Metadatos, “Nuestras Huellas Online”

Ezequiel Sallis CISSP/CEH/MBCI
esallis(AT)root-secure.com
Root-Secure Director

Metadatos, Nuestras Huellas Online”

Introducción

Es común hacer referencia al término metadatos, como “los datos de los datos”, y estos tienen en las nuevas tecnologías infinidad de aplicaciones. Desde la óptica de la seguridad de la información, muchos documentos, formularios, manuales, etc. son publicados sin el tratamiento adecuado, lo que permite extraer de los mismos, múltiples y útiles datos que conforman lo que llamamos metadatos:

. •*Nombres de usuario.*
. •*Rutas a directorios donde el archivo fue almacenado.*
. •*Impresoras donde este documento fue impreso.*
. •*Sistemas operativos y software involucrado en su edición.*
. •*Otros.*

Herramientas como *Foca* [(1)] desarrollada entre otros por Chema Alonso o *Metagoofil* [(2)] desarrollada por Christian Martorella, demuestran el impacto que tienen los metadatos. El análisis al que es sometida esta información, es extremadamente útil en la etapa de reconocimiento pasivo de un análisis de seguridad, como así también, serían de gran utilidad para un potencial atacante en el desarrollo de su estrategia en pos de comprometer un sistema de información.

Los formatos asociados a las suites de ofimática (MS Office o OpenOffice), no son los únicos que contienen Metadatos, sino que estos pueden encontrarse en múltiples variedades de formatos, por ejemplo PDF, AVI, WMV y por supuesto los formatos más comunes asociados a la fotografía digital. Éstos últimos formatos, son sobre los que quiero enfocarme en este documento, con el fin de ilustrar algunos de los buenos y malos usos que alguien podría darle.

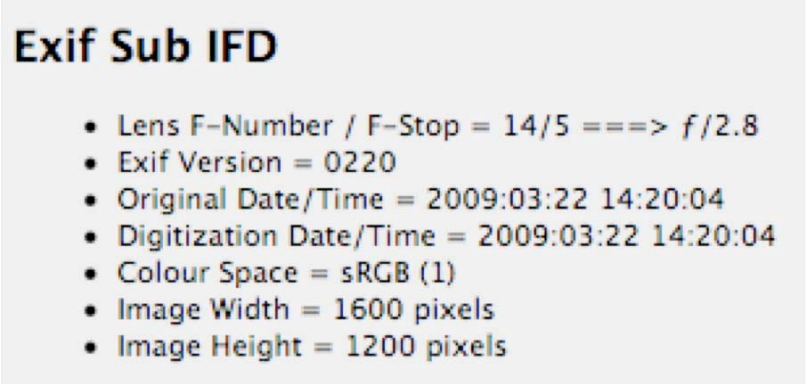
Exif (Exchangeable Image File Format)

Según *Wikipedia*[3] EXIF "es una especificación para los formatos de archivo de imagen resultantes de cámaras digitales". El mismo fue creado por *JEIDA (Japan Electronic Industry Development Association)* pero actualmente se encuentra totalmente desatendido. En resumen, es el nombre técnico con el que nos referiremos a los metadatos que se desprenden de algunos formatos asociados a la fotografía digital.

Los metadatos de una fotografía se almacenan dentro de la misma y pueden leerse con herramientas propias del sistema operativo o bien con herramientas especificas.[11] Para el caso, dependiendo del medio con el que fue generada la imagen, los datos que contienen pueden variar considerablemente. Las variables que son relevantes para que esto suceda, pueden estar asociadas tanto a la tecnología o funcionalidad de la cámara utilizada, como así también al software de edición o proceso intermedio de publicación de la imagen.

Los tipos de datos que comúnmente se pueden encontrar en los metadatos de una imagen digital son algunos de los siguientes:

- *Marca y Modelo de la Cámara Digital.*
- *Nro de Serie de la Cámara Digital.*
- *Fecha y Hora de la Fotografía.*
- *Apertura y Velocidad del Obturador.*
- *Medidor de Exposición.*
- *Sistema Operativo donde la Foto fue editada o almacenada.*
- *Versión del Software Utilizado para la edición.*
- *Datos de Geolocalizacion, es decir Latitud y Longitud donde fue tomada la fotografía.*
- *Miniatura de la Fotografía Original (Sin Edición)*
- *Otros.*



Como mencioné anteriormente, no todas las fotografías contienen toda la información que detallé, sino que muchas veces esto depende de múltiples factores. Por ejemplo, sólo contendrán información sobre la geolocalizacion de la imagen, aquellas cámaras que cuenten con un GPS integrado *(Iphone 3G o Impone 3GS, Nokia N95 y otros).*

Análisis de Metadatos Buenos y Malos Usos

Los metadatos de archivos de ofimática tuvieron su caso mediático, de la mano de *Tony Blair*[(4),] cuando se publicó un documento donde su gobierno hablaba de las armas de destrucción masiva que poseía Iraq. Éste fue publicado en el sitio Web oficial en formato Microsoft Word y en su metadata podía verse cómo fue armado desde diferentes documentos escritos por civiles, que en algunos casos tenían más de una década de antigüedad y la mayor parte del documento pertenecía a una tesis universitaria.

El Exif no queda exento de su caso mediático, y es un buen ejemplo para exponer uno de los riesgos a los que podemos quedar expuestos. Es por esto que , en el caso particular de una fotografía, uno de los datos que puede encontrarse en el Exif es la miniatura de la misma, lo que muchos conocemos como “thumbnail”, (versiones reducidas de la imagen usadas muchas veces para su fácil organización o reconocimiento) que generalmente guarda la imagen original libre de toda edición posterior. Si no, podemos preguntárselo a *Cat Schwartz*[(,5)] una celebridad que publicó fotos de ella en su blog, en las cuales en sus originales se podía ver su rostro, sin embargo, en la miniatura se la podía ver con su torso completamente desnudo. Muchos atribuyen esto al software de edición de las fotografías, en donde, si el cambio que se le realiza a la foto no es representativo, la miniatura o thumbnail no se ve alterada.

Mas allá de lo vergonzosa que pueda resultar la historia anterior, este breve análisis realizado tiene la intención de mostrar cuáles son algunos de los buenos y malos usos que alguien podría darle a esta información.

Desde la óptica de un noble uso que habitualmente se le da a los metadados alojados en las fotografías, están aquellos asociados a los análisis forenses llevados a cabo como parte de un proceso de investigación relacionado a la *pedofilia*[(6)], donde, del análisis de esta información, podría encontrarse, la fecha y hora de la fotografía, la marca y modelo de la cámara fotográfica y hasta en algunos casos, el número de serie de la misma, lo que permitiría intentar llegar con mayor facilidad a una obtención exitosa de su origen.
Desde la óptica del mal uso que alguien podría darle a una fotografía que publicamos online, surgen cientos de variables que permitirán obtener múltiples resultados. Para facilitar el entendimiento, separemos las mismas en diferentes categorías:

. •*Aspectos relacionados con la fotografía.*
. •*Aspectos relacionados con la cámara utilizada.*

Aspectos relacionados con la fotografía

De los aspectos relacionados a la tecnología, surgen datos que permitirían obtener el sistema operativo utilizado, el software de edición y el software de visualización de la fotografía entre otros. En principio, en base a estos datos, alguien podría, al igual que con los archivos de ofimática, tener una mejor visibilidad para planificar un ataque que explote debilidades del lado del

cliente. Por ejemplo, conociendo la versión exacta del software de visualización de fotografías, el atacante podría explotar debilidades asociadas a éste con el solo hecho de que el cliente intente visualizar una fotografía especialmente preparada por el atacante.

Exif IFD0

- Camera Make = Apple
- Camera Model = iPhone
- Picture Orientation = normal (1)
- X-Resolution = 4718592/65536 ===> 72
- Y-Resolution = 4718592/65536 ===> 72
- X/Y-Resolution Unit = inch (2)
- Software / Firmware Version = QuickTime 7.6
- Last Modified Date/Time = 2009:03:24 20:45:15
- Unknown tag, Tagnum 0x013c = data ===> Mac OS X 10.5.6
- Y/Cb/Cr Positioning (Subsampling) = centered / center of pixel array (1)

Por otro lado, y desde la óptica de la privacidad, es muy común ver en fotologs, foros de contenido erótico y similares, fotografías de hombres y mujeres que se exponen desnudos o en situaciones comprometedores, pero que ocultan mediante trucos de edición sus rostros o señas particulares para no se reconocidos. Sin embargo, luego del análisis de la miniatura que esta dentro del exif, acompañando la fotografía original, es enorme la cantidad de identidades o porciones de la fotografía que quedan al descubierto, recordemos el caso de la celebridad que mencionamos mas arriba.

Claro está, es bueno aclarar que esta misma razón que haría enrojecer a más de uno o una, estos metadatos alojados en la fotografía permiten en los casos en donde el abuso de menores podría está en juego, poner en evidencia las caras de las personas adultas que participan en estos.

Publicada Thumbnail

Publicada Thumbnail

Un buen ejemplo local de la información de una fotografía, fue el que expuso Leonardo Pigner en su blog[7], sobre “Los Metadatos de Cristina”.

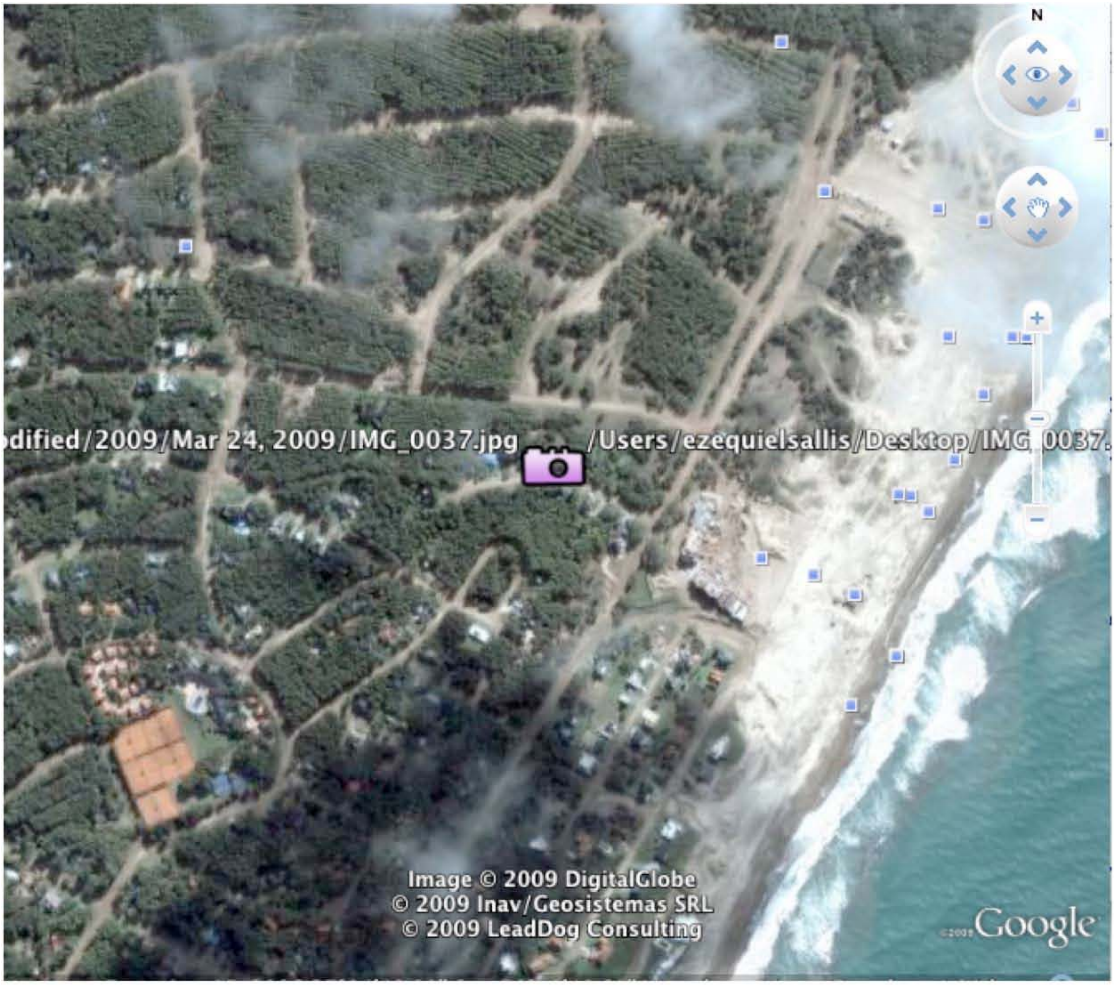
Aspectos relacionados con la cámara utilizada

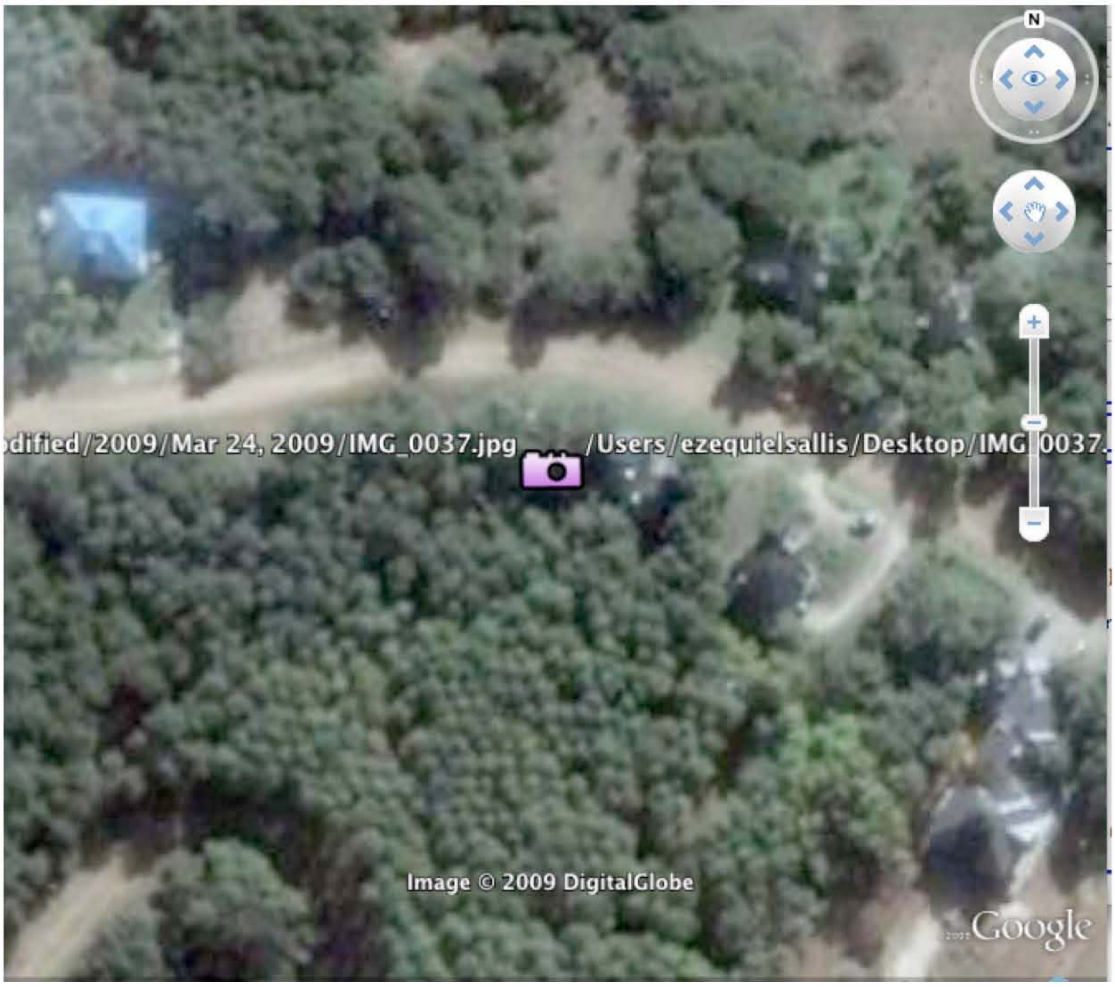
La cámara que se utilice para fotografiar es una variable importante que nos limitará, en mayor o menor medida, la información que podemos encontrar en una fotografía. Por ejemplo, lo más relevante de esto, son los metadatos asociados al *GPS*[8] (Sistema de Posicionamiento Global) que pueda tener incorporado el equipo fotográfico: Sin ir más lejos, el Iphone es la cámara más usada para la publicación de fotos en *Flickr*[,1,] según las *últimas estadísticas*[(9)0] *.*

Esta información permite obtener la latitud y la longitud de donde fue tomada la fotografía, permitiendo obtener la localización casi exacta. Volvamos al ejemplo anterior donde un productor de material pedófilo utilice este tipo de cámaras, podría ser de suma utilidad para lo investigadores contar con ese dato, pero también veamos el impacto que esto podría tener con la foto de contenido erótico que mencionamos antes o con una foto que nuestro hijo adolecerte suba a una red social con el titulo de, “En casa con amigos”. Créalo o no, fácilmente alguien podría ubicar el lugar donde se tomo esa foto.

Exif GPS IFD

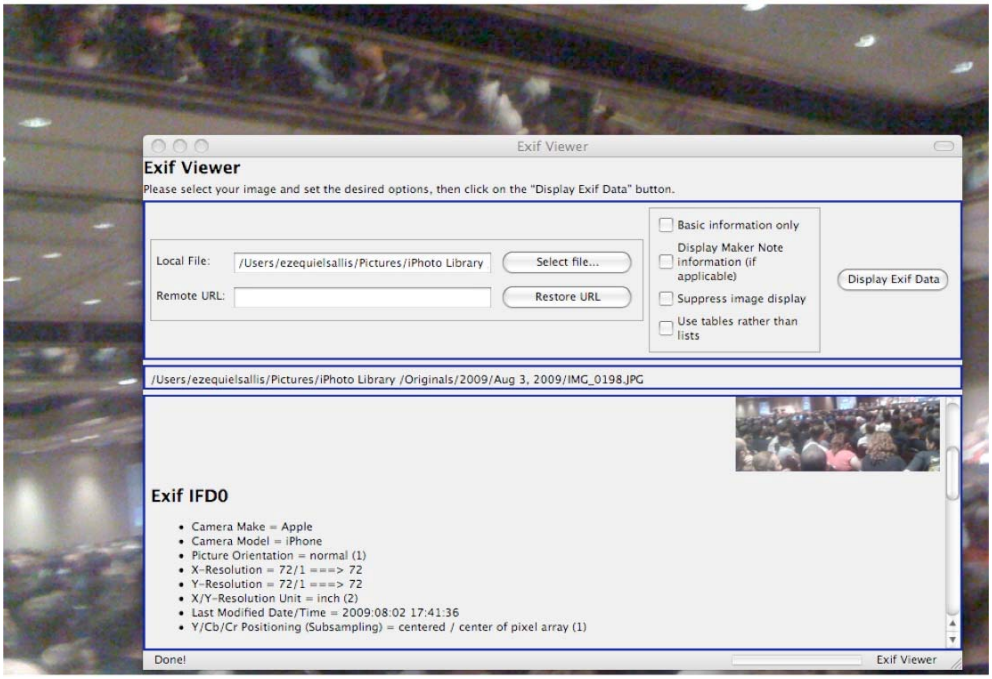
- GPS Latitude Reference = S
- GPS Latitude = 37/1,483/100,0/1 [degrees, minutes, seconds] ===> 37° 4.83′
- GPS Longitude Reference = W
- GPS Longitude = 56/1,4981/100,0/1 [degrees, minutes, seconds] ===> 56° 49.81′
- Links to online mapping websites:
    - Google™ Maps
    - Yahoo!® Maps
    - MSN® Maps & Directions
    - Mapquest®
    - Open KML data with Google™ Earth
    - Save KML data to file
    - Save KML data to file and open with Google™ Earth
- GPS Time Stamp / UTC Time = 14/1,20/1,219/100 [hours, minutes, seconds] ===> 14h 20m 2.19s


N
dified/2009/Mar 24, 2009/IMG_0037.jpg
/Users/ezequielsallis/Desktop/IMG_0037.
Image © 2009 DigitalGlobe
© 2009 Inav/Geosistemas SRL
© 2009 LeadDog Consulting
Google

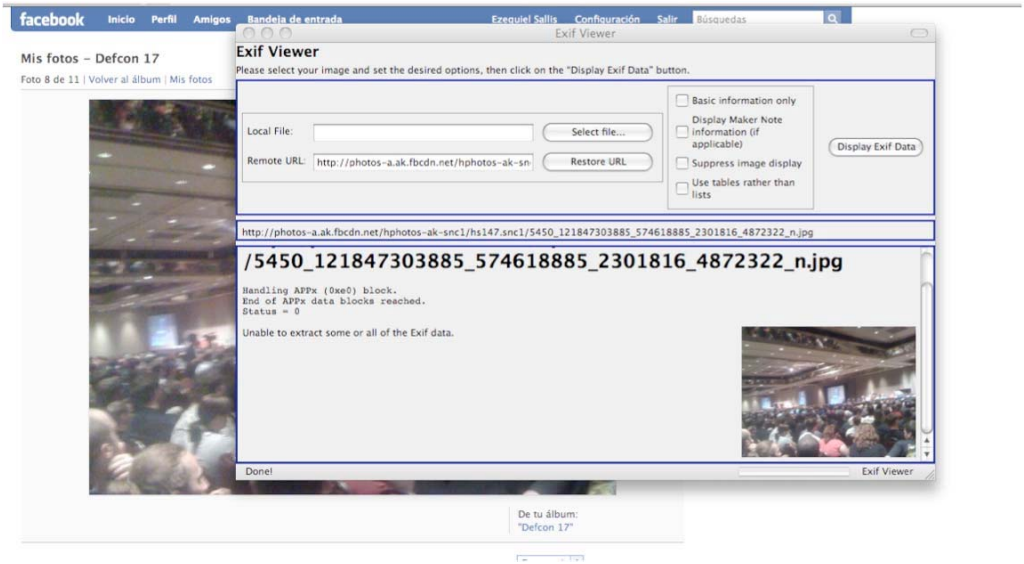

N
dified/2009/Mar 24, 2009/IMG_0037.jpg
/Users/ezequielsallis/Desktop/IMG_0037.
Image © 2009 DigitalGlobe
Google

Donde se publiquen las fotografías es un factor relevante, ya que en algunos casos, el proceso de publicación incluye la eliminación de los metadatos, como en el caso de Facebook, este no es un proceso que este relacionado con la seguridad sino mas bien con la propiedad intelectual de las imágenes que se suben a la red. En el caso por ejemplo de Flickr o de Fotolog.com.ar los metadatos quedan en su lugar, incluso en algunos casos como Picassa y Flickr existen programas desarrollados para realizar búsqueda directas sobre los datos almacenados en el exif.

Estado de los metadatos antes de la publicación en Facebook

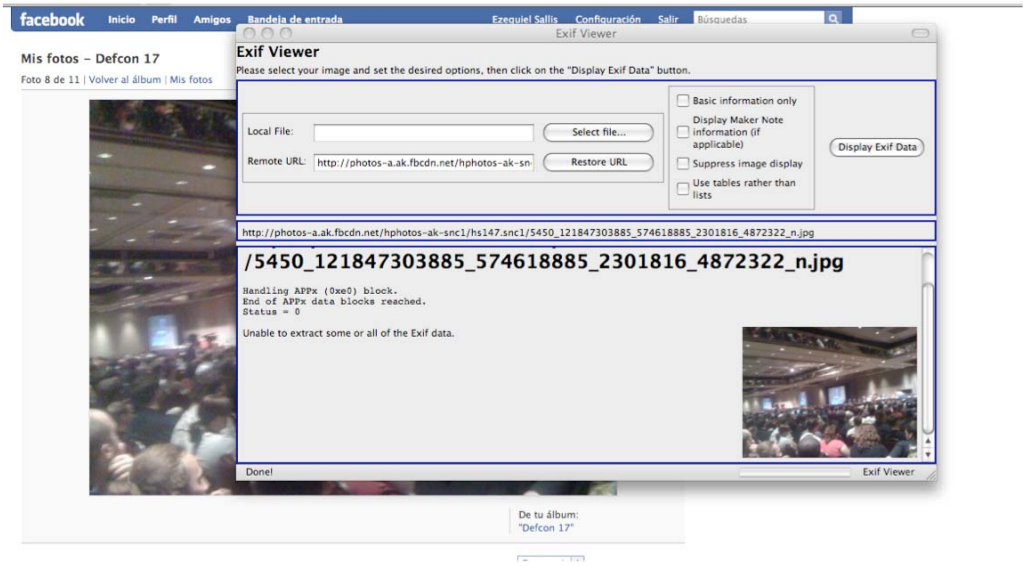


Estado de los metadatos después de la publicación en Facebook



Conclusión:

Este documento tiene como fin, poder ilustrar los riesgos que se esconden detrás de la interacción con las nuevas tecnologías y las nuevas costumbres. Para concluir aquí, les dejo algunas referencias para leer más sobre el tema y para poder obtener las herramientas necesarias para eliminar (12) los metadados de las fotografías antes de proceder a su publicación.



Conclusión:

Este documento tiene como fin, el poder ilustrar los riesgos que se esconden detrás de la interacción con las nuevas tecnologías y las nuevas costumbres. Para concluir aquí les dejo algunas referencias para leer mas sobre el tema y para poder obtener las herramientas necesarias para eliminar (12) los metadados de las fotografías antes de proceder a su publicación.

Referencias:

(1) http://www.edge-security.com/metagoofil.php
(2) http://www.informatica64.com/foca/
(3) http://es.wikipedia.org/wiki/Exchangeable_image_file_format o http://www.exif.org/specifications.html
(4) http://www.computerbytesman.com/privacy/blair.htm
(5) http://graphicssoft.about.com/b/2003/07/26/techtvs-cat-schwartz-exposed-is-photoshop-to-blame.htm
(6) http://es.wikipedia.org/wiki/Pedofilia
(7) http://kungfoosion.blogspot.com/2009/01/los-metadatos-de-cristina.html
(8) http://es.wikipedia.org/wiki/Sistema_de_posicionamiento_global
(9) http://www.celularis.com/apple-iphone/el-iphone-ya-es-la-camara-mas-usada-en-flickr.php
(10) http://www.flickr.com/
(11) https://addons.mozilla.org/en-US/firefox/addon/3905
(12) http://www.steelbytes.com/?mid=30