proyecto enfocado a persona y rellenaremos la información específica del mismo. En nuestro caso investigaremos sobre la Universidad de Alcalá.

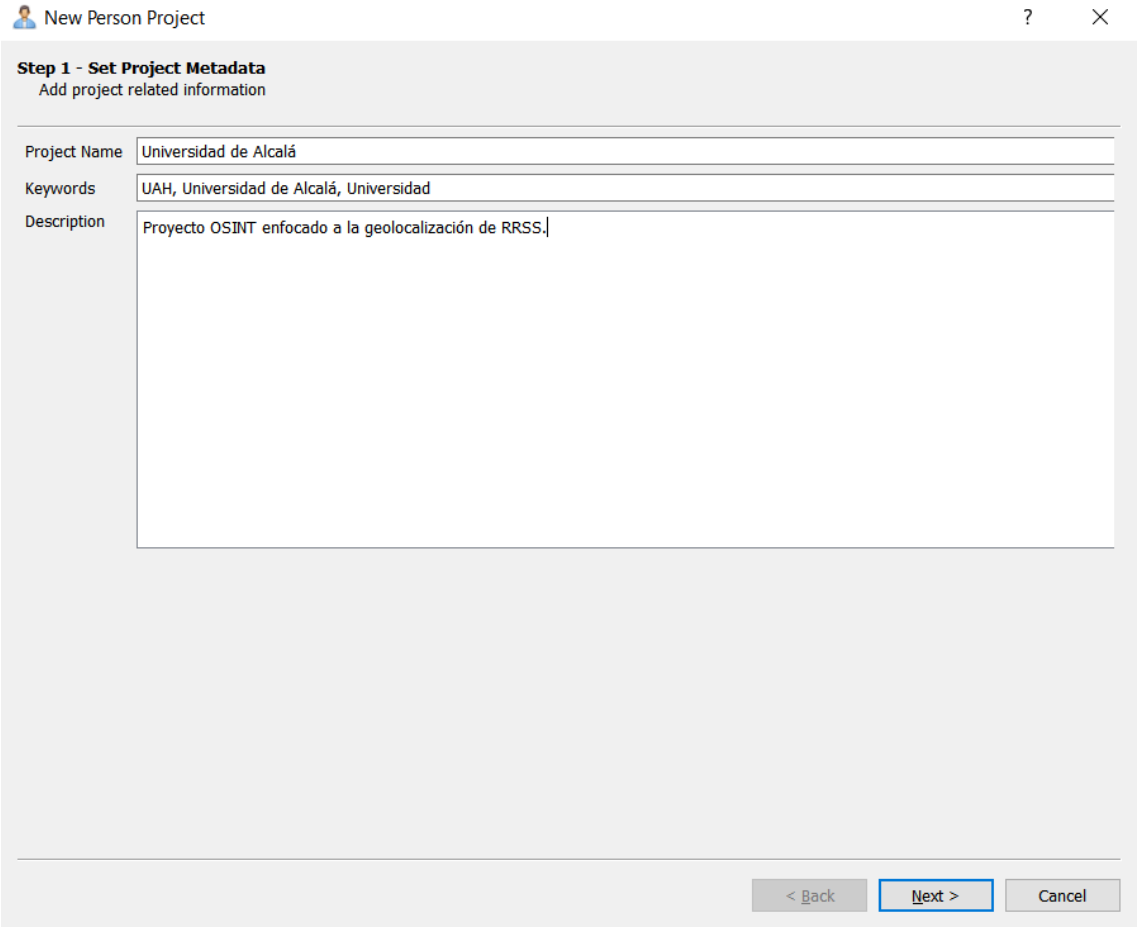


Ilustración 80. Creamos nuestro proyecto dentro de Cree.py destacando diferentes detalles de este.

Fijaremos la búsqueda en Twitter con el username de la Universidad de Alcalá ya sabemos que es legítimo tras varias comprobaciones anteriores, y es uahes.

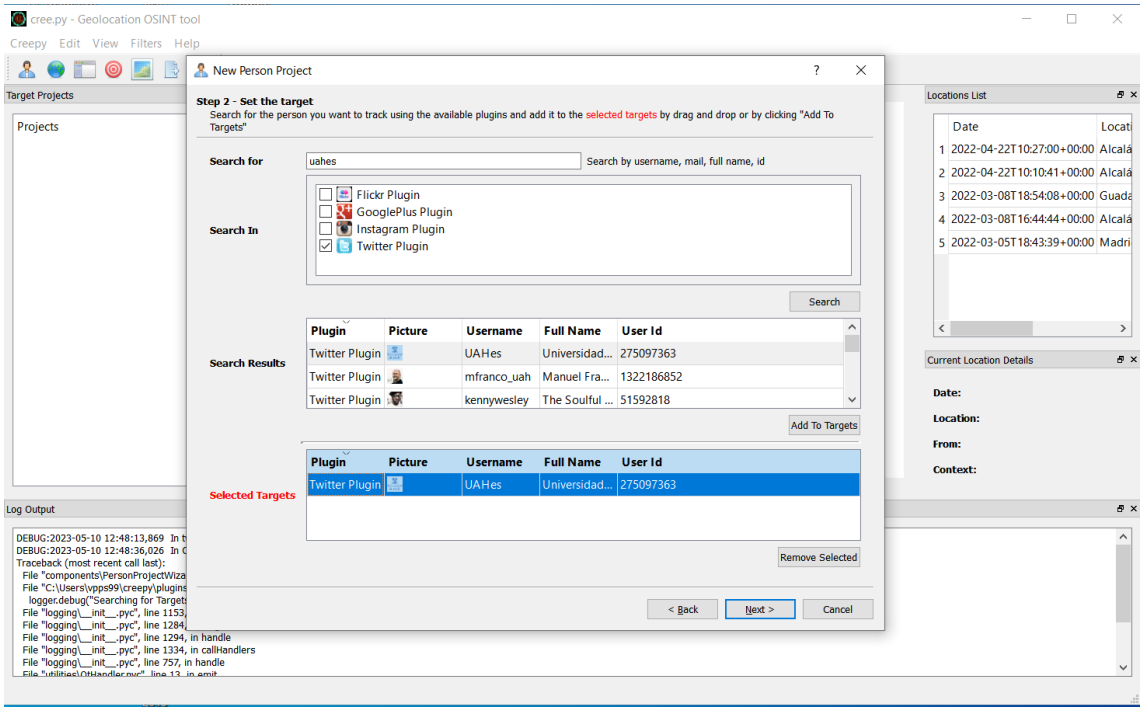


Ilustración 81. Proceso de búsqueda de la Universidad de Alcalá y selección de los objetivos a tratar.

Tras realizar la búsqueda en la red social determinada, en nuestro caso Twitter, nos aparecen diferentes resultados de búsqueda. Seleccionaremos UAHes y lo mandaremos a Selected Targets para que la herramienta lo pueda analizar e investigar. Establecemos las opciones de Twitter a emplear en la búsqueda y finalizamos.

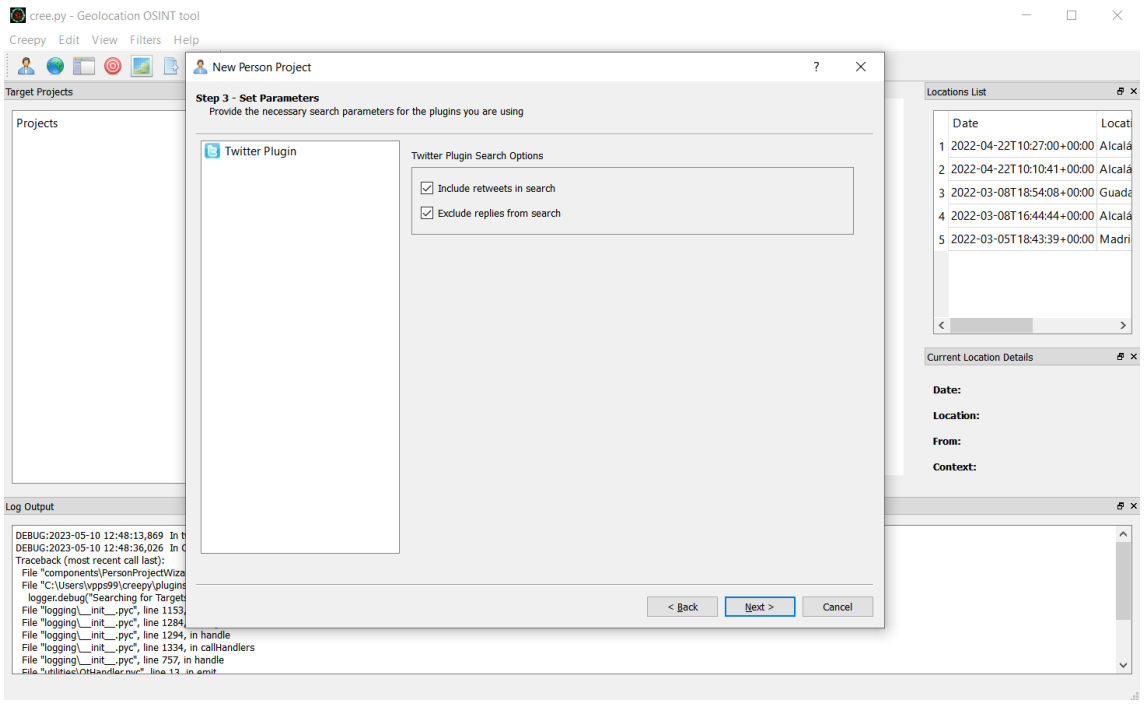


Ilustración 82. Opciones de Twitter establecidas para la investigación.

El siguiente paso será el de analizar e iniciar el proyecto creado para poder obtener la información de la geolocalización de la Universidad de Alcalá.

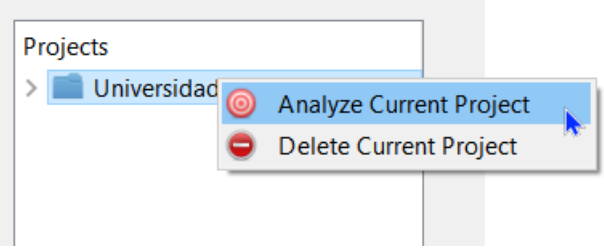


Ilustración 83. Iniciamos el análisis del proyecto creado.

A su vez, la interfaz nos ofrece unos logs de salida para que podamos saber lo que se encuentra realizando y llevando a cabo Cree.py en cada momento.

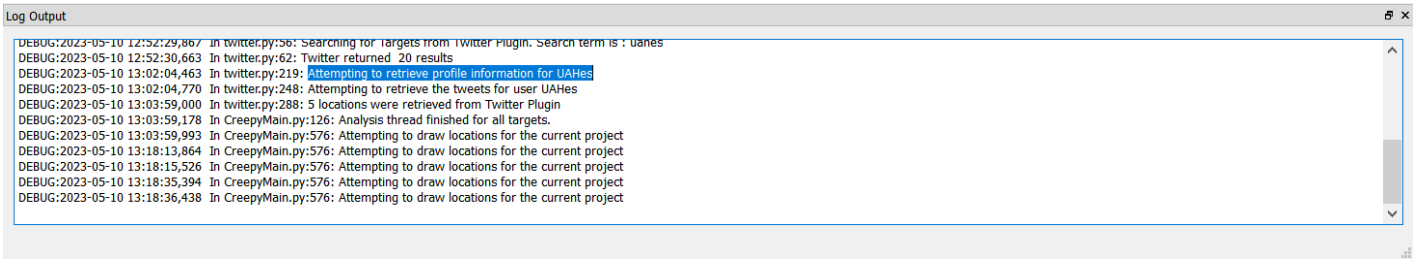


Ilustración 84. Panel de logs dentro de Cree.py donde se nos informa de todo lo que se realiza.

Como observamos en el log, tras analizar los tweets y el usuario UAHes en Twitter, se han encontrado un total de cinco localizaciones para la Universidad de Alcalá; sin embargo, de esas cinco localizaciones, tres son del mismo lugar.

Estas localizaciones se obtienen de los diferentes tweets que se publican, estas son las siguientes: Alcalá de Henares, Guadalajara y Madrid.

	Date	Location
1	2022-04-22T10:27:00+00:00	Alcalá de Henares
2	2022-04-22T10:10:41+00:00	Alcalá de Henares
3	2022-03-08T18:54:08+00:00	Guadalajara
4	2022-03-08T16:44:44+00:00	Alcalá de Henares
5	2022-03-05T18:43:39+00:00	Madrid

Ilustración 85. Localizaciones obtenidas dentro de la Universidad de Alcalá usando Twitter.

Además, Cree.py nos permite observar los diferentes detalles de cada localización obtenida tras su investigación en Twitter.

3.6.5.1. **Localizaciones obtenidas**

Localización 1: Alcalá de Henares, 2022-04-22 10:27:00

Ilustración 86. Primera localización obtenida: Alcalá de Henares, en Cree.py con Twitter como fuente principal.

https://twitter.com/i/web/status/1517449885670420481

Ilustración 87. Tweet de la Universidad de Alcalá donde se encuentra la localización de Alcalá de Henares.

Localización 2: Alcalá de Henares, 2022-04-22 10:10:41

Ilustración 88. Segunda localización obtenida: Alcalá de Henares, en Cree.py con Twitter como fuente principal.

https://twitter.com/UAHes/status/1517445778771435525

Ilustración 89. Tweet de la Universidad de Alcalá donde se encuentra la localización de Alcalá de Henares.

Localización 3: Guadalajara, 2022-03-08 18:54:08

Ilustración 90. Tercera localización obtenida: Guadalajara, en Cree.py con Twitter como fuente principal.

https://twitter.com/i/web/status/1501270055350378498

Ilustración 91. Tweet de la Universidad de Alcalá donde se encuentra la localización de Guadalajara.

Localización 4: Alcalá de Henares, 2022-03-08 16:44:44

Ilustración 92. Cuarta localización obtenida: Alcalá de Henares, en Cree.py con Twitter como fuente principal.

https://twitter.com/i/web/status/1501237493555142656

Ilustración 93. Tweet de la Universidad de Alcalá donde se encuentra la localización de Alcalá de Henares.

Localización 5: Madrid, 2022-03-05 18:43:39

Ilustración 94. Cuarta localización obtenida: Madrid, en Cree.py con Twitter como fuente principal.

https://twitter.com/UAHes/status/1500180256958861312

Ilustración 95. Tweet de la Universidad de Alcalá donde se encuentra la localización de Madrid.

3.6.5.2. Información de Twitter obtenida

Podemos observar diferente información que nos ofrece Cree.py acerca de la cuenta de Twitter de la Universidad de Alcalá.

A primera instancia la herramienta nos ofrece datos relevantes a la cantidad de tweets, seguidos, seguidores, listas… Información correspondiente a la ubicación, descripción y permiso para geolocalizar tweets.

Twitter profile information

Account was created on 2011-03-31 16:34:35

The user has tweeted 39347 times (including retweets).

Self-reported real name: Universidad Alcalá

Description: Ubicada en Alcalá de Henares, ciudad Patrimonio de la Humanidad. Es una de las instituciones con mayor tradición e historia. https://t.co/RZhGLdzZ0N

Self-reported location: España

User has enabled the possibility to geolocate their tweets.

The user has 59248 followers.

The user is following 99 users.

The user is listed in 750 public lists.

Ilustración 96. Información general relevante a la cuenta de Twitter de la Universidad de Alcalá.

Tenemos datos relevantes a los usuarios que han hecho retweet a los tweets de la Universidad de Alcalá.

User has retweeted the following users :

User screen name	Count
Cultura UAH □	146
AlumniUAH	53
OTRI Univ. de Alcalá	53
madrimasd	35
Instituto Franklin-UAH	30
Deportes UAH	28
EPS UAH	27
Ayuntamiento de Alcalá de Henares	26
IQ de las Artes del Humor	21
Facultad C. Económicas, Empresariales y Turismo	18

Ilustración 97. Información sobre los usuarios que interactúan con la cuenta de Twitter de la Universidad de Alcalá.

También, datos sobre los diferentes clientes con los cuales se ha usado esta red social por parte de la Universidad de Alcalá: Hootsuite, web, Android y iPhone.

User has been using the following clients :

Client Application	Count
Hootsuite Inc.	1909
Twitter Web App	1095
Twitter for iPhone	80
Twitter for Android	68

Ilustración 98. Información sobre los clientes que se usan con la cuenta de Twitter de la Universidad de Alcalá.

A su vez, acerca de la frecuencia horaria de la publicación de tweets.

User tweets time frequency:

Hour Period	Count
8-15	2518
9-16	2503
7-14	2342
10-17	2277
6-13	2115
11-18	1962
5-12	1839
12-19	1722
4-11	1414
13-20	1307
3-10	1162
14-21	1035
2-9	828
15-22	810
16-23	626
1-8	383
17-0	266
18-1	47
19-2	28
20-3	16
0-7	8
21-4	6
22-5	2
23-6	0

Ilustración 99. Información sobre la frecuencia horaria de publicación de la cuenta de Twitter de la Universidad de Alcalá.

Por último, cuándo la Universidad de Alcalá ha estado publicando tweets.

User has been twitting at the following times :

Hour	Count
9	445
12	425
8	375
16	360
10	334
13	276
11	252
14	227
17	219
15	184
18	19
19	12
20	10
7	8
21	4
22	2

Ilustración 100. Información sobre la frecuencia de publicación de la cuenta de Twitter de la Universidad de Alcalá.

3.6.6. Conclusiones de Cree.py

Con Cree.py hemos podido ver de qué forma podemos obtener información relevante a la geolocalización que podemos obtener desde diferentes redes sociales. Esta investigación se basa en los diferentes posts que un usuario sube a sus redes sociales y que se puede averiguar fácilmente.

Además, hay que destacar la posibilidad que existe de extraer el proyecto en diferentes formas tipos de ficheros para un análisis más exhaustivo con la información más a mano.

Por último, añadir la posibilidad de geolocalización sobre mapa que incorpora la herramienta en su interfaz, pero que no hemos podido disfrutar de ella debido a diferentes problemas de compatibilidad de navegador.

3.7. FOCA: obtención de metadatos

FOCA (Fingerprinting Organisations with Collected Archives) se trata de una herramienta que nos permite obtener información oculta y metadatos a partir de documentos; estos se pueden subir a la misma, o bien, son encontrados en dominios.

3.7.1. Instrumentos empleados

Máquina atacante: **Windows 10**
Herramientas:

- **FOCA: https://github.com/ElevenPaths/FOCA**
- **Redes sociales:**
  - **Flickr: https://www.flickr.com/groups/uah/**

3.7.2. Puesta en escena de FOCA

Para llevar a cabo el despliegue de la herramienta **FOCA**, hemos decidido implementarla en nuestro entorno habitual. Esta elección se basa en varias consideraciones estratégicas y prácticas.

Además de los métodos tradicionales de investigación, también hemos utilizado las redes sociales como *Twitter, Facebook, Instagram y Flickr* para obtener información adicional. Estas plataformas nos han brindado acceso a valiosos metadatos y datos ocultos que se encuentran en los documentos. Al explorar estas redes sociales, hemos podido recopilar información relevante y descubrir conexiones o patrones que podrían pasar desapercibidos de otra manera.

En esta puesta en escena se llevará a cabo la **búsqueda a un objetivo.**

3.7.3. FOCA: el entorno

Nos descargamos la herramienta de FOCA (Fingerprinting Organisations with Collected Archives), la desplegamos y observamos todas las opciones disponibles dentro de esta.

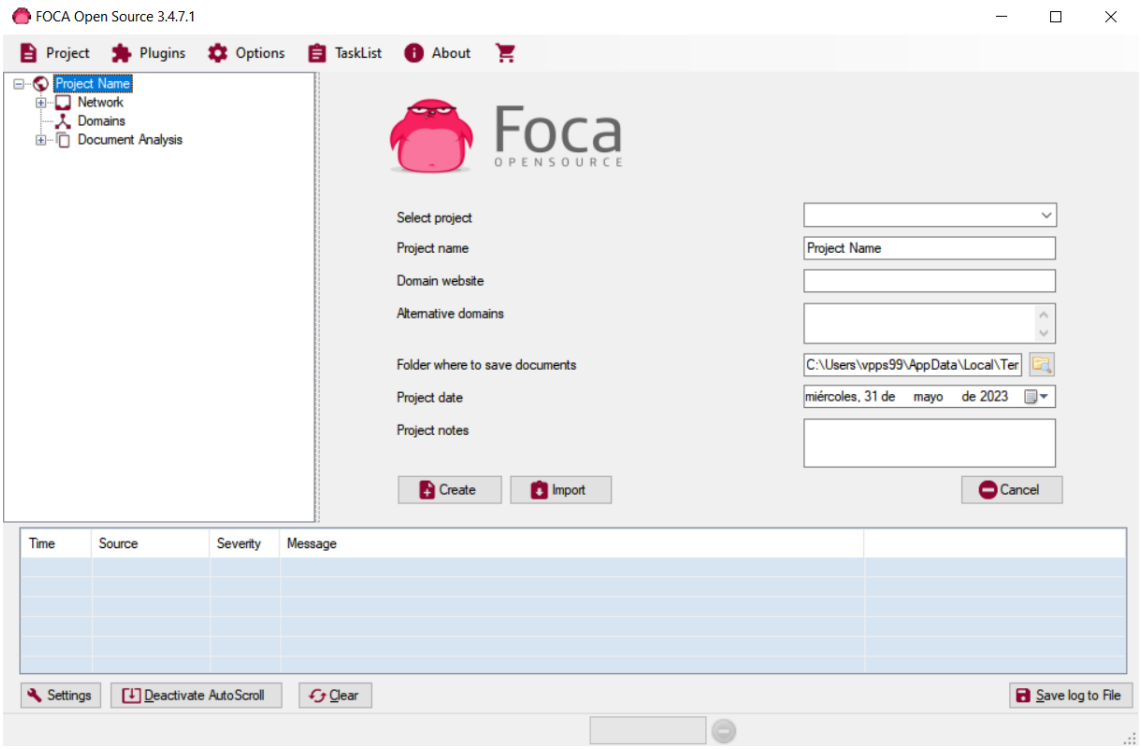


Ilustración 101. Interfaz de la herramienta de FOCA.

Como podemos observar, hay diferentes opciones dentro de la herramienta, nosotros iremos al objetivo que queremos, encontrar metadatos e información oculta a partir de documentos en redes sociales de nuestro objetivo.

La herramienta puede analizar diferentes tipos de documentos, como los de Microsoft Office, Open Office, PDF e incluso Adobe InDesign y svg. Para encontrar estos documentos, se utiliza Google, Bing y DuckDuckGo como buscadores, lo que permite obtener una gran cantidad de resultados. También es posible agregar archivos locales para extraer información EXIF de imágenes y se realiza un análisis previo de la URL antes de descargar el archivo.

Una vez que se recopila toda la información de los diferentes archivos, la herramienta FOCA busca patrones y trata de identificar qué documentos se crearon en el mismo equipo y qué servidores y clientes están relacionados con ellos. En resumen, FOCA recopila datos de varios tipos de documentos y busca conexiones entre ellos para obtener información sobre su origen y las personas o servidores asociados.

Nosotros realizaremos la investigación de la Universidad de Alcalá, gracias a la información recolectada con las herramientas expuestas y mostradas durante el trabajo de investigación.

3.7.4. Búsqueda de metadatos en dominio

Creamos nuestro proyecto, al que llamaremos Metadata UAH y estableceremos el dominio www.uah.es; de esta forma fijamos como objetivo a la Universidad de Alcalá, nuestro objetivo principal en la investigación.

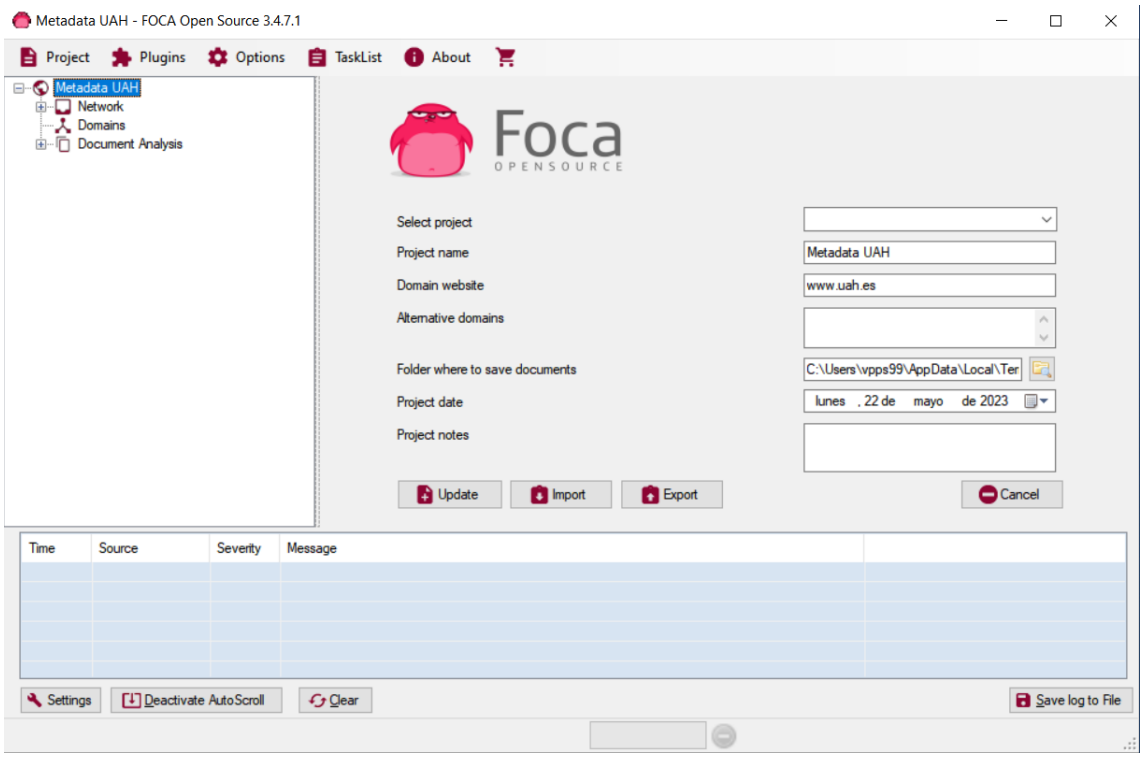


Ilustración 102. Creamos nuestro proyecto en FOCA y establecemos el dominio de la Universidad de Alcalá.

Para ello, en primer lugar, seleccionaremos los buscadores en los que queramos que FOCA realice la búsqueda y el tipo de extensión de los documentos. Fijamos y seleccionamos todos los buscadores y extensiones para de esta forma poder encontrar más información al existir mayor cantidad de datos a analizar.

Una vez seleccionadas todas las opciones disponibles, entonces procedemos a realizar la búsqueda.

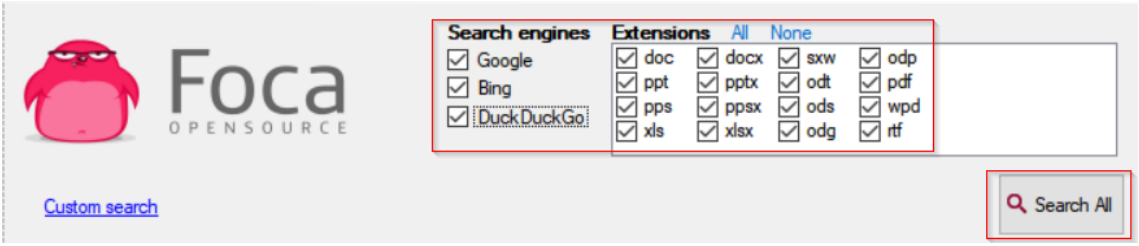


Ilustración 103. Seleccionamos todos los buscadores y extensiones disponibles dentro de FOCA para el análisis.

Mientras se está llevando a cabo la búsqueda, podremos ir visualizando los archivos que se van encontrando. Además, tenemos una consola que nos muestra los logs y todo lo que se va realizando en la herramienta.

Como podemos comprobar en la consola de logs, únicamente ha podido realizar la búsqueda mediante el buscador de Bing, donde ha encontrado un total de 70 documentos.

Los otros buscadores no han logrado encontrar ningún documento:

- Google, error 429 ocurre cuando el servidor recibe demasiadas solicitudes, lo que afecta a las API de terceros y al rastreo del sitio web por los motores de búsqueda.

- DuckDuckGo, error 403 es causado por una solicitud incorrecta del cliente y resulta en la prohibición de acceso a la página que se intenta abrir en el navegador.

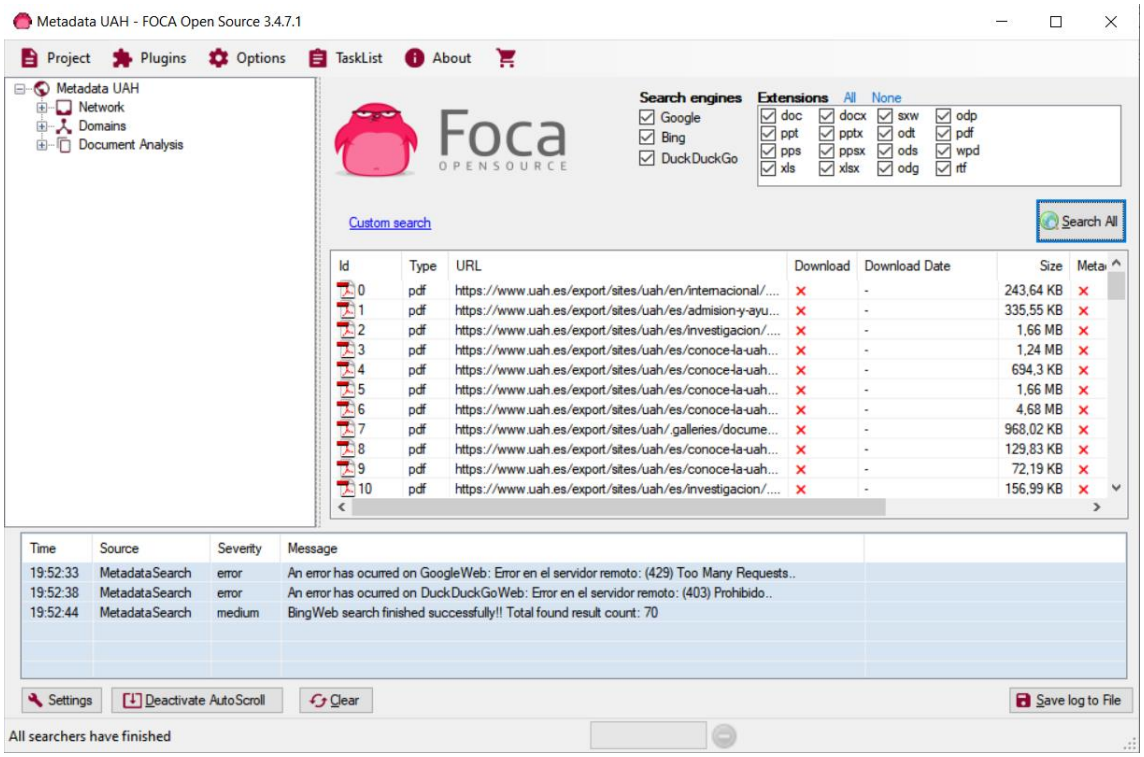


Ilustración 104. Proceso de FOCA con nuestro proyecto donde vemos los archivos encontrados y los fallos en los buscadores de Google y DuckDuckGo.

Ahora bien, deberemos de descargar todos los documentos para albergarlos en la herramienta; por ello, descargaremos los 72 ficheros que se han encontrado.

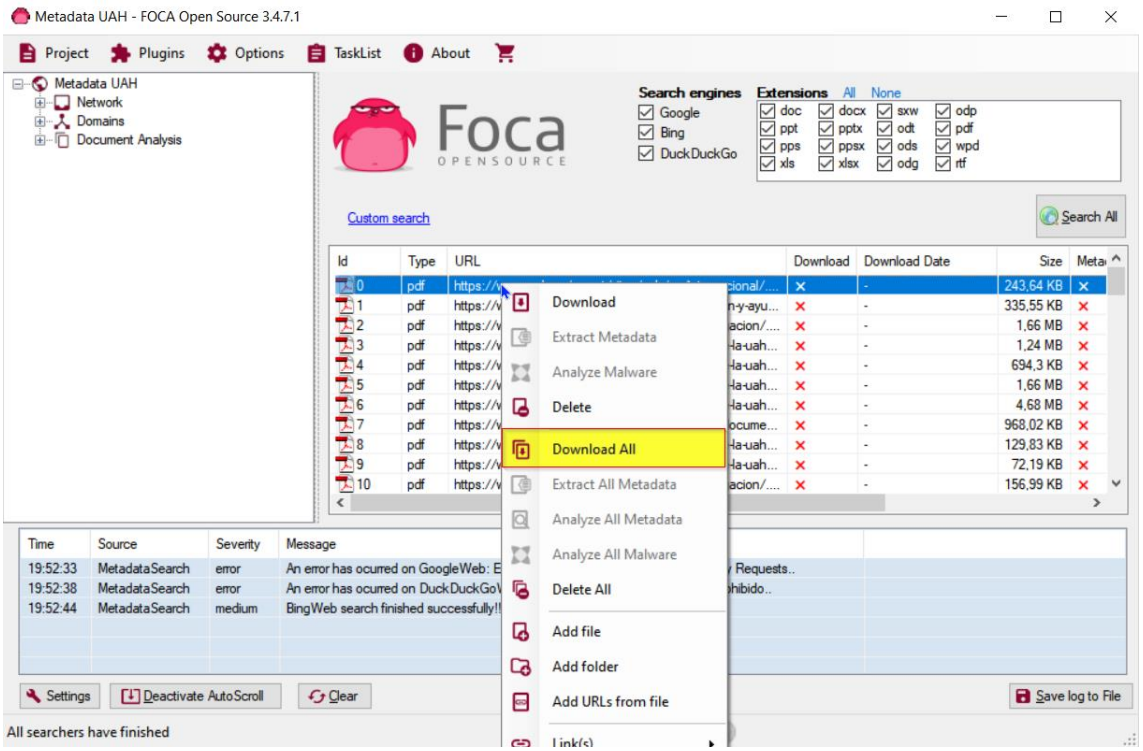


Ilustración 105. Procedemos a descargar todos los ficheros encontrados por FOCA en el dominio de la Universidad de Alcalá.

Una vez descargados los 72 ficheros encontrados, toca extraer los diferentes metadatos e información oculta que se puedan encontrar en ellos.

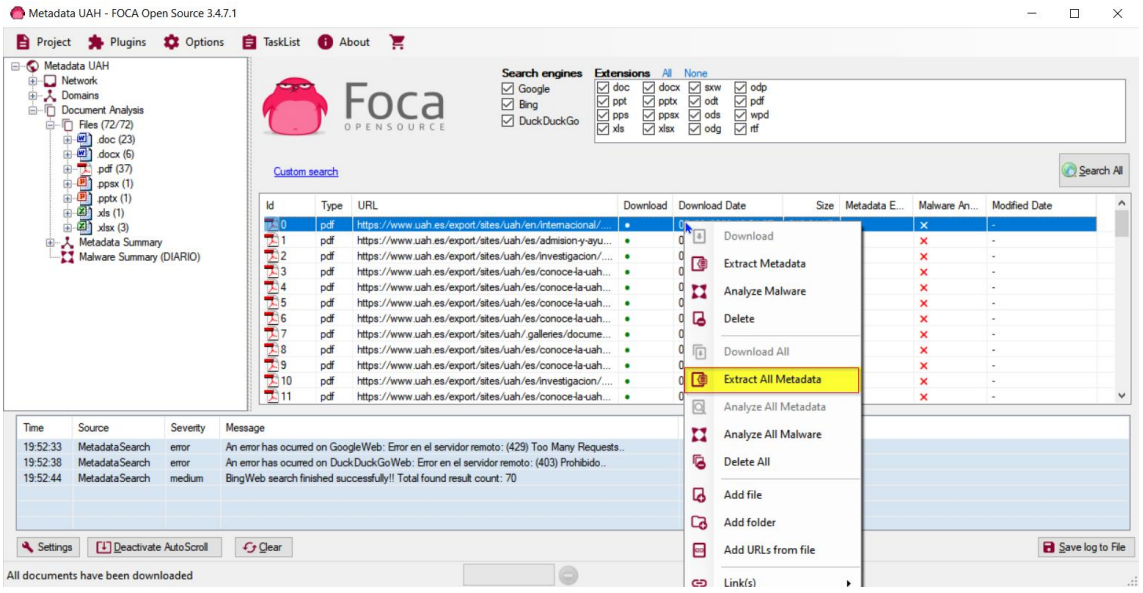


Ilustración 106. Una vez descargados los ficheros, realizamos la extracción de todos los metadatos que estos puedan contener.

En la barra lateral izquierda podemos apreciar los 72 documentos descargados diferenciados por su tipo de extensión. Más abajo, observamos los metadatos encontrados desplegados por secciones de información.

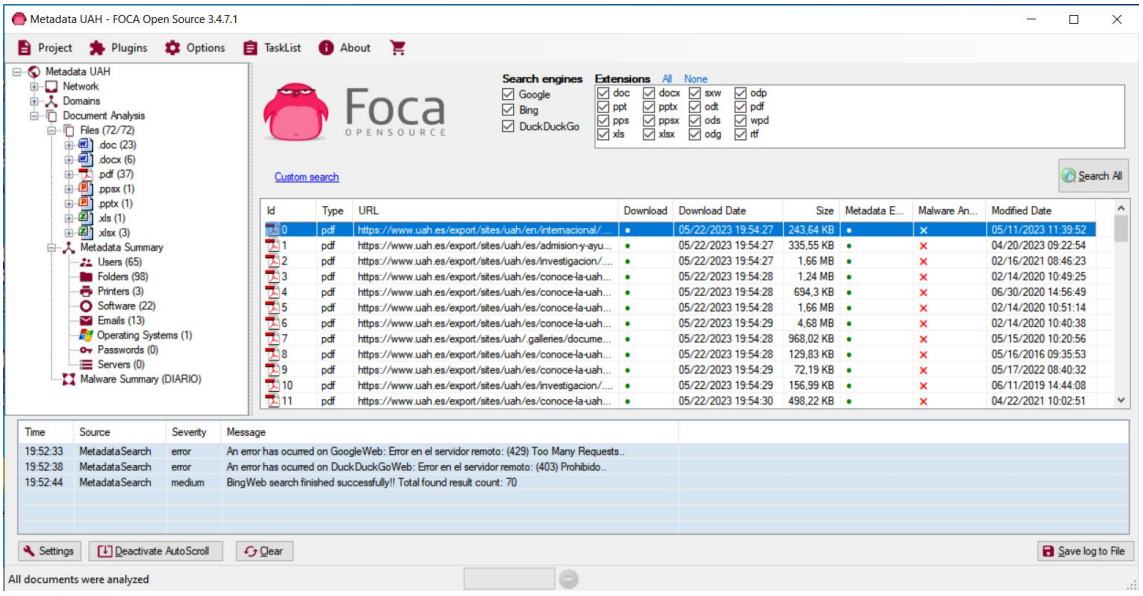


Ilustración 107. Una vez descargados y extraídos los ficheros, podemos apreciar los diferentes metadatos encontrados.

Los metadatos encontrados han sido: 65 usuarios, 98 carpetas, 3 impresoras, 22 softwares, 13 correos electrónicos y un sistema operativo.

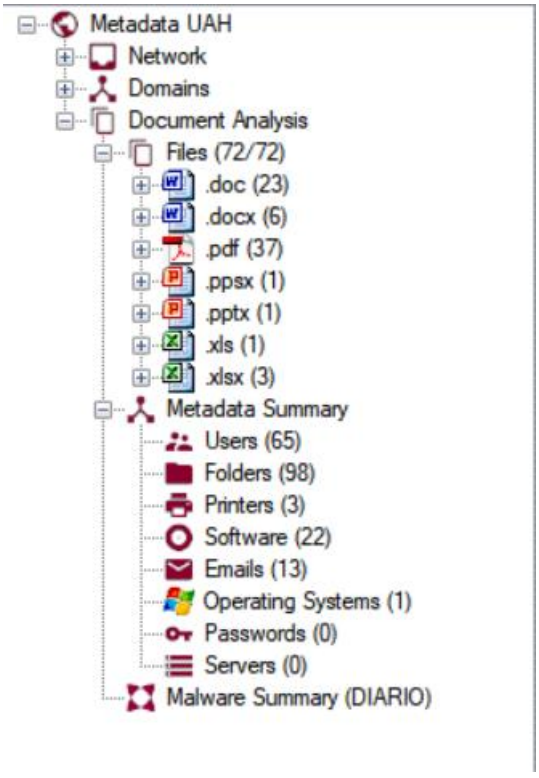


Ilustración 108. Resumen de los metadatos encontrados con FOCA sobre los archivos descargados con el dominio de la Universidad de Alcalá.

Metadatos en dominio: usuarios

Los usuarios encontrados en los diferentes documentos que se han analizado han sido:

Attribute	Value
Name	Martínez Fernández Raúl
Name	raul.martinez
Name	Gómez González Rafael
Name	Julio García Arribas
Name	UAH
Name	Raul
Name	Martínez Fernández Raúl
Name	Sandín Vázquez M. del Val
Name	usuario
Name	Usuario
Name	Buitrago Jiménez Daniel
Name	Buitrago JimÃ©nez Daniel
Name	daniel.buitrago
Name	Mangada Cañas Patricia
Name	Fernando Gomez Hermoso
Name	UNIVERSIDADES
Name	Mangada Cañas Patricia
Name	Mangada CaÃ±as Patricia
Name	Mangada Cañas Patricia

Ilustración 109. Usuarios encontrados en los metadatos extraídos de los ficheros sacados del dominio de la Universidad de Alcalá.

Metadatos en dominio: carpetas

Las carpetas encontradas en los diferentes documentos que se han analizado han sido:

Attribute	Value
Path	https://vfirma.uah.es/vfirma/code/
Path	https://www.linkedin.com/in/fernando-g%C3%B3mez-98b8154b/
Path	https://www.linkedin.com/in/david-garcia-arrate-1270b0136/
Path	https://teams.microsoft.com/l/meetup-join/19%3ameeting_YTA4NWU1MzUtY2I...
Path	https://www.ikiam.edu.ec/
Path	https://ka107.web.uah.es/
Path	https://www.uah.es/export/sites/uah/es/admision-y-ayudas/.galleries/descarg...
Path	https://www.uah.es/export/sites/uah/es/conoce-la-uah/organizacion-y-gobiern...
Path	http://www3.uah.es/ka107/Forms/
Path	http://www3.uah.es/ka107/
Path	https://www.google.com/maps/d/
Path	https://www.mdp.edu.ar/
Path	https://www.ikiam.edu.ec/
Path	https://www2.unesp.br/
Path	http://www.univ-maroua.cm/
Path	https://www.ufh.ac.za/
Path	http://www.univ-fhb.edu.ci/
Path	http://www.cu.edu.ge/
Path	http://www.gau.ge/

Ilustración 110. Carpetas encontradas en los metadatos extraídos de los ficheros sacados del dominio de la Universidad de Alcalá.

Metadatos en dominio: impresoras

Las impresoras encontradas en los diferentes documentos que se han analizado han sido:

Attribute	Value
All printers found (5) - Times found	
Printer Name	Becas
Printer Name	RICOH MP C307
Printer Name	Becas
Printer Name	Adobe PDF

Ilustración 111. Impresoras encontradas en los metadatos extraídos de los ficheros sacados del dominio de la Universidad de Alcalá.

Metadatos en dominio: software

Los softwares encontrados en los diferentes documentos que se han analizado han sido:

Attribute	Value
Software	Microsoft Office
Software	Microsoft Office
Software	Adobe InDesign 15.1 (Macintosh)
Software	Adobe PDF Library 15.0
Software	Microsoft Office
Software	Microsoft Office
Software	Microsoft Office
Software	Adobe PDF Library 15.0
Software	Microsoft Office
Software	Adobe Illustrator 24.1 (Macintosh)
Software	Adobe PDF Library 15.00
Software	bizhub C3110
Software	bizhub C3110
Software	Microsoft Office
Software	bizhub C3110
Software	Adobe InDesign CC 2015 (Macintosh)
Software	Adobe PDF Library 15.0
Software	Microsoft Office 95
Software	Adobe PDF Library 17.11.238

Ilustración 112. Softwares encontrados en los metadatos extraídos de los ficheros sacados del dominio de la Universidad de Alcalá.

Metadatos en dominio: correos electrónicos

Los correos electrónicos encontrados en los diferentes documentos que se han analizado han sido:

Email	ka107.outgoing@uah.es
Email	roberto.cuellar@uah.es
Email	barbara.navarro@uah.es
Email	otriuah@uah.es
Email	ka171@uah.es
Email	internacional.becas@uah.es
Email	erasmus.outgoing@uah.es
Email	noerasmus.incoming@uah.es
Email	erasmus.incoming@uah.es
Email	erasmus.bilateral@uah.es
Email	alba.yela@uah.es
Email	jose.lafuente@uah.es
Email	vicer.mi@uah.es

Ilustración 113. Correos electrónicos encontrados en los metadatos extraídos de los ficheros sacados del dominio de la Universidad de Alcalá.

Metadatos en dominio: sistemas operativos

Los sistemas operativos encontrados en los diferentes documentos que se han analizado han sido:

All operating systems found (0) - Times found

OS	Windows 7
OS	Windows 7

Ilustración 114. Sistemas operativos encontradas en los metadatos extraídos de los ficheros sacados del dominio de la Universidad de Alcalá.

3.7.5. Búsqueda de metadatos en redes sociales

Creamos un nuevo proyecto, METADATA RRSS UAH, y fijamos objetivo a la Universidad de Alcalá, nuestro objetivo principal en la investigación, a partir de sus redes sociales legítimas.

Después de llevar a cabo una serie de casos de prueba utilizando contenido multimedia de diversas redes sociales de la UAH, hemos realizado interesantes descubrimientos. Al analizar *Instagram, Twitter y Facebook*, hemos constatado que estos sitios no contienen metadatos ni información oculta relevante para nuestros propósitos de investigación.

Sin embargo, al explorar la plataforma de *Flickr*, hemos encontrado una fuente rica en contenido para analizar.

Seleccionamos un total de tres imágenes que hemos encontrado en Flickr, en la red social legítima de la UAH: https://www.flickr.com/groups/uah/

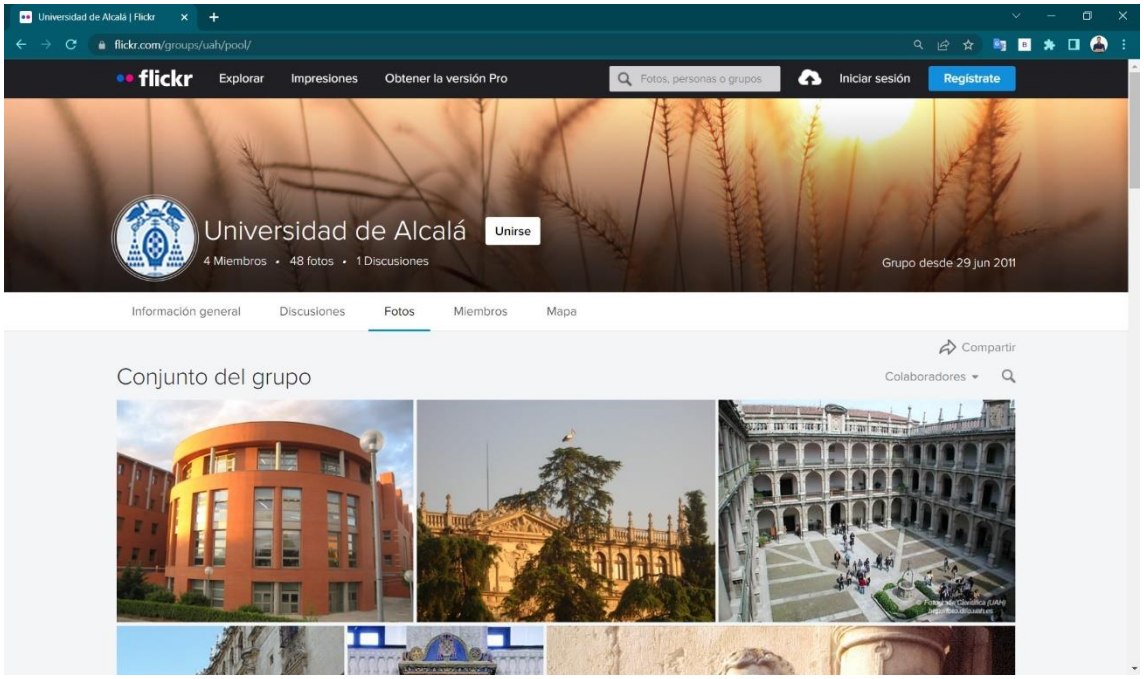


Ilustración 115. Perfil oficial de la Universidad de Alcalá en Flickr.

Las descargamos en nuestro equipo, para de forma posterior, subirlas a FOCA:

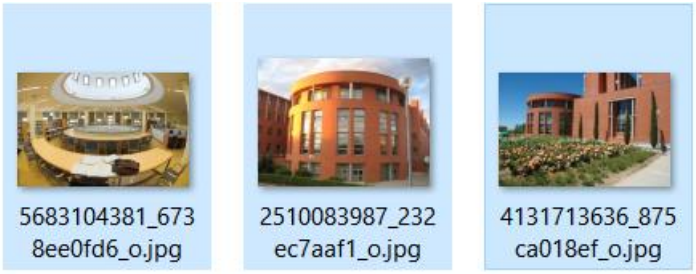


Ilustración 116. Imágenes seleccionadas para realizar la investigación y análisis con FOCA.

Subimos las tres imágenes descargadas del Flickr oficial de la Universidad de Alcalá a la herramienta. Después, toca extraer los diferentes metadatos e información oculta que se puedan encontrar en ellas.

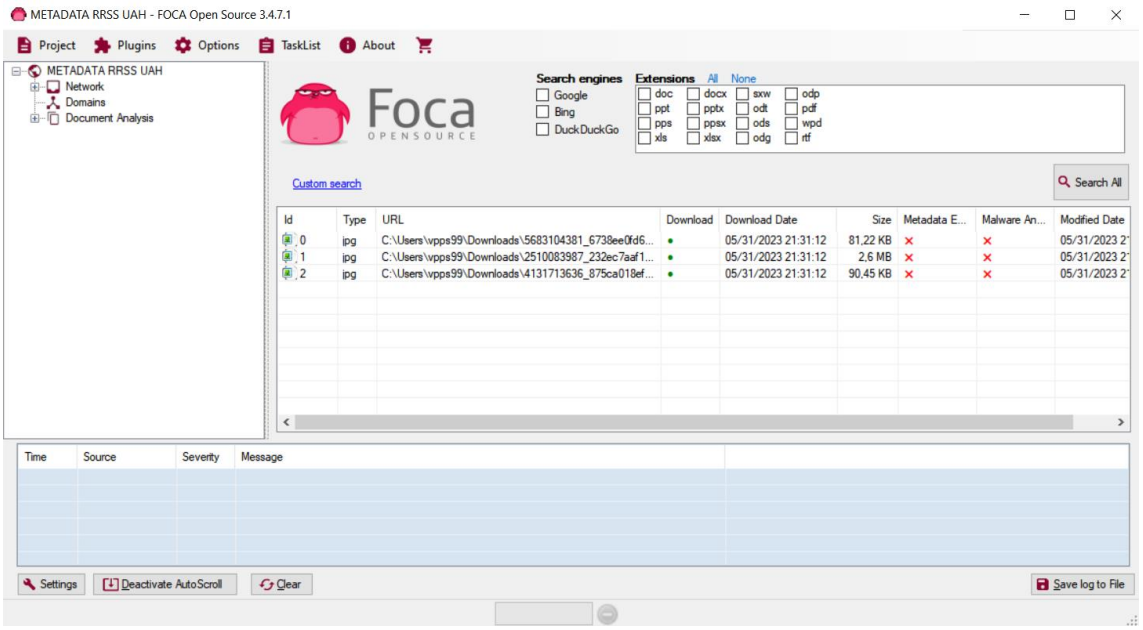


Ilustración 117. Creamos nuevo proyecto en FOCA y añadimos las imágenes seleccionadas del Flickr de la Universidad de Alcalá.

Los metadatos encontrados han sido:

- Información de la cámara: Los metadatos EXIF pueden revelar el modelo de la cámara utilizada para capturar la imagen. Esto puede ser útil para determinar las características técnicas de la cámara y cómo pueden afectar la calidad y el estilo de la imagen.

- Ajustes de exposición: Los datos de exposición, como la velocidad de obturación, la apertura y la sensibilidad ISO, permiten comprender cómo se configuró la cámara para capturar la imagen.

- Fecha y hora: Los metadatos EXIF también incluyen la fecha y la hora en que se tomó la imagen. Esto es útil para organizar y clasificar las imágenes cronológicamente y tener un registro preciso del momento en que se realizó la captura.

- Información de derechos de autor: Algunos usuarios de Flickr pueden agregar información sobre los derechos de autor y las restricciones de uso en los metadatos de la imagen.

Lo encontrado ha sido: imagen 1 (fechas, otros y EXIF - estándar de metadatos utilizado en imágenes digitales); imagen 2 (EXIF); imagen 3 (fechas, software y EXIF).

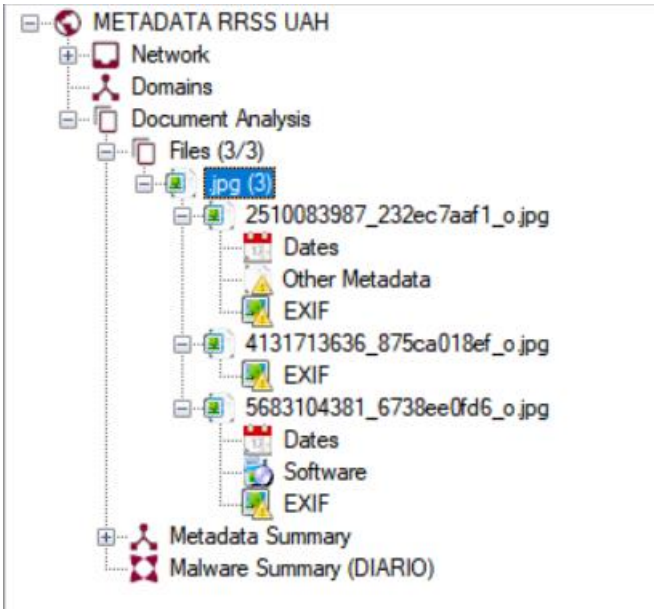
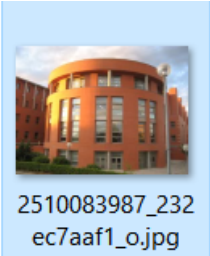


Ilustración 118. Metadatos encontrados en las imágenes analizadas sacadas del Flickr oficial de la Universidad de Alcalá.

Metadatos en redes sociales: imagen 1



Los metadatos encontrados en la imagen 1, que proviene de la plataforma de Flickr, proporcionan una valiosa información adicional sobre la imagen y su contexto.

Observamos diferente información general de la imagen, desde el origen de la propia imagen en nuestro equipo, cámara y día de cuando tuvo lugar la fotografía.

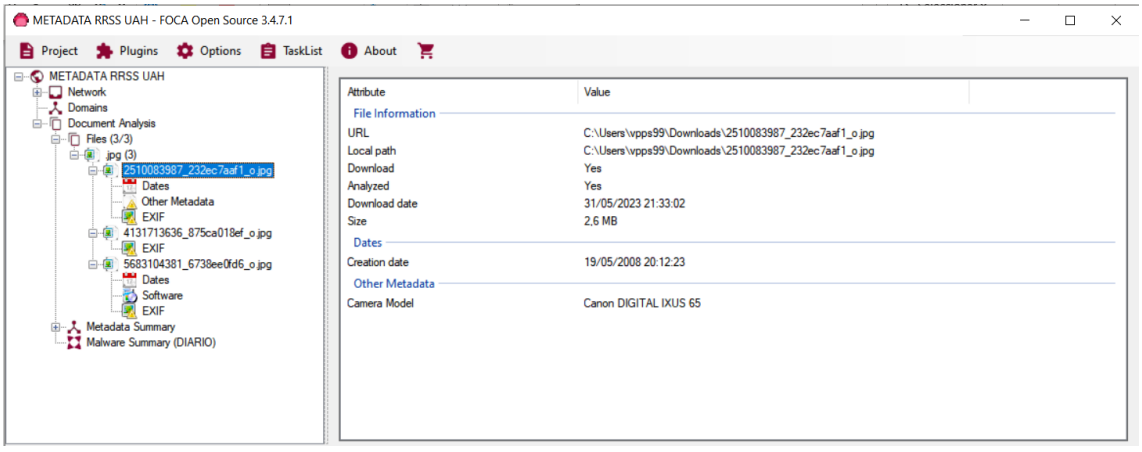


Ilustración 119. Información general de la primera imagen a analizar en FOCA.

Información relevante a fecha, nos muestra la fecha donde tuvo lugar la fotografía: 19/05/2008 a las 20:12:23h.

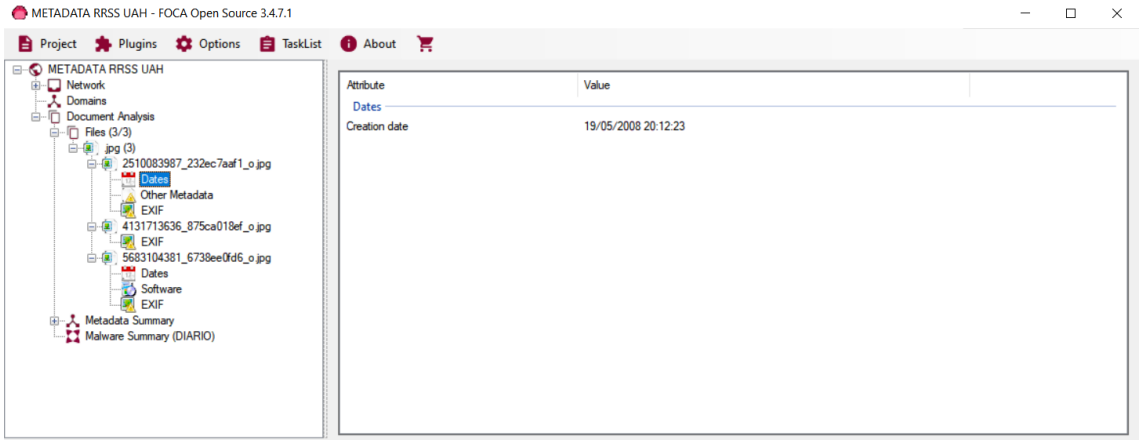


Ilustración 120. Información relevante a la fecha de la primera imagen a analizar en FOCA.

Información relevante a otros metadatos, que nos muestra el modelo de cámara con el que tuvo lugar la fotografía: Canon DIGITAL IXUS 65.

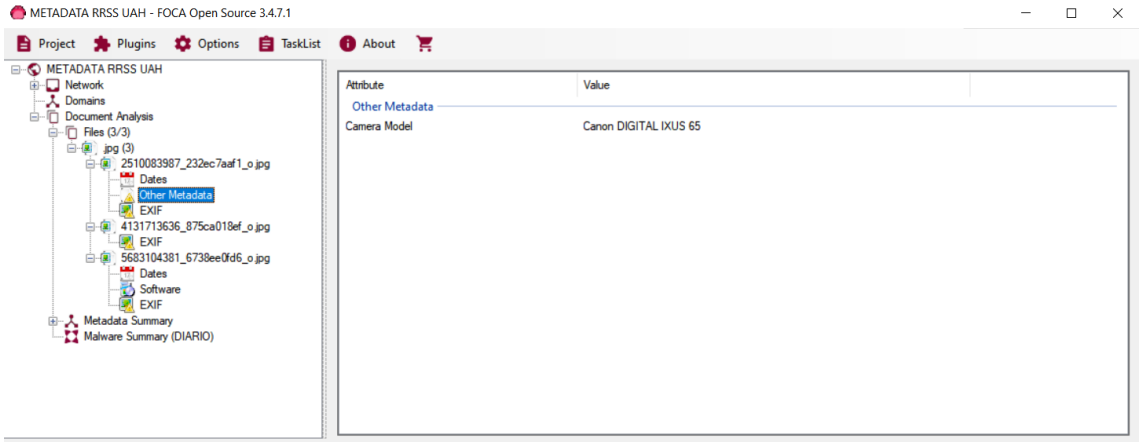


Ilustración 121. Información relevante al modelo de cámara de la primera imagen a analizar en FOCA.

Información relevante a EXIF, que nos muestra el modelo de cámara con el que tuvo lugar la fotografía: Exif Ifd0, Exif SubIFD...

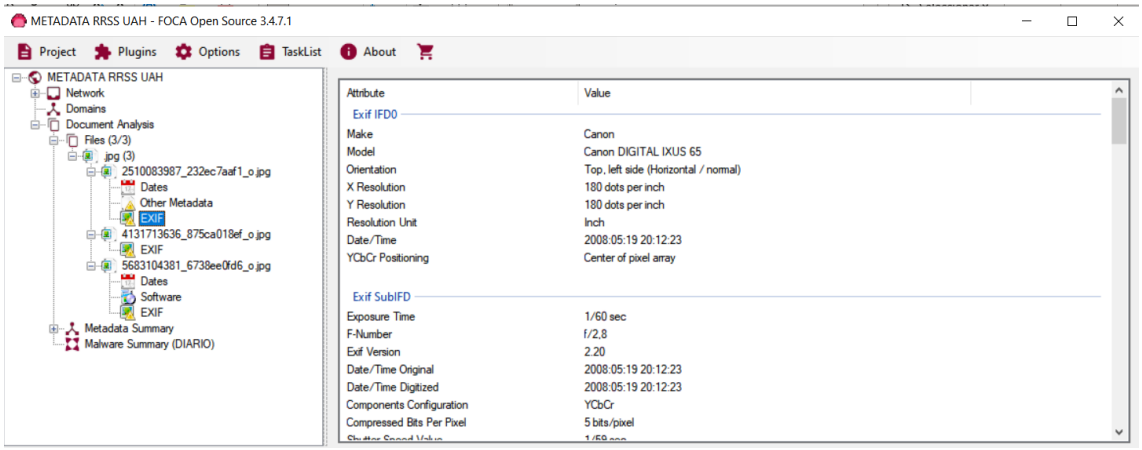
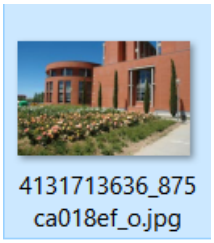


Ilustración 122. Información relevante a EXIF de la primera imagen a analizar en FOCA.

Metadatos en redes sociales: imagen 2



Los metadatos encontrados en la imagen 2, que proviene de la plataforma de Flickr, proporcionan una valiosa información adicional sobre la imagen y su contexto.

Observamos diferente información general de la imagen, desde el origen de la propia imagen en nuestro equipo, hasta el peso de memoria de esta.

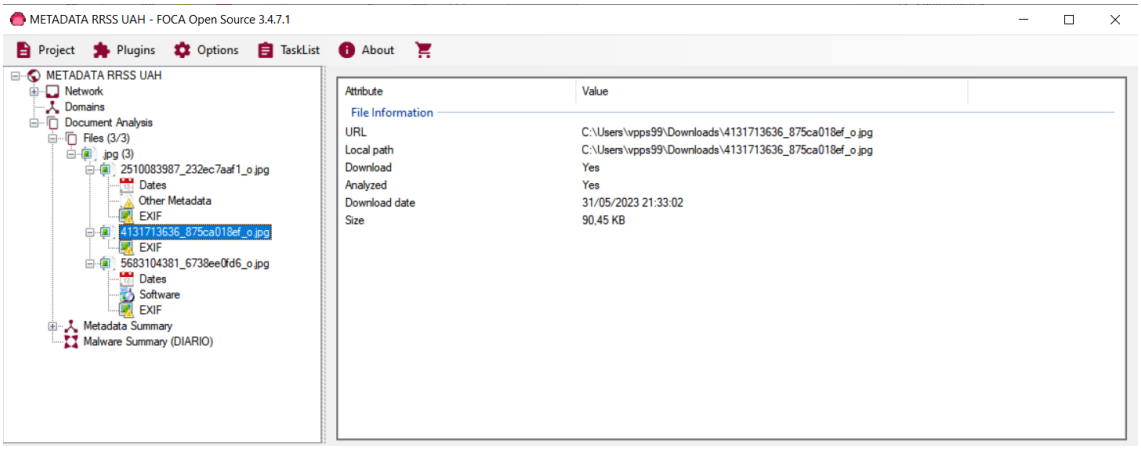


Ilustración 123. Información general de la segunda imagen a analizar en FOCA.

Información relevante a EXIF, que nos muestra el modelo de cámara con el que tuvo lugar la fotografía: Ducky y Adobe JPEG.

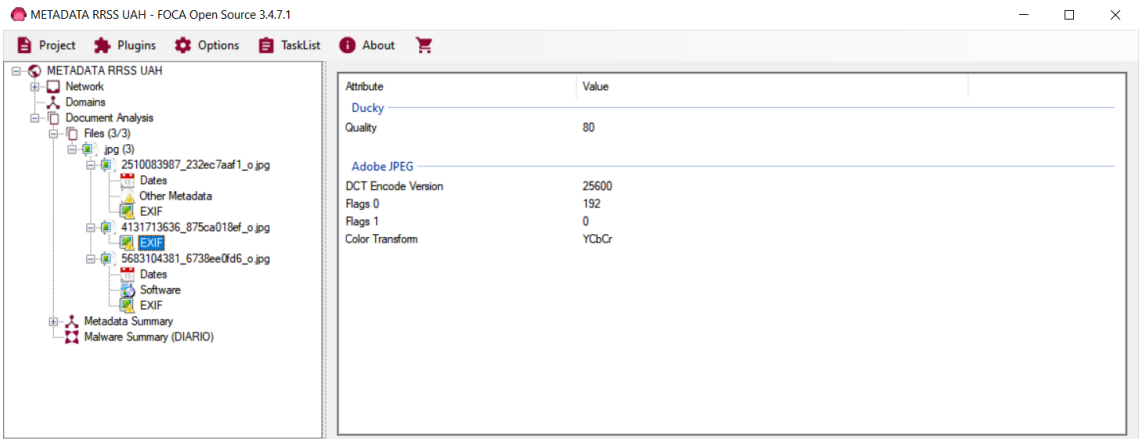
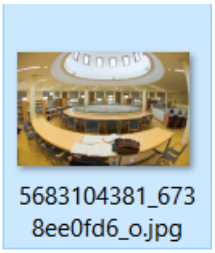


Ilustración 124. Información relevante a EXIF de la segunda imagen a analizar en FOCA.

Metadatos en redes sociales: imagen 3



Los metadatos encontrados en la imagen 3, que proviene de la plataforma de Flickr, proporcionan una valiosa información adicional sobre la imagen y su contexto.

Observamos diferente información general de la imagen, desde el origen de la propia imagen en nuestro equipo, hasta la fecha de creación de esta.

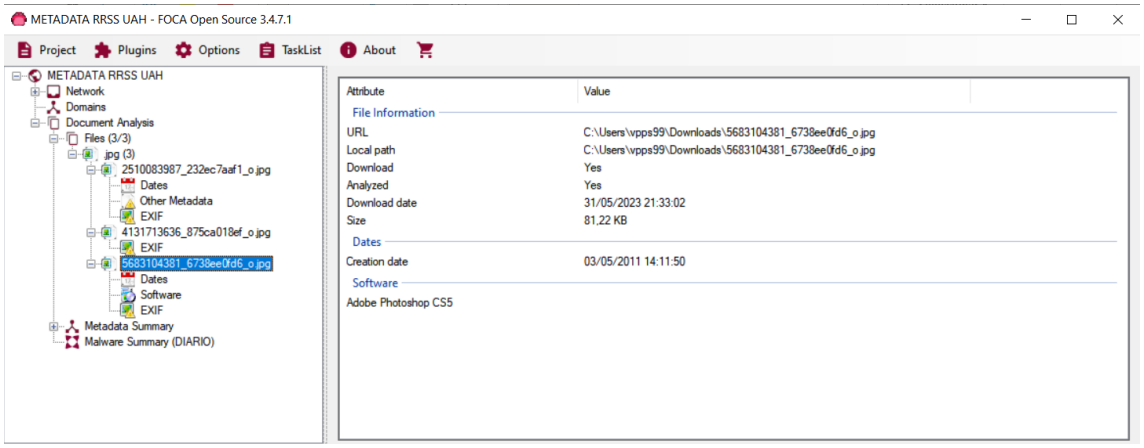


Ilustración 125. Información general de la tercera imagen a analizar en FOCA.

Información relevante a fecha, nos muestra la fecha donde tuvo lugar la fotografía: 03/05/2011 a las 14:11:50h.

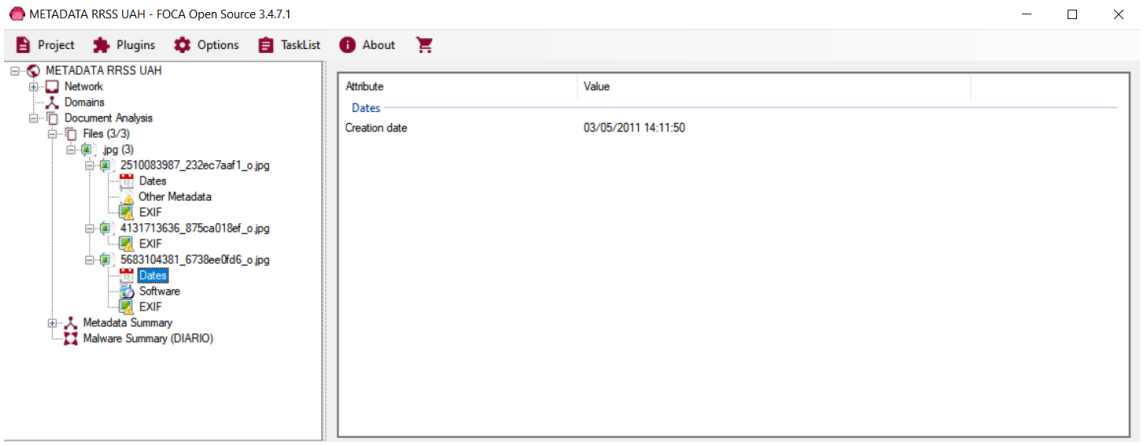


Ilustración 126. Información relevante a la fecha de la tercera imagen a analizar en FOCA.

Información relevante al software, nos la aplicación donde ha sido tratada la fotografía: Adobe Photoshop CSS.

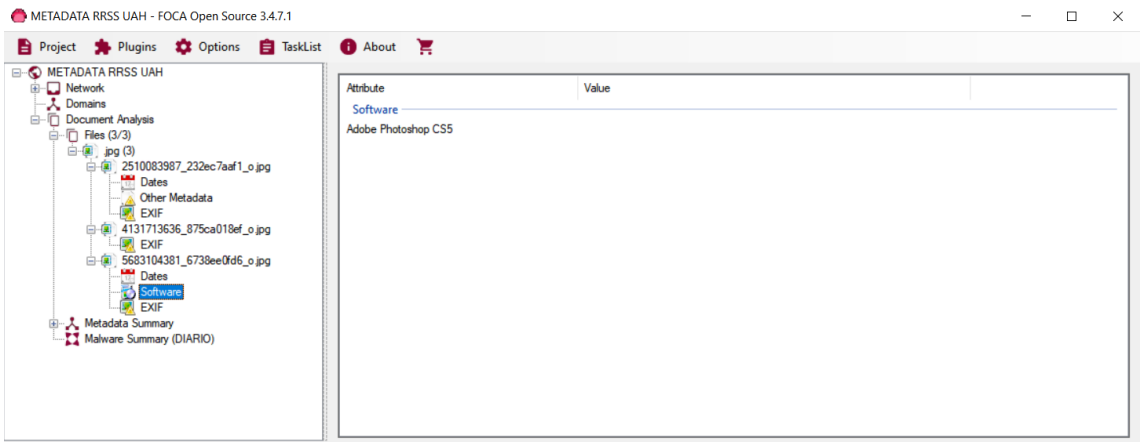


Ilustración 127. Información relevante al software de la tercera imagen a analizar en FOCA.

Información relevante a EXIF, que nos muestra el modelo de cámara con el que tuvo lugar la fotografía: Exif Ifd0, Exif SubIFD, Photoshop...

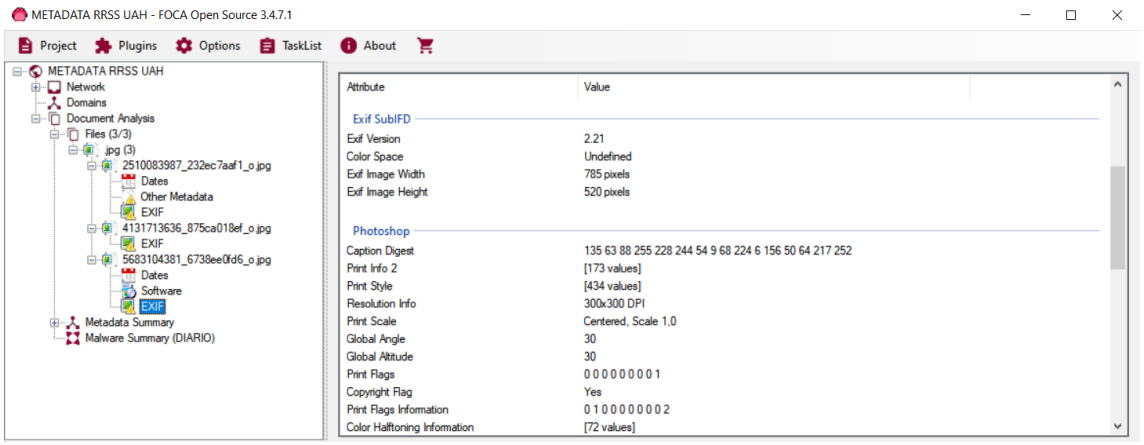


Ilustración 128. Información relevante a EXIF de la tercera imagen a analizar en FOCA.

3.7.6. Conclusiones de FOCA

Una de las conclusiones que podemos extraer tras realizar la investigación de los metadatos existentes en las redes sociales con el objetivo marcado que hemos establecido de la Universidad de Alcalá, es la eliminación y conservación de la información oculta en redes sociales.

Podemos confirmar que Twitter, Instagram y Facebook, cuando el usuario sube contenido, este contenido no contiene metadatos, se borran automáticamente; esto se debe a que su enfoque principal no se encuentra en la conservación de metadatos detallados de las imágenes o publicaciones.

No obstante, la política de privacidad de cada una detalla que pueden recolectar información correspondiente a funciones de etiquetado, geolocalización y otras opciones de personalización que permiten a los usuarios agregar información adicional a sus publicaciones.

Por otro lado, otras redes sociales como Flickr o Google+, no se encargan de eliminar los metadatos e información oculta de las publicaciones, sino que es cada usuario el que debe configurar si esta información se quiere hacer pública, o no.

Con la herramienta de FOCA hemos podido realizar e investigar toda la información oculta y metadatos que se esconden dentro de cada contenido que se encuentra expuesto en internet de forma accesible para todo usuario.

El análisis de los metadatos nos ha permitido obtener información sobre la autoría de los documentos, las interacciones entre usuarios y otras pistas valiosas para nuestra investigación. En resumen, el uso de las redes sociales ha enriquecido nuestro análisis al proporcionarnos una visión más completa y detallada de los documentos y sus contextos.

3.8. OSINTGram: análisis y evaluación de cuentas de Instagram

OSINTGram se trata de una herramienta OSINT enfocada a la red social de Instagram; se encarga de analizar el comportamiento y evaluar las diferentes cuentas de Instagram que se encuentren públicas.

3.8.1. Instrumentos empleados

Entorno virtual: **VMWare**
Máquina atacante: **Kali Linux**
Herramientas:

- **OSINTGram: https://github.com/Datalux/Osintgram**
- **Redes sociales:**
  - **Instagram.**

3.8.2. Puesta en escena de OSINTGram

Para el despliegue de la herramienta **OSINTGram**, hemos optado por lanzarla en entorno virtual.

Desde la herramienta podemos obtener diferente información sobre el objetivo que queremos encontrar en la red social de Instagram; geolocalización y direcciones registradas en las imágenes publicadas, información del usuario, recolectar correos electrónicos, teléfonos vinculados...

En esta puesta en escena se llevará a cabo la **búsqueda a un objetivo.**

3.8.3. OSINTGram: el entorno

Nos descargamos la herramienta de OSINTGram, la lanzamos y vemos en la shell información importante; donde nos muestra las opciones disponibles dentro de esta.

El comando de instalación es el siguiente:

git clone https://github.com/Datalux/Osintgram.git

Ahora bien, debemos de establecer las credenciales (credentials.ini dentro de la carpeta de config) de nuestra cuenta de investigación y configurar la herramienta para poder acceder de forma correcta a la API de Instagram.

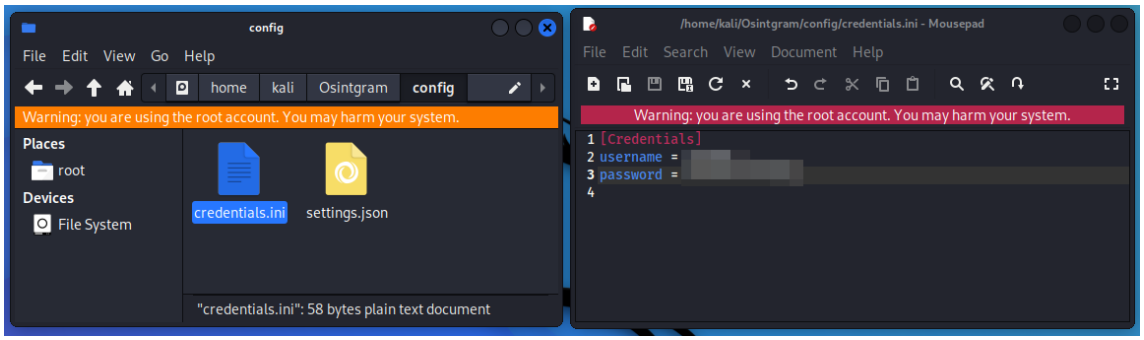


Ilustración 129. Configuración de las credenciales para la conexión de OSINTGram con Instagram y su API.

Para lanzar la herramienta, basta con ir al directorio principal donde se encuentra albergada, y la lanzamos mediante Python con su fichero main.py seguido del objetivo al que queramos investigar.

Cabe destacar que la última actualización de la API de Instagram (instagram_private_api: biblioteca de Python para acceder a la API privada de Instagram) ya que únicamente permite evaluar y analizar el comportamiento de la cuenta de usuario con la que introduces las credenciales.

```
┌──(root㉿kali)-[/home/kali/Osintgram]
└─# python3 main.py uahes

Attempt to login...
Error parsing error response: Expecting value: line 1 column 1 (char 0)
Traceback (most recent call last):
  File "/usr/local/lib/python3.10/dist-packages/instagram_private_api/client.py", line 523, in _call_api
    response = self.opener.open(req, timeout=self.timeout)
  File "/usr/lib/python3.10/urllib/request.py", line 525, in open
    response = meth(req, response)
  File "/usr/lib/python3.10/urllib/request.py", line 634, in http_response
    response = self.parent.error(
  File "/usr/lib/python3.10/urllib/request.py", line 563, in error
    return self._call_chain(*args)
  File "/usr/lib/python3.10/urllib/request.py", line 496, in _call_chain
    result = func(*args)
  File "/usr/lib/python3.10/urllib/request.py", line 643, in http_error_default
    raise HTTPError(req.full_url, code, msg, hdrs, fp)
urllib.error.HTTPError: HTTP Error 404: Not Found

During handling of the above exception, another exception occurred:

Traceback (most recent call last):
  File "/home/kali/Osintgram/main.py", line 121, in <module>
    api = Osintgram(args.id, args.file, args.json, args.command, args.output, args.cookies)
  File "/home/kali/Osintgram/src/Osintgram.py", line 48, in __init__
    self.setTarget(target)
  File "/home/kali/Osintgram/src/Osintgram.py", line 61, in setTarget
    self.following = self.check_following()
  File "/home/kali/Osintgram/src/Osintgram.py", line 1159, in check_following
    return self.api._call_api(endpoint)['user_detail']['user']['friendship_status']['following']
  File "/usr/local/lib/python3.10/dist-packages/instagram_private_api/client.py", line 527, in _call_api
    ErrorHandler.process(e, error_response)
  File "/usr/local/lib/python3.10/dist-packages/instagram_private_api/errors.py", line 135, in process
    raise ClientError(error_msg, http_error.code, error_response)
instagram_private_api.errors.ClientError: Not Found
```

Ilustración 130. Error al realizar la conexión con Instagram debido a sus medidas y políticas de seguridad y privacidad.

Por este motivo, Instagram ha tomado medidas de seguridad en su API y ha restringido este tipo de análisis en diferentes herramientas de OSINT en la cual se extraen datos para de forma posterior conseguir información de las cuentas de usuario. Esta red social, al igual que otras muchas, regularmente actualiza sus políticas y restricciones para mantener la seguridad de la plataforma y proteger la experiencia de los usuarios.

No obstante, nosotros llevaremos a cabo la evaluación y síntesis de OSINTGram, pero a diferencia del resto de herramientas tratadas, el objetivo no será la Universidad de Alcalá (@uahes), sino nuestra cuenta personal.

3.8.4. Búsqueda de un objetivo y análisis de resultados

Llevaremos a cabo la búsqueda de nuestro objetivo por medio de la herramienta. Fijamos como objetivo a nuestra cuenta de Instagram creada para pruebas, ya que no podemos realizarlo con la Universidad de Alcalá, nuestro objetivo principal en la investigación.

OSINTGram accede a la API de Instagram a partir de las credenciales insertadas en el credentials.ini, dentro de la carpeta de config. A partir de ahí, toda interacción que deberemos llevar a cabo es con la herramienta.

Como comentábamos antes, analizaremos la cuenta tfmuah99, cuenta creada para analizar la herramienta. Esta cuenta contiene una publicación con la ubicación de la Escuela Politécnica Superior, y únicamente sigue a un usuario, la uahes.

Al iniciar OSINTGram podemos observar diferentes opciones para almacenar los resultados del análisis e investigación realizados sobre la búsqueda del usuario marcado como objetivo.

Ilustración 131. Análisis de la cuenta tfmuah99 con los diferentes comandos que contiene OSINTGram.

Además, podemos observar los diferentes comandos con los que cuenta OSINTGram con los cuales podemos obtener información de alta importancia en el análisis realizado sobre la cuenta de Instagram que estamos tratando.

Algunos de los más destacados que podemos ver son: fwersnumber y fwingsnumber (obtener número de teléfono de seguidores/seguidos), fwersemail y fwingsemail (obtener correo electrónico de seguidores/seguidos), addrs (direcciones y geolocalizaciones encontradas), potos (descargar fotos) …

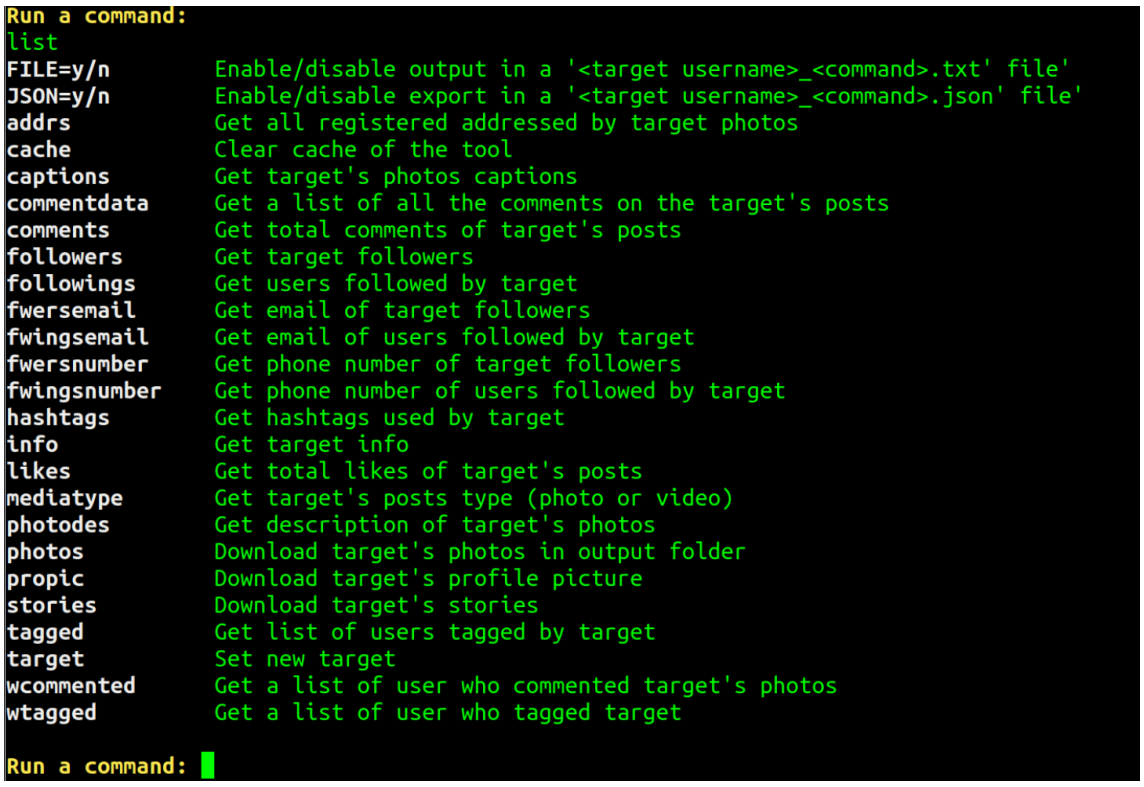


Ilustración 132. Todos los comandos con los que cuenta OSINTGram para obtener información de Instagram.

Ahora bien, procederemos a mostrar el funcionamiento de diferentes comandos con el fin de obtener datos e información de alta calidad dentro de Instagram.

Correo electrónico

Obtenemos el correo electrónico de los usuarios a los que seguimos con la cuenta de tfmuah99 por medio de **fwingsmail**. Como teníamos un único usuario seguido, uahes, el correo electrónico vinculado a esta cuenta es **socialmedia@uah.es**.

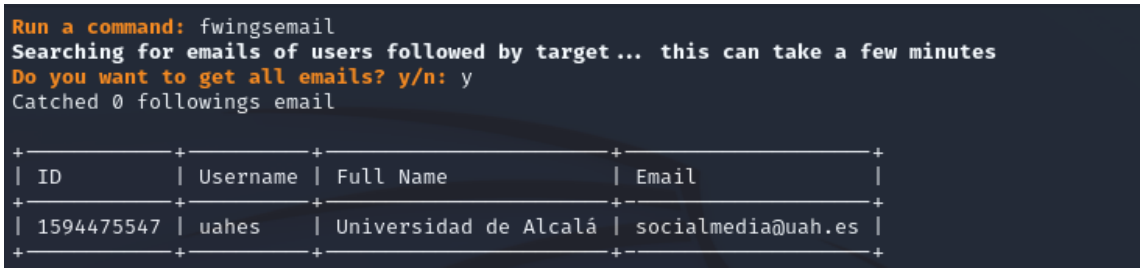


Ilustración 133. Fwingsmail, comando para obtener el correo electrónico de los usuarios a los que seguimos con la cuenta de tfmuah99.

Número de teléfono

Obtenemos el número de teléfono de los usuarios a los que seguimos con la cuenta de tfmuah99 por medio de **fwingsnumber**. Como teníamos un único usuario seguido, uahes, el correo electrónico vinculado a esta cuenta es **918855000.**

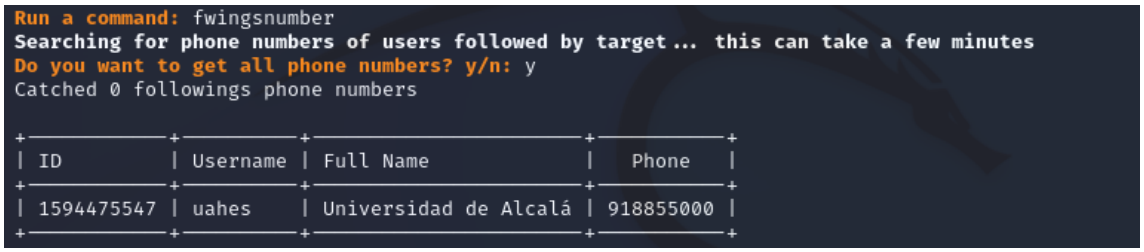


Ilustración 134. Fwingsnumber, comando para obtener el número de teléfono de los usuarios a los que seguimos con la cuenta de tfmuah99.

Direcciones

Obtenemos las direcciones y geolocalización de la cuenta que estamos investigando, tfmuah99 por medio de **addrs**.

Como teníamos una única publicación con dirección de geolocalización de la Escuela Politécnica Superior, la ubicación encontrada **Escuela Politécnica Superior – UAH, Avenida de León, Distrito IV, Alcalá de Henares, Comunidad de Madrid, 28805, España | 2023-06-03 14:14:10.**

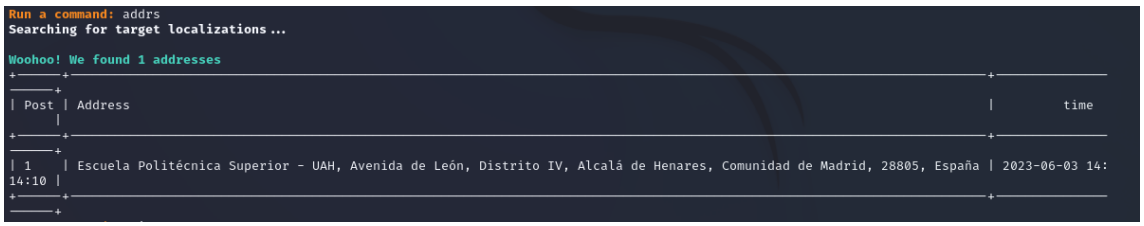


Ilustración 135. Addrs, comando para obtener las direcciones y geolocalización de la cuenta que estamos investigando, tfmuah99.

Información

Obtenemos la información de la cuenta que estamos investigando, tfmuah99 por medio de **info**. Sin embargo, podemos observar que de repente la cuenta nos dice que no existe y que establezcamos un nombre de usuario válido.

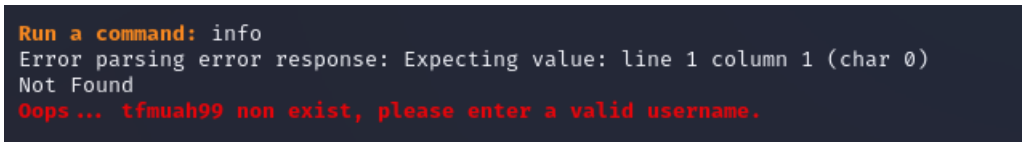


Ilustración 136. Info, comando para obtener información de la cuenta que estamos investigando, tfmuah99.

Cuenta invalida – error API Instagram

De forma consecutiva, intentamos realizar otro comando, pero al realizar la siguiente instrucción en la herramienta, esta nos dice que la cuenta de tfmuah99 con la que estamos conectados a la API de Instagram para obtener información ha sido suspendida por la red social.

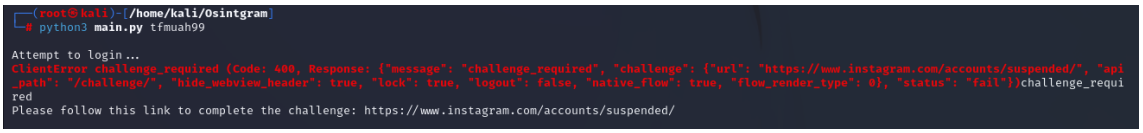


Ilustración 137. La cuenta tfmuah99 ha sido suspendida por Instagram y no podemos acceder a la API con OSINTGram.

A su vez, recibimos en el correo electrónico en el que tenemos vinculado la cuenta de Instagram un correo en el cual nos comunicaba dicha red social que tenemos que adoptar las medidas necesarias, si no perderemos la cuenta.

Además, el motivo por el que nos comentan dicha incidencia se debe a que: *"Se ha suspendido tu cuenta porque la cuenta o actividad en ella no cumplen nuestras Normas comunitarias. Si crees que hemos cometido un error, tienes hasta el 30 de noviembre de 2023 para apelar la decisión".*

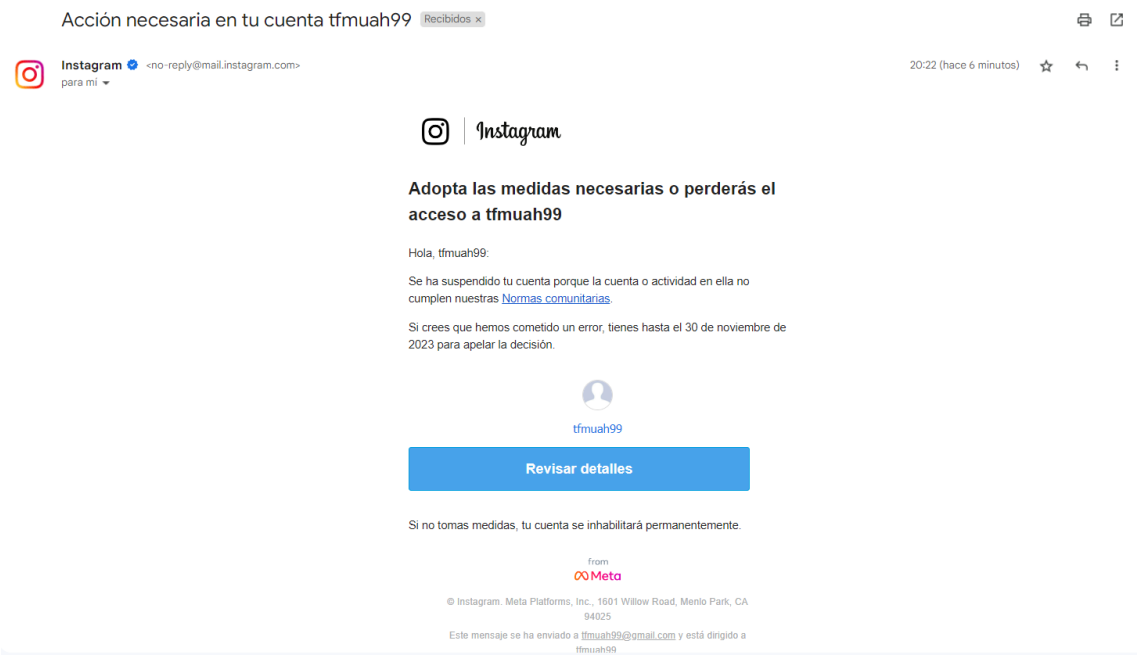


Ilustración 138. Correo electrónico recibido de Instagram comunicando la incidencia de la suspensión de la cuenta tfmuah99 en la red social.

Esto se debe a que Instagram considera que se encuentran alteradas las políticas y restricciones actuales de la plataforma en relación con esta herramienta. Es por ello, que si esta red socia detecta una determinada herramienta o actividad que está en contra de las políticas, existe la consecuencia del bloqueo y/o restricciones en la cuenta.

3.8.5. Conclusiones de OSINTGram

Después de probar la herramienta OSINTGram, hemos podido comprobar toda el comportamiento e información que se puede extraer de una cuenta en la red social de Instagram.

A su vez, hemos podido comprobar las buenas políticas y restricciones con las que cuenta esta red social, que al mínimo momento de detectar una actividad sospechosa bloquea y suspende el usuario con el cual se está accediendo a su API con el fin de obtener diferente información del objetivo marcado.

Por último, dentro de OSINTGram hemos podido comprobar la diferente información que podemos obtener con tan solo acceder a la API de Instagram y de esta forma conseguir desde números de teléfonos, correos electrónicos, geolocalizaciones...

4. El entorno virtual: Sandboxing

Vamos a llevar a cabo la elaboración de nuestra máquina virtual donde realizaremos todo nuestro proyecto de investigación; se trata algo similar a un Sandboxing, para de esta forma no arriesgar ni exponer nuestro equipo.

Previamente se hace una especificación de las tres formas de existentes para crear máquinas virtuales:

- **.OVA (archivo autoinstalable):** haciendo doble clic sobre un archivo autoinstalable .OVA podemos lanzar la auto instalación de una máquina virtual ya preconfigurada.

- **.ISO:** podemos crear nuestra máquina de la misma forma que realizamos una instalación en un equipo físico, mediante la ejecución de un archivo .ISO en una unidad de CD virtual.

- **Mediante la ejecución de una clonación de disco de una máquina anteriormente creada**: la forma más sencilla y rápida, sería ejecutando la imagen de una máquina virtual ya creada. A la hora de crear la máquina deberemos indicar donde se encuentra el archivo.

También hemos de destacar que el tipo de red que vamos a emplear en nuestra máquina virtual se trata de **NAT** (Network Address Translation); funcionalidad básica desde el sistema operativo huésped. Cómo inconveniente, tiene bastantes limitaciones si tenemos que establecer conexiones con la máquina virtual.

Emplearemos de aplicación para crear máquinas virtuales **MVware en su versión Workstation Pro**.

4.1. Kali-Linux en su última versión

En primer lugar, hemos de ir a la página web oficial de Kali Linux (https://www.kali.org/get-kali/) y descargarnos su última versión para la creación de máquinas virtuales.

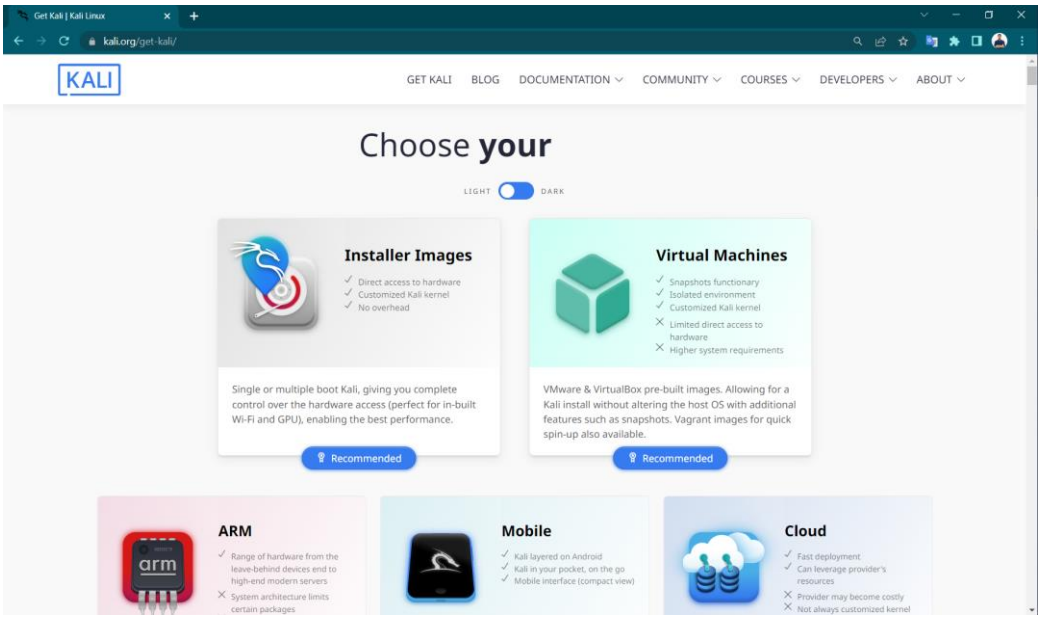


Ilustración 139. Página principal de Kali Linux

Ahora bien, abriremos nuestro VMware y procederemos a la creación de la máquina virtual con Kali Linux en su última versión.

Al tratarse el archivo de Kali Linux en un **.OVA (archivo autoinstalable)** la instalación será por medio de abrir máquina virtual y buscar nuestro archivo en la ruta indicada y después abrirlo.

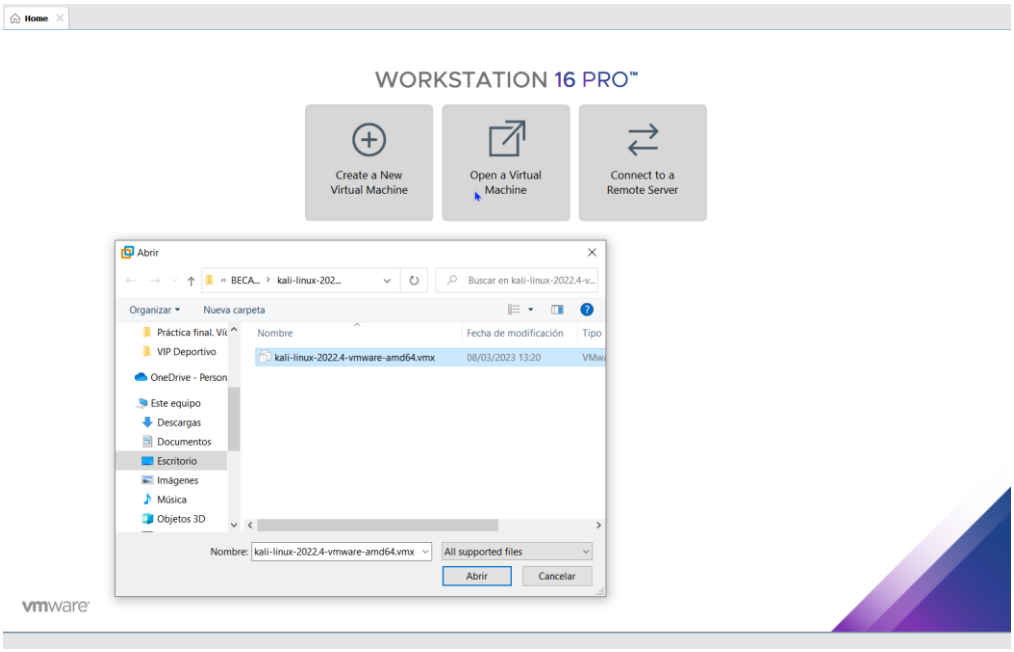


Ilustración 140. Instalación de Kali Linux dentro de VMware

Después de darle a abrir, unos instantes después obtendremos la máquina virtual creada, donde una vez lista le debemos dar a ejecutar.

Ilustración 141. Instalación de Kali Linux en VMware completado

Cuando la máquina virtual ha completado su ejecución, esta nos solicitará las correspondientes credenciales de acceso. USER: kali || PASSWORD: Kali.

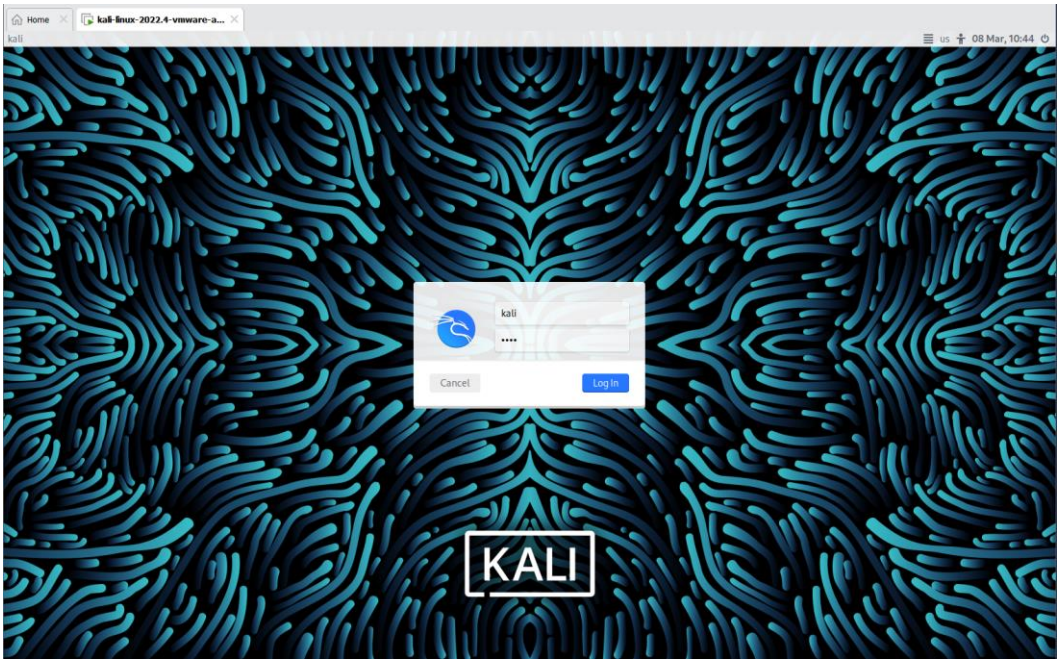


Ilustración 142. Ejecución de Kali Linux en VMware con inicio de sesión

Tras introducir las credenciales de forma correcta se nos cargará y abrirá de forma correcta nuestro entorno Kali Linux listo ya para su uso.

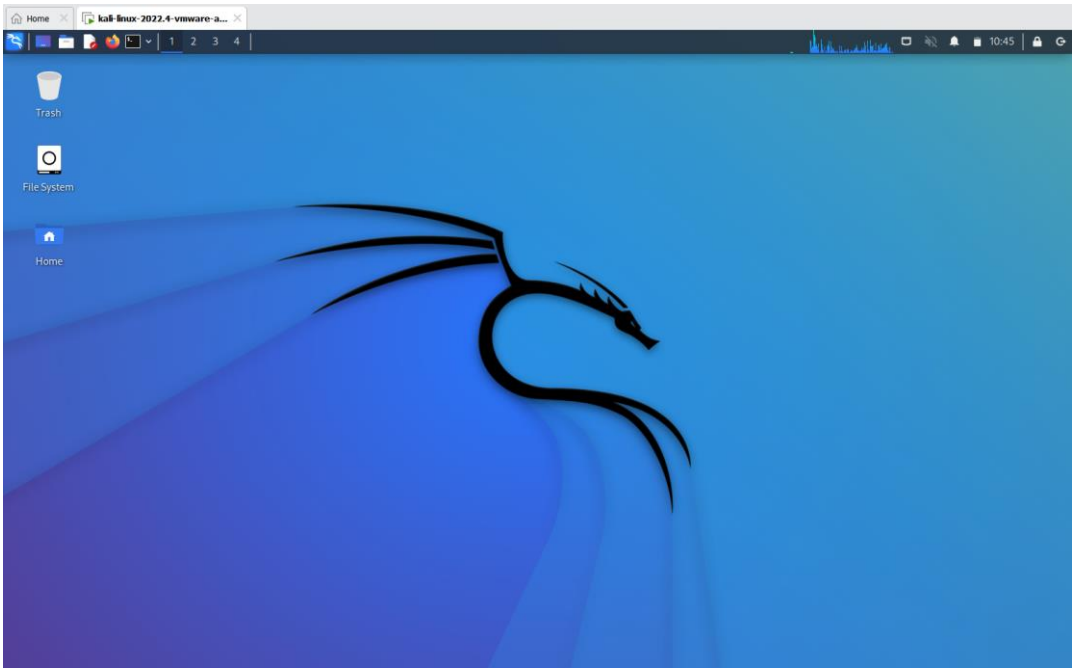


Ilustración 143. Puesta en escena y ejecución de Kali Linux en VMware

5. Conclusiones y trabajo futuro

Realizamos un análisis de las redes sociales analizadas y explotadas en nuestra investigación, donde como sabemos, estas dependen del **nivel de facilidad** para su obtención, y el **nivel de la calidad** de esta.

Nivel	Camino	Facilidad	Calidad	Red Social
Nivel 1	Autenticación No API	MEDIA	S/C	**Facebook**
Nivel 3	Autenticación API Multimedia No metadatos	MEDIA	ALTA	**Instagram**
Nivel 3	Autenticación API Multimedia No metadatos	MEDIA	ALTA	**Twitter**
Nivel 8	No autenticación API Multimedia Metadatos	ALTA	MUY ALTA	**Flickr**

Como podemos observar, la mejor red social se trata de Flickr, seguida de Twitter, después Instagram, y por último Facebook. Aunque hay que puntualizar que Instagram da problemas debido a su autenticación, API y políticas de privacidad, pero está un escalón por encima de Facebook, ya que podemos investigar algo en ella.

A su vez, en este trabajo se han visto diferentes formas de análisis y recolección de datos e información con distintas herramientas.

Con el uso de Namechk, hemos podido comprobar que tiene como fin encontrar tanto dominios como redes sociales registradas a partir del objetivo dado por el usuario.

Por medio de Sherlock, nos hemos dado cuenta de lo expuestos que estamos los usuarios en la red; con únicamente saber el nombre de usuario del objetivo que queremos buscar, podemos encontrar todas sus redes sociales.

En theHarvester, hemos podido comprobar toda la información perteneciente al dominio establecido como objetivo en el análisis respecto a los dominios y correos electrónicos.

A través de los dorks, operadores que realizan que las búsquedas se encuentren más avanzadas, hemos podido observar por medio de Google Dorks y por la herramienta de SXDork, la capacidad de estos operadores para obtener una información que en los buscadores de forma normal no podemos encontrar.

Con accountanalysis, hemos podido comprobar toda el comportamiento e información que se puede extraer de una cuenta en la red social de Twitter.

Por medio de Cree.py hemos podido ver de qué forma podemos obtener información relevante a la geolocalización que podemos obtener desde diferentes redes sociales.

A través de FOCA, los metadatos nos han permitido obtener información sobre la autoría de los documentos, las interacciones entre usuarios y otras pistas valiosas para nuestra investigación. Además de las diferentes restricciones dentro de la API y políticas de privacidad.

Por último, con OSINTGram, hemos podido comprobar toda el comportamiento e información que se puede extraer de una cuenta en la red social de Instagram.

Con este trabajo hemos podido observar la gran cantidad de datos e información que se encuentran expuestos en internet. No se trata de indagar dentro de la Dark o la Deep web, basta con estar en la Surface web y usar nuestros medios habituales.

Las personas y organizaciones se exponen de forma muy sencilla en internet, por medio en gran parte, al uso de sus redes sociales, ya que de esta forma son vulnerables gracias a ellos mismos debido al contenido que suben a la red.

Este proyecto se podría complementar siguiendo el procedimiento llevado a cabo, pero añadiendo nuevas redes sociales (TikTok, YouTube, Spotify…) con las cuales poder obtener información relevante.

Otra manera de ampliar el proyecto sería la del análisis teórico y experimental de las diferentes políticas de privacidad que contienen Facebook y Twitter; esto se debe a que como hemos visto, son las que más restricciones tienen, también a nivel de API.

Por otro lado, también se puede llevar a cabo el estudio práctico, es decir, fijar unos objetivos como víctima, y obtener un informe final con todo el conocimiento obtenido a través del OSINT en las redes sociales; en ese informe se puede plasmar la información por medio de herramientas de representación gráfica y visualización (GEPHI, SOCINT, MALTEGO…).

Además, este trabajo me sirve de base de cara al doctorado y tesis doctoral que quiero realizar, haciendo un nexo entre OSINT, la Inteligencia Artificial y la Dark Web. A su vez, la taxonomía de clasificación creada nos sirve de esquema para investigaciones futuras.

6. Referencias

Digital 2023. (2023, enero 26). We Are Social Spain. Recuperado el 6 de septiembre de 2023 de https://wearesocial.com/es/blog/2023/01/digital-2023/

Osint La Informacion Es Poder. (2014, mayo 28). Incibe.es. Recuperado el 6 de septiembre de 2023 de https://www.incibe.es/incibe-cert/blog/osint-la-informacion-es-poder

Fonte, A. (2021, marzo 8). OSINT, ¿Qué es? ¿Para qué sirve? Derecho de la Red; derechodelared. Recuperado el 6 de septiembre de 2023 de https://derechodelared.com/osint/

Alemán Romero, D. (2016). Inteligencia Basada en Efectos para la Seguridad Humana de las Naciones Unidas. TDX (Tesis Doctorals en Xarxa). https://www.tdx.cat/bitstream/handle/10803/400082/dar1de1.pdf?sequence=1

About: Open Source Enterprise. (s/f). DBpedia. https://dbpedia.org/page/Open_Source_Enterprise

NOSIC. (s/f). Com.au. https://www.nosic.com.au/

Get Kali. (n.d.). Kali Linux. Recuperado el 6 de septiembre de 2023 de https://www.kali.org/get-kali/

VMware en español ofrece la base digital para la empresa. (2023, August 28). VMware. https://www.vmware.com/es.html

La herramienta de Namechk se ha hecho uso a través de su página web oficial (https://namechk.com/).

La herramienta de Accountanalysis se ha hecho uso a través de su página web oficial (https://accountanalysis.app/).

La herramienta de Sherlock se ha descargado a través de su página web oficial (https://github.com/sherlock-project/sherlock).

La herramienta de theHarvester se ha descargado a través de su página web oficial (https://github.com/laramies/theHarvester).

La herramienta de FOCA se ha descargado a través de su página web oficial (https://github.com/ElevenPaths/FOCA).

La herramienta de Osintgram se ha descargado a través de su página web oficial (https://github.com/Datalux/Osintgram).

La herramienta de Creepy se ha descargado a través de su página web oficial (http://www.geocreepy.com/).

Universidad de Alcalá