empresas.blogthinkbig.com

Cómo analizar documentos con FOCA en diez pasos (o menos) - Think Big Empresas

6-8 minutos

Cada vez que creamos un documento ofimático, como puede ser un procesador de texto (por ejemplo, Microsoft Word), una presentación (un PowerPoint), una hoja de cálculo (un Excel), un PDF o incluso imágenes, estos almacenan por defecto **mucha más información de la que pensamos**. Existe un **contenido adicional incrustada en los ficheros que recibe el nombre de** metadatos y pueden contener datos como por ejemplo el nombre del autor, la fecha de creación/modificación o incluso el título del documento. Aunque esto ya de por sí ofrece bastante información, un análisis más profundo puedo extraer todavía más datos que van más allá de los mencionados, dando, por ejemplo, contenido muy importante sobre la infraestructura donde fue creado.

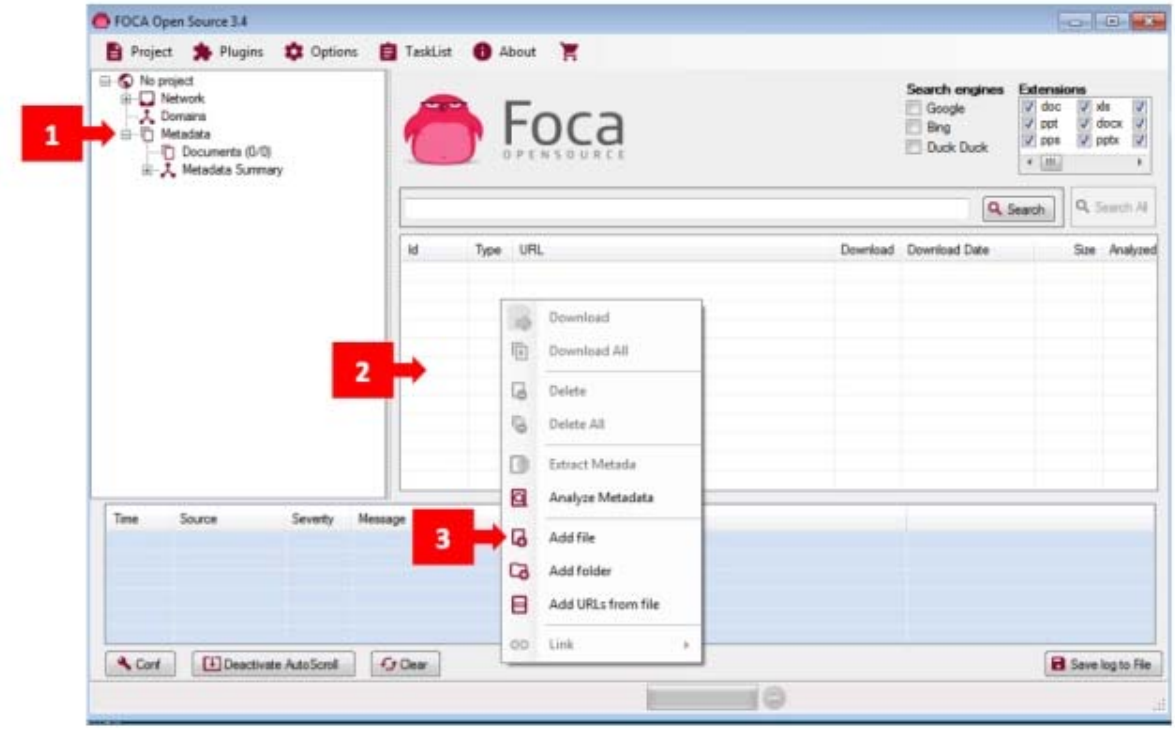
Foca
OPENSOURCE

Por ejemplo, **es posible extraer contraseñas, nombres de usuarios, nombres de carpetas, nombres de servidores, impresoras, ediciones realizadas, etc. todo esto en simple fichero ofimático**. Esta información podría poner en grave riesgo, además de nuestra privacidad, la integridad de nuestra empresa u organización, ya que ofrece muchos datos importante que un posible ciberdelicuente podría utilizar para analizar nuestra infraestructura (esta técnica se llama "*fingerprinting*") y luego lanzar algún tipo de ataque basándose en esto. En el caso de las imágenes, la información más relevante que se puede obtener es la localización geográfica desde la cual se tomó la fotografía, ofreciendo datos por ejemplo, de un itinerario que hayamos realizado. **Los metadatos son más importantes de lo que parecen a simple vista.** Quizás el caso más llamativo fue el de Tony Blair y el documento Word que teóricamente probaba que Irak tenía armas de destrucción masiva, pero una revisión de los metadatos sacó a la luz mucho contenido, como revisiones y comentarios que probaron que dicha información era falsa.

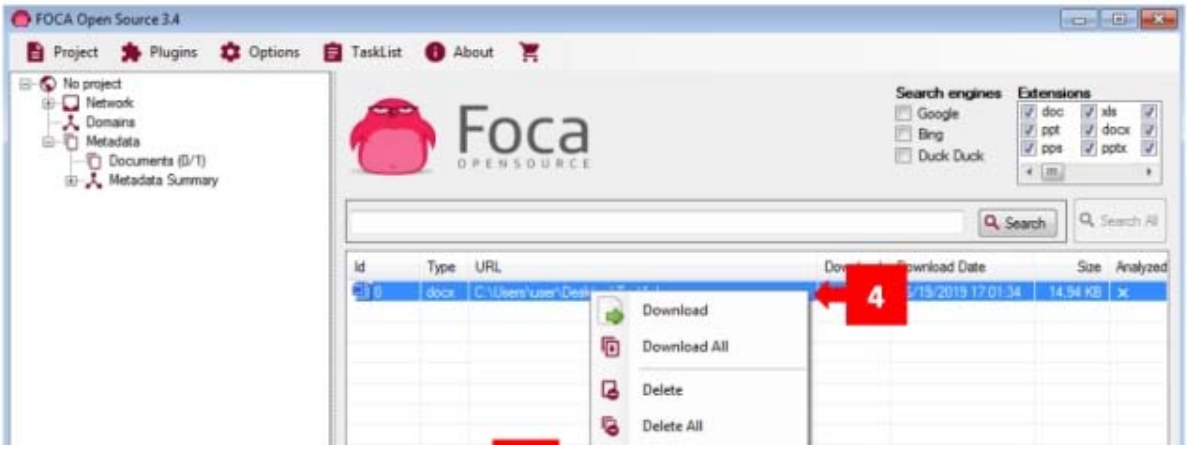
FOCA es una herramienta gratuita creada por ElevenPaths la cual es de gran utilidad a la hora de analizar los metadatos, ya sea de un documento o incluso de toda una organización. La FOCA es código abierto y disponible para su descarga desde el repositorio GitHub de ElevenPaths. Vamos a ver lo sencillo que es extraer todos los datos de un documento ofimático y también obtener los informes de metadatos de toda una organización en unos sencillos pasos.

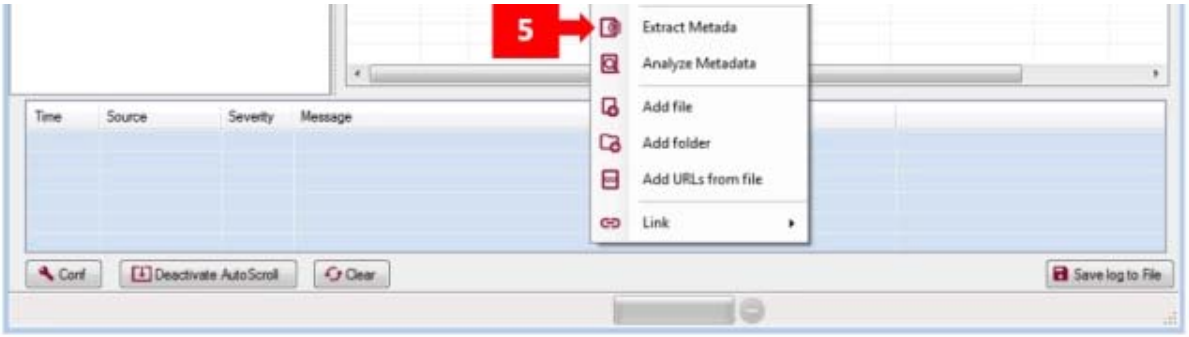
Extracción de metadatos de uno o varios ficheros locales

Paso 1: Una vez tenemos abierta FOCA, simplemente marcamos la opción "*Metadata*" **[1]** y luego, con el botón derecho del ratón hacemos *click* en la zona **[2]** que se indica en la imagen y finalmente en "*Add file*" **[3]** (si queremos analizar el contenido de una carpeta completa, utilizaremos la opción "*Add folder*") seleccionamos el fichero que queremos analizar sus metadatos (también es posible arrastrar el fichero o la carpeta directamente):

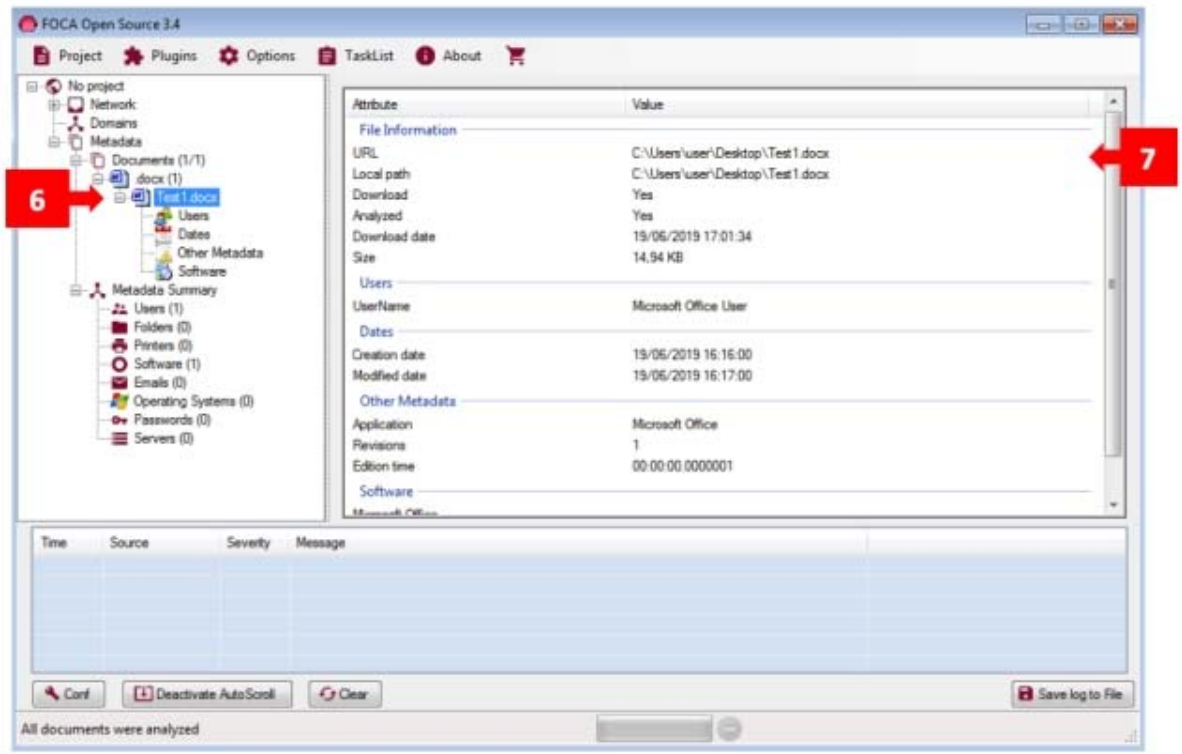


Paso 2: una vez cargado el fichero haremos *click* sobre él con el botón derecho del ratón **[4]** y luego seleccionamos la opción "*Extract Metadata*" **[5]** :





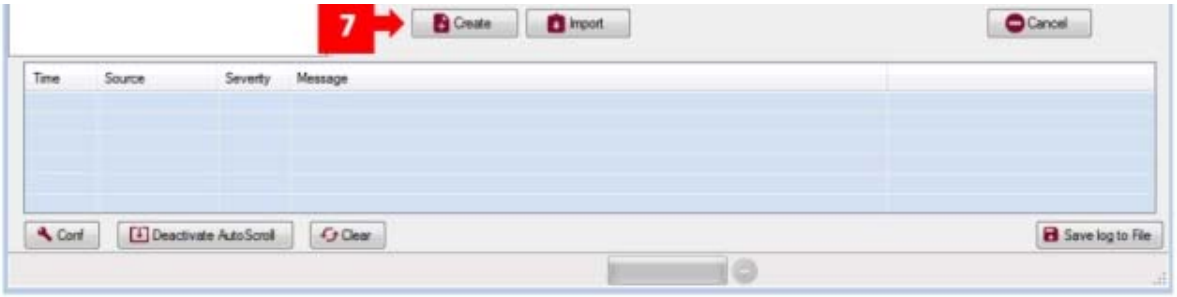
Paso 3: para visualizar los resultados, nos fijaremos en la parte izquierda del panel done aparecerá en el apartado "*Metadata*" el nombre y formato del fichero (en este caso un .docx llamado "Test1") **[6]** . Pulsando sobre él, podremos ver a la derecha un resumen de todos los metadatos extraídos **[7]** :



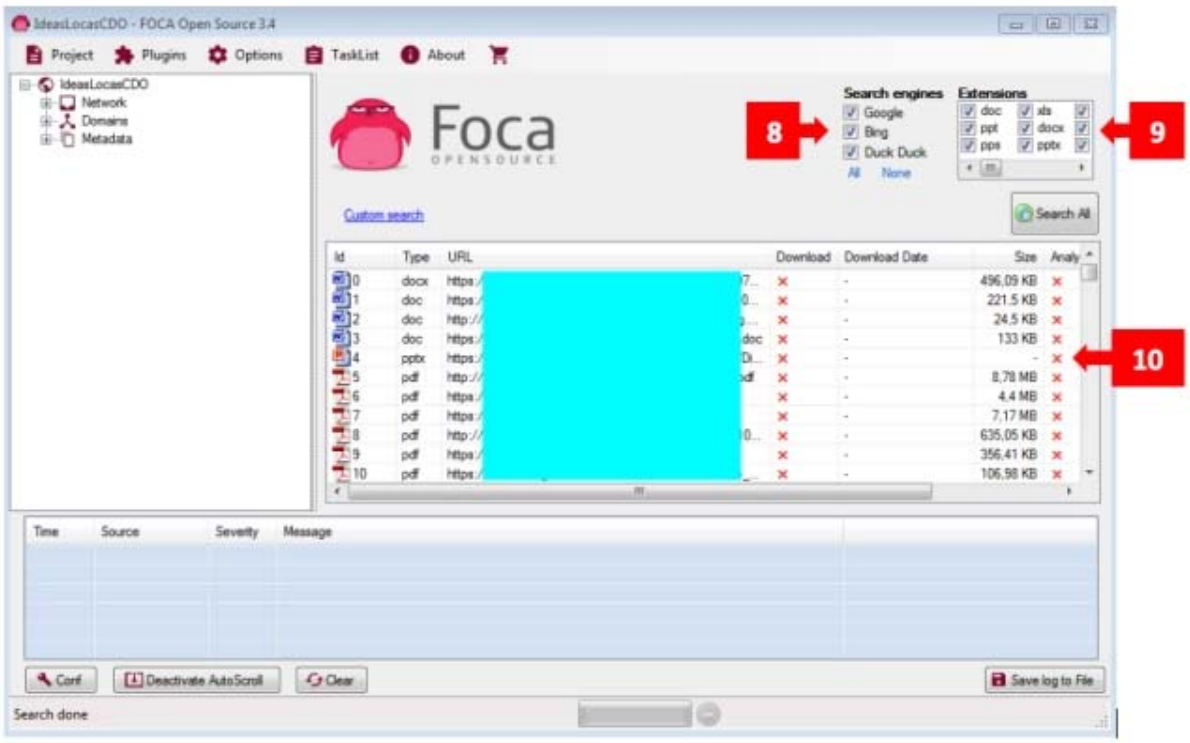
Extracción de todos los metadatos de una organización

Paso 1: el primer paso será definir un proyecto. Para ellos vamos a la sección "*Project*" y seleccionamos "*New Project*" **[1]** :

Paso 2: El apartado **[2]** "*Select Project*" lo utilizaremos si previamente ya hemos creado un proyecto y lo queremos reutilizar, en caso de crear uno desde cero, dejaremos vacía esta opción. En "*Project name*" daremos el nombre del proyecto **[3] .** "*Domain website*" **[4]** nos permite introducir la dirección URL que vamos a auditar. Si hubiera otros dominios alternativos donde queremos que FOCA también busque ficheros, es posible añadirlos en "*Alternative domains*" **[5]**. Los ficheros que vayamos descargando (luego veremos el procedimiento) se almacenarán en la carpeta que se defina en "*Folder where save documents*" **[6]** . Finalmente pulsaremos en "*Create*" **[7]** para definir nuestro proyecto.



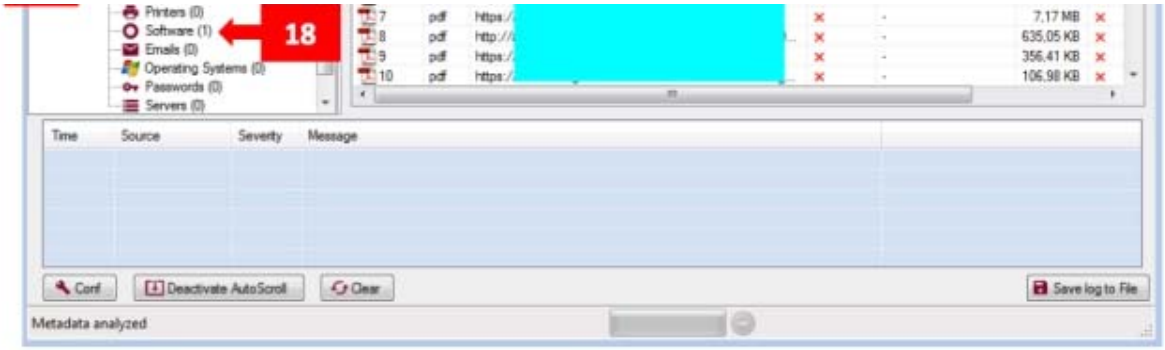
Paso 3: en este punto volveremos a estar en la pantalla de "*Metadata*". El primero de los pasos será marcar los motores de búsqueda "*Search Engines*" **[8]** (en el ejemplo, hemos marcado los tres). En el apartado "*Extensions*" tenemos la opción de seleccionar o no el tipo de fichero que queramos buscar en nuestro proyecto **[9]** . Después de pulsar "*Search All*", al cabo de un tiempo (el cual estará definido por la cantidad de ficheros localizados en la URL del proyecto) nos aparecerá un listado similar al que podemos ver en el punto **[10]** .



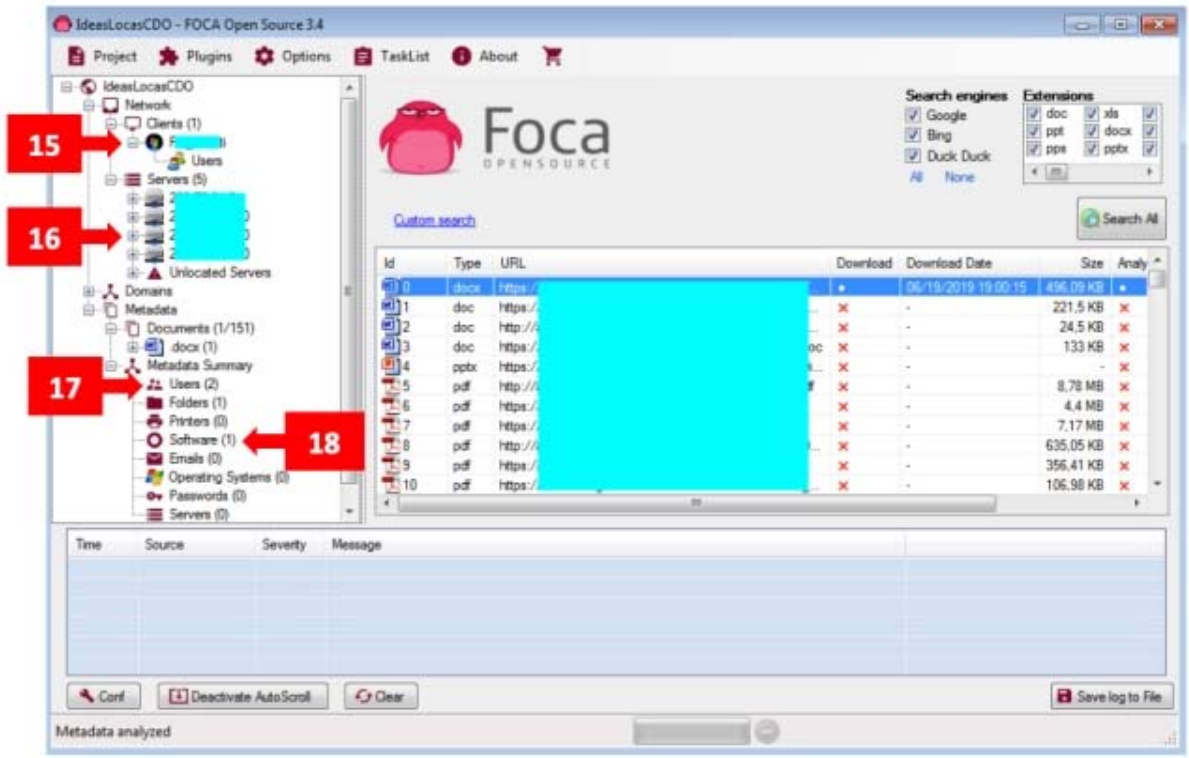
Paso 4: para analizar el fichero o los ficheros obtenidos del análisis, el proceso es similar al que realizamos en el apartado anterior para un fichero único. Pero esta vez tenemos que dar un

paso previo, y es descargarlo. Para ellos pulsamos sobre el botón derecho sobre el fichero **[11]** (también podemos seleccionar varios ficheros manteniendo la tecla “Mayúsculas” pulsada(que queramos analizar y luego la opción “*Download*” tal y como muestra el punto **[12]** (si queremos descargar todos, pulsaremos la opción “*Download All*”).

Paso 5: una vez descargado veremos a la derecha un punto y también la fecha y la hora de descarga. Ahora procederemos a extraer los metadatos con “*Extract Metadata*” [13] y luego los analizaremos [14] con “*Analyze Metadata*”:



Paso 6: finalmente, obtenemos la salida que se muestra en la imagen siguiente (hemos ocultado el contenido por motivos de privacidad) donde se puede apreciar claramente que hemos obtenido el nombre del ordenador donde se ha creado **[15]** , datos de los servidores **[16]**, el nombre de dos usuarios **[17]**, el tipo de *software* **[18]** y también información general sobre el documento como la fecha de creación, etc.



En el siguiente vídeo se muestra en detalle más info sobre FOCA, opciones y funcionamiento:

Y este libro publicado por la editorial 0xWord ofrece en detalle, cómo sacar partido a todas las opciones y posibilidades de FOCA:

Es importante **utilizar un *software* como FOCA para auditar tanto ficheros personales como los de una organización**, y de esa forma obtener una visión sobre el contenido que estamos desplegando de manera involuntaria con este tipo de ficheros y **evitar *data leaks* o fuga de información**. Si necesitas una solución profesional, recuerda que **ElevenPaths ofrece Metashield Protector, una solución empresarial** con varias herramientas destinadas a analizar, proteger y filtrar los metadatos.