

# Big Data: cómo afecta a la privacidad de los ciudadanos

**Federico Valero Valdés**  
Master Seguridad de las TIC  
BIG DATA/SEGURIDAD

**Tutor: Marco Antonio Lozano Merino**

Fecha 04/06/2018



Esta obra está sujeta a una licencia de Reconocimiento-NoComercial-SinObraDerivada [3.0 España de Creative Commons](https://creativecommons.org/licenses/by-nc-nd/3.0/es/)

<b>Título del trabajo:</b>	<i>BIG DATA: Cómo afecta a la privacidad de los ciudadanos</i>
<b>Nombre del autor:</b>	<i>Federico Valero Valdés</i>
<b>Nombre del consultor/a:</b>	<i>Marco Antonio Lozano Merino</i>
<b>Nombre del PRA:</b>	
<b>Fecha de entrega (mm/aaaa):</b>	MM/AAAA
<b>Titulación::</b>	<i>Master seguridad de las TIC</i>
<b>Área del Trabajo Final:</b>	<i>Big Data y Seguridad</i>
<b>Idioma del trabajo:</b>	<i>Español</i>
<b>Palabras clave</b>	<i>Big data, seguridad, privacidad</i>
<b>Resumen del Trabajo (máximo 250 palabras):</b> <i>Con la finalidad, contexto de aplicación, metodología, resultados i conclusiones del trabajo.</i>	
<p>El título del trabajo es BIG DATA: Cómo afecta a la privacidad de los ciudadanos, como bien dice el título, en este trabajo se va a investigar sobre como de expuestos están nuestros datos hoy en día por la red, debido al ya conocido "BIG DATA" y su tratamiento masivo de nuestra información, para que cualquiera pueda hacer uso de ellos sin que los usuarios lo sepan y cómo afecta esto a la seguridad y privacidad de los ciudadanos.</p> <p>Se analizaran casos reales de empresas que hacen uso de datos personales, y por el contrario veremos qué podemos hacer los ciudadanos para protegernos apoyándonos en la ley de protección de datos actualizada.</p> <p>Conoceremos tipos de ataques reales que los atacantes pueden realizar para obtener nuestra información y que maneras tenemos los ciudadanos de protegernos ante estos ataques.</p> <p>¿Hasta qué punto estamos dispuestos a regalar nuestra información por un servicio gratuito? ¿Realmente merece la pena?</p>	
<b>Abstract (in English, 250 words or less):</b>	
<p>The title of the work is BIG DATA: How it affects the privacy of citizens, as the title says, in this work we will investigate how our data is exposed today by the network, due to the already known "BIG DATA "and its massive treatment of our information, so that anyone can make use of them without users knowing and how this affects the security and privacy of citizens.</p> <p>Real cases of companies that make use of personal data will be analyzed, and on the contrary, we will see what citizens can do to protect us, based on the updated data protection law.</p> <p>We will know the types of real attacks that attackers can take to obtain our information and what ways we citizens have to protect ourselves from these</p>	

attacks.

To what extent are we willing to give away our information for a free service? Is it really worth it?

# Índice

1. Introducción.....	1
1.1 Contexto y justificación del Trabajo.....	1
1.2 Objetivos del Trabajo.....	1
1.3 Enfoque y método seguido.....	2
1.4 Planificación del Trabajo.....	2
1.5 Estado del arte.....	3
2.1. ¿Qué es Big Data y cómo ha evolucionado con el tiempo?.....	4
2.2. Casos Reales de éxito con el uso de nuestra información.....	8
2.2.1.Caso Obama.....	8
2.2.2.Netflix.....	9
2.3 Ley de protección de datos.....	10
2.3.1¿Qué es GDPR?.....	10
2.3.2.Nuestros derechos.....	11
2.3.3.Como ejercer nuestros derechos.....	12
2.3.4.Obligaciones de las empresas.....	13
2.4. Como obtienen nuestros datos y para que los usan.....	14
2.4.1 Como funcionan las herramientas que recogen nuestra información y como evitar que las utilicen con nosotros.....	15
2.4.2.Caso práctico. ¿Qué información enviamos y como la tratan cuando entramos a una web?.....	19
2.4.3El Gran Hermano te observa.....	22
2.5. OSINT ¿Estamos expuestos públicamente?.....	24
2.5.1 ¿Qué es OSINT?.....	24
2.5.3 OSINT FRAMEWORK.....	25
2.5.4 Proceso OSINT.....	27
2.5.5 Herramientas OSINT.....	28
2.5.6 Como podemos protegernos.....	29
2.6. Caso práctico recolección de información.....	30
2.7. Delitos con nuestra información.....	41
3. Conclusión.....	42
4.Glosario.....	43
5.Bibliografía.....	44

# 1. Introducción

## 1.1 Contexto y justificación del Trabajo

Vivimos en un mundo rodeado de datos que hasta hace relativamente poco, no se sabía muy bien cómo usarlos, de manera que aparentemente se estaba “desaprovechando” mucha información que hoy en día es mucho más valiosa de lo que nos podemos llegar a imaginar, ya empieza a considerarse estos datos como el próximo “oro negro”.

Cuando nos registramos en una red social, nos bajamos una aplicación, o simplemente algo tan cotidiano ya como es navegar por internet, las diferentes empresas almacenan multitud de información sobre nosotros que sin ni siquiera darnos cuenta, se la estamos regalando para que la usen en su propio beneficio.

A lo largo de este trabajo se pretende ver si realmente merece la pena regalar nuestra información, y que ganamos nosotros a cambio frente a lo que ganan los que utilizan nuestros datos. ¿Hay un “mercado negro de datos” detrás de todo esto?

Se pretende analizar las distintas herramientas que existen para obtener información privada, conocer como nos podemos defender ante el que usa nuestra información para fines propios.

## 1.2 Objetivos del Trabajo

El objetivo del trabajo se podría resumir en los siguientes puntos a tratar:

Primero conoceremos que es Big Data y como está evolucionando con el tiempo, en este apartado se va a profundizar en el término Big Data, e iremos viendo cómo la información que antes no se usaba como son los metadatos ahora es una información incluso más interesante que la propia información que el usuario da conscientemente.

Seguidamente trataremos el tema de la Ley de protección de datos, a raíz de la actualización de dicha ley, es muy interesante profundizar en la misma y conocer los derechos que tenemos los ciudadanos con nuestra propia información y conocer hasta qué punto nuestra información es nuestra o no y que es lo que las empresas pueden realizar con ella y a su vez conocer cómo podemos evitar que la usen.

El siguiente punto a tratar sería conocer quien usa nuestra información, para qué la necesita y que ganan con ella, en esta sección hablaremos sobre qué tipo de datos usan las empresas sobre nosotros, como consiguen esta información, que beneficios sacan ellos de nuestros datos y que recibimos nosotros a cambio. Se analizaran casos concretos como puede ser google, Facebook etc.

También trataremos el tema de ataques a nuestra privacidad, es decir, tipos de ataques a nuestra información, donde analizaremos herramientas como por ejemplo OSINT que permiten acceder de manera masiva a la información para cometer algún tipo de ciberdelito a través de las direcciones de correo encontradas, números de teléfono, etc.

Tras el estudio realizado comentaremos “El coste de nuestros datos”, es decir, una vez hemos analizado y estudiado que se hace con nuestros datos, reflexionaremos sobre si realmente merece la pena “regalar” nuestra información.

### 1.3 Enfoque y método seguido

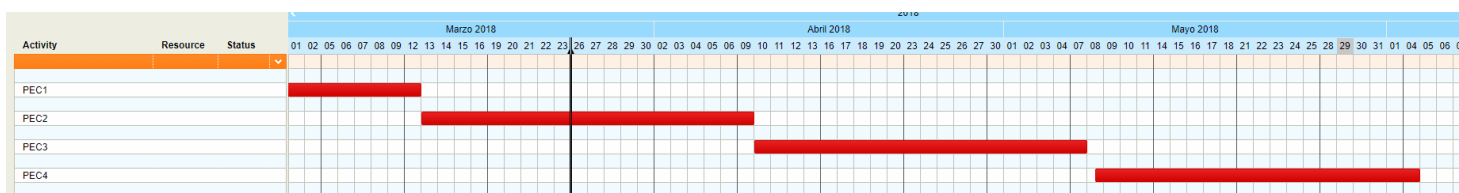
El método de trabajo, básicamente va a consistir en buscar información sobre el tema a tratar, de las máximas fuentes posibles para poder contrastar y finalmente obtener la información más próxima a la realidad posible.

Se estudiará en profundidad los casos reales que se expondrán en el trabajo, y se van a estudiar las herramientas que se utilizan para obtener información y poder aportar la máxima información al trabajo acerca de ellas.

También se ha de leer en profundidad y analizar la ley de protección de datos y conocer que artículos tienen que ver con el tema a tratar.

### 1.4 Planificación del Trabajo

El trabajo se realizara en el periodo de 21 marzo a 4 de junio, se realizaran entregas parciales, el orden de estas entregas es el siguiente, representado en un gantt:



En cada una de las entregas se pretende abordar uno de los temas comentados en el apartado anterior ‘Objetivos del trabajo’.

En la primera entrega nos centraremos en concepto de Big Data, haciendo una introducción al término y todo lo que conlleva y seguidamente abordaremos el tema de la ley de protección de datos.

En la siguiente entrega, nos centraremos en estudiar casos reales de uso de información de los ciudadanos, conocer como obtienen dicha información, como la usan y para que la necesitan.

En la última entrega hablaremos de los ataques y herramientas existentes para poder hacer uso de una manera “maliciosa” de la información privada.

## 1.5 Estado del arte

Cuando hablamos de Big data nos referimos a conjuntos de datos o combinaciones de conjuntos de datos cuyo tamaño, complejidad y velocidad de crecimiento dificultan su captura, gestión, procesamiento o análisis mediante tecnologías y herramientas convencionales, tales como bases de datos relacionales y estadísticas convencionales o paquetes de visualización, dentro del tiempo necesario para que sean útiles.

Lo que hace que Big Data sea tan útil últimamente para muchas empresas es el hecho de que proporciona respuestas a muchas preguntas que las empresas ni siquiera sabían que tenían, es decir, proporciona un punto de referencia.

Con una cantidad tan grande de información, los datos pueden ser moldeados o probados de cualquier manera que la empresa considere adecuada, al hacerlo, las organizaciones son capaces de identificar los problemas de una forma más comprensible, de ahí que la información hoy en día se pueda llegar a considerar el “oro negro”.

La recopilación de grandes cantidades de datos y la búsqueda de tendencias dentro de los datos permiten que las empresas se muevan mucho más rápidamente, sin problemas y de manera eficiente. También les permite eliminar las áreas problemáticas antes de que los problemas acaben con sus beneficios o su reputación.

Como todas las cosas en esta vida, puede tener un buen uso o usarse para propósitos no tan buenos, lo primero que llama la atención es el tema de la privacidad, ya que cada vez más detalles de nuestras vidas son almacenados y analizados por empresas y gobiernos, por supuesto. Esto no es algo que nos debemos tomar a la ligera, pero a medida que siga avanzando la tecnología, habrá que ir adaptando las leyes y regulaciones para proteger a las personas.



## 2.1 ¿Qué es Big Data y cómo ha evolucionado con el tiempo?

Comenzaremos definiendo qué es Big data, es un término que hace referencia al gran volumen de datos ya sean estructurados o no estructurados, que generan los negocios cada día.

Aunque ya no es la cantidad de datos lo que es importante, si no lo que importa con el Big Data es como y para que las organizaciones hacen uso de los datos. Big Data se puede analizar para obtener ideas que conduzcan a mejores decisiones y movimientos de negocios estratégicos.

De manera más técnica podríamos definir el Big Data como un conjunto de datos o combinaciones de conjuntos de datos cuyo tamaño (volumen), complejidad (variabilidad) y velocidad de crecimiento (velocidad) dificultan su captura, gestión, procesamiento o análisis mediante tecnologías y herramientas convencionales, como por ejemplo las bases de datos relacionales y estadísticas convencionales o paquetes de visualización, dentro del tiempo necesario para que sean útiles.

Principalmente la complejidad de la naturaleza del Big Data viene de que los datos no están estructurados, datos como por ejemplo: Logs, la identificación por radiofrecuencia (RFID), los sensores incorporados en dispositivos, la maquinaria, los vehículos, las búsquedas en Internet, las redes sociales como Facebook, computadoras portátiles, teléfonos inteligentes y otros teléfonos móviles, dispositivos GPS y registros de centros de llamadas, de ahí la importancia de conseguir dar utilidad a datos difícilmente procesables.

Mucha de la información que hoy gracias al Big Data se está pudiendo explotar, realmente ya la teníamos, lo que hace que Big Data sea tan útil para muchas empresas es el hecho de que proporciona respuestas a muchas preguntas que las empresas ni siquiera sabían que tenían. En otras palabras, proporciona un punto de referencia. Con una cantidad tan grande de información, los datos pueden ser moldeados o probados de cualquier manera que la empresa considere adecuada. Al hacerlo, las organizaciones son capaces de identificar los problemas de una forma más comprensible.

El análisis de Big Data ayuda a las organizaciones a aprovechar sus datos y utilizarlos para identificar nuevas oportunidades por lo que esto conduce a movimientos de negocios más inteligentes, operaciones más eficientes, mayores ganancias y clientes más felices. Las empresas con más éxito con Big Data consiguen valor de las siguientes formas:

- **Reducción de coste.** Con tecnologías como Hadoop y el análisis basado en la nube, aportan importantes ventajas en términos de costes cuando se trata de almacenar grandes cantidades de datos, además de identificar maneras más eficientes de hacer negocios.
- **Más rápido, mejor toma de decisiones.** Con la velocidad de Hadoop y la analítica en memoria, combinada con la capacidad de analizar nuevas fuentes de

datos, las empresas pueden analizar la información inmediatamente y tomar decisiones basadas en lo que han aprendido.

- **Nuevos productos y servicios.** Con la capacidad de medir las necesidades de los clientes y la satisfacción a través de análisis viene el poder de dar a los clientes lo que quieren. Con la analítica de Big Data, más empresas están creando nuevos productos para satisfacer las necesidades de los clientes.

No solo es útil para empresas tecnológicas, el Big Data es muy útil en todo tipo de sectores, como por ejemplo:

- **Turismo:** El análisis de Big data ofrece a estas empresas la capacidad de recopilar datos de los clientes, aplicar análisis e identificar inmediatamente posibles problemas antes de que sea demasiado tarde con el fin de mantener contento al cliente y que viva una buena satisfactoria experiencia.
- **Publicidad:** La proliferación de teléfonos inteligentes y otros dispositivos GPS ofrece a los anunciantes la oportunidad de dirigirse a los consumidores cuando están cerca de una tienda, una cafetería o un restaurante, esto abre nuevos ingresos para los proveedores de servicios y ofrece a muchas empresas la oportunidad de conseguir nuevos prospectos.
- **Cuidado de la salud:** El Big Data aparece en grandes cantidades en la industria sanitaria. Los registros de pacientes, planes de salud, información de seguros y otros tipos de información pueden ser difíciles de manejar, pero están llenos de información clave una vez que se aplican las analíticas. Es por eso que la tecnología de análisis de datos es tan importante para el cuidado de la salud. Al analizar grandes cantidades de información - tanto estructurada como no estructurada - rápidamente, se pueden proporcionar diagnósticos u opciones de tratamiento casi de inmediato.
- **Administración:** La administración se encuentra ante un gran desafío: mantener la calidad y la productividad con unos presupuestos ajustados. Esto es particularmente problemático con lo relacionado con la justicia. La tecnología agiliza las operaciones mientras que da a la administración una visión más holística de la actividad.
- **Retail:** El servicio al cliente ha evolucionado en los últimos años, ya que los compradores más inteligentes esperan que los minoristas comprendan exactamente lo que necesitan, cuando lo necesitan, el Big Data ayuda a los minoristas a satisfacer esas demandas.
- **Empresas manufactureras:** Estas despliegan sensores en sus productos para recibir datos de telemetría. A veces esto se utiliza para ofrecer servicios de comunicaciones, seguridad y navegación. Ésta telemetría también revela patrones de uso, tasas de fracaso y otras oportunidades de mejora de productos que pueden reducir los costos de desarrollo y montaje.

Para ver el principio de donde empezó a surgir el concepto de Big Data nos podríamos remontar mucho años atrás, incluso hay quienes lo sitúan en el paleolítico con una lógica que relaciona el término con el primitivo interés de los seres humanos por lograr y procesar la información.

Como dato curioso, algunas fuentes indican que en el **18000 AEC**, en el Paleolítico Superior se empleaban rudimentarios métodos de almacenamiento de datos con el empleo de palos o muescas en huesos. Con este sistema, se podía llevar cuenta de provisiones, realizar cálculos básicos e incluso predecir necesidades de comida para el grupo. Quizá sea demasiado incluirlo en la historia del Big Data, pero es el primer momento documentado en el que la humanidad se interesa por los datos: el germen de todo lo que viene después. Si bien las cantidades no podían ser muy grandes, es la primera evidencia del interés por recopilar, contar y guardar datos.

En lo que a nosotros nos interesa no queda tan lejos como el paleolítico, podríamos comenzar la cronología por el año **1989** cuando Erik Larson habla por primera vez de Big Data en el sentido que conocemos la expresión hoy en día. La revista Harpers Magazine recoge su artículo, en el que especula sobre el origen del correo basura que recibe. En torno a este año se empiezan a popularizar las herramientas de business intelligence para analizar la actividad comercial y el rendimiento de las operaciones.

En el año **1991** con el nacimiento de internet nace también la gran revolución de la recolección, almacenamiento y análisis de datos. Tim Berners-Lee establece las especificaciones de un sistema de red con interconexiones a nivel mundial accesible para todos en cualquier lugar.

Pocos años después, en **1993**, se funda QlikTech, germen de la actual Qlik, que crea un sistema revolucionario de business intelligence.

En el **1996**, los precios del almacenamiento de datos empiezan a ser accesibles con un coste eficiente en lo que es una de las grandes revoluciones en la historia del Big Data. El libro La evolución de los sistemas de almacenamiento, de 2003, establece esta fecha como el primer año en el que el almacenamiento digital es más barato que el papel.

Tras el nacimiento de Google, en **1997** este lanza su sistema de búsqueda en internet y en los siguientes años será de largo el primer lugar al que acudir en busca de datos en internet. Este mismo año, se publica el estudio ¿Cuánta información hay en el mundo?, de Michael Lesk. La conclusión es que hay tanta y crece a tal velocidad, que gran parte de ella no será vista por nadie jamás.

Más tarde empezamos a utilizar en termino Big Data en **1999**, término Big Data es analizado por primera vez en un estudio académico. La Asociación de Sistemas Informáticos recomienda centrarse en el análisis de información ya que existe gran cantidad de datos y no todos son útiles. Recuerdan el propósito de la computación, que es el entendimiento, no los números.

Doug Laney, de Gartner en el año **2001** define las 3 V's del Big Data. Este es un hito clave en la historia del big data. Se trata de tres conceptos que definen el término: volumen, velocidad y variedad.

En **2005** nace la Web 2.0, una web donde predomina el contenido creado por los usuarios. Este mismo año se crea Hadoop, un entorno de trabajo Big Data de software libre.

Dos años más tarde en **2007**, la revista Wired publica un artículo que lleva el concepto de Big Data a las masas.

Hace aproximadamente 8 años, por el **2010**, según Eric Schmidt (Google) los datos que se generan en dos días equivalen a la cantidad de datos generados desde el inicio de la civilización hasta 2003.

En el **2013** se sabe que el archivo de mensajes públicos de Twitter en la Biblioteca del Congreso de Estados Unidos llega a los 170 billones de mensajes, creciendo a ritmo de 500 millones al día. Según la institución que alberga algunos de los documentos históricos más importantes del mundo, dicho archivo ofrece una imagen más amplia de las normas culturales, diálogos, tendencias y eventos de hoy en día. De este modo, contribuye a una mejora de la información en procesos legislativos, educación, definición de autoría de nuevos trabajos y otras cuestiones.

En **2014** coincidiendo ya con el boom de los smartphones, los móviles superan a los ordenadores en accesos a internet. La conexión casi continua contribuye a generar muchos más datos y mejora la conectividad con otros dispositivos.

Hace pocos años por el **2016** el Big Data se convierte en la palabra de moda. Se generaliza la contratación de expertos en Big Data, el Machine Learning llega a las fábricas y el "Internet de las Cosas" empieza a impregnarlo todo.

En la actualidad los datos llegan a las masas. La gente controla sus patrones de descanso con pulseras, sabe en qué se gasta el dinero con aplicaciones móviles etc. Los datos están en todas partes y la población está ya predispuesta a usarlos.

Futuro. ¿Qué nos deparará el futuro? Muy difícil de pronosticar, pero seguramente un aumento de datos y la consiguiente necesidad de tecnología para recogerlos, adaptarlos, almacenarlos y analizarlos. La computación cuántica está a la vuelta de la esquina y la historia del big data sigue avanzando.

## 2.2. Casos Reales de éxito con el uso de nuestra información

Por ir entrando en el tema de la privacidad en Big Data, considero interesante analizar casos reales de éxito mediante el uso de Big Data, es decir mediante el análisis de nuestra información, y ser conscientes del alcance de este concepto “Big Data” y lo importante que es nuestra información, tan importante es que a partir de unos análisis, el Big Data ayudo a Obama a ganar las elecciones de Estados Unidos.

### 2.2.1 Caso Obama

Fue Obama en 2012 el primer candidato a unas elecciones presidenciales que decidió utilizar una combinación de la base de datos analíticos HP Vertica MPP con modelos predictivos, con la finalidad de obtener un mayor margen de competencia frente a sus rivales.

¿Cuál era la intención principal de usar este mecanismo? Básicamente la idea era sencilla, conseguir convencer a los posibles votantes y asegurarse los votantes pasados, según Chris Wegzyn, director de Data Architecture en el Comité Nacional del Partido Demócrata pretendían tres objetivos aplicando Data Analytics a dicha campaña:

1. **Registros.** Aumentar el número de votantes inscritos para la votación.
2. **Persuasión.** Convencer de que el voto fuera para Obama.
3. **Asistencia.** Conseguir que el número de votantes el día de la votación fuera el más alto posible. .

Para llevar a cabo este proyecto, tuvo que multiplicar por cinco el número de trabajadores destinados a este proyecto. El equipo de campaña de Obama decidió recopilar toda la información que los ciudadanos estadounidenses publicaban en la red, de esta manera podían saber quién estaba a favor de qué medidas propuestas por el futuro presidente y quién no, con el fin de ir mejorando de esta manera sus propuestas y su enfoque con tal de acercarse a más gente.

Por otro lado, el proyecto destacó además por conocer mejor a los diferentes segmentos en los que sus votantes estaban divididos, especialmente a aquellos sectores más indecisos, y poder convencerles en los medios en los que prevalecían. Una vez detectado el sector de los indecisos, entre otras medidas, por ejemplo, decidieron anunciarse en las pausas publicitarias de Walking Dead o en la revista Reddit, ya que según los estudios ahí se encontraban los segmentos a los que tenían que convencer.

Otra anécdota interesante mediante el uso de Big data en estas elecciones históricas fue básicamente para lidiar con el complicado estado de Ohio debido a la geolocalización, se hizo un estudio mediante la información de los habitantes de este lugar, conociendo finalmente cuáles son las inquietudes de sus habitantes y sabiendo utilizarlas ya no sólo para saber en qué mejorar, sino también para saber en qué publicidad invertir con el fin de que ésta les llegue a los habitantes de Ohio.

Para recaudar fondos se analizó los gustos de los ciudadanos, de ahí surgió la famosa campaña “Obama, Cloonie y tu”, la cena de George Clooney con el que fué presidente de Estados Unidos llevó a las mujeres de entre 40 y 49 años a invertir mucho dinero en

la campaña electoral, como dato, con esta campaña tras analizar los gustos de los ciudadanos, se consiguió ganar 15 millones de dólares.

### 2.2.2 Netflix

Por poner un ejemplo de éxito más actual, si el Big Data es capaz de lograr ganar unas elecciones de Estados Unidos, ¿qué no podrá hacer en el campo del marketing?

El caso de Netflix, la plataforma de vídeo por suscripción ha sabido aprovechar como nadie las posibilidades que le ofrece esta tecnología.

No decide sus contenidos como lo hacen las cadenas tradicionales, consultando simplemente la opinión de unos directivos, sino que estudia el comportamiento y los hábitos de sus consumidores para saber qué es lo que los que realmente van a consumir el producto, quieren ver.

Netflix lo tiene realmente fácil para acceder a la información de sus espectadores ya que esta información es el mismo usuario el que se la está facilitando, a diferencia de medios como la televisión, que se basan en cuotas de pantalla, en el caso de Netflix el proveedor de vídeo puede extraer los datos de todos sus usuarios como por ejemplo:

- Qué búsquedas realizan.
- Qué dispositivos usan
- Cuál es su día preferido
- Cuánto tiempo emplean en el servicio y en cada uno de los contenidos
- Si ven los capítulos enteros o parcialmente e, incluso, qué fragmentos vuelven a visionar
- En qué momento abandonan el visionado y si lo recuperan o abandonan
- Las valoraciones de los consumidores
- Qué preferencias tienen en común con sus amigos o con la audiencia de su misma zona geográfica
- La información de sus perfiles en redes sociales...

La información que han recopilado durante casi 20 años es lo que les permite decidir qué películas y series incluir en su parrilla. Y de esta forma la firma estadounidense arrasa en todos los países donde se implanta, es por esto que en cada uno de ellos, ofrece una programación distinta adaptada a los datos que ha recogido, almacenado y analizado de los espectadores de ese lugar.

Podríamos afirmar que gracias al Big Data, Netflix ha logrado garantizar el éxito de sus productos. Para reconocer el éxito, basta con comparar los resultados de las cadenas convencionales con los de la plataforma.

Netflix invirtió mucho dinero en la serie House of Cards, unos 100 millones de dólares, el éxito no les cogió por sorpresa ya que habían estudiado los intereses de los ciudadanos y sabían que iba a ser un éxito, así fue que en el primer trimestre

consiguieron nada más y nada menos que 3 millones de suscriptores en su plataforma, utilizaron técnicas como por ejemplo conociendo el sector al que iba a ir dirigido el tráiler de la serie, se les ofrecía uno u otro, se llegaron a hacer 10 diferentes adaptados para los distintos sectores.

### 2.3. Ley de protección de datos

A modo de resumen, Ley Orgánica de Protección de Datos es una ley orgánica española que tiene por objeto garantizar y proteger, en lo que concierne al tratamiento de los datos personales, las libertades públicas y los derechos fundamentales de las personas físicas, y especialmente de su honor, intimidad y privacidad personal y familiar. Fue aprobada por las Cortes Generales el 13 de diciembre de 1999. Esta ley se desarrolla fundamentándose en el artículo 18 de la constitución española de 1978, sobre el derecho a la intimidad familiar y personal y el secreto de las comunicaciones. Su objetivo principal es regular el tratamiento de los datos y ficheros, de carácter personal, independientemente del soporte en el cual sean tratados, los derechos de los ciudadanos sobre ellos y las obligaciones de aquellos que los crean o tratan. Esta ley afecta a todos los datos que hacen referencia a personas físicas registradas sobre cualquier soporte, informático o no.

Actualmente se ha actualizado dicha ley, la cual se debe cumplir desde mayo de 2018 a partir de este momento tendremos que dar nuestro consentimiento inequívoco para que las empresas puedan usar nuestros datos si eres ciudadano europeo. Es más, te tendrán que decir qué datos están utilizando, cómo los están tratando, para qué y quién es la persona responsable de los mismos.

A partir de este momento entra en vigor en toda Europa una nueva ley de protección de datos: GDPR (General Data Protection Regulation). Una normativa que afecta a todas aquellas empresas que traten datos de los ciudadanos europeos aunque sean de Estados Unidos, como Google o Facebook.

Las grandes multas a las que se enfrentan quienes no cumplan con ella son uno de los puntos más controvertidos y mediáticos. Pero detrás de estas siglas también se esconde una nueva manera de informar a los usuarios sobre qué información cedemos y para qué se usa.

#### 2.3.1 ¿Qué es GDPR?

GDPR, por sus siglas en inglés, General Data Protection Regulation, o RGDPD por sus siglas en español, Reglamento General de Protección de Datos, es la nueva normativa que regula la protección de los datos de los ciudadanos que vivan en la Unión Europea. El reglamento entró en vigor el 24 de mayo de 2016, pero será de obligado cumplimiento a partir del 25 de mayo de 2018.

Esta nueva normativa determina que todas las empresas, independientemente de su país de origen o de actividad, deberán cumplirla si recogen, guardan, tratan, usan o

gestionan algún tipo de dato de los ciudadanos de la Unión Europea. Es decir, que Apple o Amazon (por poner algunos ejemplos) también están sujetas a ella.

Y, por supuesto, nos afecta a todas las personas que vivimos en la Unión Europea.

### 2.3.2 Nuestros derechos

Este reglamento recoge y reconoce, por tanto, entre otros, derechos como al olvido y el derecho a la portabilidad.

En cuanto al derecho al olvido, establece que los ciudadanos podemos solicitar y lograr que nuestros datos personales sean eliminados cuando, entre otros casos, estos ya no sean necesarios para la finalidad con la que fueron recogidos, cuando se haya retirado el consentimiento o cuando estos se hayan recogido de forma ilícita.

El derecho al olvido como tal ya existía anteriormente, desde la sentencia del Tribunal de Justicia de la Unión Europea, pero ahora se recoge en este Reglamento. El problema que tiene este derecho es que no es un derecho absoluto, existe confrontación con otros derechos como por ejemplo libertad de información, por lo que hay que ponderar cada caso concreto según unos criterios.

Por otro lado, el derecho a la portabilidad te permite que, si tus datos se están tratando de modo automatizado, puedas recuperarlos en un formato para cederlos a otro responsable. Estos datos deben estar en un formato estructurado, de uso común y lectura mecánica (por ejemplo un excel) para que pueda transmitirlos fácilmente a otro responsable y facilitar así un cambio de proveedor, por ejemplo.

Hay que aclarar que en principio, este derecho sólo se aplicaría a los que has aportado en cada web, es decir, no las segmentaciones que realizan o los tratamientos inferidos posteriores.

Por ejemplo, Facebook **solo** estaría obligado a darte los datos que tú has facilitado, no la información que hayas ido dejando con tus acciones en la red social. Sin embargo, en el futuro seguramente veamos muchas resoluciones de las autoridades de control perfilando este tema.

Otro de los nuevos derechos interesantes que contempla GDPR es el de acceso. Así, podrás pedir a las empresas que te confirmen si tus datos se están procesando, dónde y con qué propósito. Si lo haces, puedes pedir también una copia de tus datos personales sin que se te cobre por ello.

Con la entrada en vigor de GDPR, la solicitud de consentimiento debe darse en una forma inteligible y de fácil acceso, con el propósito del procesamiento de datos adjunto a ese consentimiento. Es decir, que el consentimiento debe ser inequívoco, claro y distinguible de otros asuntos. Las empresas tendrán que mostrar sus condiciones de forma inteligible y de fácil acceso, usando un lenguaje claro y sencillo.

Esto implica dos cosas. Por un lado, que cuando te des de alta en un servicio, web, aplicación o producto, el aceptar los términos de uso irá por un lado y, por otro, todo lo relativo al tratamiento de tus datos.



Por otro lado, dejaremos de ver casillas premarcadas para el envío de publicidad, consentimientos tácitos para comunicaciones comerciales o cesiones de datos. A partir del 25 de mayo, las empresas están obligadas a informarte sobre cada finalidad del tratamiento de los datos, "la legitimación para su recogida y uso, y en su caso, la obtención del consentimiento por separado para cada finalidad", por lo que empezaremos a ver más casillas para que aceptes cada uno de estos aspectos

Por otro lado, también con la entrada en vigor de GDPR las empresas deberán informar en un plazo de 72 horas de que han sufrido un incidente de seguridad. Y no sólo deberán dar parte a las autoridades competentes (en el caso de España, la Agencia de Protección de Datos), sino también a todos aquellos usuarios cuyos datos se hayan podido ver comprometidos.

### 2.3.3 Como ejercer nuestros derechos

Los datos personales que te identifican a ti, concretamente, como persona física son tuyos, y en la Ley de Protección de Datos vienen recogidos una serie de derechos que los usuarios podemos ejercer.

En primer lugar podemos ejercer el derecho más básico de todos, que es el **Derecho de consulta** en el Registro de Protección de Datos, y que establece que cualquier persona puede conocer la existencia de tratamientos de sus datos de carácter personal, la finalidad de los mismos y los responsables del tratamiento.

Como puedo ejercer este derecho, para ello, bastará con acceder a la web de la Agencia Española de Protección de Datos (<https://www.aepd.es>), y buscar en sus ficheros, introduciendo tus datos, cuentan con ficheros de titularidad pública y de titularidad privada, y podrás realizar la búsqueda en el que desees.

Por otro lado tenemos **derecho de acceso**, es decir, podemos solicitar y obtener gratuitamente información acerca de nuestros datos de carácter personal sometidos a tratamiento, el origen de estos y las comunicaciones que se han hecho o van a hacer de los mismos.

Este derecho, a diferencia del anterior, se ejercita frente al responsable del fichero, que es una empresa privada, por ejemplo nuestra compañía de teléfono, y con este obtenemos los datos personales concretos que el responsable del tratamiento tiene de nosotros, pero este tiene ciertas limitaciones, ya que sólo podremos realizar una consulta a la misma empresa cada doce meses, a no ser que acreditemos un interés legítimo. La empresa, por su parte, tiene que resolver esta solicitud de acceso en un plazo máximo de un mes desde que recibió la solicitud, y, en caso de estimarla, debe proporcionar dichos datos en un plazo de diez días.

También tenemos **derecho de rectificación**, esto significa que los ciudadanos tenemos derecho a corregir errores y a modificar datos inexactos o incompletos para garantizar la certeza de la información, pero esto es siempre y cuando los datos que queramos modificar sean, como hemos dicho, inexactos e incompletos, como por ejemplo el año de nacimiento, dirección etc.

Para ejercer este derecho lo que debemos de hacer es dirigirnos a la empresa en cuestión que trata tus datos y redactar un escrito en el que se señale claramente el dato erróneo, y la corrección que debe realizarse, acompañado de la documentación que lo justifique. El responsable del fichero tendrá que atender la solicitud en un plazo de diez días desde la recepción de la misma, incluso aunque no tenga datos del afectado, en cuyo caso deberá comunicarlo.

Para ello, tenemos disponibles los formularios disponibles en la web de la AEPD. Incluidos también los que se deben presentar en caso de querer solicitar la tutela de tus derechos en caso de que no te respondan dentro del plazo.

Otro de los derechos importantes que tenemos los ciudadanos y podemos ejercerlo es el **Derecho de cancelación**, permite que podamos defender nuestra privacidad, pudiendo solicitar la eliminación de nuestros datos personales, pero solo en caso de que los mismos sean inadecuados o excesivos.

Esta cancelación implica el bloqueo de esos datos, es decir, se identificarán y reservarán los mismos para impedir que se traten, excepto para la puesta a disposición de los correspondientes órganos judiciales, en este caso, el responsable del fichero o del tratamiento debe hacer efectivo el derecho de cancelación en un plazo de diez días. A no ser, claro está que exista un deber legal de conservar dichos datos, aunque también podrá denegar dicha cancelación cuando la conservación sea necesaria para cumplir obligaciones contractuales que le vinculen con la persona que solicite dicha cancelación.

Cabe destacar que la Unión Europea reconoce también el derecho de cancelación, pero con otro nombre, "Derecho a la supresión". Es decir, el también llamado Derecho al olvido, que nos permitiría obtener del responsable de los datos la supresión de los mismos si no son necesarios para la finalidad que justifique el tratamiento, o bien porque, simplemente, queramos revocar el consentimiento.

#### 2.3.4 Obligaciones de las empresas

A modo de resumen, algunas de las obligaciones más destacables que las organizaciones tendrán que cumplir con la llegada de GDPR son las siguientes:

- **Legalidad, legitimidad y transparencia.** Las empresas que gestionen datos personales tienen la obligación de informar a los usuarios acerca de cómo se procesará su información en cumplimiento con la normativa.
- **Limitación de uso.** La información personal sólo puede recogerse con un fin explícito y legítimo, y su uso no puede expandirse más allá del consentimiento del usuario.
- **Minimización de los datos.** Existe una clara tendencia por parte de la mayoría de organizaciones de maximizar los datos. Con la llegada del GDPR, los datos personales recogidos deben limitarse únicamente a lo que es necesario en relación con los objetivos para los cuales fueron recogidos y tratados.

- **Precisión.** La información personal debe ser precisa. Los usuarios deben tener derecho a solicitar correcciones que deben ser atendidas a la mayor brevedad posible.
- **Limitación de almacenamiento.** Las organizaciones están obligadas a no retener los datos personales más tiempo del necesario para el uso explícito y legítimo autorizado por el usuario.
- **Integridad y confidencialidad.** Las empresas que gestionen datos deben garantizar un nivel adecuado de seguridad que incluye la protección frente a tratamiento sin autorización o ilegal, y frente a pérdidas, destrucción o daños accidentales.

#### 2.4. Como obtienen nuestros datos y para qué los usan

Como ya hemos comentado anteriormente nosotros mismos, somos los que ponemos en peligro nuestros datos, consciente o inconscientemente, pongamos un ejemplo de un día normal de una persona cualquiera.

Esta persona cualquiera se despierta para ir a trabajar mira la pantalla de su teléfono móvil, pone Elche (Alicante) [revela su localización]. Abre la aplicación de su pulsera para ver cómo ha dormido gracias a su pulsera inteligente [documenta todos sus movimientos y horas de sueño].

A continuación mientras desayuna, lee los mensajes que le han llegado esta noche vía Whatsapp [debido a sus ajustes de la propia aplicación estos mensajes están en la nube].

Se prepara un bol de cereales y prepara la foto perfecta para subir a Instagram y menciona a la marca mediante un hashtag [avisa a la marca de que es consumidora de sus productos].

Mientras tanto, busca en Google sobre las nuevas zapatillas que se quiere comprar [las empresas de moda ya tienen un nuevo cliente al que enviar anuncios].

A continuación sale a la calle para dirigirse al trabajo, y como es ya habitual lee en la pantalla de su teléfono móvil: "Con las condiciones actuales tardará unos 30 minutos en llegar a su trabajo" [ha registrado automáticamente sus rutinas diarias, ya saben dónde vive y donde trabaja, la hora a la que entra a trabajar...]. Mientras espera al autobús, decide comprar una nueva aplicación por internet [sus datos bancarios están almacenados en el servidor] y se registra en él de forma automática con

Facebook [acaba de dar acceso a la app a toda la información que tiene en la red social]. Cuando la abre, le pide acceso a sus contactos, al calendario, a la galería... Acepta las condiciones sin apenas prestar atención [su agenda y actividades están en manos de terceros].

Solo con el uso diario normal de nuestro terminal, esta persona está ofreciendo gratuitamente muchísima información, ese ciudadano, podría ser una de las miles de personas que no se preocupan por el uso que terceros hacen de sus datos. Pero su información tiene un precio que muchos pagan.

Hace ya unos cuantos años George Orwell escribió “El Gran Hermano te observa” (1984), pero como vemos la realidad acaba superando a la ficción, estamos más cerca que nunca de 1984, la fantasía distópica de George Orwell tan solo 34 años después, Tanto es así que su libro, una crítica despiadada del estalinismo y el fascismo publicada en 1949, se ha convertido en el sorprendente best seller de la primera semana de la América de Donald Trump. Es número uno en las listas de Amazon.com y el 16 en su versión española.

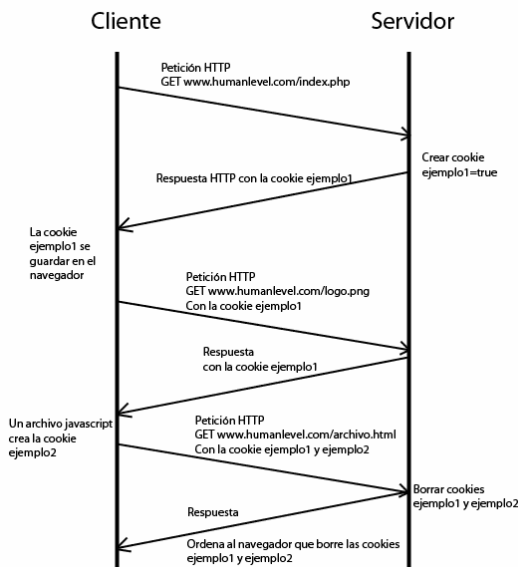
Lo saben todo de nosotros y pueden anticipar qué haremos o incitarnos a hacerlo. Con el caso que comentábamos en el caso anterior, buscamos unas zapatillas por ejemplo en una tienda online, pero no las compramos. Durante días, los anuncios de las páginas web que visitamos no hacen más que mostrar esas zapatillas de distintas tiendas, precios etc, es más, Amazon sabe qué zapatillas vamos a comprar antes de que lo busquemos y lo tiene listo para enviarlo. ¿Cómo es posible? Gracias a los datos que vamos dejando cada día. Las conocidas cookies. Estas galletas informáticas se almacenan en nuestro navegador y permiten a las páginas conocer qué sitios hemos visitado antes.

#### 2.4.1 Como funcionan las herramientas que recogen nuestra información y como evitar que las utilicen con nosotros

Hay muchos métodos o herramientas que se utilizan continuamente para recopilar información sobre los usuarios, ya adelantábamos antes las famosas galletas o **cookies**, entremos un poco más en detalle.

Una cookie es un archivo con información que se almacena en el navegador del usuario cuando visita un sitio web y en el que se suelen guardar las configuraciones y preferencias del usuario o el estado de la sesión de navegación. Las cookies pueden ser persistentes o no persistentes. Son persistentes, si tienen un tiempo de expiración y si no lo tienen, se borran en el momento en que se cierra el navegador. Las cookies no persistentes son las que se usan para mantener abierta la sesión del usuario, ya que el servidor sabe por el identificador de la cookie a qué usuario pertenece cada sesión abierta en memoria.

Un ejemplo de funcionamiento es el siguiente:



¿Cómo puede afectar las cookies a nuestra privacidad?, bien, las cookies no pueden por sí solas determinar información acerca de un usuario, pero si sirven para rastrear las acciones que lleva a cabo una persona cuando navega por Internet, cuando rellena formularios, se registra en un sitio, cuando hace pagos en línea, etc. Este tipo de acciones se usan habitualmente en tiendas y otros sitios de comercio electrónico, pero al mismo tiempo, todo está protegido en los servidores del sitio para evitar algún hackeo, sin embargo, este tipo de rastreo usando cookies está siendo utilizado para crear un perfil del usuario y así saber qué le gusta y qué le disgusta, y así usar esa información para obtener alguna ganancia. A esto se le llama ad tracking o seguimiento de anuncios.

¿Podemos evitar que se almacene información de este tipo?

Desde 2012 una directiva de la Unión Europea adaptada por la agencia de datos de cada país, regula y obliga a las páginas web a instalar un aviso de instalación de 'Cookies' cuando accedemos a ella. Su intención con esta normativa era proteger la privacidad del usuario cuando la web a la que se accedía instalaba algún tipo de Cookie en el navegador para mostrar publicidad relacionada, control de los visitantes que reciben y otras muchas cuestiones menos obvias pero necesarias para el funcionamiento normal de los sitios.

La Comisión Europea se ha dado cuenta por fin de que este tipo de avisos son contraproducentes y propone limitarlos como parte de una nueva regulación de protección de datos pero ¿qué es lo que proponen en ella exactamente? En propias palabras de la Comisión: "Se simplificará la denominada "disposición de cookies", que ha resultado en una sobrecarga de solicitudes de consentimiento para los usuarios de Internet. Las nuevas reglas permitirán a los usuarios tener más control de sus configuraciones, proporcionando una forma fácil de aceptar o rechazar el rastreo de cookies y otros identificadores en caso de riesgos de privacidad. La propuesta aclara que no se necesita el consentimiento de cookies intrusivas que no sean de privacidad para mejorar la experiencia de Internet (por ejemplo, para recordar el historial del carrito de compras). Las cookies establecidas por un sitio web visitado que cuente el número de visitantes a ese sitio web ya no requerirán consentimiento."

Todo esto se traduce en que a partir de entonces, serán eliminados los avisos de cada sitio web y que serán sustituidos por un aviso general del navegador solicitando al usuario si autoriza automáticamente a aceptarlas o no. Una buena noticia para los usuarios, pues podremos navegar sin tanta molestia.

Por lo que el uso de las cookies es inevitable y prácticamente en todas las webs vas a tener que asumir que se está haciendo uso de esto, se puede pensar que haciendo uso de las ventanas de incognito estamos “saltándonos” el uso de las cookies, bien tanto Google como Firefox explican en sus respectivas webs que estos modos únicamente evitan que se almacenen las cookies de navegación y se guarde el historial de la misma. Es decir, que si otra persona abre ese navegador tras una sesión privada no conocerá el contenido de la misma, pero esta información sí estará almacenada en los servidores de la empresa que nos da acceso a internet, las propias páginas que se visitan y como recuerda Google, la empresa para que la que trabajas si se usa desde un ordenador corporativo.

¿Cómo puede entonces garantizarse una navegación completamente anónima? Lo cierto es que parece que nadie se atreve a garantizar un uso totalmente privado de internet, pero sí puede elevarse el listón de la privacidad empleando los servicios VPN que cifran el contenido que sale del ordenador, y por descontado, utilizar los modos privados si vamos a utilizar ordenadores públicos. Apple va más allá y sugiere que si no se quiere dejar rastro, no hay que olvidar borrar a mano el historial de navegación y las cookies.

#### **Web Beacon** o baliza web, ¿Qué es esto?

Básicamente una baliza web es un archivo que ayuda a hacer un seguimiento de tu navegación por las webs. Estas balizas normalmente los utilizan las webs que recurren a una tercera parte para monitorizar el tráfico y hacer seguimiento de los servicios. Estas también pueden ser utilizados conjuntamente con las cookies para comprender cómo los usuarios de una web navegan y procesan el contenido almacenado en ella.

Algunos ejemplos de uso de web beacon es por ejemplo para:

- a) Contar el número de visitantes de estas páginas web
- b) Verificar si el usuario ha interactuado con correos electrónicos o ha hecho clic en determinados enlaces, por ejemplo, cuando éstos sean condiciones de participación en concursos
- c) Determinar el éxito de una campaña de marketing o concurso
- d) Establecer niveles de interés para determinados apartados de estas páginas web o con relación a ciertos productos y servicios disponibles
- e) Evaluar si propagandas o competiciones resultan en ventas
- f) Comprender la popularidad de un producto o servicio, así como la variación de niveles de interés.

La información por tanto sobre el usuario que se puede obtener mediante el uso de esta tecnología es entre otras:

- IP del usuario o del proxy.
- Programa cliente utilizado.
- Sistema operativo del usuario.
- Fecha y hora de la conexión.
- Permanencia en el sitio web.
- Ubicación del usuario.

Esto puede tener serias implicaciones en la privacidad de los usuarios, ya que en el caso de las páginas web permiten monitorizar su actividad en la red, y en el caso del email permiten conocer si visualizó un correo y desde qué IP, así como otros datos acerca del sistema de correo electrónico.

Este seguimiento puede desactivarse mediante diversas técnicas:

1. **Página web:** para prevenir el rastreo mediante balizas en páginas web, muchos navegadores permiten limitar el acceso a imágenes externas, desactivar las cookies o mostrar las páginas en modo texto. También se puede utilizar un navegador basado en texto.
2. **Correo electrónico:** en correos electrónicos el seguimiento vía web bug se puede evitar mediante el uso de clientes de correo electrónico que no descarguen las imágenes automáticamente. Para ello se puede utilizar o bien un lector de correo basado en texto que no interprete imágenes, o bien un cliente gráfico de email que ofrezca la opción de desactivar el HTML de los mensajes, mostrando así todos los mensajes como texto sin formato, de tal manera que únicamente visualizará las imágenes si el usuario lo decide explícitamente.

Cabría destacar la existencia de la funcionalidad **Do Not Track** de los navegadores, esta es una acción que realizas de manera proactiva, excepto en aquellos navegadores en los que viene activado por defecto. Al activar la opción, cada vez que entras en una página, la información que envía tu ordenador mediante protocolo HTTP incluye un encabezamiento llamado "DNT:1" que solicita a dicha página que tus datos no sean registrados ni compartidos con otros servicios.

Como ventajas y desventajas, el beneficio del uso de este sistema básicamente es dejar de entregar información que no sabías que estabas cediendo a empresas que pueden hacer con ella lo que quieran, desde modificar tu experiencia de navegación a lucrarse vendiéndosela a terceros.

Esta información, realmente se aprovecha para personalizar las ofertas y anuncios que vemos por Internet, por lo tanto si lo activamos seguiríamos viendo publicidad en las páginas pero no tendrá por qué ser afines a nuestros gustos.

En conclusión lo que decides usando esta tecnología es si deseas o no entregar información básica a cambio de una experiencia de navegación más personalizada

#### 2.4.2 Caso práctico. ¿Qué información enviamos y como la tratan cuando entramos a una web?

Además de las "galletas de terceros", los sitios webs utilizan diferentes técnicas para poder registrar las visitas de los diferentes usuarios y comprender en profundidad sus patrones de uso por ejemplo, para poder ver qué páginas web son populares, el efecto de las campañas de publicidad en línea etc..

Para saber qué técnicas existen de seguimiento, existe un complemento llamado "Ghostery" mediante el cual podremos saber mucha más información sobre los datos que estamos dando, a quien y para qué.

Analicemos por ejemplo la propia web de la UOC, al acceder a esta web vemos que tenemos 3 rastreadores los cuales van a estar recopilando nuestra información, veamos cuales son:

##### Google tags Manager:

Este servicio recopila información desde encontrar elementos básicos como el idioma que hablamos, a cosas más complejas, como qué anuncios le resultarán más útiles, las personas que más le interesan en línea o qué videos de YouTube nos podrían gustar, afirma que "La información que recopila Google y cómo se usa esa información depende de cómo utilice nuestros servicios y de cómo administre sus controles de privacidad".

Cuando no hemos iniciado sesión en una cuenta de Google, indican que "almacenamos la información que recopilamos con identificadores únicos vinculados al navegador, la aplicación o el dispositivo que está utilizando. Esto nos ayuda a hacer cosas como mantener sus preferencias de idioma en todas las sesiones de navegación".

Cuando inicias sesión, "también recopilamos información que almacenamos en tu cuenta de Google, que tratamos como información personal."

¿Qué tipo de información recopilan sobre nosotros?

Recopilan información sobre las aplicaciones, los navegadores y los dispositivos que usamos para acceder a los servicios de Google, lo que nos ayuda a proporcionar funciones como actualizaciones automáticas de productos y atenuar su pantalla si la batería se agota. La información que recopilan incluye identificadores únicos, tipo y configuración del navegador, tipo y configuración del dispositivo, sistema operativo, información de la red móvil, incluido el nombre y número de teléfono del proveedor, y el número de versión de la aplicación. También, información sobre la interacción de sus aplicaciones, navegadores y dispositivos con nuestros servicios, incluida la dirección IP, los informes de fallos, la actividad del sistema y la fecha, hora y URL de referencia de su solicitud. Recopilamos esta información cuando un servicio de Google en su dispositivo se comunica con los servidores, por ejemplo, cuando instala una aplicación desde Play



Store o cuando un servicio busca actualizaciones automáticas. Si usa un dispositivo Android con aplicaciones de Google, su dispositivo se comunica periódicamente con los servidores de Google para proporcionar información sobre su dispositivo y la conexión a nuestros servicios. Esta información incluye elementos como el tipo de dispositivo, el nombre del operador, los informes de fallos y las aplicaciones que ha instalado.

A parte de esto guarda información sobre nuestra actividad:

Términos que busca

Videos que mira

Vistas e interacciones con contenido y anuncios Información de voz y audio cuando usa funciones de audio

Actividad de compra

Personas con quienes se comunica o comparte contenido

Actividad en sitios de terceros y aplicaciones que usan nuestros servicios Historial de navegación de Chrome ve sincronizado con su cuenta de Google

Información sobre nuestra ubicación cuando utiliza nuestros servicios, lo que permite ofrecer funciones como indicaciones para llegar en coche a un destino o horarios para ver películas que se encuentran cerca.

Datos de ubicación como:

Dirección IP

GPS

Datos del sensor de su dispositivo

Información sobre cosas cercanas a su dispositivo, como puntos de acceso Wi-Fi, torres de telefonía móvil y dispositivos habilitados para Bluetooth

Además cabe destacar que:

Los datos agregados se comparten con terceros.

Los datos anónimos se comparten con terceros.

Los datos PII se comparten con terceros.

Los datos confidenciales se comparten con terceros.

[Google Analytics](#), este es otro de los rastreadores que se usan en la web de la UOC,

Al ser también de google como el anterior, nos hacemos una idea de los datos que son capaces de recolectar simplemente accediendo a una web.

Cabe destacar que también se comparte con terceros los datos recogidos.

El tercer rastreador que se usa es [HOTJAR](#)

Se definen como “Hotjar es una herramienta nueva de análisis y retroalimentación que revela el comportamiento en línea y los comentarios de los visitantes de un sitio web. Al

combinar A) Análisis y B) las herramientas de retroalimentación, Hotjar brinda a nuestros clientes una visión general de cómo mejorar la experiencia de su sitio y su rendimiento”.

En este caso indican que los datos recopilados se almacenan electrónicamente en Irlanda, Europa en la infraestructura de Amazon Web Services, centro de datos eu-west-1. Como controlador de datos, usted es el único con acceso directo a sus datos, que nunca podrá ser utilizado por un tercero para ningún otro fin que no sea el consentimiento directo. Los tiempos de retención de datos no superan los 365 días, lo que garantiza que no se almacenen datos innecesariamente.

También indica que tienen un acuerdo de procesamiento de datos que se puede firmar, que brinda una transparencia total sobre cómo se almacenan y mantienen los datos.

Hotjar se basa en información anónima y no en datos personales. A los visitantes del sitio se les asigna un identificador de usuario único, de modo que Hotjar puede realizar un seguimiento de los visitantes que regresan sin depender de ninguna información personal. Al recopilar datos con grabaciones, Hotjar también cuenta con varias funciones de supresión automatizadas y los datos se suprimen del lado del cliente, en el navegador del visitante, lo que significa que la información de identificación personal nunca llega a nuestros servidores y mantiene su sesión privada.

Aunque para la mayoría de los usuarios creen que pasan desapercibidos, los rastreadores o trackers son utilizados por más del 75% de los sitios web. Estos scripts o fragmentos de código, son la herramienta utilizada por las plataformas de publicidad digital programática para dirigir sus campañas a targets específicos.

Un análisis reciente de más de 144 millones de páginas web en más de una docena de países, realizado por Cliqz y Ghostery encontró que el 77.4% de todos los sitios web contenía al menos un rastreador de terceros.

Además, el análisis reveló que una minoría considerable de sitios web examinados (16.2%) tenía 10 o más trackers funcionando simultáneamente. Según Cliqz y Ghostery, los rastreadores de Google son con diferencia los más utilizados. Los trackers relacionados con su servicio de Google Analytics aparecieron en el 46.4% de las páginas web examinadas. Los rastreadores relacionados con sus otros servicios, como DoubleClick y AdSense, también fueron relativamente frecuentes.

Facebook fue la otra gran empresa recolectora de datos, con su servicio Facebook Connect en más de una quinta parte (21.9%) de las páginas.

### ¿Cómo impedir que los rastreadores colecten datos y elaboren perfiles?

Un primer paso debería ser una visita a la configuración del explorador. Aquí se puede activar en muchos programas la opción “Do not track” ya mencionada anteriormente.

La navegación privada o el modo incógnito del explorador también comentado, les complica el trabajo a los rastreadores. Y cuando se desactivan cookies en el explorador, al menos se pone freno a los rastreadores que trabajan con los minidatos. Sin embargo, ninguna de estas soluciones ofrece una protección total. Y también es cierto que si quieres protegerte, también pierdes comodidad.

Otra posibilidad de protección son las llamadas listas de protección de rastreo (TPL, en sus siglas en inglés) para el explorador de Internet, como las del Instituto Fraunhofer. Para otros exploradores hay numerosas extensiones contra rastreadores, entre ellas NoScript, ShareMeNot y Disconnect. Una de las más populares es Add-on Discovery, que reconoce y bloquea más de 1.900 rastreadores, según los desarrolladores.

Durante un ensayo, Ghostery encontró en la compañía de comercio electrónico Amazon nada menos que 11 rastreadores diferentes. Además, está claro que sin los colectores no sólo mejora la protección de datos: como en segundo plano se transmiten menos informaciones y muchos anuncios publicitarios ni siquiera aparecen sin un rastreador, las páginas muchas veces se abren bastante más rápido con Ghostery. Sin embargo, según la revista "Technology Review", el desarrollador Evidon también recolecta y vende datos, aunque los hace anónimos. Además, la función correspondiente, Ghostrank, está desactivada después de la instalación del add-on.

### 2.4.3 El Gran Hermano te observa

A diferencia de la novela comentada, en el mundo real no hay un solo ente vigilante, sino que los datos pasan por muchas manos.

De aquí, nace así una nueva profesión, la del Data Broker, este no tiene que ver con los brokers de bolsa, este broker, es intermediario en operaciones de compra-venta de datos personales.

En España gracias a la Ley Orgánica de Protección de Datos de Carácter Personal (LOPD), que entró en vigor el 13 de diciembre de 1999 y que se ha estado modificando tal como hemos visto anteriormente, los usuarios están más protegidos que en otros países.

En España no se permite el ejercicio de los Data Brokers sin el consentimiento del usuario. Pero el problema viene a partir de las condiciones de uso, ¿quién se lee todas las condiciones de una aplicación antes de aceptarlas? Hoy en día es muy sencillo obtener el consentimiento cuando los usuarios están acostumbrados a darle a 'Aceptar' sin mayores problemas, es la cosa más sencilla del mundo, por lo que en la práctica muchas veces estamos dando la autorización para que se exploten nuestros datos con fines comerciales, a través de interminables políticas de privacidad, sin siquiera ser conscientes de ello.

La pregunta que todos nos hacemos ante esta situación es ¿Realmente tanto vale nuestra información?

Analicemos un caso real de un ciudadano estadounidense, Federico Zannier, decidió graduarse de comunicación computacional en la Universidad de Nueva York con un proyecto que consiste en violar su propia privacidad y vender toda la información en Internet.

Zannier mediante un arsenal de herramientas de espionaje digital alojadas anónimamente en una computadora, "un tercero" (que podría ser Facebook o alguien de dudosas intenciones) puede conocer, mediante una extensión en Chrome, las direcciones web y los documentos vistos en línea, con otro software, conocer su

localización exacta en GPS, además de aplicaciones que toman fotos con la webcam y screenshots cada 30 segundos, incluyendo la posición del mouse cuando un nuevo tab se abre en el navegador y todos los clicks que hizo en cada sitio.

Federico Zannier dió con la idea al investigar sobre las cookies instaladas en su navegador. "La idea del proyecto era tratar de arrojar luz sobre un asunto que ocurría cada día y nadie prestaba atención." Así que decidió convertir su computadora en una máquina de vigilancia, que pudiera monitorear y grabar todos sus movimientos (tal como nosotros olvidamos las cookies instaladas en nuestras computadoras, que a su vez vigilan los nuestros.)

La idea principal del ciudadano estadounidense era publicar la información que almacenó durante todo el año pero luego se lo pensó mejor y llevó la broma a Kickstarter, donde ofrecía vender un día de información suya, a quien quisiera comprarla. El resultado fue que 213 usuarios compraron sus datos mediante mecenazgo en Kickstarter se calcula un valor de algo más de 2.500 euros el joven visionario espera que la gente se vuelva más responsable por la cantidad de información que comparte en línea y sobre los derechos que sobre nuestra propia privacidad en línea nos deshacemos al dar "Aceptar" sin leer las cláusulas de servicios como Facebook, Instagram o Twitter.

Si hablamos de casos reales más conocidos, España por ejemplo, Movistar compró Tuenti en 2010 pagando 9,62 euros por usuario. Una cifra que de forma individual no es muy alta, pero si se multiplica por los casi 8 millones de usuarios que tenía la red social, hacen una cantidad considerable: 70 millones de euros.

Como curiosidad a parte del valor económico de los datos de una persona, estos pueden revelar cosas sorprendentes de la vida de uno. Para bien y para mal. Una muestra de esto la tenemos el caso de los supermercados Target de EEUU. Hace unos años, un hombre acudió enojado con varios cupones a una de las tiendas, en Mineápolis. Este quería ver al gerente porque los descuentos, destinados a su hija adolescente, eran de ropa de bebe. Le parecía aberrante que incitaran a una joven a ser madre. Lo que no sabía era que en realidad su hija sí estaba embarazada. Target se había anticipado basándose en un análisis de las compras de la joven.

Como decíamos, en España hay menos peligro pero, como todos sabemos, hecha la ley, hecha la trampa, la legislación española no permite la venta de datos personales entre empresas, por lo que muchas se hacen con ellos de dos formas.

1. La primera es mediante una joint venture en la que una empresa paga a otra para que envíe publicidad de ella a sus usuarios registrados.
2. Para la segunda sirve el ejemplo de Movistar: una compañía se hace con otra para quedarse con sus usuarios sin ánimo de seguir con el desarrollo de la actividad.

Como se puede ver, el uso de los datos de usuario no siempre es perjudicial para el ciudadano. Hay muchos casos que demuestran lo contrario: la búsqueda de personas en catástrofes, la monitorización cardíaca (un estudio de la Universidad de Standford

reveló que las pulseras inteligentes son capaces de detectar problemas de salud antes de que el cuerpo manifieste los síntomas) o la posibilidad de ajustar de forma personalizada las primas de los seguros.

Pero como contrapartida, también existen efectos negativos: que se etiquete a los usuarios y que estos se vean rechazados a la hora de buscar un seguro, que unas antiguas fotos subidas a la red supongan un impedimento para encontrar trabajo o que se reciba de forma masiva publicidad sin poder evitarlo.

La Era Digital ha supuesto que los ciudadanos pierdan su intimidad. Un estudio del que se hizo eco el Wall Street Journal argumentaba que las aplicaciones móviles llegaban a solicitar 6.200 veces a la semana la localización del usuario. La tecnología avanza y también lo hacen las formas de beneficiarse de ella con fines maliciosos. En 2014 se dio el conocido Celebgate: decenas de celebridades veían comprometidas sus fotos íntimas cuando un hacker se hizo con ellas a través de iCloud y las publicó en Internet. Experian es otro ejemplo. En 2015 sufrió un hackeo que expuso los datos de 15 millones de personas. Joaquín Cuenca, extrabajador de Google y cofundador de Panoramio, explica que es frecuente ver cómo los ciberdelincuentes extorsionan a los usuarios: “Desde hace unos años es muy común que los virus de ordenador encripten todos los datos y pidan un soborno para descifrar esa información (suelen ser unos 1.000 euros, pagados con la moneda digital bitcoin). Se han visto chantajeadas asesorías, notarías e incluso comisarías de policía. Y no queda más remedio que pagar o perder los datos si no hay copia de seguridad externa”.

## 2.5. OSINT ¿Estamos expuestos públicamente?

### 2.5.1 ¿Qué es OSINT?

OSINT es un acrónimo anglosajón de Open Source Intelligence, el cual se usa para referirse a los recursos de información libremente accesibles no solo en Internet, sino en la vida off-line.

La recopilación y análisis de la información contenida en esa tipología de fuentes puede permitir alcanzar inteligencia utilizable (inteligencia de seguridad interior, militar, competitiva, económica y vigilancia tecnológica).

Por ejemplo, OSINT es un periódico online, un periódico tradicional, un blog, las redes sociales etc.

Por lo que en base a esto, internet es lo más cercano a una máquina perfecta de vigilancia que el mundo ha conocido (y la vigilancia tecnológica explota esta peculiaridad). Todo lo que se publica en Internet es en algún lugar registrado; los servicios en la nube son el primer objetivo: no se precisa arrebatar la información a su propietario, ya que este ya la ha cedido libremente.

### ¿Qué se puede conseguir con esto?

Veamos algunos datos:

- Un estudio realizado por investigadores de la Universidad de Cambridge (Reino Unido) en colaboración con Microsoft Research Cambridge advierte que las

preferencias mostradas haciendo clic en los «Me gusta» son suficientes para trazar un detallado perfil del usuario.

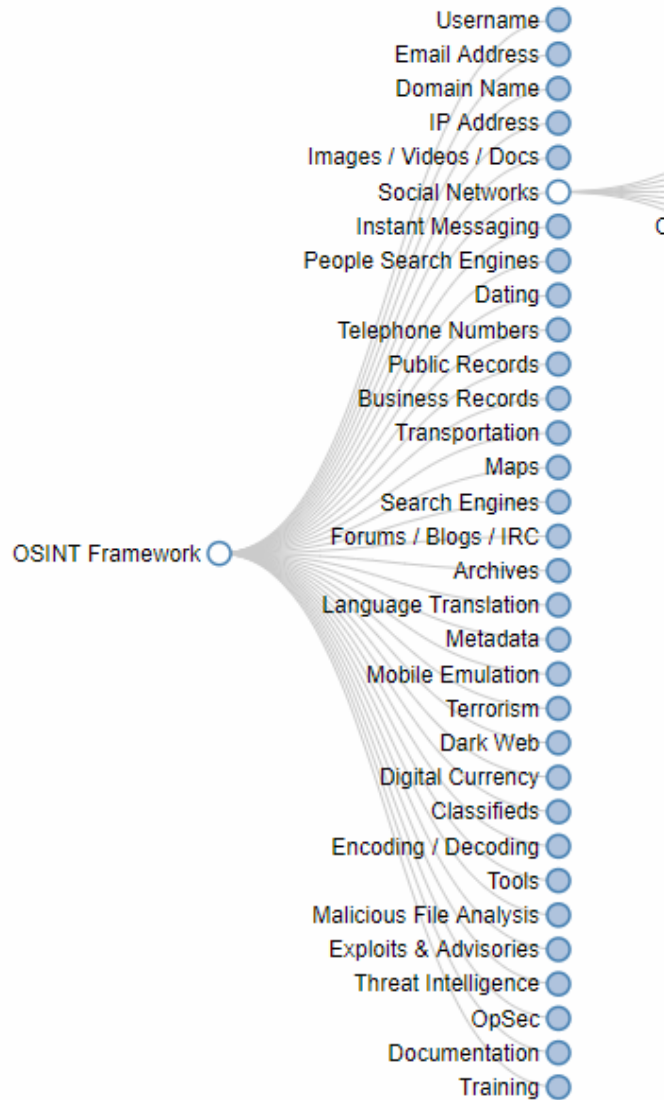
- Investigadores de la Universidad de Pensilvania, en Estados Unidos, tomando con fuente de información las actualizaciones de estado de 75.000 personas en Facebook han conseguido predecir su edad, sexo e incluso el tipo de personalidad basándose únicamente en las palabras que usaron.
- Alessandro Acquisti y Ralph Gross, de la Universidad Carnegie Mellon, realizaron un estudio en el que usaron información de diversas fuentes de carácter público, incluyendo perfiles de redes sociales, informaron que pudieron predecir con precisión el número de afiliado de la seguridad social del 8,5% de las personas nacidas en Estados Unidos entre 1989 y 2003, prácticamente cinco millones de personas.
- Las universidades de Sevilla y Alicante están desarrollando una plataforma que analiza las opiniones de la web y de los medios sociales para ayudar a las instituciones o empresas a tomar decisiones estratégicas.
- Investigadores de la Universidad Carnegie Mellon concluyen cómo la información compartida a través de las redes sociales puede llevar a la discriminación en la contratación.
- Distintos proyectos europeos tratan de obtener y explotar información de redes sociales, que mejoren la gestión total e integrada de todos los intervinientes en situaciones de crisis y emergencias, dentro del programa de Seguridad 2013 «Topic SEC-2013.6.1-1 The impact of social media in emergencies».

Podríamos definir OSINT como un arma de doble filo, depende en las manos y objetivo para el cual se utilice. Por lo tanto como vemos, mediante OSINT nuestra información está prácticamente expuesta al resto del mundo y cualquiera podría conseguirla.

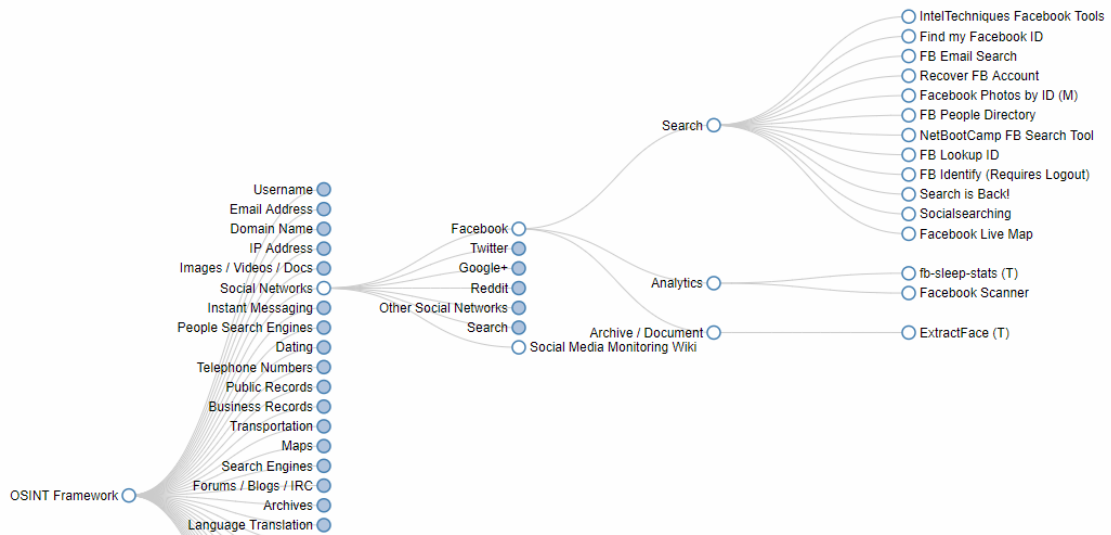
## 2.5.2 OSINT FRAMEWORK

### ¿Cómo un atacante podría obtener información sobre nosotros?

OSINT se lo pone bastante fácil, observando el árbol que OSINT Framework proporciona al usuario podemos ver cómo existe una clasificación que comienza orientada a la persona. Si pinchamos en cada nodo se desplegarán los sitios dónde se podrá acceder a la información.



La búsqueda de usuarios, la búsqueda de direcciones de correo electrónico, direcciones IP, recursos multimedia o perfiles en redes sociales son casos típicos de OSINT. Si vamos indagando vemos infinidad de formas de obtener información sobre los usuarios.



Como vemos en la imagen anterior, algo que puede proporcionar mucho juego son las redes sociales. Solo con Facebook y Twitter tenemos diferentes funcionalidades y fuentes de información disponibles, en función de lo que se quiera en cada momento. Por ejemplo, en Facebook tenemos la posibilidad de buscar desde fotografías, cuentas, archivos, personas a través del directorio de personas de Facebook. En el caso de Twitter, la posibilidad de buscar patrones, usuarios, localizaciones y geolocalizaciones, etcétera, hacen que se pueda extraer gran cantidad de información a través de estos sitios.

Por otro lado las categorías de nombres de dominio y direcciones IP proporcionan sitios dónde el usuario podrá obtener gran cantidad de información. Por ejemplo, en el caso de las direcciones IP podemos ver cómo existen fuentes de información sobre reputación, geolocalización, puertos abiertos, direccionamiento IPv4 e IPv6, incluso, mapas de redes wireless.

En el caso de los nombres de dominio tenemos otras subcategorías interesantes y que proporciona información sobre los dominios. Por ejemplo, el Whois, todo un clásico, la reputación de los dominios, subdominios almacenados de los propios dominios, etc.

#### 2.5.4 Proceso OSINT

El proceso de OSINT consta de las siguientes fases:

1. Identificar fuentes de información relevante
2. Obtención de la información
3. Procesar la información recopilada para posteriormente analizarla
4. Análisis de los datos recopilados y procesados para llegar a alguna conclusión significativa.
5. Presentar la información obtenida de una manera eficaz, potencialmente útil y comprensible, de manera que pueda ser correctamente explotada.





Dentro de lo que sería el proceso general OSINT las fases típicas serían, una primera fase de requerimientos, de identificación, saber cuáles son los objetivos que perseguimos, cual es la información que queremos obtener y el tiempo que vamos a necesitar que va a ser decisivo.

Seguidamente vendría una fase de identificación de las fuentes relevantes que vamos a utilizar para obtener esa información y a continuación, una vez hemos identificado las fuentes, lo que hacemos es adquirir esa información.

Se adquiere la información, se procesa para darle algún tipo de formato que sea fácilmente interpretable y a partir de ahí generar esa inteligencia, se generan esos datos o se interpretan los datos que hemos obtenido de diferentes herramientas para garantizar que esa información sea lo más fiable posible.

### 2.5.5 Herramientas OSINT

Hay infinidad de herramientas que facilitan el trabajo mediante el uso del OSINT a continuación se presentan algunas:

#### MALTEGO

Maltego es una de las herramientas más completas y mejor implementadas que existen actualmente en el mercado enfocada sobre todo en la recolección de información y minería de datos, su valor añadido con respecto a las herramientas existentes en el mercado actualmente: La representación de la información en una forma simbólica, es decir, la información es presentada en distintos formatos de forma visual y enseñan las distintas relaciones encontradas entre la información presentada, por otro lado Maltego permite enumerar información relacionada con elementos de red y dominios de una forma bastante comprensible, así como también permite enumerar información relacionada con personas, datos tales como direcciones de email, sitios web asociados, números de teléfono, grupos sociales, empresas asociadas, etc.

## Ready or Not?

Esta web se encarga de obtener y analizar la información de la cuenta de Twitter o Instagram de la persona que se quiere analizar. Entre los datos que podemos obtener se encuentra un mapa de calor con las zonas desde las que suele twittear o subir fotos a Instagram y los días de la semana y la hora a la que suele hacerlo.

## FOCA

FOCA (Fingerprinting Organizations with Collected Archives) es una herramienta utilizada principalmente para encontrar metadatos e información oculta en los documentos que examina. Estos documentos pueden estar en páginas web, y con FOCA se pueden descargar y analizar. Los documentos que es capaz de analizar son muy variados, siendo los más comunes los archivos de Microsoft Office, Open Office, o ficheros PDF, aunque también analiza ficheros de Adobe InDesign, o svg por ejemplo.

## WhoIS

WHOIS es un protocolo TCP basado en petición/respuesta que se utiliza para efectuar consultas en una base de datos que permite determinar el propietario de un nombre de dominio o una dirección IP en Internet. Las consultas WHOIS se han realizado tradicionalmente usando una interfaz de línea de comandos, pero actualmente existen multitud de páginas web que permiten realizar estas consultas.

Esta herramienta suele resultar más útil para buscar información de una empresa o un servicio web, que de un particular

## 2.5.6 Como podemos protegernos

Como vemos es muy fácil que alguien pueda obtener información sobre nosotros, y a pesar de los riesgos que conlleva publicar información personal de forma pública en Internet, seguimos creando nuevas huellas digitales en Internet todos los días sin darnos cuenta (tal vez es nuestra única opción si queremos vivir en una sociedad globalizada)

Aun así se puede conseguir controlar esta fuga de información, con medidas bastante lógicas que no todo el mundo tiene en cuenta:

**Sentido común:** no introducir fechas de nacimiento completas en los perfiles de usuario, no anunciar cuándo nos vamos de vacaciones, no incluir datos o fotos de menores de edad, no dejar que éstos usen las redes sociales sin la supervisión de un adulto, comprobar con los motores de búsqueda que no aparezcan más datos sobre nosotros que los deseados, etc.

**Control de acceso a recursos:** utilizar contraseñas fuertes, usar un doble factor de autenticación (envío de SMS al móvil, Google Authenticator, Latch, Authy, etc.), utilizar los mecanismos de privacidad proporcionados por la aplicación que estemos usando, y el uso de técnicas criptográficas son algunas de las cosas que pueden ayudarnos a evitar que un tercero tenga acceso a datos o recursos no deseados relacionados con nuestra identidad digital.

**Perturbación de datos:** consiste en modificar los datos personales para aumentar su ambigüedad (generalización), reducir el nivel de detalle de la localización del usuario, etc. Es decir: dar los datos mínimos imprescindibles (quizá solo el nombre y no los apellidos, o el uso de un seudónimo) y no tener activada la función de geolocalización, por ejemplo.

Como ya se comentaba anteriormente, también creamos huellas pasivas es decir huellas que dejamos en Internet sin ser conscientes de ello como las ya comentadas cookies, a parte de las medidas anteriores también podemos controlarlas mediante la utilización de canales anónimos como los siguientes:

1. Uso de un nodo central de confianza: VPN comerciales, proxies del tipo Crowds y UUP
2. Mix Networks: Como por ejemplo Mixmaster y Mixminion, consisten en un grupo de nodos interconectados que forman una red en la que cada uno oculta la entrada y salida de información mediante técnicas criptográficas
3. Navegar por la red TOR: red de comunicaciones distribuida de baja latencia que mantiene el anonimato de las IP origen y destino en los nodos intermedios mediante transmisiones cifradas por capas.
4. I2P: A diferencia de TOR, en vez de utilizarse circuitos de nodos intermedios, se utilizan túneles unidireccionales: el cliente establece un túnel de salida que se comunicará con el túnel de entrada del servidor; éste enviará la respuesta utilizando su propio túnel de salida que se comunicará con el túnel de entrada del cliente.

## 2.6. Caso práctico recolección de información

Hemos visto de forma más o menos teórica lo fácil que puede resultar obtener información sobre alguien en internet, en esta sección se pretende demostrarlo buscando información sobre alguien en concreto.

En este caso voy a hacer uso de herramientas para obtener toda la información que hay en la red sobre mí.

Empezaremos con algo básico que tal vez la mayoría de la gente no sepa ¿Que sabe google sobre mí?

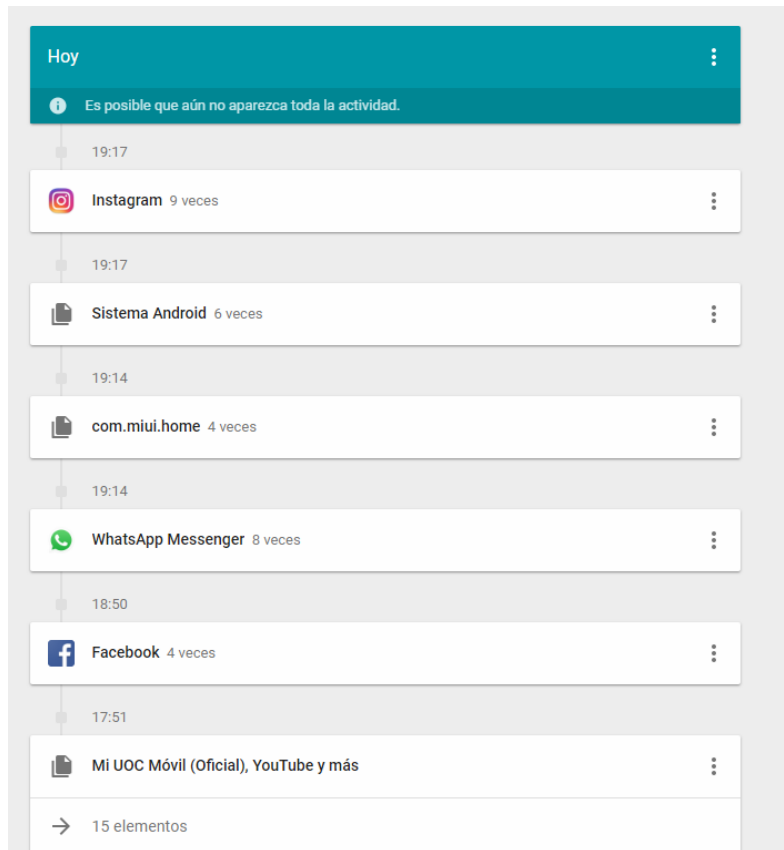
Si hacemos una simple búsqueda en google poniendo mi nombre, ya tenemos acceso a mi perfil de LinkedIn

A screenshot of a Google search page. The search bar contains the text "federico valero valdés". Below the search bar, there are navigation tabs: "Todo" (selected), "Maps", "Noticias", "Imágenes", "Vídeos", "Más", "Configuración", and "Herramientas". The search results show "Aproximadamente 7.820 resultados (0,50 segundos)". The first result is titled "Federico Valero Valdés | Berufsprofil - LinkedIn" with the URL "https://es.linkedin.com/in/federico-valero-valdés-377642116". The description reads: "Sehen Sie sich das Profil von Federico Valero Valdés auf LinkedIn an, dem weltweit ... Desarrollador en el departamento Business Intelligence en Teralco."

Podemos acceder a mi facebook

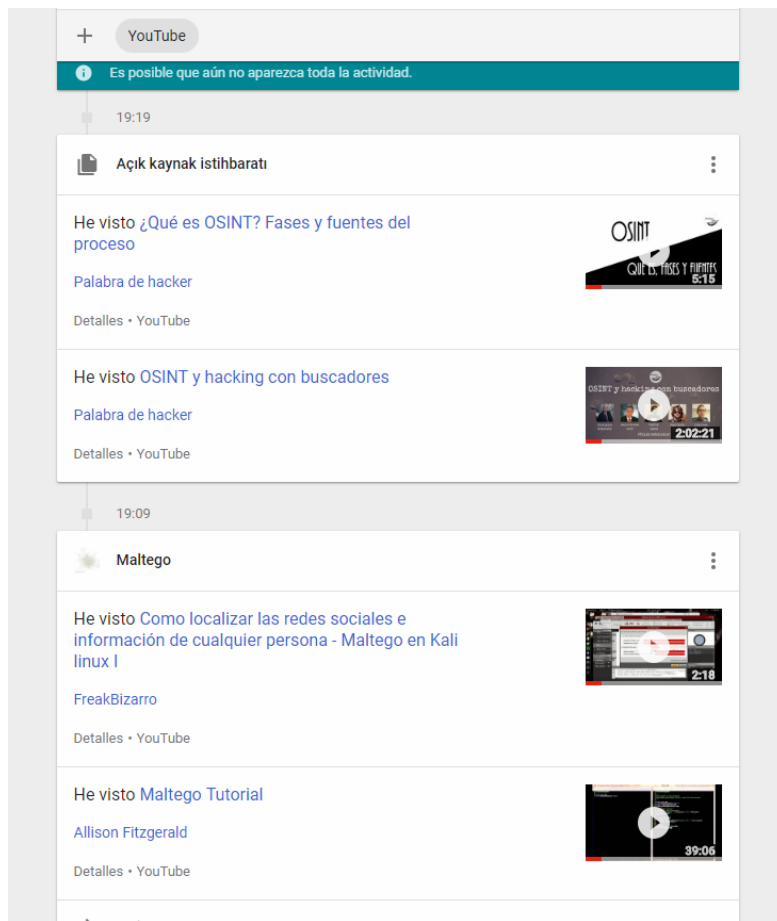
A screenshot of a Google search page. The search bar contains the text "fede valero facebook". Below the search bar, there are navigation tabs: "Todo" (selected), "Imágenes", "Noticias", "Vídeos", "Maps", "Más", "Configuración", and "Herramientas". The search results show "Aproximadamente 442.000 resultados (0,36 segundos)". The first result is titled "Fede Valero | Facebook" with the URL "https://es-es.facebook.com/fede.valero.1". The description reads: "Fede Valero está en Facebook. Únete a Facebook para conectar con Fede Valero y otras personas que quizá conozcas. Facebook da a la gente el poder de..."

Hasta aquí nada que no sepamos, pero si accedemos a las propias herramientas que google nos facilita podemos ver que google sabe incluso el uso que le hemos dado a nuestro móvil:

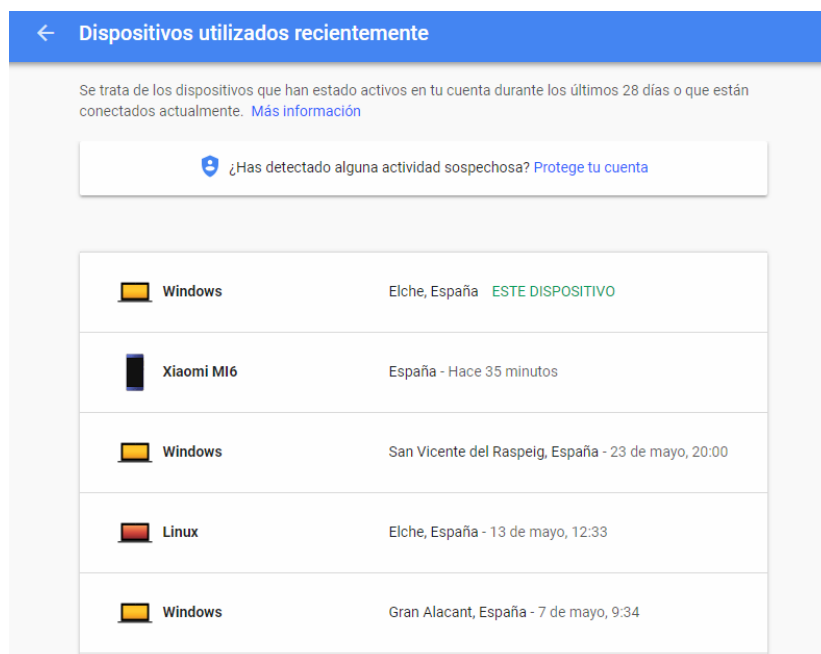


Google ya sabe lo que he estado haciendo con mi móvil en cada momento, como vemos, he entrado en Instagram, en Facebook, he usado WhatsApp, he accedido a la aplicación de la UOC. A su vez obviamente ha quedado registrado todos los likes que he dado en Instagram, todo lo que he visto en Facebook... ect.

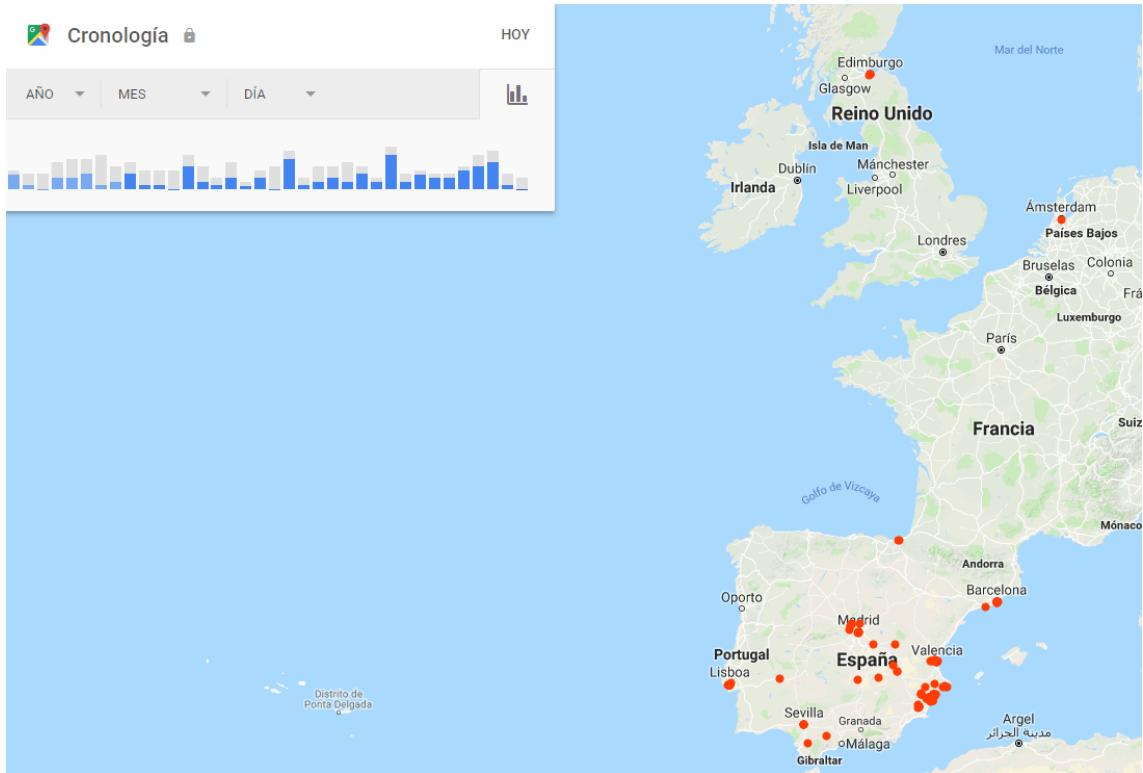
Como es lógico también ha guardado el historial de todo lo que he visto en YouTube:



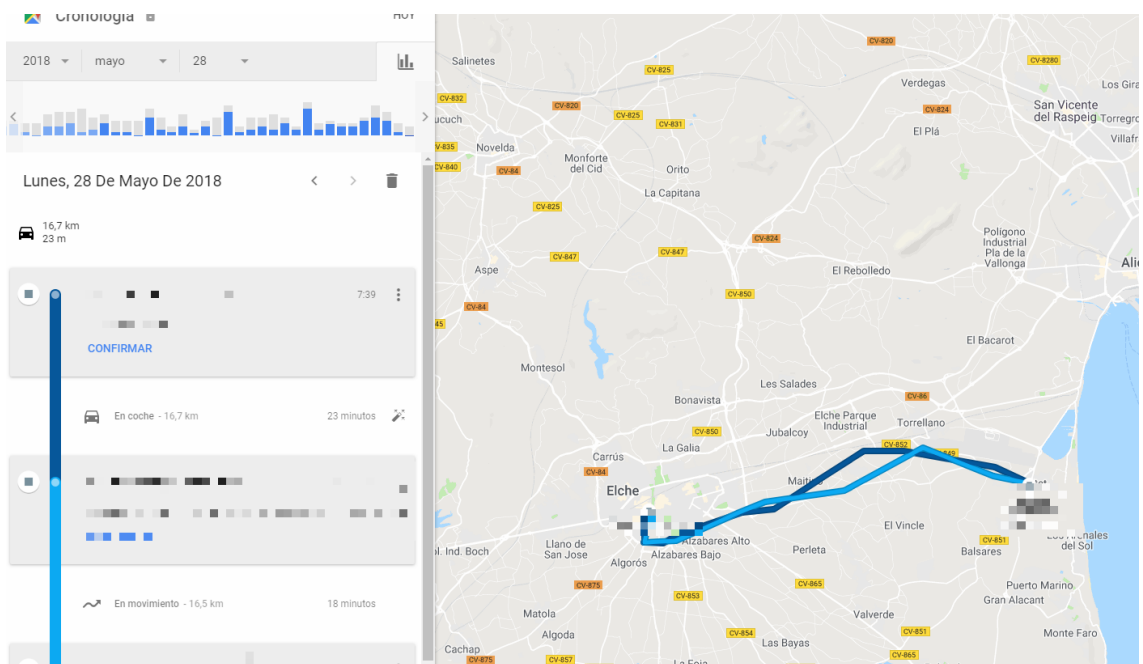
Si seguimos investigando, vemos que Google también sabe donde hemos estado hoy, y donde estamos en este momento.



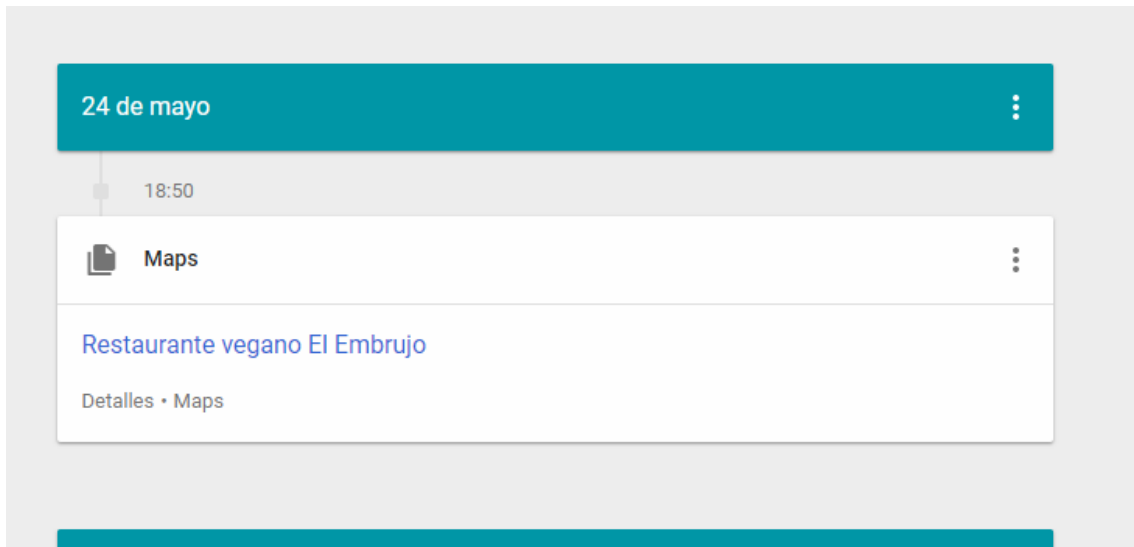
Es más Google recuerda seguramente incluso mejor que yo los sitios a los que he viajado últimamente:



Incluso sabe dónde vivo, y donde trabajo y es capaz de darme la ruta que he hecho esta mañana para ir y volver de casa al trabajo e indicarme el tiempo que he tardado incluso decirme donde he aparcado el coche.



Podemos ir aún más allá y es que también sabe dónde has cenado, lugares que has visitado etc:



Sigamos por ejemplo con algo que hoy en día casi todos usamos, y es cualquier red social, por ejemplo facebook.

No todo el mundo sabe que esta plataforma nos permite bajarnos un archivo con toda la información que almacena sobre nosotros, analicemos todo lo que sabe sobre mí:

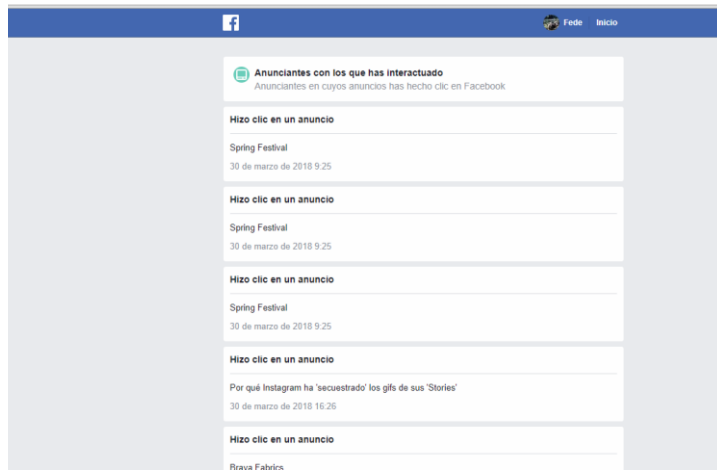


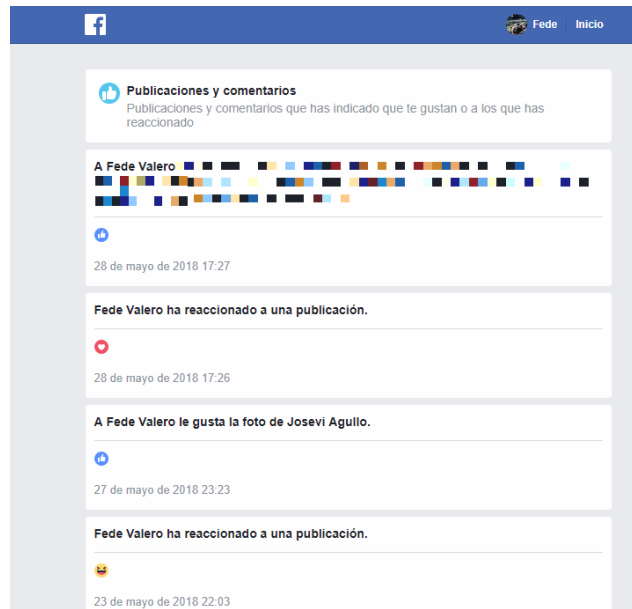
Al descargar esta información, a parte de las fotos videos y material que directamente sabemos que le estamos cediendo a Facebook vemos cosas interesantes como por ejemplo información sobre nuestros contactos y sus números de teléfono, por lo que Facebook ya no solo sabe información nuestra sino también sobre nuestros amigos:





Por otro lado, guarda los eventos que me han interesado, likes, comentarios...





Con todos los likes que damos y la información que nosotros mismos proporcionamos, Facebook sabe prácticamente todo sobre nosotros:

**Información del perfil**  
 Tu información de contacto, los datos que has incluido en la sección "Información" de tu perfil y tus acontecimientos importantes

<b>Nombre</b>	Fede Valero
<b>Perfil</b>	
<b>Fecha de registro</b>	viernes, 3 de julio de 2009 a las 22:42 UTC+02
<b>Direcciones de correo electrónico</b>	
<b>Dirección</b>	
<b>Neighborhood</b>	
<b>Teléfonos</b>	
<b>Fecha de nacimiento</b>	
<b>Sexo</b>	Hombre
<b>Ciudad Actual</b>	Elche
<b>Hometown</b>	Elche
<b>Formación</b>	 University of Alicante
<b>Workplace</b>	
<b>Activities</b>	

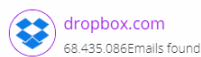
Del mismo modo, todas las redes sociales, recopilan el mismo tipo de información sobre nosotros, ya sea Facebook, twitter, Instagram... etc.

Si vamos un poco más allá analicemos las políticas de alguna red social y veamos que hacen con nuestros datos, por ejemplo de otra aplicación que todos usamos, por no analizar solo Facebook, la famosa aplicación de mensajería WhatsApp, si leemos los términos, vemos que En el caso de WhastApp, trabaja y comparte información con otras empresas, aunque actualmente WhatsApp solo comparte ciertos tipos de información con las empresas de Facebook.

Estos datos son: el número de teléfono que verificaste al registrarte en WhatsApp, algunos datos de tu dispositivo (el identificador de tu dispositivo, la versión del sistema operativo y de la aplicación, datos sobre la plataforma, tu código de país y de red, e indicadores que permitan realizar un seguimiento de la aceptación de actualizaciones y las opciones de control que elijas) y algunos datos sobre el uso (cuándo se ha utilizado WhatsApp por última vez, la fecha en la que se realizó el registro de la cuenta, la frecuencia con la que se utilizan las funciones y el uso que se les da).

Si leemos los términos de cualquier servicio veremos que no siempre los datos se quedan para uso de la misma plataforma sí que se ceden a terceros.

Bien, hasta aquí, todo el mundo más o menos es consciente de que los grandes como Google o Facebook guardan este tipo de información sobre nosotros, a continuación complicaremos un poco más la búsqueda y vamos a ver por ejemplo si mi correo electrónico ha sido comprometido.



Highlighted leaks where  has been compromised

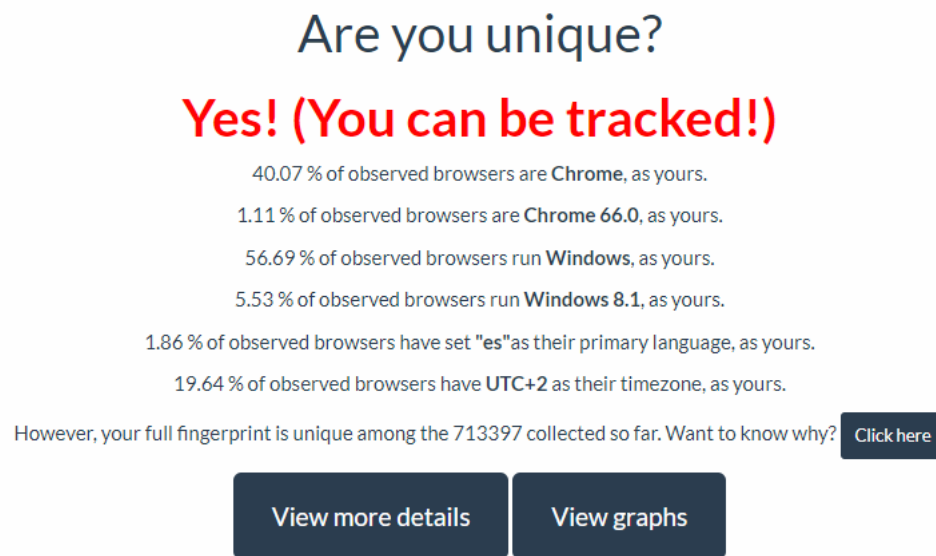
Date	Title	Network	Emails found
Sep 2017	runator.com	darknet	115.724
Sep 2016	dropbox.com	darknet	68.435.086

Por alguna vulnerabilidad en una aplicación, mi correo electrónico (en este caso es el que uso para registros en sitios no fiables) fue comprometido en septiembre de 2017, por lo que seguramente haya habido terceras personas con acceso a él y hayan podido obtener información de mi correo, como cuentas de mis contactos, ver los correos recibidos etc.

También quedo comprometido por una vulnerabilidad de la famosa plataforma de almacenamiento en la nube Dropbox.

Si seguimos indagando, ya comentábamos anteriormente que es el browser fingerprinting o huella dactilar, esto consistía en recopilar datos sobre la configuración del navegador y el sistema de un usuario cuando este usuario visita un sitio web.

Hay herramientas que analizan y nos dan información sobre nuestra huella dactilar, vamos a ver la mía:



Como vemos en el resultado del analisis, tengo una huella digital única, es decir que entre todos las huellas que se han analizado no hay ninguna como la mia por lo tanto se me puede identificar.

Es imposible esquivar las huellas digitales de los navegadores de Internet por completo, y es que los gestores de los servidores web reciben las características transmitidas automáticamente en la cabecera HTTP durante el fingerprinting pasivo pero se puede intentar reducir al mínimo el valor de reconocimiento del cliente para que las digital fingerprints no tengan un carácter único e impidan realizar el seguimiento. La solución más simple es recurrir a extensiones del navegador, de modo que estas bloqueen contenidos activos como aplicaciones JavaScript, Flash o Silverlight y no puedan transmitir ningún dato al servidor. Aparte de los bloqueadores de scripts, ya solo queda renunciar a la personalización del sistema y del navegador. Es recomendable en este caso decidirse por un navegador que se utilice a menudo, por ejemplo Firefox, y recurrir a la configuración predeterminada, lo que también es de aplicación para el sistema operativo utilizado.

Podemos saber también que es lo que sabe el navegador sobre nosotros, por ejemplo usando herramientas como <http://webkay.robinlinus.com/> vemos que:

Sabe nuestra ubicación

## Location



Geo Coordinates: 3

El software que uso:

### Software

#### Operating System

Windows 8.1

#### Browser

Chrome 66.0.3359.181

#### Browser Plugins

Chrome PDF Plugin  
Chrome PDF Viewer  
Native Client  
Widevine Content Decryption Module

#### Prevention:

To prevent your browser from leaking information about your software use **NoScript**.

El hardware:

### Hardware

#### CPU:

Win32, 8 Cores

#### GPU:

Vendor: Google Inc.  
Renderer: ANGLE (Intel(R) HD Graphics 4600 Direct3D11 vs\_5\_0 ps\_5\_0)  
Display: 1920 x 1080 - 24bits/pixel

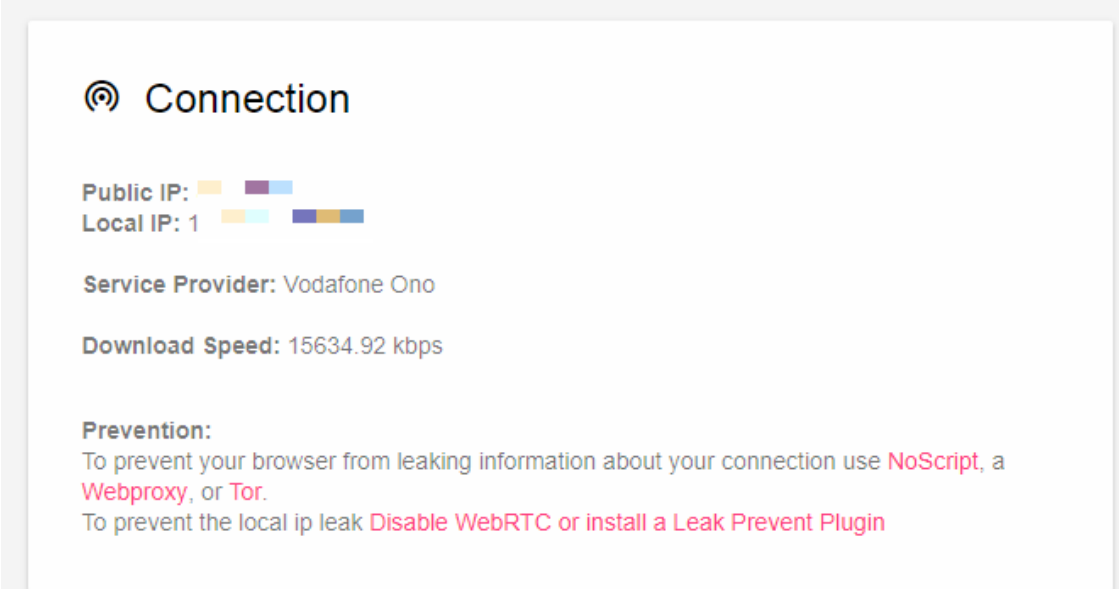
#### Battery

Charging: charging  
Battery Level: 100%  
Charging Time: 0h

#### Prevention:

To prevent your browser from leaking device information use **NoScript**.

Información sobre mi proveedor de internet, velocidad, ips etc:



**Connection**

Public IP: [Colorful bar]  
Local IP: 1 [Colorful bar]

Service Provider: Vodafone Ono

Download Speed: 15634.92 kbps

**Prevention:**  
To prevent your browser from leaking information about your connection use **NoScript**, a **Webproxy**, or **Tor**.  
To prevent the local ip leak **Disable WebRTC** or **install a Leak Prevent Plugin**

En resumen, sin necesidad de entrar en herramientas muy técnicas, ni avanzadas, he conseguido en muy poco tiempo recopilar prácticamente toda la información sobre mí, donde vivo, donde trabajo, donde he estudiado, quienes son mis amigos, mis datos personales, donde he viajado, donde estoy en cada momento, software que utilizo, hardware que utilizo, eventos a los que he ido, eventos a los que quiero ir, lugares que frecuento, tipo de artículos que me gustan, incluso ya no solo sobre mí, ya que si seguimos indagando, hay casos que podremos saber quiénes son los familiares, incluso sabremos donde viven, como se llaman, donde trabajan...

Con esto, lo primero que se nos viene a la cabeza es el mal que se puede hacer con toda esta información, en la siguiente sección analizaremos los problemas que esto puede generar.

## 2.7. Delitos con nuestra información

Son muchos los tipos de delitos que se pueden cometer a partir de conocer tanta información sobre las personas, pero en este caso lo más fácil es pensar en los perfiles falsos que se pueden crear en nuestro nombre, con datos tan reales, como los que hemos podido obtener en el ejercicio anterior.

Cada vez ocurre con más frecuencia que nuestros datos personales, fotografías y demás contenido personal, es utilizado por terceras personas.

Esta utilización de nuestros datos y contenidos personales nos pueden producir tanto problemas personales como financieros.

En cuanto a los problemas personales, se pueden producir porque se desvela información personal a terceras personas, que no deseamos dar a conocer. También se puede producir que se desvelen datos personales de contacto, que no deseamos que se den a conocer.

En cuanto a daños económicos, uno de los casos que hemos comprobado que se producen con mayor frecuencia es la utilización de datos bancarios que son utilizados por terceras personas y que nos puede producir deudas.

### 3. Conclusión

A lo largo de este trabajo se ha intentado concienciar y dar a conocer lo expuestos que estamos mediante el Big Data y el tratamiento que hoy en día se les da a nuestros datos, como vemos nada es gratis, es más, detrás de lo que aparentemente para nosotros es gratuito, hay un gran negocio de datos en el que se “trafica” con nuestra información que probablemente estamos regalando sin prácticamente darnos cuenta, o no queriendo darnos cuenta de ello.

Todos los servicios que utilizamos a diario gratuitamente, realmente lo estamos pagando con el precio de nuestra información, la pregunta es, si en el mundo digital que vivimos realmente estaríamos dispuestos a dejar de hacer uso de aplicaciones que nos facilitan nuestro día a día, por el hecho de ser anónimos.

Deberíamos actuar con cabeza, conocer lo que hay detrás de todo esto y valorar hasta qué punto estoy dispuesto a ceder mi información y en el caso de aceptar cederla, plantear que tipo de información estoy dispuesto a compartir.

Al fin y al cabo la mayoría de la información que se puede recoger sobre nosotros por internet es información que nosotros mismos hemos publicado, o hemos permitido que se publique, por lo que es tan sencillo como decidir que quiero que sepan de mí, y que no quiero que sepan.

Por lo que el primer paso sería este, y es lo que está al alcance de cualquier ciudadano preocupado por su privacidad, por otro lado tenemos otro tipo de acciones que se escapan más de nuestro alcance y es que las empresas vendan datos y lo que tu habías cedido a una empresa esta se lo da otra de forma que pierdes el control de quien sabe qué sobre ti.

Ante esto, en los términos que aceptamos sin leer, es bastante probable que se indique si la información que cedemos se va a compartir con terceros, aquí de nuevo debemos de pensar que preferimos, calidad en nuestra vida mediante el uso de un servicio o ser anónimos.

Por último siempre podemos recurrir a la ley, y ejercer nuestros derechos como ciudadanos sobre nuestros datos que están almacenados en sitios de terceros.

En definitiva, como opinión personal, considero que hoy es prácticamente imposible ser anónimo y esto incluso en parte tiene sus ventajas, debemos de ser conscientes de lo potente que es lo que hay detrás de lo que somos capaces de ver a simple vista como ciudadano, beneficiarnos de las ventajas que nos ofrece el hecho que parte de nuestra información este controlada por terceros, siendo conscientes obviamente de los inconvenientes que puede haber.

En cuanto al trabajo en sí, creo que se ha llegado al objetivo principal que era que el lector tome consciencia de la exposición de sus datos, tenga conocimiento de lo grave que puede ser un exceso de publicación personal en la red debido a las consecuencias que esto puede conllevar, y a su vez ayudar o dar a conocer las formas que tenemos de proteger nuestra privacidad en internet conociendo como se obtienen nuestros datos, de esta manera podemos evitar que estos sean obtenidos por terceros si nosotros no lo deseamos.

#### 4. Glosario.

**Big data.** *Es un concepto que hace referencia a un conjunto de datos tan grandes que aplicaciones informáticas tradicionales de procesamiento de datos no son suficientes para tratar con ellos y los procedimientos usados para encontrar patrones repetitivos dentro de esos datos*

**Fingerprints.** *Es el proceso de recopilación de información que permite identificar el sistema operativo en el ordenador que se tiene por objetivo*

**Cache.** *Una caché es un componente de hardware o software que almacena datos para que las solicitudes futuras de esos datos se puedan atender con mayor rapidez; los datos almacenados en un caché pueden ser el resultado de un cálculo anterior o el duplicado de datos almacenados en otro lugar, generalmente, de velocidad de acceso más rápido*

**Cookie.** *Una cookie, galleta o galleta informática es una pequeña información enviada por un sitio web y almacenado en el navegador del usuario, de manera que el sitio web puede consultar la actividad previa del usuario.*

**Framework.** *Es un conjunto estandarizado de conceptos, prácticas y criterios para enfocar un tipo de problemática particular que sirve como referencia, para enfrentar y resolver nuevos problemas de índole similar.*

**Mix Networks.** *Son protocolos de enrutamiento que crean comunicaciones difíciles de rastrear mediante el uso de una cadena de servidores proxy conocida como mezclas que captan mensajes de múltiples remitentes, los mezclan y los envían de vuelta en orden aleatorio al siguiente destino (posiblemente otro nodo de mezcla) .*



**VPN.** Una red privada virtual (RPV), en inglés: Virtual Private Network (VPN) es una tecnología de red de computadoras que permite una extensión segura de la red de área local (LAN) sobre una red pública o no controlada como Internet.

**RED TOR.** Tor es la sigla de The Onion Router (en español: Enrutador de Cebolla). Es un proyecto cuyo objetivo principal es el desarrollo de una red de comunicaciones distribuida de baja latencia y superpuesta sobre internet, en la que el encaminamiento de los mensajes intercambiados entre los usuarios no revela su identidad, es decir, su dirección IP (anonimato a nivel de red) y que, además, mantiene la integridad y el secreto de la información que viaja por ella.

**Bitcoin.** Bitcoin es un protocolo y red P2P que se utiliza como criptomoneda, sistema de pago y mercancía.

**Hadoop.** Apache Hadoop es un framework de software que soporta aplicaciones distribuidas bajo una licencia libre. Permite a las aplicaciones trabajar con miles de nodos y petabytes de datos

## 5. Bibliografía

1. 02/03/2018. <http://osintframework.com/>
2. 02/03/2018. <http://www.elladodelmal.com/2016/05/osint-framework-donde-buscar-datos-de.html>
3. 02/03/2018. <https://www.cybrary.it/0p3n/find-juicy-data-targets-using-osint-framework/>
4. 02/03/2018. <http://cso.computerworld.es/social-security/redes-sociales-beneficios-y-riesgos-para-la-privacidad-de-los-datos>
5. 02/03/2018. <https://www.xataka.com/empresas-y-economia/uber-aclara-las-acusaciones-que-es-el-fingerprinting-con-el-que-monitorizaban-a-sus-ex-usuarios>
6. 18/03/2018. <https://economictimes.indiatimes.com/definition/digital-fingerprinting>
7. 18/03/2018. <https://www.humanlevel.com/articulos/desarrollo-web/que-son-y-como-funcionan-las-cookies.html>
8. 18/03/2018. <http://www.whatarecookies.com/>
9. 06/04/2018. <https://aticser.wordpress.com/2014/02/08/beacons-balizas-la-nueva-revolucion-tecnologica/>

10. 06/04/2018. <https://www.elevenpaths.com/labstools/foca/index.html>
11. 06/04/2018. <http://app.teachingprivacy.org/>
12. 02/05/2018. <https://www.ghostery.com/>
13. 02/05/2018. <http://informatizarte.com.ar/blog/?p=1739>
14. 02/05/2018. <https://panopticlick.eff.org/results?aat=1&t=111&dnt=111#finger-printTable>
15. 02/05/2018. <https://www.whois.net/>
16. 02/05/2018. <http://thegreatbigbrother.blogspot.com/>
17. 02/05/2018. <https://www.xataka.com/legislacion-y-derechos/gdpr-rgpd-que-es-y-como-va-a-cambiar-internet-la-nueva-ley-de-proteccion-de-datos>
18. 02/05/2018. <http://forbes.es/emprendedores/7560/como-el-big-data-ayudo-a-obama-a-ganar/>
19. 02/05/2018. <http://www.ticbeat.com/bigdata/campana-big-data-dio-victoria-obama/>
20. 02/05/2018. <http://www.datacentric.es/blog/insight/exito-netflix-datos/>
21. 02/05/2018. <https://policies.google.com/privacy?hl=en>
22. 02/05/2018. <https://www.whatsapp.com/legal/?eea=1#how-we-process-your-information>
23. 02/05/2018. <https://www.isecauditors.com/curso-tecnicas-osint>
24. 02/05/2018. <https://rootear.com/windows/ver-cookies-chrome>
25. 02/05/2018. <https://www.google.com/analytics/tag-manager/>
26. 02/05/2018. [http://noticias.juridicas.com/base\\_datos/Admin/lo15-1999.html](http://noticias.juridicas.com/base_datos/Admin/lo15-1999.html)
27. 02/05/2018. <https://www.1and1.es/digitalguide/paginas-web/derecho-digital/el-rgpd-normativa-europea-de-proteccion-de-datos/>