Universitat Oberta de Catalunya

uoc.edu

¿Qué sabe Internet de nosotros?

Alumno: Vicente Esbrí González

Trabajo Final de Máster: ¿Qué sabe Internet de nosotros?

Plan de Estudios del Estudiante: Máster Interuniversitario de Seguridad en las Tecnologías de la Información y las comunicaciones (MISTIC)

Área del trabajo final: TFM-Ad hoc

Directora: Angela María García Valdés

Profesor responsable de la asignatura: Víctor García Font

Empresa: Instituto Nacional de Ciberseguridad (INCIBE)

Fecha de entrega: 12/2018

FICHA DEL TRABAJO FINAL

Título del trabajo:	*¿Qué sabe Internet de nosotros?*
Nombre del autor:	*Vicente Esbrí González*
Nombre del consultor/a:	*Angela María García Valdés*
Nombre del PRA:	*Víctor García Font*
Fecha de entrega:	12/2018
Titulación:	*Máster Interuniversitario de Seguridad en las Tecnologías de la Información y las comunicaciones (MISTIC)*
Área del Trabajo Final:	*TFM-Ad hoc*
Idioma del trabajo:	*Español*
Palabras clave	*Identidad, Privacidad, Reputación*

Resumen

El objetivo del presente trabajo es estudiar qué tipo de información puede ser recopilada sobre una persona debido al uso de Internet y de las nuevas tecnologías en general, es decir, en qué consiste y cómo se genera la identidad o huella digital, cómo se recopilan los datos de los usuarios, para qué pueden ser utilizados, y cómo puede influir la identidad digital en una persona, describiendo las medidas que se pueden tomar para evitarlo o para potenciar una reputación digital positiva.

En un mundo cada vez más digitalizado e interconectado, este concepto es de vital importancia. La capacidad de identificar a un individuo es clave en la sociedad de la información y del conocimiento, bien para ofrecer servicios a los ciudadanos, bien para obtener un beneficio económico o social. Conocer estos usos, así como estrategias de gestión de la reputación en línea, se presenta como una habilidad valiosa para los ciudadanos del futuro.

Para responder a estas preguntas se investiga el contenido de la huella digital y los mecanismos activos y pasivos de generación, así como usos legítimos e ilegítimos que pueden hacerse de nuestros datos. A su vez se contrasta la preocupación de los ciudadanos con su desconocimiento manifiesto acerca de los mecanismos de protección a su alcance, por lo que se proponen algunos procedimientos para mejorar la reputación digital.

Abstract

The present paper aims to study what kind of information can be gathered about a person as a consequence of using the Internet and new technologies in general, in other words, what is the digital footprint (or digital identity) and how is it generated, how users data is compiled, what can this data be used for, and finally how can digital identity influence people's life, describing which measures can be taken to avoid adverse effects or, on the contrary, to promote a favourable digital reputation.

In a world where digitalisation and interconnection are steadily increasing, this concept is of crucial importance. The ability to identify an individual is key to the information and knowledge society, either to offer services to citizens or to get an economic or social benefit. Getting to know these applications, as well as digital reputation managing strategies, seem to be valuable abilities for future citizens.

To answer those questions, we are going to study the digital footprint content and its generation mechanisms, active and passive, as well as legitimate and illegitimate uses of our data. At the same time, citizens concerns stand in contrast to their absolute lack of knowledge about the protection mechanisms available. Therefore, we will suggest some procedures to improve digital reputation.

CONTENIDO

1. INTRODUCCIÓN

1.1. CONTEXTO Y JUSTIFICACIÓN

En un mundo cada vez más digitalizado e interconectado, el concepto de identidad o huella digital cobra cada vez mayor relevancia. Ya en el mundo físico, la capacidad de identificar a un individuo puede cumplir con numerosas y variadas finalidades, como pueden ser aplicaciones legales, administrativas, de seguridad, etc. Podemos pensar en ejemplos cotidianos, como el uso del documento nacional de identidad para identificarse ante una mesa electoral, el uso del carnet de conducir para garantizar que hemos superado un requisito que nos habilita para ello, la posesión de una llave que nos da acceso a una vivienda como legítimos propietarios, o el conocimiento de una contraseña secreta que nos daba acceso a la casa del árbol en un juego de nuestra infancia. El proceso de identificación es continuo y muchas veces inconsciente, ¿cuántas veces hemos recibido una llamada telefónica de un número desconocido e intentado identificar a nuestro interlocutor a través de su voz? ¿Quién no ha dudado ante una cara familiar, procurando ponerle nombre y contexto antes de admitir que hemos olvidado a su propietario? Estos mecanismos innatos del ser humano, de hecho, se intentan replicar con las nuevas tecnologías de reconocimiento facial o de voz, añadiéndolos a la lista de características biométricas que los algoritmos del futuro serán capaces de medir.

También inherente al ser humano subyace la idea de obtener un provecho de dicha identificación, por lo que no sorprende encontrar personas o empresas interesadas en identificar a otros individuos y recopilar una serie de características de interés para, por ejemplo, dirigir campañas de marketing individualizadas o, al menos, enfocadas a distintos *target*.

Como decíamos al comienzo, en un mundo cada vez más digitalizado e interconectado, la capacidad de identificar a un individuo se revela como un elemento clave en el funcionamiento de la sociedad de la información y del conocimiento, ya sea para ofrecer servicios a los ciudadanos, ya sea para obtener información de utilidad sobre los mismos con finalidades muy variadas. En la actualidad son múltiples las vías en que se recopila y muy variadas las maneras en que se utiliza nuestra huella digital, no siendo a menudo el usuario plenamente consciente ni de lo primero ni de lo segundo, lo que provoca su manifiesta indefensión. Conocer dichas vías de recopilación y explotación debería ser un primer paso para asumir el control de nuestra propia identidad digital y, con una estrategia adecuada, proyectar una huella digital acorde a nuestros intereses, beneficiosa al fin y al cabo, aumentando nuestra reputación digital, concepto íntimamente ligado al primero.

Así pues, se estudiará qué tipo de información puede ser recopilada sobre una persona debido al uso de Internet y de las nuevas tecnologías en general, es decir, cómo se genera la identidad o huella digital, investigando de dónde se puede obtener esta información, cómo se recopilan datos de usuarios, para qué pueden ser utilizados; y reflexionando sobre cómo puede influir la identidad o huella digital en una persona y describiendo las medidas que se pueden tomar para evitarlo.

De manera general, se estudiarán las dos vías en que generamos la huella digital, la primera activa, mediante la publicación voluntaria de contenidos, participación en redes sociales, emisión de comentarios públicos, etc., y la segunda pasiva, a través tanto de mecanismos de rastreo sencillos como de otros más sofisticados.

Una vez analizada la generación de la identidad digital, pondremos el foco en el uso de la huella digital por parte de terceros y en las consecuencias que dicho uso puede representar para el usuario, tanto positivas como negativas, así como en las medidas de protección activas y pasivas que se pueden adoptar. Finalmente pondremos especial énfasis en cómo tomar el control de nuestra identidad digital, y cómo construir una reputación digital positiva.

1.2. OBJETIVOS

El principal objetivo del presente trabajo consiste en estudiar la identidad o huella digital de las personas derivada del uso de las nuevas tecnologías, así como las consecuencias que ésta puede generar. Para lograrlo, se tratará de dar respuesta a los siguientes objetivos:

- Definir el concepto de identidad o huella digital. Averiguar qué es y cómo se genera.
- Descomponer la huella digital en sus partes, esto es, reconocer y clasificar toda aquella información que puede ser utilizada para identificarnos.
- Valorar el riesgo percibido por el usuario y los riesgos reales de cada uno de los datos que conforman la huella digital.
- Determinar los principales mecanismos de constitución de la huella digital del individuo, tanto activos o directos como pasivos o indirectos.
- Valorar el grado de control que un usuario puede tener sobre los mecanismos de constitución de su huella digital.
- Analizar las vías de explotación de la información contenida en la huella digital del individuo, tanto legales como maliciosas.
- Reflexionar sobre si el usuario es consciente de toda la información que volcamos en Internet y cómo puede afectarle.

- Definir el concepto de reputación digital. Realizar un estudio sobre posibles pautas y recomendaciones para tener una identidad digital acorde a nuestros deseos.

1.3. ENFOQUE Y MÉTODO SEGUIDO

Estamos principalmente ante un trabajo de investigación, de manera que la metodología consistirá en primer lugar en la lectura de todas aquellas fuentes de información que analicen la huella digital, qué tipos de datos la componen, cómo se genera, cuáles son las últimas técnicas desarrolladas para obtener información de los usuarios, etc.

Dado que estamos ante un tema actual y cercano, procederemos a estudiar casos reales, como la configuración de seguridad y privacidad de varias redes sociales, la huella dejada por un navegador web corriente, los complementos existentes contra el rastreo, la legislación actual y sus implicaciones en la gestión de *cookies*, etc.

Una vez recogida la información, se procederá a un posterior análisis, descomponiendo la huella digital en sus distintas partes, analizando el riesgo que entraña para el usuario cada uno de los datos que la componen, así como sus posibles beneficios. Parte del análisis de riesgos consistirá en analizar qué posibles usos podría darle un tercero a nuestros datos.

Finalmente, se analizarán las estrategias existentes para tratar de proteger nuestra identidad digital o darle valor, y se reflexionará sobre posibles estrategias nuevas. En concreto se estudiará la huella digital de algún personaje famoso, y qué medidas tomadas contribuyen a una reputación digital positiva o, por el contrario, la penalizan.

1.4. PLANIFICACIÓN

Las etapas necesarias para cumplir con los objetivos marcados para el presente trabajo son las siguientes:

Plan de trabajo.

En esta etapa introductoria se describe el problema que se pretende resolver, el trabajo concreto que se llevará a cabo, la metodología utilizada, y la descomposición de este trabajo en tareas y metas temporales, justificando la importancia del área estudiada.

Hito *– Entrega de la PEC 1.*

Tarea 1 – Concepto de huella o identidad digital.

Es esta etapa se define qué se considera una huella digital en la literatura existente, resaltando su importancia e impacto en la sociedad actual y futura.

Tarea 2 – Análisis del concepto de huella digital.

En la segunda fase, se analizará el concepto anteriormente definido para descomponerlo en sus partes. Entre otros, se estudiarán datos de geolocalización, metadatos, redes de contactos, actividad en redes sociales, *cookies*, etc. Se finalizará con un análisis del riesgo percibido y del riesgo real de las partes de la huella digital.

Tarea 3 – Generación de la huella digital.

En el tercera etapa se estudiarán los métodos más comunes de recolección de datos a través de los distintos dispositivos conectados a Internet, desde *smartphones* hasta *wearables*, pasando por los tradicionales equipos sobremesa y portátiles. El objetivo en este punto es determinar las principales vías de filtración de información, desde los permisos que concedemos a las *APPs* móviles hasta las publicaciones que realizamos de manera activa.

***Hito** – Entrega de la PEC 2.*

Tarea 4 – Descomposición y estudio de los métodos de generación de la huella digital.

Se trata de una fase de profundización, en la que se clasificarán los métodos anteriormente vistos en directos o activos e indirectos o pasivos, así como los datos publicados por terceros. Además se realizará una descripción pormenorizada de dichos métodos.

Tarea 5 – Usos de la huella digital.

Esta tarea se centrará en realizar un estudio pormenorizado de los posibles usos que se pueden dar a la huella digital de un individuo, ya sean legales como el marketing, como maliciosos. Se procurará realizar una descripción de ejemplos del mundo real.

***Hito** – Entrega de la PEC 3.*

Tarea 6 – Huellas digitales reales.

En esta etapa se llevará a cabo una investigación sobre ejemplos de huellas digitales existentes de personajes conocidos. En concreto, se buscarán casos de huellas digitales de perfil predominantemente positivo, así como ejemplos del caso contrario, cuya identidad digital haya puesto en un compromiso a su propietario.

Tarea 7 – Estrategias.

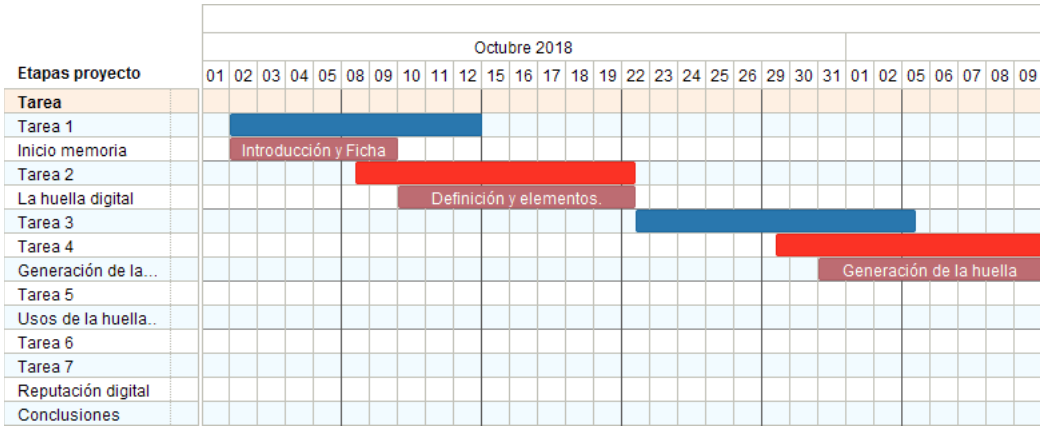
Analizadas huellas digitales existentes, estudiaremos las estrategias que favorezcan el desarrollo de una huella digital positiva, así como aquellas que la penalicen. Clasificaremos dichas estrategias, describiéndolas en profundidad.

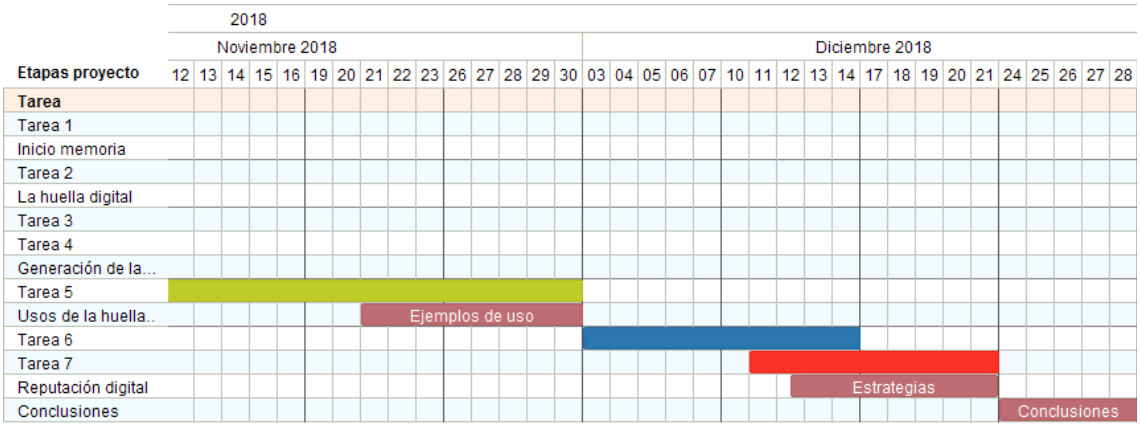
Conclusiones y trabajo futuro.

Como cierre se presentarán las conclusiones, valorando si se han logrado o no los objetivos propuestos, las dificultades encontradas y los mecanismos de corrección aplicados. También se identificarán posibles líneas de trabajo que no se hayan podido tratar en este trabajo.

***Hito** – Entrega de la PEC 4 o memoria final.*

Se presenta a continuación una planificación temporal del trabajo:





1.5. PRODUCTOS OBTENIDOS

Este Trabajo Final de Máster se puede dividir en varias entregas parciales, que se describen brevemente a continuación:

- PEC1: El plan de trabajo, donde se define el problema a abordar, los objetivos a cumplir, así las metas temporales para su correcta ejecución.
- PEC2: Se trata de un análisis del concepto de huella digital, así como la descripción pormenorizada de sus partes, y los riesgos reales y percibidos de cada una de ellas. También se introducen a grandes rasgos las vías de generación de la huella digital.

- PEC3: Se analizan en profundidad las vías de generación de la identidad digital de un individuo, para a continuación estudiar los posibles usos que de estos datos podría hacer un tercero.
- PEC4: Tras estudiar algunos casos reales, así como el concepto de reputación digital, se presentan estrategias para lograr el control de la propia huella digital y proyectar una reputación positiva. Se completa el trabajo con las conclusiones finales, trabajo futuro, glosario y anexos si los hubiere.
- Presentación/Vídeo: Se presenta una síntesis del trabajo realizado sobre una presentación de diapositivas, de forma análoga como se haría en una exposición presencial, en formato vídeo y con la voz en *off* del autor.

1.6. CONTENIDO DE LA MEMORIA

El contenido de la memoria es el que sigue:

- En el **capítulo 2** analizamos el concepto de huella digital, dando una definición y una pequeña reflexión sobre el impacto que esta puede tener en nuestra vida. Posteriormente, se analizan los principales datos que la suelen constituir para, finalmente, realizar un breve análisis de riesgos de estos.
- En el **capítulo 3** estudiamos la formación de la identidad digital, esto es, los mecanismos por los que nuestros datos son recogidos, almacenados y procesados. Se establece una clasificación según dicho mecanismo sea activo, pasivo, o fruto de la intervención de terceros.
- En el **capítulo 4** discutimos por posibles usos que se pueden realizar de los datos contenidos en la huella digital de un individuo, ya sean legítimos o ilegítimos, incluyendo un punto de vista ético. También ofrecemos una primera aproximación a la utilidad que puede tener la publicidad de sus datos para el propietario de estos.
- Utilizamos el **capítulo 5** para describir un concepto relacionado con la identidad digital, el concepto de reputación *online*. Una vez entendida su importancia, estudiamos tres casos de reputación digital, una neutra, una positiva y una última negativa.
- En el **capítulo 6** establecemos un plan de acción recursivo incremental para lograr una reputación *online* acorde a nuestros intereses, controlando el contenido de nuestra huella digital. Describimos los cuatro pasos básicos: investigación, saneamiento, limitación y proyección; dando ejemplos de herramientas y configuraciones óptimas.

2. LA HUELLA DIGITAL

2.1. DEFINICIÓN

La huella o identidad digital de un individuo está constituida por el rastro que éste deja al utilizar las redes, más concretamente Internet. Cualquier actividad en línea, como realizar una publicación o comentario en redes sociales, proporcionar *feedback* positivo o negativo a una publicación, realizar una llamada o participar en una conversación, el uso de aplicaciones móviles, la orientación o registro de ubicación mediante GPS, el envío o la recepción de un correo electrónico, o la simple navegación por un sitio web... todo ello termina formando parte de nuestro historial en línea y, potencialmente, puede ser visto por terceras personas o almacenado para su tratamiento o uso posterior, generalmente en bases de datos o ficheros de *log*.

Estamos ante una definición ampliamente aceptada aunque, como puede observarse, demasiado general. Conviene que en los siguientes apartados analicemos en profundidad qué impacto puede tener la huella digital en nuestra vida cotidiana, qué datos componen por regla general la huella digital del individuo, así como el nivel de riesgo de dichos datos.

2.2. IMPACTO DE LA HUELLA DIGITAL

Como ya se indicaba en la introducción, en este mundo digitalizado e interconectado la capacidad para identificar a un individuo constituye los cimientos mismos de la sociedad de la información y del conocimiento, ya sea para ofrecer servicios a los ciudadanos, ya sea para obtener información de utilidad sobre los mismos con distintas finalidades.

Por lo que se refiere a los servicios que se pretende ofrecer al ciudadano, pensemos en un ejemplo tan sencillo y a la par tan complejo como el proceso de compra en una tienda *online* extranjera. A través de las *cookies*, pequeños ficheros de información que se depositan en nuestro ordenador a través del navegador web, podemos ser identificados como un comprador concreto y único. A partir de dicha identificación, la página de compra puede mostrarse por ejemplo en nuestro idioma, en esta y futuras visitas, así como almacenar un carro de la compra con distintos productos, manteniendo incluso la persistencia de datos entre diferentes sesiones. Este rastro digital, esta huella que vamos dejando con nuestras acciones, puede ser utilizada de manera útil para el usuario, facilitando o incluso haciendo posibles procesos que serían imposibles de otro modo.

En cuanto a otras finalidades, de momento baste con decir que numerosas compañías reconocen abiertamente haber rastreado Internet en busca de información sobre aquellos candidatos que se presentaron a algún proceso de selección para acceder a un puesto de trabajo ofertado, bien de

manera previa a dicho proceso, bien de manera inmediatamente posterior a éste. En similar medida, algunas compañías comprueban de manera periódica el historial en la red de sus empleados, en busca de posibles comportamientos que no se ajusten a las políticas de la empresa.

Si este ejemplo debe servirnos de algo es para confirmar lo que todos sospechamos: que en un mundo en el que cada vez hay más información pública y, en cierta manera, permanente sobre nosotros, y en el que el acceso a dicha información está a un mero clic de distancia, el contenido y la calidad subjetiva de nuestra identidad digital puede afectar a nuestra vida personal, familiar, social, laboral, económica, etc. de maneras que de momento solo podemos empezar a imaginar, y que muy probablemente se van a ir ampliando en el futuro.

Volveremos sobre este punto más adelante, cuando analicemos los posibles usos que diferentes actores pueden realizar de nuestra huella digital. Mas, para entender dichos usos, conviene tener una idea más detallada de los elementos que pueden constituir la identidad digital de cualquiera de nosotros.

2.3. COMPONENTES DE LA HUELLA DIGITAL

Múltiples datos pueden formar parte de la huella digital de un individuo y contribuir por lo tanto a su identificación, definiendo a su vez una imagen más o menos cercana a la realidad de éste.

Proponemos a continuación una lista no exhaustiva de los datos que acostumbran a formar parte de la identidad digital.

2.3.1. DIRECCIÓN IP

Cualquier dispositivo, independientemente de su complejidad o tamaño, requiere de una dirección IP para conectarse a Internet y deja así un rastro que puede formar parte de la huella digital. Esto es obvio para equipos de sobremesa o portátiles conectados a una red de área local, o para una tableta digital conectada a un punto de acceso inalámbrico. Un *smartphone*, incluso si está asociado a una red 3G/4G, necesita conectarse a un APN para salir a Internet. Y en el Internet de las cosas y el mundo de los *wearables*, conviven dispositivos con capacidad de conexión directa a Internet con otros que necesitan hacerlo a través de un tercero, por ejemplo estableciendo una conexión Bluetooth con un teléfono móvil. Todos ellos, en el mismo instante en que intercambian información con la red, lo hacen usando una dirección IP.

Esta dirección, además del número en sí, puede proporcionar información como el proveedor de servicios de Internet contratado (ISP), así como ciertos datos de localización geográfica, como el país o la ciudad desde la que se realiza la conexión[1].

2.3.2. DATOS PERSONALES

Una parte especialmente sensible de nuestra huella digital viene conformada por nuestros datos personales, tales como nombre y apellidos, sexo biológico, dirección postal, dirección de correo electrónico, teléfono… Todos ellos son datos que cedemos gustosamente o, al menos, con bastante ligereza, por lo que podríamos decir que generalmente no los consideramos datos críticos.

A poco que utilicemos algunos servicios de Internet, sin embargo, esta información se puede ver ampliada con datos que quizá no cederíamos si no lo impusieran así las condiciones del servicio, y que no nos gustaría ver comprometidos. Estamos hablando de datos de salud, financieros, contraseñas personales, situación sentimental, datos académicos de formación y resultados, datos profesionales como nuestra empresa y cargo, etc.

De igual manera, el uso de servicios en línea de tipo agenda o calendario deja un rastro sobre nuestra vida cotidiana, incluyendo todo aquello que anotemos: citas al médico, viajes, eventos sociales...

2.3.3. PREFERENCIAS Y GUSTOS

Nuestra huella digital conoce perfectamente nuestros gustos y preferencias, como demuestran las páginas de gestión de preferencias de personalización de anuncios de Google[2] y Facebook[3].

Conoce por consiguiente nuestra música favorita, si es clásica o reggaeton, si nos gusta leer y qué tipo de libros, nuestras películas y series favoritas, si somos aficionados o practicamos algún deporte, nuestras preferencias culinarias, restaurantes favoritos, si nos gusta o no viajar, qué tipo de ropa llevamos, en qué invertimos nuestro tiempo de ocio, si estamos estudiando algún idioma extranjero, si tenemos mascotas en casa, qué aplicaciones utilizamos con mayor frecuencia, qué tipo de noticias leemos, e incluso nuestra orientación sexual, ideológica, política y religiosa, datos estos últimos considerados especialmente protegidos por algunas de las legislaciones de protección de datos existentes.

2.3.4. DATOS DE GEOLOCALIZACIÓN

Es bastante probable que nuestra huella digital tenga un historial bastante detallado de los lugares en los que hemos estado, cuándo y por cuánto tiempo. Tomemos como ejemplo lo que sabe el archiconocido gigante Google[4], que almacena dicho historial por defecto.

Por si dicha información no fuera suficientemente sensible por sí sola, de los datos brutos se pueden inferir con relativa facilidad otros más útiles, como nuestro lugar de residencia y nuestro lugar de trabajo, comparando a qué horas y qué días se está en determinados lugares; los lugares a los que hemos viajado, de lo que también se puede intentar colegir nuestro nivel adquisitivo; nuestras rutinas deportivas, actividades de ocio, hábitos de compra...

2.3.5. DATOS DE NAVEGACIÓN Y BÚSQUEDA

Una parte bastante obvia de nuestra huella digital viene conformada por nuestros datos de navegación, esto es, el historial de sitios web que hemos visitado, lo que incluye no solo la URL del sitio, sino la fecha y la hora de la visita, cuánto ha durado ésta, o con qué frecuencia se realiza. Dicha información también incluye las búsquedas realizadas. No hablamos aquí de la información que almacena localmente nuestro navegador web, sino de los datos en posesión de empresas tecnológicas que operan en la red, como el historial que almacena Google[5].

Los datos de búsqueda no se limitan a la información recopilada por los buscadores web propiamente dichos, sino que en muchos casos hay que considerar que la huella digital contendrá información de otras muchas aplicaciones, como pueden ser plataformas de comercio *online*, servicios de *streaming* de audio y/o vídeo, o el popular servicio de vídeo bajo demanda *Youtube*, que almacena no solamente los vídeos buscados, sino aquellos que hemos reproducido.

No debemos olvidar, además, que no hablamos únicamente de texto introducido en un buscador, sino que cada vez hay más dispositivos con capacidad de escucha, que generan un historial de nuestras consultas de voz, extremo éste que muchos usuarios desconocen.

2.3.6. DATOS DE SENSORES

No solamente estamos rodeados de dispositivos con capacidad de escucha. Cada vez es más habitual contar con pequeños aparatos inteligentes dotados de sensores para medir nuestra actividad, o para tomar decisiones en base a los valores medidos. Algunos ejemplos son los *wearables* específicos de monitorización, o ciertos dispositivos electrónicos tradicionales a los que se les ha añadido la capacidad de conectarse a Internet, en lo que se conoce como el Internet de las cosas.

Los datos de actividad deportiva, por ejemplo, se comparten de manera pública o privada en muchos casos, dando lugar a un perfil de datos de salud que podría resultar de alto interés para determinadas compañías que operan en el ámbito de la medicina o de los seguros. Datos como el tipo de actividad física que realizamos, la frecuencia y la duración de ésta, o incluso nuestra frecuencia cardíaca, pasan de esta manera a formar parte de nuestra huella digital.

De igual manera, dispositivos inteligentes en el mundo del ocio, la domótica, etc. ofrecen a día de hoy la posibilidad de compartir información en las redes, por ejemplo la temperatura de nuestro hogar, la cantidad de lluvia medida por un pluviómetro, o si hemos utilizado nuestra aspiradora robot. Todos ellos ofrecen, en cierto sentido, información sobre nosotros.

2.3.7. DATOS DE DISPOSITIVOS

Nuestra huella digital contiene numerosa información sobre nuestros dispositivos electrónicos. Del apartado anterior se deduce fácilmente que contiene no solamente datos sobre nuestra actividad o el resultado de realizar una medición a través de un sensor, sino también sobre el propio dispositivo que la realiza.

Sin embargo, los dispositivos que mayor rastro dejan al ser utilizados son aquellos de los que hacemos un uso más activo: ordenadores personales, teléfonos inteligentes, tabletas digitales… A modo de ejemplo, una vez más el gigante tecnológico Google es conocedor de qué dispositivos poseemos, y desde qué dispositivos nos hemos conectado a nuestra cuenta[6], lo cual puede emplearse para aumentar nuestra seguridad, por ejemplo.

Al utilizar un navegador web, por ejemplo, dejamos un rastro con cuantiosa información sobre el dispositivo desde el que estamos accediendo, como qué navegador utilizamos y qué complementos o *plug-ins* hay instalados, nuestra zona horaria, nuestro sistema operativo, datos de *hardware* como la CPU y la GPU, la resolución de nuestra pantalla, qué fuentes tenemos instaladas, etc.

Por otra parte, un teléfono inteligente deja un rastro sobre el número de teléfono, IMEI, dirección MAC, número de serie de la tarjeta SIM, identificador Bluetooth...

2.3.8. DATOS DE CONTACTOS

No toda nuestra huella digital gira en torno a nosotros: también incluye numerosa información sobre nuestro entorno social. Con quién nos relacionamos en línea deja un rastro sobre esta red de vínculos. Teniendo como punto de partida nuestra agenda en un programa de correo o en nuestro *smartphone*, numerosas aplicaciones tratan de establecer un mapa de contactos. Dicho mapa tiene especial interés cuando se logra además categorizar a cada una de estas personas, asignándoles un rol en calidad de pareja sentimental, compañeros de trabajo, familiares, amigos o meros conocidos.

Así pues, Internet conoce con sumo detalle a qué comunidades de individuos pertenecemos, con quién nos relacionamos con más frecuencia, el grado de familiaridad de dichas relaciones, y puede incluso predecir, según algunos estudios, cuándo una relación sentimental ha empezado[7] o se acaba de romper[8].

2.3.9. DATOS DE COMUNICACIONES

Consecuencia de lo anterior, esto es, que no estamos solos en Internet sino acompañados de nuestra red de contactos, surge la parte de la huella digital conformada por todas las interacciones que realizamos en la red.

Así pues dejamos un rastro al usar aplicaciones de mensajería, por ejemplo, tales como el correo electrónico o la mensajería instantánea. También cuando respondemos a un mensaje en una red social, ya sea con un texto o con un *feedback* de tipo emocional (me gusta, me divierte, me disgusta…), o cuando reenviamos un mensaje publicado por alguno de nuestros contactos. Así mismo, supone un mecanismo de comunicación que deja una huella digital indeleble el hecho de etiquetar a un contacto en una fotografía, o el hecho de seguir a alguien en una red social.

También ayuda a conformar la huella digital el hecho de aglutinar a varios de estos contactos en un grupo o en una lista, a los que además ponemos un nombre identificativo y, generalmente, significativo.

2.4. ANÁLISIS DE RIESGOS

Los diferentes datos que conforman nuestra identidad digital pueden suponer un riesgo para nosotros, como también una oportunidad si son vigilados y tratados adecuadamente. Sin ánimo de profundizar, ya que volveremos sobre este punto más adelante, conviene realizar un somero análisis de los riesgos percibidos y reales de las diferentes categorías de datos descritas. La percepción de riesgos se ha medido con una encuesta cuyos resultados pueden consultarse en el Anexo I:

- **Dirección IP**: la gran mayoría de los usuarios desconocen en su totalidad o en parte este concepto, de manera que no hay una percepción de riesgo clara (la encuesta muestra una enorme dispersión que interpretamos como desconocimiento). Si nos preocupamos estrictamente de la huella digital, y no de otras amenazas cibernéticas, probablemente el riesgo real es bajo, exceptuando que realicemos actividades delictivas y seamos rastreados por algún cuerpo de seguridad del estado.
- **Datos personales**: en esta categoría encontramos gran variedad de datos. El dato menos relevante para los usuarios es su sexo, que no les importa que se conozca en Internet, seguido de los datos académicos, nombre y apellidos, y correo electrónico. Por algún motivo, se muestran bastante más reticentes a que se conozca su relación sentimental, y dan la misma importancia a que sea accesible su dirección y número de teléfono que datos acerca de su salud, a pesar de que éstos últimos son considerados sensibles tanto por la Ley Orgánica de Protección de Datos como por el Reglamento Europeo de Protección de Datos. El rechazo es muy acusado si hablamos de publicar datos de contraseñas o datos financieros.

  Es evidentemente que muchos de estos datos están presentes en Internet (un servicio no puede validar nuestra identidad si no dispone de nuestra contraseña para contrastarla), de manera que el usuario

presupone una relación de confianza con la empresa a la que cede los datos (un banco, por ejemplo), que no debería compartirlos con terceros. Sin profundizar en consideraciones legales, la ley ampara la cesión de dichos datos cuando "la finalidad sea la prestación de un servicio al responsable del tratamiento"[9].

En cuanto al potencial riesgo de estos datos, debemos considerar en primer lugar la ingeniería social. Un criminal que desee obtener un beneficio de nosotros tratará en primer lugar de obtener toda la información posible al respecto, para usar técnicas de manipulación psicológica que faciliten su tarea. Los datos de salud pueden ser utilizados para negar a un usuario el acceso a un seguro (Manulife ya ha introducido la idea de utilizar los datos recogidos por *wearables* para ajustar los precios de sus pólizas de seguro[10]), y los datos financieros pueden ser empleados para negar un préstamo o una hipoteca a un potencial consumidor. Nuestra situación sentimental nos puede hacer susceptibles de ser víctimas del *romance scam*, y nuestro correo de cualquier tipo de *spam*. Finalmente, un criminal puede aprovechar su conocimiento de nuestra dirección postal para, combinando esta información con otras, perpetrar un robo o asalto.

- **Preferencias y gustos**: es probablemente este el apartado donde los usuarios encuestados se encuentran más cómodos compartiendo información en Internet. Consideran de bajo riesgo dar datos sobre sus gustos musicales, cinematográficos, restaurantes favoritos… y, en un orden algo inferior pero igualmente aceptable, sus actividades de ocio y viajes. Sorprende que éste último dato resulte de similar nivel de riesgo que la orientación sexual, religiosa o política (solamente un punto menos), y a un nivel similar que el nombre y los apellidos o el correo electrónico, pese a que ideología, religión, creencias y orientación sexual se consideran datos especialmente protegidos por la legislación vigente.

  En cuanto al riesgo de estos datos, dejando a un lado nuevamente la ingeniería social, los datos especialmente protegidos se consideran como tales porque son fuente de posible discriminación. Esto implica que el acceso a estos datos por parte de un posible contratante, por poner un ejemplo, podría influir en un proceso de selección de personal, impidiéndonos el acceso a un puesto de trabajo.

- **Datos de geolocalización**: la muestra de usuarios encuestada manifiesta un fuerte rechazo a que se almacene un registro de ubicaciones, lo que entra en conflicto con la aceptación de publicar datos sobre viajes, actividades de ocio o restaurantes favoritos. Probablemente debamos atribuirlo a una sensación de "ser seguidos y vigilados" en todo momento. Igualmente, se muestran menos preocupados por que se conozcan las fotos que publican en redes

sociales, lo que puede significar que desconocen el funcionamiento de los metadatos y el *geotagging*.

Las implicaciones de seguridad y riesgos de los datos de ubicación son amplios en caso de filtración, ya que revelan mucha información sobre hábitos y rutinas. Un caso reciente, en este caso ligado al *big data*, fue el escándalo de la filtración en Strava[11] de la ubicación de bases militares secretas.

- **Datos de navegación y búsqueda**: los encuestados presentan un rechazo a que se conozca su historial de Internet superior a que se sepa su orientación política, pese a que este sea probablemente uno de los rastros más públicos y sobre el que los usuarios ejercen menor control. El rechazo es todavía mayor si planteamos la pregunta en términos de grabación de voz.

  A día de hoy, esta información se utiliza principalmente con fines comerciales y de estudio de mercado, en algunos casos llegando a un nivel de *microtargeting*. Los dispositivos de escucha activa, por el contrario, son bastante nuevos y es pronto para conocer todas las implicaciones de seguridad, aunque de momento las principales preocupaciones giran en torno a los mismos aspectos[12].

- **Datos de sensores**: los dispositivos *wearables*, quizás por su novedad, se consideran de bajo riesgo por los usuarios. Sin embargo, como ya se ha comentado, pueden recoger datos de salud y de localización, con las implicaciones de seguridad conocidas.

- **Datos de dispositivos**: no existe tampoco gran preocupación sobre los datos de los dispositivos que los usuarios poseen, pese a que se pueda deducir el nivel adquisitivo a partir de estos, o a que los ciberdelincuentes puedan aprovechar este conocimiento para realizar un ataque posterior contra nuestra intimidad o datos.

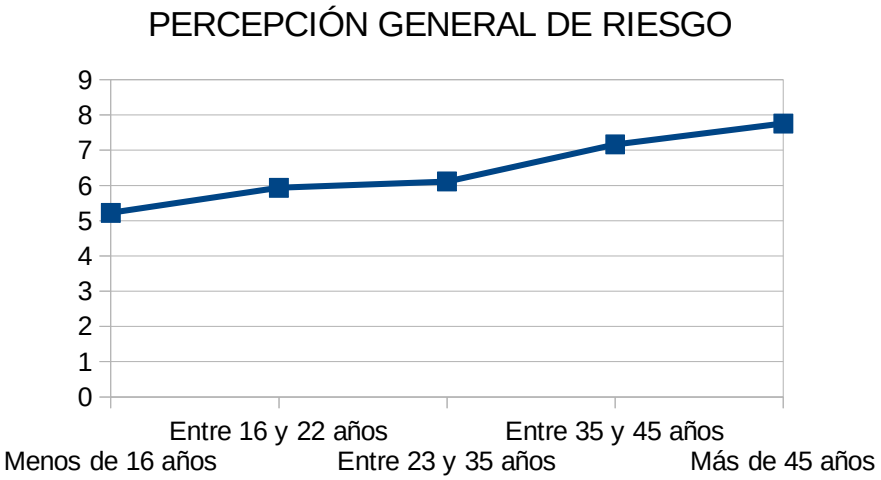
- **Datos de contactos**: un dato que los encuestados aseguran no querer filtrar es el contenido se su red de contactos. Huelga decir que esto no se corresponde con el comportamiento habitualmente observado, ya que generalmente se acepta el acceso a la agenda de cualquier aplicación que se instale en el teléfono y lo solicite, incluso cuando el objeto último de la *app* no tenga ninguna relación con la comunicación entre individuos.

  Nuevamente, el principal uso que se está dando a esta información en la actualidad es de marketing, segmentación de mercado, etc. Conocer datos de contactos significa añadir potenciales clientes a una lista.

- **Datos de comunicaciones**: según lo esperado, los usuarios se muestran muy recelosos de que los contenidos de sus comunicaciones formen parte de su huella digital.

No obstante, se observa una tendencia clara hacia el desplazamiento de métodos de comunicación que podemos llamar tradicionales (telefonía de voz, correo postal) hacia servicios de mensajería. Durante mucho tiempo no ha habido excesiva percepción de riesgo, incluso cuando un servicio dominante como *WhatsApp* no ofrecía cifrado de las comunicaciones extremo a extremo. Respecto a los riesgos de que quede constancia escrita o grabada de todas nuestras comunicaciones, baste con citar un caso mediático como la sentencia por violación de los cinco miembros de "La Manada", donde el contenido de éstas fue determinante[13].

Finalmente, de la muestra estadística analiza, si desagregamos los datos por rangos de edad se aprecia una tendencia que se adivinaba, y es la menor percepción de riesgo o menor valoración de la intimidad a medida que la edad desciende. Los más jóvenes consideran datos especialmente públicos e inocuos su nombre y apellidos, sexo, *email*, gustos, actividades de ocio y viajes, datos de *wearables* y dispositivos, fotos publicadas y, dato llamativo, perciben menor riesgo en que se almacenen comandos y búsquedas de voz que el historial de búsquedas en Internet.



Para finalizar, cabe destacar que los encuestados consideran que les preocupa mucho o bastante su privacidad al usar Internet (72%), pero desconocen las medidas de protección a su alcance o les resultan difíciles de aplicar en casi un 45% de los casos, considerándose solo un 10% de los encuestados como personas bien formadas en materia de seguridad. Esta preocupación se refleja en que más del 70% de los encuestados cree que su reputación digital puede tener repercusiones en su vida personal o privada.

3. GENERACIÓN DE LA HUELLA DIGITAL

Una vez definida la identidad digital como la información acerca de nosotros mismos que dejamos al utilizar Internet, procede ahora estudiar los mecanismos por los que nuestros datos pasan a formar parte de dicha huella digital. Las vías de recolección de información por parte de la red son múltiples, y pueden categorizarse de varias formas, sin embargo podemos establecer una clasificación sencilla distinguiendo entre métodos activos, métodos pasivos, y datos proporcionados por terceros.

Atendiendo a esta clasificación, los métodos activos serían todos aquellos en los que el usuario, de manera voluntaria o al menos consentida, en aceptación de las condiciones de uso de un servicio y en general para beneficiarse de él, introduce o entrega ciertos datos personales de forma consciente.

Por el contrario, los métodos pasivos serían el conjunto de técnicas por las que un dispositivo, producto o servicio, recolecta información acerca del usuario de manera inconsciente para este o, por lo menos, no clara y transparente, sin su intervención explícita. Para ello se suelen realizar medidas sobre el uso que el usuario realiza del producto o servicio para inferir determinada información.

Finalmente, los datos proporcionados por terceros son aquellos sobre los que el usuario no tiene control alguno, ya que no participa en su generación de manera activa ni pasiva, sino que son entregados por un agente externo, en numerosas ocasiones sin su conocimiento ni consentimiento. Es por ello que se trata de un mecanismo especialmente peligroso y potencialmente dañino, al poder llevar aparejadas variables como la ignorancia o la mala fe.

3.1. MÉTODOS ACTIVOS

En los métodos activos, que también podemos llamar explícitos o directos, siempre se produce una de las siguientes dos situaciones: o bien al usuario se le requiere determinada información, en general como condición *sine qua non* para poder hacer uso de determinado dispositivo, producto o servicio; o bien el usuario publica de manera voluntaria y por iniciativa propia cierta cantidad de información en la red.

Se puede afirmar que, en principio, el usuario tiene cierto control sobre lo que se publica a través de estos métodos, si bien es cierto que a menudo no tiene más remedio que aceptar la entrega de sus datos como “pago” por el uso del dispositivo, producto o servicio. En cualquier caso, está en su mano negarse a cederlos, o incluso manipularlos para que resulten falsos o inexactos, aunque con ello pueda perder los beneficios esperados.

Veamos a continuación algunos de los métodos activos más habituales.

3.1.1. FORMULARIOS DE REGISTRO

Son numerosos los servicios de Internet que requieren del registro previo del usuario para su adecuado funcionamiento. Pensemos por ejemplo en una plataforma web de comercio electrónico. Para poder proveer del servicio básico que el cliente espera, deberá recoger como mínimo datos como el nombre del usuario, su dirección postal, y datos financieros como un número de tarjeta de crédito. No tiene sentido para el usuario tratar de ocultar o manipular esta información, ya que se arriesga a no recibir el producto deseado o a no ser capaz de realizar el pago. El usuario puede tratar de minimizar los datos entregados, por ejemplo escogiendo métodos de pago como el contrareembolso, sin embargo en general estos mecanismos suponen un coste adicional. Podemos decir, por consiguiente, que el cliente acepta la cesión de una pequeña parcela de su intimidad a cambio de disfrutar del servicio.

Estos datos, que inmediatamente pasan a formar parte de nuestra huella digital, se entiende que deberían estar exclusivamente bajo el control del propietario del producto o servicio. Conviene no obstante leer los términos y condiciones ya que, en especial en caso de grandes empresas con múltiples actividades, puede que se estén cediendo a terceros o a otras secciones del mismo grupo.

Cabe señalar también que, en algunos casos, los datos no se entregan directamente *online* por parte del cliente. Pensemos por ejemplo en el cliente de un banco tradicional. En algún momento ha cedido sus datos a través de algún impreso o formulario físico a la entidad financiera, quien los tiene dentro de su sistema informático cerrado. Si el banco ofrece al cliente la opción de operar a través de la red y este acepta, los datos que ya constaban en la entidad pasarán a ser accesibles a través de un entorno abierto como es Internet. Debemos considerar de todos modos que se trata de una cesión de datos activa, ya que es el usuario en todo momento quien rellena y entrega sus datos, y acepta explícitamente el uso del servicio en la red.

3.1.2. INTERACCIÓN SOCIAL

Tras el registro en una plataforma o red social, durante el que el usuario entrega los datos suficientes para generar su perfil, se produce una fase de uso más o menos activo de dicha red, hasta el momento en que el usuario decida abandonarla o darse de baja.

El tipo de interacción dependerá, por supuesto, de la categoría o tipología de la red social en cuestión. Por ejemplo, en una red de *blogging* como Blogger o *microblogging* como Twitter, el usuario realizará publicaciones más o menos largas sobre el tema deseado, a menudo acompañadas de material multimedia. Por supuesto, de las publicaciones pueden deducirse gustos, intereses u otros datos del usuario, si bien como ya hemos comentado, al tener el usuario el control de lo que publica, estos pueden ser ficticios,

parciales o interesados. En una red social como Instagram, el contenido principal a publicar es de tipo audiovisual, como fotografías y vídeos que pueden ir acompañadas de texto. Facebook permite la publicación tanto de texto como de material audiovisual, y Youtube solamente vídeos y el texto descriptivo que queramos añadir.

Estas plataformas sociales potencian la interacción entre usuarios, de manera que también dejamos una huella digital en forma de comentarios acerca de publicaciones propias o de terceros, *feedback* positivo o negativo a las distintas publicaciones, reenvío de contenido a otros usuarios… También nos animan a crear listas de intereses y gustos, y nos sugieren otros usuarios a los que seguir. En cuanto a esto, también es una huella que dejamos de forma activa, pues al seguir o dejar de seguir a ciertos usuarios, creamos una red de contactos a nuestro alrededor que conforma nuestro círculo próximo.

Todas estas plataformas incluyen controles de privacidad que permiten configurar, en mayor o menor medida, con qué personas queremos compartir nuestras publicaciones. De esta forma, pueden ser desde totalmente públicas y accesibles por cualquiera, hasta privadas para solo aquel círculo de personas que nosotros escojamos. Sin embargo, como ha quedado demostrado más de una vez, una vez publicamos un contenido en línea perdemos el control sobre él, con lo que un tercero puede reenviarlo, descargarlo, o publicarlo en otro lugar sin nuestro consentimiento. Además, en más de una ocasión los usuarios tienen dificultades para establecer unos controles efectivos sobre con quién y de qué manera se comparten sus contenidos, ya sea por desconocimiento o por la complejidad de la plataforma en cuestión.

A día de hoy, además de las plataformas puramente sociales, podemos asegurar que se ha generalizado un modelo de servicios de Internet que impulsa y anima al usuario a adoptar un rol activo. Podemos nombrar numerosos ejemplos, desde la opción de realizar comentarios en periódicos digitales en línea, hasta publicar un contenido en nuestra red social favorita con un solo clic. También aplicaciones cuya finalidad principal no es social sino deportiva, como por ejemplo Strava o Garmin Connect, favorecen la publicación de actividades en línea y la interacción social entre usuarios. O plataformas de contenido multimedia, como Netflix o Spotify, solicitan la valoración del usuario de los contenidos que consume, o la creación y compartición de listas de reproducción. Numerosas aplicaciones móviles reclaman nuestra valoración y calificación *online*, y muchos juegos piden permiso para publicar nuestros logros o el mero hecho de estar jugándolos. Cuando compramos un producto *online*, con mucha frecuencia se nos requiere una opinión sobre el mismo o sobre el proceso de compra. Y el programa Google Local Guides premia a sus usuarios con un sistema de gamificación a base de medallas y logros a cambio de publicar reseñas de todo tipo de lugares, desde restaurantes hasta museos, pasando por centros comerciales. Toda esta actividad deja una huella digital y se genera de forma activa.

3.1.3. COMUNICACIONES

También nuestras comunicaciones en línea dejan una huella digital, como ya hemos comentado. Todo aquello que decidamos escribir, adjuntar o reenviar es añadido de manera activa por nosotros mismos, ya sea a través de un correo electrónico, como mediante una aplicación de mensajería instantánea.

Es razonable pensar que el usuario espera un cierto grado de intimidad en sus comunicaciones, esto es, que solamente sean accesibles por los usuarios a los que van destinadas, de modo similar a lo esperado en el caso del correo postal o de una conversación telefónica. Huelga decir que esto no es estrictamente así, ya que existen numerosas vías para que el contenido de nuestras conversaciones llegue a un destinatario no deseado, empezando por el reenvío intencionado de nuestro mensaje hacia un tercero. También es conocido que, hasta 2017, Google leía el contenido de nuestro correo electrónico[14] para personalizar anuncios en la plataforma Gmail. A día de hoy, todavía se interpreta de alguna manera el contenido de nuestro correo de Google para añadir eventos a nuestro calendario de manera automática[15], tales como reservas de hotel o información de vuelos. Finalmente, una orden judicial o una intervención policial pueden interrumpir, lógicamente, el secreto de nuestras comunicaciones.

3.1.4. OTROS SERVICIOS

Existen numerosos servicios en los que nuestros datos no son un prerrequisito para realizar la prestación, sino que dan sentido último a éste. Dada la tendencia actual a ofrecer servicios en la nube, accesibles desde cualquier dispositivo, en cualquier momento y en cualquier lugar, cada vez encontramos más aplicaciones en Internet que trabajan con nuestros datos.

Dado que el ánimo de este apartado no es realizar un listado exhaustivo, baste nombrar servicios como los de calendario, donde publicamos todo tipo de eventos y recordatorios; aplicaciones de ofimática o de almacenamiento de ficheros en la nube, capaces de trabajar con todo tipo de documentos y susceptibles por consiguiente de guardar datos sensibles; o plataformas orientadas a actividades profesionales como por ejemplo Google Classroom o Moodle para la docencia, donde habrá datos de profesores, alumnos, actividades, notas, etc.

El uso de todas estas aplicaciones –como tantas otras que hemos visto– presupone una relación de confianza entre el prestador de servicios y el cliente, esto es, el usuario confía en que sus datos solo serán accesibles por él mismo, y por otros usuarios que él autorice. Cuando se tratan datos privados y sensibles, la legislación es muy exigente en cuanto a requisitos de seguridad y privacidad para la empresa prestadora; no debe olvidarse sin embargo que, a la postre, estamos dejando nuestra huella digital con el uso de estos servicios.

3.2. MÉTODOS PASIVOS

Los métodos pasivos, implícitos o indirectos, se caracterizan porque recopilan datos del usuario sin su intervención explícita. No significa esto que siempre se haga de manera secreta, a espaldas del usuario, sino que se realiza de manera no intrusiva, tomando datos y estadísticas sin que seamos necesariamente conscientes de qué se está midiendo, cómo o con qué finalidad. Es evidente que seguimos teniendo que realizar una actividad en red, de lo contrario no habría datos que medir, pero generalmente se trata de acciones que no consideramos ligadas a la introducción o publicación de datos, así que la huella digital que dejamos no resulta obvia.

Se puede asegurar por consiguiente que el usuario tiene poco o ningún control sobre la huella digital que deja a través de estos métodos. Algunos mecanismos, como las *cookies*, están fuertemente legislados en territorios como la Unión Europea[16], de forma que el usuario conserva cierta capacidad de gestión de sus preferencias. Por el contrario otros mecanismos como el *browser fingerprinting* o la lectura de metadatos nos exigen un grado de conocimiento y unas técnicas de protección que hacen muy complejo el control de nuestro rastro digital.

Vamos a analizar algunos de los métodos pasivos de recogida de datos más comunes.

3.2.1. COOKIES

Una *cookie* no es más que un fichero, generalmente de muy pequeño tamaño, enviado por un sitio web que estamos visitando, y que se almacena en el contexto de nuestro navegador (a la postre, en algún directorio del disco duro). El objetivo es que el sitio web pueda obtener información sobre la actividad previa que hemos realizado en dicho sitio. Este objetivo no es, en sí mismo, malicioso o malintencionado, por el contrario las *cookies* nacieron con el objeto de ofrecer un mejor servicio de manera sencilla. Así pues, gracias a las *cookies* no es necesario que introduzcamos una y otra vez nuestras preferencias de idioma o nuestra ciudad favorita para conocer la previsión meteorológica cuando volvemos a visitar un sitio web, o podemos conservar los productos de nuestra cesta de la compra *online* incluso entre diferentes sesiones. Se puede simplificar señalando que las *cookies* no implementan otra cosa que mecanismos de persistencia.

Sin embargo, las *cookies* también tienen implicaciones en nuestra privacidad y anonimato. Así pues, además de proveernos de las funcionalidades ya mencionadas, permiten saber con qué frecuencia visitamos un sitio web, por ejemplo. Sin embargo, las *cookies* siempre vienen acompañadas de un contexto, de manera que cuando un sitio web deposita una en nuestro navegador, solamente este mismo sitio web será capaz de leer su contenido posteriormente. Además las *cookies* también pueden ser borradas

por parte del usuario, dificultando el rastreo. Muchas compañías, en su afán por trazar un perfil lo más fiel posible de los potenciales clientes, han buscado mecanismos para eludir estas restricciones: las *zombie cookies*, las *HSTS supercookies* y las *third-party cookies*.

Una *zombie cookie* es una *cookie* que se regenera tras ser borrada, de manera que aunque el usuario desee borrar el rastro de su navegación, esto no es posible. Un ejemplo de funcionamiento es la aplicación JavaScript llamada *Evercookie[17]*, que almacena la información en diferentes lugares como en forma de LSO (Flash *cookies*), valores RGB de un fichero de imagen PNG almacenado en caché a la fuerza, etc., y si detecta que el usuario ha borrado alguno de los tipos de *cookie* lo vuelve a generar, haciendo casi imposible su borrado, salvo que el usuario navegue en modo Privado o Incógnito.

Una *HSTS supercookie*, por el contrario, es capaz de saltarse también este método de protección. Para ello utiliza HSTS[18], un mecanismo que permite a una web indicarle al navegador que se conecte siempre usando HTTPS en lugar de HTTP, mediante unos *flags*. Abusando de estos *flags*, el sitio web puede guardar un identificador único utilizando el valor de Max-AGE, que además podrá ser leído desde cualquier pestaña y desde cualquier navegación, incluso si esta se hace desde otra instancia del navegador en modo Privado o Incógnito [19].

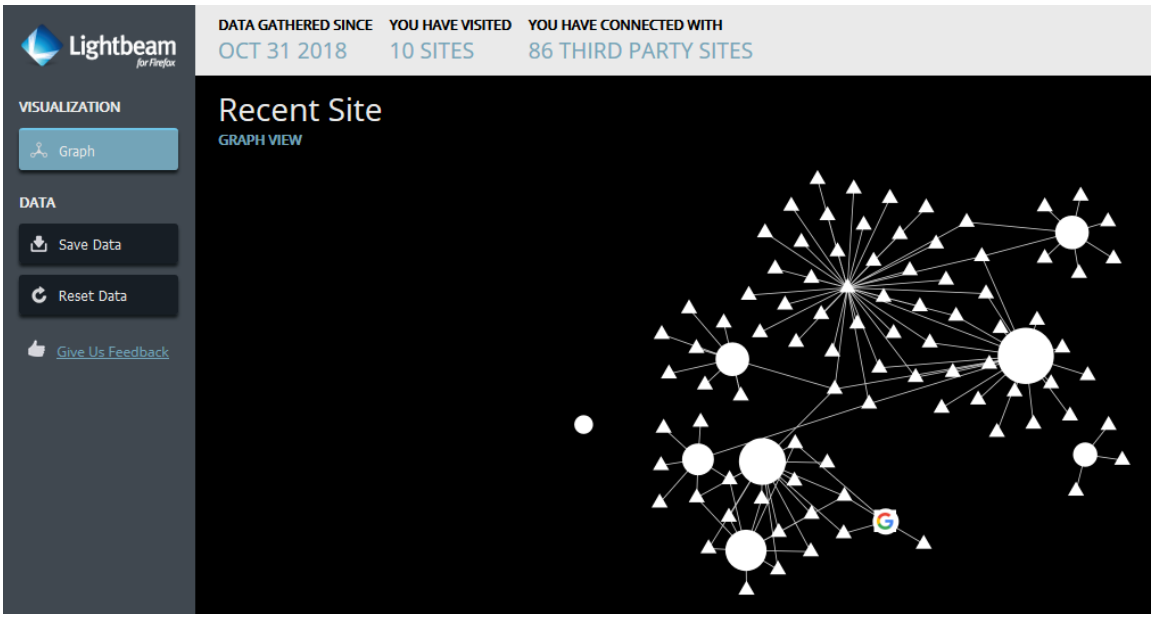
El más antiguo, habitual y sencillo modo de rastreo, sin embargo, es a través de lo que se conoce como *cookies* de terceros, o *third-party cookies*. Aunque las *cookies* solo se envían al servidor que las creó o a otro perteneciente al mismo dominio, una página web puede contener imágenes y otros componentes almacenados en servidores de otros dominios. Las *cookies* que se crean durante las peticiones de estos componentes se llaman *cookies* de terceros[20], y no pertenecen al dominio que el usuario está visitando, sino que normalmente están adscritas al dominio de una empresa de comercio electrónico, marketing, etc. Esta práctica, habitual entre las empresas cuya actividad principal es la publicidad, emplea las *cookies* para realizar un seguimiento de los usuarios a través de distintos sitios web, colocando elementos como imágenes publicitarias y *web bugs* o balizas web, que no son ni siquiera visibles para el usuario. El objetivo es generar un perfil detallado del usuario, no solo a partir de la visita y comportamiento en un determinado lugar web, sino agregando los datos como resultado de múltiples visitas (*cross-site tracking*), pudiendo así dirigir campañas publicitarias más eficaces y rentables.

El funcionamiento es el que sigue: el usuario visita un dominio **A**, por ejemplo dominioA.com. Este dominio contiene contenidos de una empresa publicitaria, alojado en el dominio publicidad.com, y durante la visita a dominioA.com deposita en el navegador una *cookie* perteneciente al dominio publicidad.com. Ahora bien, como la empresa publicitaria tiene alojada publicidad o elementos invisibles en múltiples páginas, cuando el usuario visita

el dominioB.com que también incluye contenido suyo, tiene acceso a la *cookie* anterior, pues pertenece a su dominio publicidad.com y por consiguiente está dentro de su contexto permitido. De esta manera sabe que se trata del mismo usuario, y puede añadir la información para ir construyendo el perfil.

Este modelo de negocio es el que predomina actualmente en Internet. Ya no solo hay páginas que publican contenidos y empresas que proporcionan publicidad, sino que aparece un tercer actor destacado, los agregadores de contenidos o empresas de *tracking*. Las empresas de *tracking* tienen como finalidad analizar y medir cuántas personas visitan una web y de dónde vienen dichas visitas, realizar estudios de mercado, crear perfiles de usuarios o grupos, proporcionar publicidad a dichos sitios, integrar funciones sociales, ofrecer contenidos que requieren un gran ancho de banda como vídeos e imágenes, ofrecer servicios de comentarios como Disqus, o permitir la interacción entre usuarios.

Algunas empresas dedicadas al *tracking* de usuarios con fines de publicidad y que lo hacen a través de *cookies* de terceros son Google (se calcula que Google Analytics por ejemplo rastrea el 45% del tráfico web, y no es el único *tracker* que utiliza Google), Twitter, Facebook, comScore, Cloudflare, Amazon, Yandex, o Adobe*[21]*. Según whotracks.me, un sitio web tiene una media de 8 *trackers*, y cada visita a una página implica de media 12 peticiones cuya única finalidad es la de rastrearnos. En una prueba sencilla real, un complemento de Firefox como Lightbeam nos muestra como, con solamente 10 peticiones a sitios web, 86 sitios de terceros han intervenido para recolectar información, algunos de ellos en más de un sitio, pudiendo por consiguiente rastrearnos entre dominios.



Otro proyecto relacionado es trackography.org, que nos muestra de manera interactiva los *trackers* incluidos en numerosos medios de comunicación, así como qué tipo de información recopilan sobre nosotros.

3.2.2. BROWSER FINGERPRINTING

Ante la creciente preocupación por lo que a la invasión de la privacidad se refiere, numerosos estados y gobiernos han establecido regulaciones en sus territorios sobre el uso de *cookies*, intentando que el usuario esté plenamente informado y siempre tenga la última palabra para aceptar o rechazar estos ficheros de información, en especial en el caso de *cookies* de terceros. Sin embargo, las *cookies* no son el único mecanismo para rastrear la huella de un usuario a través de diferentes sitios web. Otro método muy utilizado es el *browser fingerprinting*.

Las técnicas de *browser fingerprinting* tienen como finalidad identificar usuarios o dispositivos, y rastrearlos a través de diferentes sitios web, incluso si las *cookies* se han desactivado. Para ello crean una "huella" del cliente, que servirá como identificador único, a partir de los datos que logre averiguar del navegador y del dispositivo que está realizando la navegación. El secreto reside en que la cantidad de información que se puede recolectar es tan grande, y sus valores tan variados, que la identificación unívoca del usuario es posible en un 99% de los casos[22]. Esta información incluye, entre otros datos[23], la resolución de la pantalla, el sistema operativo, navegador utilizado, zona horaria, localización, configuración regional, *plug-ins* instalados, CPU, GPU, información sobre la batería, ISP… Efectuando un test en la máquina en la que se está escribiendo este trabajo, pese a que el navegador está configurado para bloquear anuncios y *trackers* invisibles, el equipo tiene una huella única entre 2 millones y medio de dispositivos testeados en el último mes y medio[24].

Una de las características más interesantes de esta técnica es que es resistente a muchas de las recomendaciones de seguridad básicas, como enmascarar la dirección IP a través de una VPN, o eliminar y bloquear las *cookies*. Para ello, el *browser fingerprinting* obtiene información de nuestro dispositivo usando una variedad de técnicas. La primera fuente de información es el *User Agent[25]*, una cabecera empleada en el protocolo HTTP que proporciona información sobre el navegador, sistema operativo, versión del software, etc. del cliente al servidor web destino, con ánimo de ofrecer una experiencia de navegación adaptada y óptima. Para otra gran cantidad de información se usa JavaScript, que por desgracia se ha hecho ubicuo y casi indispensable en Internet, quien se encarga de leer propiedades como *language*, *doNotTrack*, *width*, *height*, *colorDepth*, *hardwareConcurrency*, *timeZone*, *platform*, etc. de diferentes objetos accesibles en el contexto del navegador. La ubicación casi exacta se puede obtener empleando la Google GeoLocation API. Aparte de la lista de *plug-ins* instalados en el navegador, otros proveedores de un gran volumen de información son la lista de fuentes instaladas en el sistema, y el resultado de realizar un *canvas fingerprinting[26]* y un *WebGL fingerprinting[27]*. Ambas técnicas usan funciones de renderizado para dibujar gráficos en el navegador en tiempo real. Las diferencias entre

dibujos en distintos dispositivos, aunque sutiles, son suficientes para generar una función criptográfica *hash* única que identifica al usuario. Estas diferencias dependen del navegador, tarjeta gráfica, *driver* y fuentes instalados…

De forma similar, se pueden usar propiedades de la tarjeta de sonido para añadir información al *browser fingerprinting*, usando el API AudioContext. A esta técnica se la conoce como *AudioContext fingerprinting[28]*.

3.2.3. SESSION REPLAY SCRIPTING

Un *script* de *session replay* tiene como finalidad reproducir el comportamiento de un usuario en una web o aplicación. Para ello se puede grabar su pantalla, sus *inputs* de teclado y ratón, y *logs* de eventos de red o consola[29]. Sobre el papel, el objetivo es mejorar la experiencia de usuario, identificando obstáculos que el cliente encuentra al utilizar el servicio. También se puede usar para estudiar la usabilidad de un sitio web, el comportamiento de los clientes, o detectar conductas fraudulentas.

El riesgo de estas prácticas es que graban toda nuestra actividad sin discriminación alguna, lo que puede incluir información sensible como credenciales de *login*, números de tarjeta de crédito o información relativa a nuestra salud. De hecho, no es necesario ni siquiera que pulsemos el botón de envío de un formulario para que la información se grabe, almacene y procese. Se trata, por tanto, de una huella que dejamos de forma inconsciente y sin nuestro consentimiento, y que muchos sitios web envían a servidores de terceros para su análisis. Para abundar más aún en el problema, a veces este envío a terceros se realiza sobre conexiones HTTP no cifradas, con lo que la comunicación podría ser espiada con relativa facilidad. Y, según un estudio, no es infrecuente que dicha información se almacene sin anonimizar, enlazada a nuestra identidad real[30]. La lista de sitios que realizan prácticas de *session recording* es elevada[31].

3.2.4. CONTROL DE UBICACIÓN

Nuestros dispositivos están constantemente diciéndole a terceros dónde estamos. Esto es especialmente cierto en el caso de los *smartphones*, que van a donde quiera que nosotros vayamos, almacenando nuestra ubicación incluso aunque no estemos conectados a Internet, pero no es menos cierto para aparatos como ordenadores o *tablets*.

Además de los datos de ubicación que compartimos de forma activa, existen múltiples mecanismos a través de los que se puede extraer información acerca de nuestra ubicación, como la dirección IP desde la que nos conectamos o las ya mencionadas APIs de geolocalización. Por otro lado, nuestros *smartphones* se comunican constantemente con las torres de telefonía móvil para poder enviar y recibir llamadas y mensajes. Esta actividad es monitorizada y registrada por nuestro proveedor de servicios, con lo que

puede identificar dónde estamos y dónde hemos estado en todo momento[32]. Obviamente dejar el GPS del teléfono encendido por descuido u omisión es una fuente inagotable de información de geolocalización.

Quizás no es tan evidente el uso que puede darse al historial de conexiones WiFi para rastrear nuestra ubicación. Cuando nuestro dispositivo tiene esta tecnología activada, está continuamente escaneando el espectro en busca de redes conocidas, esto es, redes a las que se ha conectado en el pasado. Para ello realiza un *broadcast* con los nombres de las redes que conoce, información que puede ser vista por cualquier punto de acceso dentro del rango de alcance de nuestro dispositivo. Esta información puede no parecer demasiado relevante, al fin y al cabo el SSID no es más que un nombre que en muchos casos no tiene excesivo significado, y esto podría llegar a ser en parte verdad si no fuera por casos como el de los coches de Google Street View*[33]*. Además de mapear las calles, los coches de Google escanean los nombres SSID de redes WiFi a su alcance, así como su dirección MAC, y los geolocalizan, de manera que cuando nos conectamos a cualquiera de estas redes sabrá inmediatamente dónde nos encontramos sin necesidad de utilizar servicios de GPS.

En la barrera entre la publicación activa y la pasiva, encontramos el etiquetado social que realizamos en algunas publicaciones, como por ejemplo Instagram. Con frecuencia, tras una publicación se nos sugiere adjuntar la ubicación, y también de forma habitual se nos sugiere de manera aproximada qué ubicación puede ser esta. Para realizar esta función, muchas plataformas hacen uso de los metadatos, que analizamos en el siguiente apartado.

3.2.5. METADATOS

Un metadato es un dato que ofrece información sobre otros datos. Un ejemplo de metadatos es la información acerca del autor, el lenguaje de programación empleado o la plataforma de desarrollo de un sitio web; otro ejemplo puede ser el *abstract* y las palabras claves de un *paper* científico, junto con su fecha de publicación; un último ejemplo puede ser la información sobre los cantantes, músicos e intérpretes de cada pista de un CD de música. El objetivo de estos metadatos es ayudar a los usuarios a encontrar información relevante y descubrir recursos relacionados, así como ayudar en el proceso de organización, identificación y archivo de los recursos digitales.

Por supuesto, los metadatos también pueden ser una fuente de información para nuestra huella digital. Con una herramienta como FOCA[34] se pueden extraer multitud de datos de los ficheros que subimos a la red, datos que ni siquiera sabíamos que estaban allí. Son numerosos los programas, como por ejemplo las *suites* ofimáticas, que añaden metadatos con información personal a nuestros ficheros relativa a nosotros o a nuestros dispositivos.

Otro tipo de ficheros que suelen llevar metadatos aparejados son las fotografías e imágenes. Ejemplos de metadatos típicos en una fotografía tomada desde nuestro *smartphone* son el tamaño de la imagen, la resolución, la fecha de la captura, la velocidad de obturación, el dispositivo desde el que ha sido tomada, y la geolocalización de la fotografía. Así, cuando publicamos una fotografía en una red social por ejemplo, no es difícil deducir que también estamos indicando dónde estábamos, cuándo, y qué dispositivo poseemos.

3.2.6. PERMISOS DE APLICACIONES

Los permisos que damos a nuestras aplicaciones también son fuente de datos para configurar nuestra huella digital. Pese a que el procedimiento de otorgar los permisos es activo, una vez concedidos rara vez se vuelve a consultar al usuario sobre si desea o no compartir determinada información, con lo que durante la mayor parte del ciclo de vida de la aplicación en el dispositivo se estarán compartiendo los datos sin nuestra intervención.

Existe una preocupación con bases fundadas acerca del elevado número de permisos que solicitan ciertas aplicaciones móviles, por ejemplo, muchos de los cuales no estarían justificados por la actividad que se supone que desempeñan. Además, en muchas ocasiones el usuario desconoce el riesgo que corre, y autoriza todos los permisos como el peaje a pagar por el uso de la aplicación. Los peligros son claros, pues incluso si la intencionalidad de los programadores no fuera maliciosa, un error en el código podría dar acceso a gran cantidad de información a una tercera parte con motivaciones más dudosas. En todo caso, como reza la famosa frase, “si algo es gratis, el producto eres tú”. Esto tiene implicaciones en el modelo de negocio actual de las aplicaciones y servicios de Internet, donde podemos distinguir entre:

- Servicios o aplicaciones con **pago por descarga**: el usuario paga una cantidad fija para poder usar el producto a partir de ese momento.
- Servicios o aplicaciones con **pago por suscripción**: el usuario paga una cantidad mensual o anual para poder usar el producto mientras ésta esté activa.
- Servicios o aplicaciones con **compras *in-app***: el usuario parte de un servicio gratuito, y realiza micropagos para conseguir funciones *premium* u otras mejoras.
- Mezcla de alguno de los anteriores: se paga por un servicio básico, y se pueden añadir funciones con compras *in-app*.
- Servicios o aplicaciones con ***in-app advertising***: el usuario no realiza pagos, los ingresos vienen en forma de publicidad dentro de la aplicación o servicio. Abre la puerta a rastrear al usuario entre distintas aplicaciones, puesto que la publicidad puede aparecer en más de una. Es este caso no se utilizan *cookies*, ya que este mecanismo no está

disponible en las *apps* móviles, sino otros indicadores nuevos como el Google Android ID o el IFA de Apple[35].

- Servicios o aplicaciones **gratuitos**: el usuario no paga ni recibe publicidad, de manera que la empresa realiza *branding,* u obtiene bases de datos de usuarios, perfiles, etc. para realizar estudios de mercado; es decir, el usuario paga con sus datos.

Para ver qué datos del usuario puede obtener una aplicación gratuita, veamos por ejemplo algunos permisos del sistema operativo Android[36]:

1. Lectura/escritura en almacenamiento USB externo: la aplicación podría obtener los ficheros que tuviéramos almacenados.
2. Acceso a ubicación: la aplicación será capaz de conocer nuestra posición si el GPS está activado.
3. Permisos sobre las funciones telefónicas: bajo esta categoría encontramos diversos permisos que permitirían conocer nuestro número de teléfono, IMEI, IMSI, números de teléfono de llamadas entrantes, acceso al registro de llamadas… De forma similar, varios permisos permiten controlar el acceso a mensajes de texto SMS.
4. Acceso a los contactos: la aplicación tendría acceso a nuestra lista de contactos, y usar esta información de manera no lícita.
5. Acceso al calendario: se podría leer el contenido del calendario del usuario, con todos sus recordatorios y eventos.
6. Acceso a la cámara: la aplicación podría tomar imágenes y vídeos. Recordemos que, una vez concedido el permiso, las aplicaciones no necesitan volver a solicitarlo para ejercerlo, de manera que el usuario podría ser fotografiado o grabado sin su conocimiento.
7. Acceso al micrófono: al igual que en el apartado anterior, las conversaciones del usuario podrían ser grabadas sin su conocimiento.
8. Acceso a los sensores corporales: la aplicación podría conocer los valores medidos por nuestros dispositivos de monitorización de actividad.
9. Permisos sobre las comunicaciones: la aplicación podría usar, conectar, desconectar, etc. las redes WiFi, redes de datos, Bluetooth, NFC…

Estos son solo algunos de los permisos que podemos conceder a nuestras aplicaciones y que, como puede deducirse, permitirían obtener información del usuario para trazar su perfil, ampliando el contenido de su identidad digital. Otros sistemas operativos tienen esquemas de permisos similares, con problemáticas del mismo orden.

El problema de conceder permisos a aplicaciones y servicios, dando acceso a nuestros datos personales, no se limita al campo de los *smartphones*.

Es bastante habitual que cada vez más servicios y juegos gratuitos (t*hird-party applications*) nos soliciten permisos para conectarse a nuestra cuenta de Facebook o Google, por poner dos ejemplos. Parece razonable pensar que el objetivo principal de esta conexión es realizar minería de datos, como en el reciente caso de Cambridge Analytica que analizaremos más adelante. El reclamo para el usuario suele ser obtener funcionalidades sociales, o simplemente poder utilizar cierto juego, acceder a un cuestionario… A veces incluso se utiliza un elemento como cebo, un supuesto vídeo sorprendente por ejemplo, para forzar a que el usuario interactúe con alguna de sus cuentas o redes, y acceder de esta manera a sus datos.

3.2.7. HISTORIAL DE ACTIVIDAD

Muchos usuarios son conocedores de que su navegador web almacena un historial detallado de sus búsquedas, sitios visitados, fecha de la visita, etc. De estos, una parte conoce la posibilidad de borrar este historial local para aumentar su privacidad y proteger su intimidad, o la opción de no almacenarlo haciendo uso de la navegación Privada o Incógnito. Lo que quizás muy pocos saben es que, independientemente de estas precauciones, los motores de búsqueda y sitios web almacenan un historial *online* con nuestra actividad y visitas, que escapa por completo a nuestro control.

Así pues, no debe sorprendernos que todo servicio *online*, sitio web o aplicación registre nuestra actividad, no solo de manera local, de la cual podríamos tener el control, sino en sus propios servidores, cuyo control queda lejos de nuestro alcance y se limita a la posibilidad de que nuestros datos sean borrados cuando nos demos de baja del servicio… o quizá no. Este historial puede referirse a nuestras búsquedas, nuestra ubicación, nuestras compras en un portal web, nuestras transacciones en una aplicación bancaria, las redes a las que nos hemos conectado, los vídeos reproducidos o las canciones escuchadas, los dispositivos desde los que nos hemos conectado a Internet, las aplicaciones que hemos instalado, con quién nos hemos comunicado y con qué frecuencia… cualquier acción realizada es susceptible de ser almacenada en una base de datos o fichero de *log*, y luego compartida con terceros, procesada para prever nuestro comportamiento futuro, agregada a otros datos para realizar un estudio de mercado, etc.

El interés por conocer nuestra actividad hace que las empresas utilicen todo tipo de herramientas para saber más acerca de nuestro historial, como muchas de las ya mencionadas. Existen muchas más como, por ejemplo, la lectura del atributo *window.getComputedStyle[37]* que, hasta 2010, permitía obtener el valor del selector de CSS *:visited*, descubriendo de esta forma en qué enlaces había hecho clic un usuario dentro de una página web.

Con la avalancha de nuevos dispositivos inteligentes, capaces de grabar nuestros comandos de voz e interpretarlos, dando respuesta a nuestras necesidades; la aparición de los coches autónomos; los *gadgets* que

monitorizan nuestra actividad, etc., nuestro historial de actividad es cada vez más detallado y es recogido desde las fuentes más diversas.

3.2.8. PROMOCIONES

Finalmente, no cabe olvidar un método importado del comercio tradicional, la recogida de datos que tiene como cebo una promoción, descuento, etc. Consideramos que es un método pasivo porque, aunque es el usuario quien de manera activa registra sus datos personales en un formulario o similar, su finalidad última no tiene ninguna relación con la información, sino que es beneficiarse de la participación en un sorteo, la obtención de un cupón de descuento, conseguir los privilegios propios de pertenecer a un determinado club, etc. La empresa, por otra parte, no tiene como finalidad ofrecer ningún beneficio, sino recolectar datos.

En general, el usuario no se molesta en comprobar qué uso se va a realizar de sus datos, ni si van a ser cedidos a terceros para su tratamiento. Incluso datos que, según nuestra encuesta, los ciudadanos y las ciudadanas no desean compartir en la red, son cedidos a través de este procedimiento con asombrosa facilidad.

La tira[38] de XKCD de Randall Munroe ilustra el concepto anterior.

> -Serán 23,03 dólares. Pero te pagaré 24 céntimos por tu apellido, 35 céntimos por una lista de los miembros de tu familia, 79 céntimos por tu número de teléfono, y 1,20 dólares si me pasas tu teléfono y me dejas echar un vistazo a tus publicaciones de Facebook.
>
> ---
>
> *Tas tarjetas de fidelización y las ofertas en cuentas de recompensa se vuelven mucho más raras si piensas en ellas como transacciones individuales.*

3.3. DATOS DE TERCEROS

Además del rastro que dejamos durante el uso de Internet nosotros mismos, de manera activa o pasiva, no debemos olvidar que terceras personas, empresas u organizaciones pueden publicar datos sobre los distintos individuos sin su conocimiento y consentimiento, lo que repercute en su huella digital. Como ya hemos comentado, estos datos escapan en general al control del usuario, por lo que su impacto puede ser impredecible.

Podemos proporcionar múltiples ejemplos, empezado por el uso que hacen nuestros familiares, amigos y conocidos de las redes sociales. Estos pueden realizar publicaciones acerca de nosotros, incluidas las fotografías, e incluso etiquetarnos para facilitar nuestra identificación. Una publicación en Facebook o Twitter puede ser compartida o retuiteada, lo que entra dentro del funcionamiento de la plataforma, o descargada y publicada en otra plataforma distinta, incluso tras alterar el mensaje original.

Otros datos se proporcionan en forma de filtración no intencionada o por descuido. Por ejemplo, al dar acceso a su perfil en una red social a aplicaciones de terceros, alguien podría dar acceso a su lista de contactos, en la que estaríamos incluidos, filtrando parte de nuestros datos. Un descuido muy común se produce cuando alguien realiza un envío de correo electrónico masivo incluyendo nuestra dirección en el campo **para**, lo que la filtra a un gran número de contactos.

En numerosas ocasiones los hackers[39] han publicado largas listas de información confidencial sobre usuarios de servicios de Internet, tras realizar algún ataque exitoso. Pero no es necesario ser víctima de un pirata informático, las administraciones públicas ponen a disposición de los ciudadanos listados con información personal en procesos de selección de personal, etc.

4. USOS DE LA HUELLA DIGITAL

Una vez aclarados los contenidos de la huella digital y la procedencia de los datos que la componen, surge inmediatamente una pregunta, ¿para qué pueden usarse los datos de nuestra identidad digital? Esto abre a su vez otros interrogantes, ¿estamos en riesgo?, ¿todos los usos que se realizan de nuestros datos son legítimos?, ¿podemos protegernos?, ¿cómo?

No todas estas preguntas tienen fácil respuesta. La declaración acerca del uso que una empresa pretende hacer de nuestros datos suele estar recogida en los términos y condiciones que aceptamos al suscribirnos a un producto o servicio. Dichos términos y condiciones son en general muy extensos, están escritos en lenguaje jurídico y contienen numerosos términos técnicos de difícil comprensión, son vagos y demasiado generales y, en resumidas cuentas, no solemos leerlos y, cuando lo hacemos, no los entendemos. Podría decirse que las empresas no animan a leerlos, sino que incluso desincentivan nuestro interés por ellos. Además como ya hemos visto, en numerosas ocasiones se recogen nuestros datos sin contrato o suscripción previa. Algunos de estos supuestos se han regulado, tras la alarma social por prácticas abusivas de empresas del sector, como en el caso de las *cookies*, pero la lectura de los términos en que estas son utilizadas y la configuración de nuestra privacidad en todos y cada uno de los sitios que visitamos, con textos farragosos y procesos lentos y poco claros, parecen una vez más diseñados con la finalidad de desalentar al ciudadano, el cual no desea perder diez minutos diseñando un plan de privacidad para acceder a un sitio en el que quizás no permanezca ni uno. Finalmente, el control del uso y la finalidad del tratamiento de nuestros datos se torna inasumible cuando la empresa encargada de la recogida los cede a un tercero.

Parece claro que, en el uso que se puede realizar de nuestros datos, existe una dimensión económica, una dimensión legal y una dimensión ética. Lo que quizás no es tan evidente es que también puede haber una dimensión personal, en la que la finalidad al proyectar nuestra huella digital puede ser la de conseguir un beneficio como individuos. Atendiendo a estas dos posibilidades, estudiaremos los usos que un tercero puede hacer de nuestros datos, frente al uso que podemos realizar nosotros mismos.

4.1. USO POR TERCEROS

4.1.1. REGULACIÓN EN LA UNIÓN EUROPEA

El uso que terceros pueden realizar de nuestros datos está fuertemente regulado en algunos territorios, especialmente en la Unión Europea. Es importante tener en cuenta este aspecto pues es razonable pensar que los operadores de servicios en Internet no están sometidos a las mismas normas si

están establecidos en estos territorios que si tienen su sede en otras economías emergentes[40]. Para dejarlo claro: no existe ninguna ley de privacidad o de protección de datos que se aplique universalmente alrededor del mundo. Dado que Internet es una red mundial, y que el acceso a productos y servicios de cualquier parte está a un solo clic de distancia, hay que evaluar qué riesgos corremos al ceder nuestros datos en según qué países, menos restrictivos en cuanto al uso que se podrá hacer de ellos. Y, dado que la cesión de datos a terceros está a la orden del día, y la recolección de información se realiza de maneras tan variables y fuera de nuestro control, la precaución es aconsejable. Para habilitar los flujos transfronterizos protegiendo al mismo tiempo la privacidad, algunos grupos de países han suscrito acuerdos, algunos vinculantes y otros que no lo son, como son algunas directrices de la OCDE, el marco de privacidad y el sistema de reglas de privacidad transfronteriza de APEC, los marcos de *safe harbour* entre EEUU y la UE, o algunas normas corporativas vinculantes de la UE para empresas multinacionales. Dicho esto, y con ánimo de centrarnos en nuestro entorno más cercano, hablaremos de la regulación en la Unión Europea.

Antes de empezar, cabe señalar que el concepto de datos personales *online* ha evolucionado desde el comienzo de Internet, debido principalmente a que cada vez hay mayor conciencia del impacto potencial que cualquier información puede tener para tratar a un individuo o grupo de manera diferenciada, incluso si éste no puede ser identificado a nivel convencional de nombre y apellidos[41].

La normativa europea[42] aplica a compañías y organizaciones, públicas y privadas, de la Unión Europea, y aquellas de fuera que ofrecen productos y servicios dentro de su territorio. La normativa de protección de datos, conocida bajo las siglas GDRP, describe los siguientes supuestos como los únicos en los que una compañía u organización puede recoger y utilizar datos personales:

- Cuando existe un contrato entre las partes, es decir con el usuario.
- Para cumplir con una obligación legal.
- Cuando procesar los datos tiene como finalidad proteger un interés vital del usuario.
- Para completar un procedimiento público de una administración como un hospital, una escuela o un ayuntamiento.
- Cuando hay un interés legítimo.

Es en este último punto donde surgen casi todas las dudas, pues no está claro qué se entiende por interés legítimo. ¿Es legítimo que tu banco compruebe tus datos personales para ver si eres un buen candidato para ofrecerte determinado producto o servicio? Es a través de este apartado que las empresas tratan de hacer un uso no siempre ético de nuestra huella digital, bajo el escudo difuso del interés legítimo.

En cualquier caso, en todas las situaciones se debe solicitar al usuario su conformidad con que se recojan y se traten sus datos, lo que se conoce como consentimiento. En su solicitud de consentimiento, la compañía u organización debe informar al usuario de forma clara y comprensible de quién será el responsable de tratar sus datos, con qué finalidad, durante cuánto tiempo, qué derechos tiene el usuario sobre sus datos, y detalles sobre cualquier tercera parte a la que se cederán los datos para su tratamiento.

El consentimiento del usuario es fundamental para que una empresa pueda llevar a cabo la recolección, el uso y la divulgación de datos personales de sus usuarios con ánimo de ofrecer un producto o servicio. La normativa europea pone en manos del usuario la decisión sobre otorgar este consentimiento, pero los ciudadanos casi nunca tienen la información o el nivel de comprensión necesarios para tomar una decisión informada. Además, muchas veces las opciones que se le presentan son binarias, esto es, debe seleccionar "sí" o de lo contrario no obtendrá el servicio. Los usuarios reciben información incompleta sobre las consecuencias de su consentimiento y, cada vez más, se les pide que divulguen datos personales de otras personas, como por ejemplo el contenido de su lista de contactos.

A pesar de estas pegas, gracias al concepto de consentimiento expreso, prácticas como el *browser fingerprinting* son ilegales y, si se detectaran y fueran denunciadas, los proveedores de servicios implicados podrían ser fuertemente sancionados.

4.1.2. USO PUBLICITARIO

Es de sobra conocido que el principal uso que se hace de nuestros datos es con fines de estudio de mercado. El principal motor económico de Internet, y de la sociedad de la información por extensión, son los ingresos publicitarios. Al igual que sucede en el mundo *offline*, las empresas tratan de maximizar sus beneficios, y un método de lograr esto es dirigir su publicidad a colectivos de personas específicos, más susceptibles de estar interesados en sus productos. En el lenguaje de las agencias de medios, a este colectivo potencialmente interesado en un producto se le conoce como *target* comercial. Así, es más fácil en televisión ver anuncios de automóviles a determinadas horas y tras cierto programa o película, mientras que los anuncios de juguetes suceden en otro horario y tras un bloque de dibujos animados, por ejemplo. Así, la parrilla televisiva se divide en franjas, según el tipo de espectador que se espera que esté viendo los programas emitidos en ciertas cadenas a determinadas horas. Igualmente, no son iguales los anuncios que podemos encontrar en las páginas salmón de un periódico que en un suplemento de motor, o en la sección de deportes.

Internet abre de repente un mundo casi infinito de posibilidades, pues además de los mecanismos tradicionales de segmentación según los contenidos, se puede estudiar a cada usuario de manera individualizada y

dirigir publicidad a pequeños grupos de características muy concretas, que se piensa que serán más propensos a adquirir determinado producto. Incluso la forma en que se ofrece el anuncio puede adaptarse a la audiencia, conociendo sus preferencias. Esto es lo que se conoce como *microtargeting* o microfocalización. El *big data* ha contribuido a esta tendencia, pues la posibilidad de recoger ingentes cantidades de datos de los usuarios, planificar campañas focalizadas en pequeños grupos, y medir el *feedback* resultante en forma de incremento de ventas, ha dado un poder nunca visto a las agencias de medios y publicidad.

Evidentemente todo esto tiene unas consecuencias económicas. Las grandes empresas van a intentar por todos los medios conocer a los ciudadanos: sus gustos, tendencias, expectativas de respuesta a ciertos estímulos de compra… Recogidos los datos, agregados y medidos, diseñarán mecanismos de captación de compradores más focalizados y eficientes, reduciendo costes y maximizando el beneficio. A cambio, el usuario no recibirá nada, o como mucho obtendrá un servicio gratuito cuyo valor es inferior al beneficio logrado por la empresa. Se abren aquí cuestiones éticas, si bien organizaciones y compañías suelen actuar conforme a la legalidad.

Sin embargo no hay que olvidar que en multitud de escenarios el usuario no es consciente de que se están captando sus datos, o si lo sabe desconoce de qué datos se trata o la finalidad del tratamiento que se va a realizar. En otras situaciones lo sabe, pero tampoco tiene alternativa real a ceder un poco de su intimidad. Y en general no tiene sospecha de que sus datos son a menudo vendidos directamente a terceros, como un bien de consumo cualquiera, desconoce a qué precio se hace esto y con qué fin, y no puede confiar en que la cadena de custodia de su privacidad va a ser robusta y efectiva en todos sus eslabones. No sería la primera vez que datos personales que deben ser custodiados por empresas y servicios fiables, como debería ser un banco[43], son expuestos al público, incluyendo números de cuenta bancaria, contraseñas, números PIN, etc. En algunos casos se trata de empresas multinacionales[44] con millones de euros de inversión en seguridad[45], ¿qué esperar entonces de compañías que compran nuestros datos, pero tienen una infraestructura tecnológica deficiente o una ética cuestionable?

Un posible consuelo es el hecho de que, en muchos casos, estas empresas guardan datos personales anonimizados, esto es, sin vinculación directa con nosotros. A la hora de estudiar un segmento de mercado, realizar mediciones y planificar estrategias, no es infrecuente que importe poco el nombre y los apellidos de los ciudadanos implicados en el estudio, con lo que los datos se desligan de las cualidades puramente identificativas para este propósito. No obstante, a estas alturas tenemos bastante claro que a un usuario se le puede identificar de muchas maneras, y esta anonimización es

tan efectiva como falta de interés se tenga en averiguar quién es el propietario de la huella digital en cuestión.

4.1.3. USO LEGÍTIMO

Gran parte de los usos que se realizan de los datos recogidos en nuestra huella digital son legítimos, puesto que están relacionados con proveer un servicio al usuario que no podría ofrecerse de otro modo. Por ejemplo el acceso a nuestros datos de ubicación es necesario para situarnos en un mapa, sugerirnos la mejor ruta hasta destino, recomendarnos restaurantes según nuestra posición e historial de visitas anterior, o para analizar una actividad deportiva. Los datos de contacto pueden ser absolutamente imprescindibles para ofrecer un servicio, como la compra y la entrega a domicilio de un producto, mientras que los datos sobre nuestros dispositivos pueden permitir detectar y corregir errores de *hardware* o de *software*, o prevenir un robo de credenciales de acceso. Almacenar nuestras comunicaciones puede facilitar su recuperación posterior, y la interpretación del contenido de un correo electrónico por parte de un algoritmo nos puede ahorrar tiempo al proporcionarnos una respuesta automática. Así mismo, muchas aplicaciones recolectan datos sobre nuestros contactos para que podamos comunicarnos con ellos rápidamente. Interpretar nuestras preferencias en un sitio web nos ahorra repetir su configuración en sucesivas visitas, y datos como la IP son simplemente imprescindibles para el funcionamiento de Internet.

El concepto se puede llevar mucho más lejos: bajo el paraguas legislativo del uso legítimo de los datos, multitud de compañías recolectan nuestros datos para luego hacer un uso de ellos de lo más variopinto. Pensemos en nuestro banco *online*. Sabe todos nuestros movimientos, nuestros ingresos, el tipo de compras que realizamos, incluso la casa que poseemos. Si utilizara esta información para ofrecernos información, por ejemplo, de un crédito personal que puede ser de nuestro interés, beneficioso para nuestra situación, ¿se trata de un uso legítimo? ¿Quién sería el principal beneficiado?

Recientemente, Google ha llegado a un acuerdo[46] con la compañía iRobot, especializada en el diseño y fabricación de robots aspiradores. La finalidad del acuerdo es que Google Assistant pueda acceder a los mapas del interior de la vivienda generados por los robots Roomba, lo que a la larga permitiría dar órdenes a los robots desde dispositivos como Google Home. Además, se pretende que otros dispositivos inteligentes del hogar tengan acceso a esta información. En un mundo con cada vez más dispositivos conectados a Internet, la interrelación entre ellos parece un beneficio para el usuario, que quizá podrá cómodamente gestionarlos todos desde un punto centralizado, lo que parece un uso perfectamente legítimo. Sin embargo, una brecha de seguridad en el dispositivo más débil de la cadena puede exponer nuestros datos. Y, aunque Google asegura que no tendrá acceso a dichos

datos, situaciones similares van a ser cada vez más frecuentes, y en cada una de ellas debemos preguntarnos dónde empieza el uso legítimo y donde termina el mismo. Dispositivos como Google Home Hub, Amazon Echo Show o Facebook Portal pretenden hacerse un hueco en el corazón de nuestro hogar, y con la excusa de ofrecer respuestas a nuestros problemas y asistirnos en nuestro día a día, recogen y comparten nuestros datos con terceros[47], los venden al fin y al cabo. La idea es interconectarse además con todos los dispositivos del Internet de las cosas que vayan apareciendo, tejiendo una red de sensores que sabrán todo sobre nuestros hábitos, preferencias y posesiones. Se puede argumentar que hay un uso legítimo en la recogida de datos, para mejorar los servicios que se nos ofrecen, pero a las implicaciones de seguridad se añaden dudas más que razonables sobre otras finalidades que se puedan buscar.

Siguiendo con la misma compañía, Google está desarrollando un algoritmo para alertarnos de los restaurantes en los que es más probable sufrir una intoxicación alimentaria[48], para lo que usa el historial de ubicaciones y el historial de búsquedas. Parece un uso legítimo, dado el servicio de salud pública que ofrece, pero quizás los restaurantes que pudieran ser señalados por error por el algoritmo no estén totalmente de acuerdo. En cualquier caso, es difícil que los usuarios cuyos datos se están usando en el algoritmo, aunque estén debidamente anonimizados, sean remotamente conscientes del uso que se está haciendo de su información personal.

Ya en 2016, la compañía aseguradora Admiral[49] reconoció la intención de usar la información disponible en las cuentas de Facebook de sus asegurados para detectar rasgos asociados a los de un buen conductor, como ser organizado o prudente. El usuario cedería sus datos de forma voluntaria, y solamente con la finalidad de poder beneficiarse de un descuento, nunca para ver el precio de su seguro aumentado. Sin embargo, una vez proporcionada la información, ¿quién garantiza que no se va a penalizar a usuarios evaluados como conductores de riesgo, incluso negándoles un seguro? Como se puede ver, la línea que separa el uso legítimo es muy difusa.

A día de hoy, la posibilidad de que nuestro coche mida todos los parámetros de un viaje realizado, como la velocidad máxima alcanzada o el uso o no del cinturón, abre la puerta a que la policía pueda investigar lo sucedido tras un accidente, o a que una compañía aseguradora se niegue a pagar los daños tras conducción manifiestamente temeraria[50]. Las posibilidades son incontables, y los beneficios sociales de reducir la tasa de accidentes, o que los costes sean asumidos por aquellos con un comportamiento contrario a las normas, son opciones tentadoras que ponen el debate sobre la mesa. Igualmente ya hemos hablado de la medición de datos de pulseras de actividad para ofrecer ventajas en seguros de salud, caso muy similar.

El interés por acceder e interpretar nuestros datos ya existía en el mundo analógico, y noticias como que se quiera dar acceso a la policía a los datos de nuestros dispositivos y servicios, desencriptando las comunicaciones bajo demanda en caso de sospecha de actividad delictiva o amenaza contra la seguridad del estado[51], parece del todo razonable. Un ejemplo reciente es la solicitud de las grabaciones de un dispositivo Amazon Echo[52] con la finalidad de esclarecer un posible caso de homicidio, pese a que Amazon asegura que el dispositivo solo envía comandos de voz al servidor en caso de ser activado con la palabra clave. Otro asunto es que se abra la puerta a abusos o usos ilegítimos, como el control de los ciudadanos por parte de autoridades o gobiernos con el fin de manipular la opinión pública o reprimir ideales contrarios a regímenes autoritarios. No hace falta irse a los extremos, recientemente se ha aprobado en nuestro país la Ley Orgánica de Protección de Datos Personales y garantía de los derechos digitales[53] que, en su polémico Artículo 58 bis, legitima a los partidos políticos para recopilar información de nuestra huella digital y crear perfiles ideológicos de los ciudadanos. Se tratará por lo tanto de una práctica legal y legítima, aunque algunos expertos en derecho digital y de propiedad intelectual discrepan sobre si es ética. En resumen, nunca había habido tantos datos disponibles como en la actualidad, y la tentación de hacer usos ilegítimos es grande, como veremos a continuación.

4.1.4. USO ILEGÍTIMO

Pese a que el concepto de uso legítimo es ambiguo, y a él tratan de acogerse muchas compañías y organizaciones para recoger, acceder, almacenar y tratar nuestros datos, existen situaciones en las que está fuera de toda duda que se ha rebasado cierta frontera, causando escándalo social y la apertura, en algunos casos, de procedimientos judiciales o investigaciones para depurar responsabilidades. Cojamos el caso de la información sobre la raza de un individuo, que en muchas legislaciones es un dato especialmente protegido por el uso discriminatorio que se puede hacer de él. Durante mucho tiempo, Facebook permitió en la configuración de anuncios que los anunciantes discriminaran a su audiencia, dirigiendo su publicidad solamente a determinados grupos étnicos[54]. Por supuesto, Facebook defendió que sus políticas prohíben a los anunciantes usar las opciones de focalización con fines discriminatorios, de acoso o denigrantes, pero al mismo tiempo ponía en sus manos las herramientas necesarias para hacerlo, esto es, el perfil digital de millones de usuarios y opciones de configuración de anuncios discriminatorias. A día de hoy Facebook ha rectificado en algunas de estas políticas[55], pero sirva este primer ejemplo para darnos cuenta de cómo las empresas tecnológicas no siempre miden las repercusiones morales de sus prácticas, operando bajo la óptica de que si algo puede ser útil y es técnicamente viable, se lleva a cabo. Parece ingenuo pensar que una vez creada la herramienta, una vez dado el poder sobre nuestros datos a un tercero, este siempre hará un uso legítimo y ético de estos.

Un caso escandaloso en nuestro país es el la aplicación para móvil de la Liga Profesional de Fútbol[56]. Esta aplicación cuenta con una función opcional a través de la que, en caso de activarla, los usuarios dan permiso para acceder al audio y al sistema de geoposicionamiento del teléfono, así como el nombre de la red de Internet a la que está conectado. La finalidad de esta función es, como reconoce el propio organismo, "obtener información desde qué lugares se ven partidos de equipos que integran LaLiga y detectar posibles utilizaciones que infrinjan los derechos de propiedad intelectual de LaLiga por parte de establecimientos públicos" lo que, a todas luces, no constituye un uso legítimo, pues no trata los datos personales con la finalidad de proveer un servicio al usuario. Aunque se solicita el consentimiento del usuario para utilizar el micrófono (lo contrario sería espionaje), este es un claro ejemplo de que éste no basta para convertir un uso en legítimo. Además, en este caso se pueden vulnerar derechos de terceros, que pueden ser grabados y no han dado su consentimiento. Este caso nos remite, por cierto, al exceso de permisos que concedemos a muchas aplicaciones al ser instaladas en nuestros *smartphones*. Este ejemplo también nos sirve para medir hasta qué punto existe torpeza o ignorancia en muchas empresas en cuanto al tratamiento de datos personales, pues aseguran que éste no se lleva a cabo al no almacenar los audios asociados al nombre y los apellidos del usuario. En cambio, lo asocian a su IP y a un ID único y específico, lo que constituye un dato personal en el mismo grado que el nombre[57]. Estamos ante un nuevo caso de la idea de que, si es técnicamente viable y puede resultarme beneficioso, ¿por qué no hacerlo? Calcular cuántas empresas u organizaciones tienen esta falta de ética resulta aventurado, pero debe poner en alerta al ciudadano acerca de la conveniencia o no de compartir o proteger su identidad digital.

Un caso bastante polémico de uso ilegítimo fue el de Cambridge Analytica*[58]*, en el que la empresa se dedicó a trazar perfiles de usuarios de Facebook para personalizar campañas políticas. Esto no constituye un delito en sí mismo, pero cabe analizar cómo se hicieron con los datos, con qué finalidad y si se trataba de un uso legítimo como marca la legislación o no. En este caso, Cambridge Analytica hizo uso de datos personales recogidos por un profesor de psicología a través de un test de personalidad en formato de aplicación de Facebook, por el que se pagaba entre 2 y 5 dólares a cada usuario. Cabe destacar que, como en muchas otras aplicaciones de la red social, los usuarios también daban permiso para acceder a los datos de sus amigos, además de datos sobre actualizaciones de estado, me gusta, y hasta mensajes privados. El acceso a datos de amigos hizo que, aunque solo 277.000 personas realizaron el test, se recopilara información de 87 millones de usuarios[59]. El profesor traspasó estos datos a Cambridge Analytica, que estaba trabajando en el desarrollo de técnicas que pudieran influir en los votantes. Resulta difícil pensar que, cuando los usuarios de Facebook accedieron a los términos y condiciones para realizar el test del profesor de psicología, supieran que

estaban cediendo sus datos con esta finalidad. También parece dudoso que el acceso y tratamiento de los datos que hizo Cambridge Analytica corresponda a un fin legítimo tal cual lo define la legislación europea, dado que el test original aseguraba tener fines académicos, no comerciales o políticos. De hecho, la propia Facebook señaló que la transferencia de información del profesor Kogan a la empresa violó sus normas sobre datos personales (solo pueden ser usados para propósitos de la misma aplicación y no pueden ser transferidos o vendidos), y el director ejecutivo de la firma Alexander Nix fue suspendido por la junta directiva de Cambridge Analytica. Si se puede extraer alguna lección de este caso es que, independientemente de las condiciones de uso que se acepten, el riesgo de que nuestros datos se usen con fines ilegítimos es real, y debe hacer plantearse a los usuarios qué precio le ponen a su intimidad y a su privacidad. La segunda lección es que no importa cómo configuremos nuestras opciones de privacidad mientras nuestros familiares, amigos y conocidos tengan en su mano el poder de compartir datos de nuestra identidad digital sin nuestro consentimiento o conocimiento.

Si llevamos al extremo los usos ilegítimos que se pueden realizar de los datos de nuestra huella digital, podemos encontrarnos con multitud de escenarios. La ingeniería social trata de hacerse con los datos personales de sujetos con el fin de manipularlos y obtener algo de ellos, ya sea un beneficio económico, la revelación de un secreto, el acceso a un recurso… Internet y nuestra huella digital han simplificado enormemente este paso, pues con un mínimo esfuerzo pueden obtenerse ingentes cantidades de información susceptibles de ser utilizadas con distintas finalidades. Filtraciones de datos sobre números de tarjeta de crédito y código CVV pueden permitir a un cibercriminal la sustracción de dinero de nuestra cuenta. Publicar información personal y familiar, como datos sobre nuestros hijos, que luego pueden a su vez conducir a perfiles sociales de personas de nuestro entorno, pueden facilitar estafas como simulaciones de secuestros. En este y otros tipos de estafa, los delincuentes tratan de averiguar la mayor cantidad posible de información sobre la víctima, para convencerla de la veracidad de la amenaza. Revelar datos sobre nuestra información sentimental, gustos e intereses, nos puede llevar a ser víctimas del *romance scam*. Unas fotos estando de viaje pueden prevenir a un ladrón de que nuestra vivienda, geolocalizada, se encuentra sin vigilancia. Datos sobre nuestras rutinas de ejercicio, combinados con datos sobre la ubicación de nuestra vivienda, pueden facilitar la tarea de depredadores sexuales. Y por descontado, una foto de contenido sexual compartida en un entorno de intimidad, puede ser redistribuida a través de otros canales arruinando nuestra reputación, haciéndonos víctimas de *bullying*, etc. En algunos casos incluso, basta que parte de la información sea verdadera para que se fabrique otra falsa capaz de hacernos daño[60]. Los ejemplos son innumerables y el único patrón reconocible parece ser que, a más información en nuestra huella digital, mayor riesgo.

4.2. USOS PARA EL PROPIETARIO DE LOS DATOS

Por supuesto, el propietario de los datos puede beneficiarse del contenido de su huella digital si realiza un uso razonable e inteligente. Existe un primer beneficio consistente en poder recordar nuestra actividad pasada. Las fotos publicadas en redes sociales pueden constituir un álbum *online* al que acudir para rememorar un momento pasado. El historial de búsqueda de nuestro navegador nos puede facilitar encontrar una página web cuya dirección no recordamos. Nuestro historial de actividades deportivas puede suponer una motivación para entrenar más y mejor.

Estos ejemplos, sin embargo, se quedan en la mera anécdota si pensamos en la posibilidad que tiene el usuario de controlar el contenido de su propia huella digital, proyectando una imagen acorde a sus intereses. Por ejemplo, la aplicación de citas Tinder permite asociar información de otras redes sociales como Instagram o de *streaming* como Spotify. Muchos usuarios utilizan esta capacidad de la aplicación para incrementar su visibilidad, mostrando un perfil perfectamente planificado de actividades, gustos e intereses, con la finalidad de conseguir más *matches*. Así mismo, se sabe que muchas empresas investigan a los candidatos a un puesto de trabajo justo antes o después de la entrevista personal[61] a la caza de rasgos que inclinen la balanza en una u otra dirección. Por ejemplo, comentarios discriminatorios o sexistas, críticas públicas a anteriores empleadores, fotos o vídeos inapropiados o señales de uso de drogas, pueden desincentivar la contratación del candidato, mientras que una imagen positiva y saludable, que pone en relieve rasgos como la amabilidad, la capacidad de diálogo y la reflexión, puede ser el primer paso hacia el acceso a la empresa.

Muchos usuarios utilizan herramientas como el porfolio *online* para dar visibilidad a su trabajo, especialmente en el caso de artistas gráficos, aunque no exclusivamente. En este caso, una selección cuidadosa de los trabajos más representativos puede facilitar el ser contratado para realizar alguna tarea. Los *Youtubers* también buscan dar visibilidad a sus contenidos para generar ingresos publicitarios, para lo que crean una imagen de marca que los define; esto requiere de una alta planificación y de mantener una identidad digital consistente y acorde con la imagen que se desea transmitir. Otros usuarios utilizan redes sociales como Instagram, Facebook o Twitter para ofrecer sus servicios o captar clientes, nuevamente se requiere mantener una huella digital controlada que muestre exactamente lo que el usuario desea proyectar al exterior con la finalidad de maximizar su éxito.

Como podemos ver, nuestra huella digital puede ser el escaparate en el que proyectamos nuestra imagen en Internet y, como todo escaparate, debe mostrar aquello más atractivo, lo que primero queremos que se vea sobre nosotros. De esta manera podemos no solo reducir el riesgo, sino obtener un beneficio.

5. REPUTACIÓN DIGITAL

Existe un concepto íntimamente ligado a la huella o identidad digital, y este no es otro que la reputación digital. Mientras la huella digital solamente hace referencia a los datos que existen en Internet sobre un individuo, sin juicio de valor alguno, la reputación *online* se podría definir como el prestigio o la imagen que ese mismo individuo tiene en Internet, proyectada a partir de la información de su huella. Podríamos decir que la identidad digital es un concepto objetivo, relacionado con datos medibles y observables, mientras que la reputación tiene carácter subjetivo y se basa en la interpretación que se realice de dicha información.

No hace falta remontarse muchos años atrás para encontrar que, en casi todas las charlas sobre los peligros de Internet y las redes sociales, se pusiera el foco en la protección de la privacidad y de la intimidad. Esto, aunque no haya perdido vigencia, se ha relativizado con el tiempo. En la sociedad actual no solamente es harto complicado mantenerse ajeno a la actividad en las redes, sino que puede ser incluso contraproducente. Muchas compañías tecnológicas descartan candidatos a un puesto de trabajo si no tiene presencia *online*. Así pues, en la actualidad el debate está más centrado en las herramientas y estrategias para controlar el contenido de nuestra huella digital y, como consecuencia, conseguir proyectar una reputación *online* positiva.

A día de hoy las búsquedas *online* influencian muchas tomas de decisiones: en qué restaurante comer, dónde alojarnos en nuestras vacaciones, qué compañía de teléfono contratar, o qué modelo de *smartphone* adquirir. Sin embargo no solamente buscamos bienes y servicios, sino que es extraordinariamente común realizar búsquedas sobre personas. Como ya hemos comentado, el responsable de una entrevista de trabajo puede rastrear la web en busca de información sobre los candidatos, pero no es el único caso. También podemos ojear Internet para averiguar cosas sobre una persona que nos gusta, o seguir a algún famoso al que admiramos en una red social. En todos los casos sucede algo revelador, y es que si lo que encontramos no es atractivo y motivador, lo más probable es que perdamos el interés y no dediquemos nuestras energías a averiguar nada más. La conclusión es que para tener una buena reputación *online*, los primeros datos que se obtengan sobre nosotros en una posible búsqueda deben tener estas cualidades, y no hacernos parecer problemáticos, cuestionables o simplemente raros[62]. A la velocidad a la que se mueve la sociedad actual, no es de extrañar que quien busque un empleado, un becario, una pareja, etc. utilice Internet para realizar un juicio de valor y de potencial basado en la primera impresión que produzcan sus datos *online*. Una huella digital descuidada, o la ausencia total de ella, puede hacer que seamos descartados sin darnos siquiera la opción de demostrar nuestros verdaderos valores, creencias, capacidades y

personalidad. En un estudio de 2015 de CareerBuilder sobre 2000 responsables de recursos humanos de empresas en Estados Unidos, más de la mitad reconocieron usar motores de búsqueda y redes sociales para averiguar datos sobre los candidatos a puestos de trabajo, y de estos la mitad confesaron haber descartado a aspirantes a partir de la información consultada[63]. No todo es negativo, una tercera parte contrataron a candidatos a partir de información positiva encontrada, como mostrar una imagen profesional, tener valores en línea con los ideales de la empresa, o demostrar buena capacidad de comunicación o creatividad.

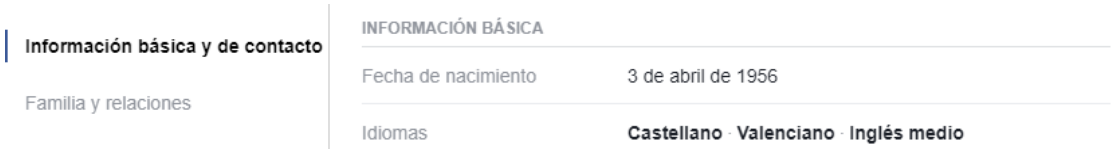
El mismo estudio de CareerBuilder al que hacíamos referencia muestra demás que no tener presencia *online* es tan perjudicial como una reputación negativa, así que eliminar toda nuestra información no parece ser una opción, pues refleja un rechazo a participar del modelo de interacción social imperante en la actualidad. Como veremos en el próximo capítulo, controlar nuestra reputación digital es crucial, y esto implica eliminar contenido cuestionable, publicar contenido positivo, y mantener el control de lo que otros publican sobre nosotros. Cuando trabajamos sobre nuestra huella digital para que proyecte una reputación positiva y, sobre todo, consistente, hablamos de que estamos creando una marca personal.

Con todo esto en mente, vamos a analizar brevemente la huella digital de tres personas para observar las posibles consecuencias de una reputación *online* positiva, negativa y neutra.

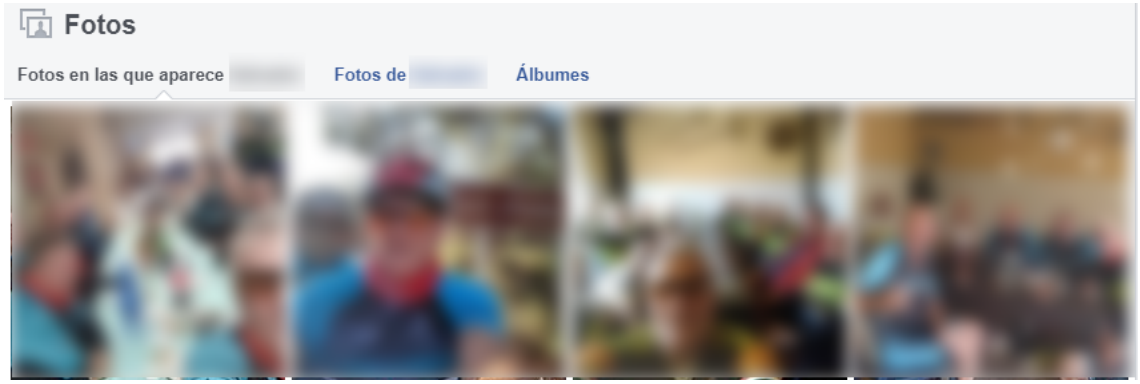
5.1. REPUTACIÓN DIGITAL NEUTRA

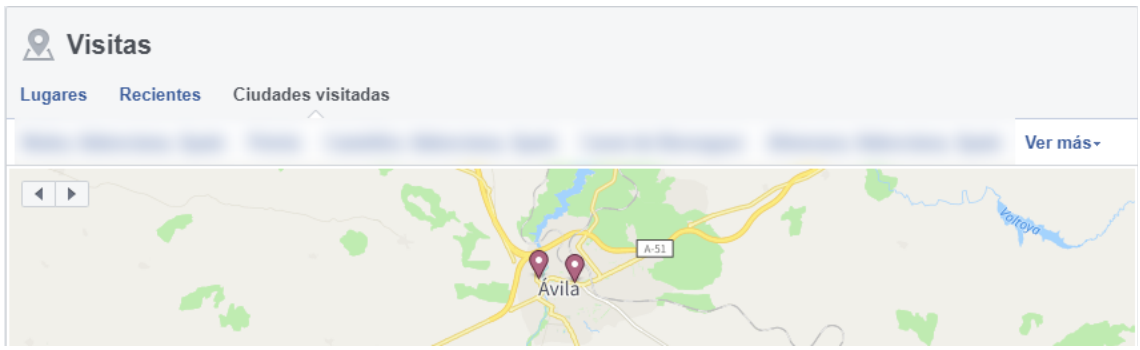
Para la confección de este apartado escogimos un ciudadano al azar, simplemente uniendo un nombre con dos apellidos cualesquiera. Cabe destacar que el experimento podría haber salido mal, haciéndonos escoger otro por falta de datos, pero el resultado fue satisfactorio al primer intento. Aunque toda la información de la que se va a hablar puede encontrarse en Internet, protegeremos la intimidad de este ciudadano al presentar los resultados.

El primer resultado nos llevó a su página de Facebook. Hay que decir que nuestro ciudadano, al que llamaremos con el pseudónimo de Ramón, es muy activo en esta red social, lo que proporciona información sobre su edad, situación sentimental, lugar de residencia, lugar de nacimiento, etc.



También publica numerosas fotos, con lo que se deduce rápidamente una lista de lugares en los que ha estado como Cuba, Asturias, Murcia, Ávila, Soria... También que es aficionado al ciclismo. Y de sus vídeos sabemos que trabaja o ha trabajado en un cuerpo de bomberos.

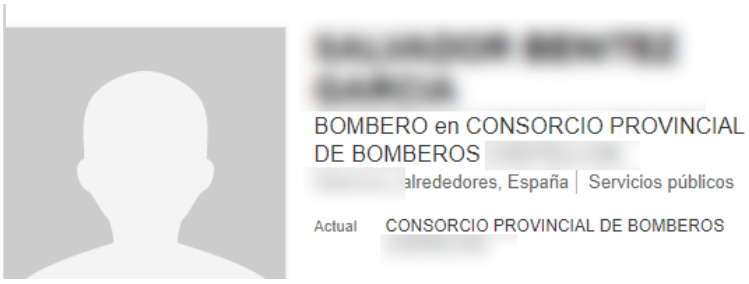




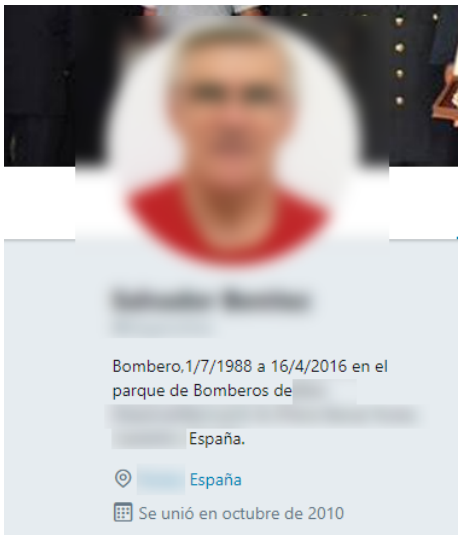
Sus gustos musicales pueden pasar un poco desapercibidos, al igual que sus intereses cinematográficos o literarios, pero su lista de programas favoritos apuntan a tertulias de corte progresista, de manera que podría tener una ideología cercana a la izquierda.

A partir del Facebook de su mujer sabemos que Ramón tiene una hermana, una hija y un hijo, una cuñada y una sobrina, como mínimo.

Una consulta a todos estos perfiles, o al de amigos y conocidos, nos daría mucha más información, pero Ramón también tiene perfil en Youtube que nos confirma que trabajó como bombero, como acredita LinkedIn, aunque está jubilado conforme acredita una página de Facebook.

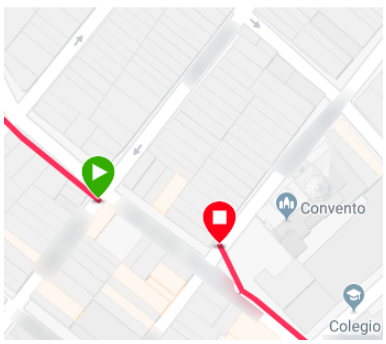


Google+ nos lleva a su página de Edmondo, donde descubrimos que además de ciclismo sale a correr. Podemos ver su actividad en 2017, como la periodicidad de sus salidas y la duración, pero será Twitter la red que nos dará información más relevante.



Esta red social no solo confirma alguna información que ya habíamos encontrado, sino que la amplía:

- Se trata de una persona de ideología de izquierdas, además de activa en redes y plataformas de reivindicación.
- Además de ciclismo y atletismo, practica la natación.
- A finales de 2017, probablemente como regalo de navidad, adquirió un GPS de Garmin, y tiene cuenta en Garmin Connect.



La cuenta de Garmin Connect es fuente de mucha información, pues dados sus patrones de actividad se puede deducir dónde vive. Además publica fotos que lo relacionan con un club ciclista, con su nieto del que obtenemos el nombre, y que podemos confirmar en la cuenta de Instagram de su hijo.

Creemos que no es necesario seguir para entender la idea. Desde el punto de vista de la reputación digital, podríamos decir que Ramón presenta un perfil neutro. La mayor parte de información es de corte personal, y sus gustos e intereses no destacan en especial.

Su actividad deportiva podría ser vista como algo positivo en determinados entornos o situaciones, mientras que su activismo y su perfil ideológico pueden caer de cualquiera de los dos lados de la balanza, según quién consulte el perfil.

La pregunta que subyace en este caso es si Ramón ha pensado realmente en las consecuencias de todas sus publicaciones, y en la gran cantidad de información que está compartiendo. Puede que a nivel de reputación no haya nada que pueda perjudicarlo, pero quizá parte de estos datos podrían suponer una amenaza para él o para algún miembro de su familia. ¿Podría parte de esta información usarse en su contra? Desde nuestro punto de vista, sus datos de ubicación y familiares, especialmente habiendo menores, son datos sensibles y la recomendación sería no compartirlos.

5.2. REPUTACIÓN DIGITAL POSITIVA

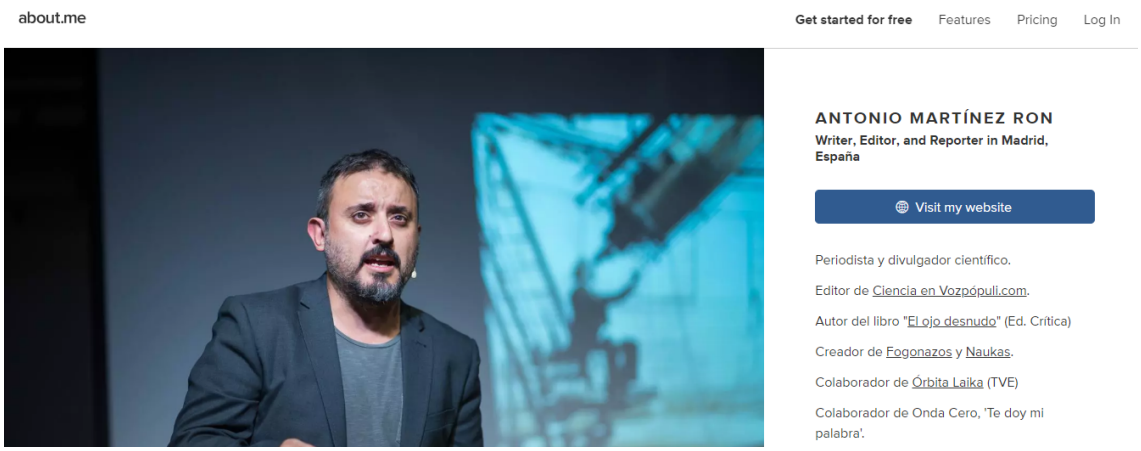
Para analizar un caso de reputación digital positiva hemos escogido al periodista y divulgador científico español Antonio Martínez Ron. Los primeros cinco resultados que devuelve un buscador son:

- Su cuenta de Twitter @aberron, donde tiene 65.000 seguidores y en cuyo perfil se destaca su faceta profesional y se proporciona un enlace a su entrada en la Wikipedia.
- La página «Sobre mí» de su blog de divulgación Fogonazos, donde el propio periodista resume su trayectoria profesional, a modo de currículo *online*, destacando su trabajo como editor de ciencia, su protagonismo en proyectos digitales de divulgación, sus roles en prensa, radio y televisión, obras publicadas y premios.
- La entrada a la Wikipedia. La mayor parte de la información aquí contenida sale de la página anterior que hemos comentado, luego está controlada por el propio periodista, y de artículos de prensa.
- Su perfil como editor de ciencia en el medio de comunicación *online* Vozpópuli.com.
- Su perfil como autor y sus obras en la editorial PlanetadeLibros.

Dejando de lado de momento Twitter, que analizaremos al final, podemos observar que el contenido del escaparate que se muestra está bajo el control de Martínez Ron, que ha escogido proyectar una imagen que resalta su

faceta profesional. Esto se debe a que Antonio es *freelance*, y utiliza Internet como herramienta para mostrar su trabajo y, quizás, conseguir otros. Su objetivo no es mostrar toda su personalidad o su vida íntima, sino aquello que tiene una utilidad para conseguir su fin. También es evidente que es más fácil hacer esto siendo su trabajo muy reconocido y premiado.

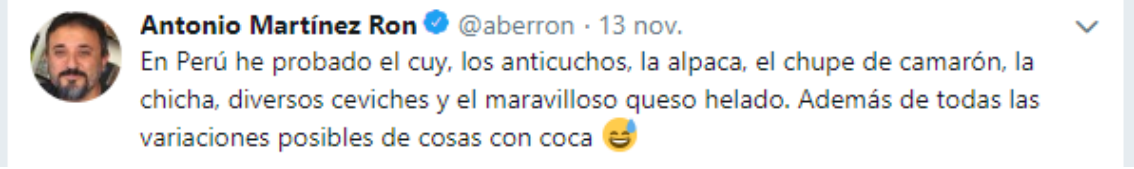
Los siguientes resultados de búsqueda combinan reseñas positivas de sus obras o trabajos, entrevistas en varios medios, artículos escritos por él, y otros proyectos en los que ha estado involucrado como el *podcast* de divulgación Catástrofe Ultravioleta. También encontramos un detallado perfil profesional en Linkedin, y una página a modo de breve currículo en about.me, uno de sus primeros intentos de construir y centralizar una imagen de marca.



El único dato personal que se puede encontrar con cierta facilidad es que tiene una hija llamada Laura, y esto es porque en su último libro consta como coescritora. Incluso en la búsqueda por imágenes encontramos con que se proyecta una imagen profesional (a veces divertida), contemplamos entregas de premios, lo vemos realizando experimentos o acompañado de personas más o menos famosas, pero no lo encontramos en situaciones familiares o personales, y mucho menos en contextos negativos.

En resumen, estamos ante una reputación digital no solamente positiva, sino que persigue un propósito claro y definido, y cuyo contenido está en general bajo el control del propietario.

Hemos dejado para el final la participación de Martínez Ron en Twitter. Es en esta red social donde el periodista se deja llevar un poco más y, aunque mantiene su vida personal y familiar lejos de sus contenidos, y se centra principalmente en divulgar su propio trabajo y noticias de trasfondo científico, tras más de 100.000 tuits hay lugar para el humor o para la opinión personal.



Si bien podemos decir que no parecen ejemplos claramente capaces de dañar una reputación *online*, queda patente que las redes sociales, donde se produce una interacción en un ambiente aparentemente distendido, pueden llevar al usuario a la sensación totalmente equívoca de que se encuentra entre amigos, en un entorno seguro. También puede ocurrir que, ante el *feedback* positivo de miles de lectores, el usuario se vea reforzado en su actitud crítica, humorística o incluso provocadora, sin medir las consecuencias. Se trata probablemente del lugar donde más difícil resulta mantener nuestra reputación *online* bajo control, como veremos al analizar el siguiente caso.

5.3. REPUTACIÓN DIGITAL NEGATIVA

Guillermo Zapata Romero es guionista de televisión, cortometrajista y novelista. Sin embargo, a partir de mayo de 2015, probablemente cualquiera que busque su nombre en Internet obtendrá otra información que poco o nada tiene que ver con dicha trayectoria profesional. Fue uno de los promotores de Movimiento por la Democracia, creado para movilizar a la sociedad y generar una corriente que derivara en un proceso constituyente capaz de recoger los cambios reclamados por la sociedad en el movimiento 15M[64], y uno de los fundadores de Ganemos Madrid, que tras confluir con el partido político Podemos formó Ahora Madrid, que se presentó a las elecciones municipales de la capital de España con Manuela Carmena a la cabeza. Tras alcanzar este partido la alcaldía, Zapata fue nombrado concejal del área de Cultura y Deportes. Y entonces, debido a una huella digital descuidada, su vida se puso del revés.

Si realizamos un ejercicio paralelo al del apartado anterior, las cinco primeras búsquedas de Guillermo Zapata en Internet nos devuelven:

- Su perfil de Twitter @gzapatamadrid, creado en junio de 2015. Esta fecha es importante, como luego comentaremos.
- La entrada a la Wikipedia. Casi toda la información sale de la prensa, en este caso, sin control directo por parte del interesado.
- La página del ayuntamiento de Madrid, donde consta su cargo actual como Concejal, así como las Comisiones Permanentes de las que forma parte, etc.

- Noticias relacionadas con Guillermo Zapata en dos medios de comunicación, El País y La Sexta. Las siguientes tres entradas son de tres medios de comunicación más, y la tendencia continua en las sucesivas páginas del buscador.

Las noticias que encontramos tienen que ver con posibles irregularidades durante su gestión a cargo de la Concejalía, y con procesos judiciales. Nos tenemos que remontar al resultado 26 para encontrar su ficha en el Círculo de Bellas Artes de Madrid. Podríamos pensar que su reputación digital no es más que un fiel reflejo de su mala reputación personal *offline* por casos de corrupción política, pero si nos remontamos a 2015 descubriremos que esto no es exactamente así.

Tras asumir su cargo de Concejal, sus rivales políticos escarbaron en su perfil de Twitter @casiopeaexpres en busca de mensajes comprometedores. Entre algunos que sacaron a la luz[65] se encuentran los siguientes:

Como se puede observar los tuits eran muy anteriores a su toma de posesión en el cargo de Concejal, pero eso poco importa en Internet. Tras salir públicamente a restarle importancia, Zapata borró los tuits polémicos, bloqueó su perfil y terminó por cerrarlo. De ahí que ahora escriba desde otro perfil, @gzapatamadrid, desde el que trata de mantener una imagen puramente profesional, dentro de su ideología política obviamente. Pero borrar un perfil en una red social no es garantía de nada, pues fue denunciado por el delito de humillación a las víctimas del terrorismo. Zapata fue finalmente absuelto un año más tarde[66], pero desde el punto de vista de la reputación digital ya no había nada que hacer a esas alturas, pues había perdido totalmente el control de su imagen *online*. Casi nada de lo que se dice, se publica o se opina en Internet sobre el político está relacionado con la imagen que él desearía proyectar.

El caso Zapata recuerda al caso James Gunn, director, guionista y productor de cine estadounidense que, en un momento que parecía aclamado y reconocido por todos, fue atacado por antiguas publicaciones en la misma red social, Twitter, en la que efectuaba chistes de mal gusto, algunos de contenido pederasta. Hay quien opina que detrás de este intento de descrédito se encuentra el hecho de que Gunn realizaba críticas a la derecha, y en concreto a Donald Trump, a través de las redes sociales[67]. Esto condujo a su despido por parte de Disney, y a que Internet se llenara de información sobre sus tuits de contenido desagradable, actuando como una cámara de amplificación. Lo que podía ser una broma más o menos privada, entendiendo que en Internet no hay nada privado ni efímero, se convirtió en el escaparate de cualquiera que quisiera saber algo sobre Gunn.

Si estas historias esconden alguna lección es justamente esa, que no existe ni privacidad ni verdadero derecho al olvido en Internet, y que aunque no sea fácil hay que medir las consecuencias de lo que se publica. Lo que hoy puede ser un chiste con más o menos gracia, mañana puede ser un arma arrojadiza contra nosotros. Lo mismo puede decirse de una foto con actitudes sexuales, de consumo de sustancias prohibidas, o en estado de embriaguez. Aquello que en el ámbito privado puede ser admisible, en otro contexto (o en ausencia de éste) y a ojos de un entrevistador para un puesto de trabajo, puede ser demoledor. Y lo que es peor, en numerosas ocasiones es lo más dañino y lo más controvertido aquello que más interés despierta, de manera que terminará por ocupar posiciones más relevantes en los resultados de una búsqueda. Zapata y Gunn perdieron el control de su reputación digital y pagaron un precio en su vida personal, pero nos puede pasar a cualquiera de nosotros si no tomamos ciertas precauciones que estudiaremos a continuación.

6. CONTROL DE LA HUELLA DIGITAL

De todo lo anterior es fácil deducir la importancia de controlar nuestra identidad digital, pues de ella se infiere una reputación que puede ser favorable o dañina para nuestros intereses. Además las consecuencias para nuestra intimidad, privacidad y seguridad son múltiples, como hemos razonado, y a menor control de la huella digital mayor riesgo asumiremos.

Sin embargo, gestionar el contenido de nuestra huella digital conlleva tiempo y esfuerzo, y sobre todo realizar una reflexión casi permanente sobre las consecuencias de nuestras decisiones y acciones en línea. Por lo general, en una sociedad acelerada en la que el tiempo es un recurso escaso y nuestras energías limitadas, la inercia nos lleva a aceptar las opciones por defecto que proponen los navegadores web, las aplicaciones y los servicios que utilizamos, y los dispositivos que adquirimos. Estas opciones, por regla general, erosionan nuestra privacidad e incrementan la información de nuestra huella digital de forma ajena a nuestro control. Además, empresas y organizaciones disponen de muchos más recursos tecnológicos, económicos y un tiempo casi ilimitado para intentar captar nuestros datos con la finalidad de obtener beneficios. Se trata por lo tanto de una lucha desigual.

Llegados a este punto, cabe tener en cuenta que no partimos de cero. La mayor parte de nosotros ya llevamos un tiempo utilizando las redes y construyendo una identidad digital, normalmente con escaso control. Por lo tanto no bastará con plantear estrategias para el futuro, es preciso analizar qué rastro hemos dejado en el pasado. Nuestra propuesta para aproximarnos al problema se descompondría en 4 pasos ordenados, a saber:

1. **Investigar** el contenido de nuestra huella digital actual.
2. **Sanear** o limpiar aquellos contenidos que sean perjudiciales o dañinos.
3. Planificar nuestras acciones futuras y configurar nuestros dispositivos, aplicaciones y servicios para **limitar** la información que revelan sobre nosotros.
4. Reflexionar antes de publicar contenidos en la red con la finalidad de potenciar aquel contenido que **proyecte** una reputación digital positiva.

Tactical Tech y Mozilla produjeron un Data Detox Kit[68] para el Glass Room London 2017 con la finalidad de concienciar a la ciudadanía sobre la necesidad de recuperar el control de su identidad digital, con un plan de 8 días durante el cual conocer y limpiar su huella digital para a continuación generar una adaptada a sus necesidades. Tactical Tech es una organización que trabaja en las intersecciones de tecnologías, derechos humanos y libertades civiles y cívicas, fomentando la reflexión sobre seguridad digital, privacidad y ética en relación a los datos. Mozilla es una organización sin ánimo de lucro,

creadora del navegador web Firefox, y que ha expresado en numerosas ocasiones su preocupación por crear un Internet más seguro y accesible.

Según este trabajo, todos podemos encontrar motivos para reducir los datos disponibles sobre nosotros en la red, como evitar ser perseguidos por anuncios personalizados, impedir un posible robo de identidad y la realización de acciones en nuestro nombre, disminuir las posibilidades de que se nos deniegue un crédito o de que se nos cobre más por un seguro, eliminar información textual o en forma de imágenes que nos resulta vergonzosa, o simplemente recuperar el control de nuestra vida digital.

6.1. INVESTIGACIÓN

Lo primero para lograr una identidad digital limpia y acorde a nuestros intereses es conocer qué dice ahora mismo la red cuando le preguntamos por nosotros. Y el primer paso para saber qué dice Internet de nosotros suele ser buscar nuestro nombre en Google. Sin embargo, diferentes personas obtienen diferentes resultados cuando utilizan un motor de búsqueda. Lo ideal es cerrar sesión en todas las cuentas de correo, redes sociales, etc., así como borrar el historial de navegación y *cookies*, así obtendremos los resultados que obtendría un ciudadano anónimo.

Además de escribir nuestro nombre en Google, que nos puede aportar muchos resultados, es conveniente ampliar nuestra búsqueda usando otros motores, como DuckDuckGo, StartPage, Yahoo, Baidu, Yandex o Bing. También es interesante el uso de operadores para refinar las búsquedas, como el operador '–' para ignorar las páginas que contienen un término, las dobles comillas para forzar por ejemplo que ambos apellidos se encuentren juntos y en el orden correcto, o el asterisco como comodín. Si nuestro nombre es muy común se puede añadir a la búsqueda la ciudad donde vivimos o donde nacimos, nuestro lugar de trabajo, nuestra escuela, etc. para que los resultados tengan relación con nosotros. Además de los resultados que ofrezca el buscador, es interesante ver qué imágenes devuelve asociadas a nuestra persona. Y hablando de imágenes, podemos comprobar si tenemos alguna foto de perfil o avatar que pueda estar circulando por la red, incluso asociada a otras personas o situaciones que nada tienen que ver con nosotros, simplemente usando un servicio de búsqueda inversa de imágenes como TinEye[69] o Google Images[70]. También puede ser este un buen momento para establecer una alerta de Google[71] para que nos avise cada vez que nuestro nombre sea publicado en la red.

Si alguien está realizando una investigación concienzuda sobre nosotros, no se detendrá en los resultados que arroje un buscador al escribir nuestro nombre. Si obtiene nuestro perfil en una red social, accederá también al perfil de nuestros contactos, a ver nuestra actividad y a comprobar si hay información sobre nosotros. Por supuesto, anotará toda la información

relevante que encuentre como fecha de nacimiento, correo electrónico, lugar de trabajo y estudios, alias o *nicknames* utilizados… Debemos realizar a partir de esta información nuevas búsquedas, con la finalidad de obtener más datos que a su vez puedan conducir a nueva información, en un proceso de refinado incremental que solo debemos detener cuando ya no obtengamos más resultados relevantes. Además, debemos hacer un ejercicio de memoria y buscar todos aquellos servicios y redes que hemos utilizado en algún momento del pasado, como un blog personal olvidado en algún rincón de la red, o todas las webs de compra o venta *online* en las que nos hayamos registrado.

Durante todo este proceso debemos anotar principalmente dos cosas, aquella información innecesaria que en algún momento compartimos en exceso, pero que no aporta nada de valor a nuestra identidad digital, y aquella información definitivamente perjudicial para nosotros que desearíamos que no constara en Internet. Todo ello será de utilidad en el siguiente paso de nuestro plan.

6.2. SANEAMIENTO

Nuestro siguiente paso debería ser eliminar toda aquella información que hayamos determinado que puede resultar perjudicial o de valor nulo, ya que en este caso puede provocarnos más problemas que beneficios. Podemos establecer una división de dicha información en dos grandes grupos, según si está en nuestras manos su borrado o modificación, o si por el contrario tenemos que acudir a un tercero para solicitarlo.

6.2.1. ELEMENTOS BAJO NUESTRO CONTROL

Puede parecer una obviedad, pero nuestras cuentas antiguas en servicios que ya no utilizamos pueden ser una importante fuente de filtración de datos, origen de información desactualizada, inexacta, o que ya no nos representa; además de un riesgo, ya que alguien podría tomar el control y suplantar nuestra personalidad sin que lo detectáramos, puesto que ya no accedemos con regularidad.

La principal recomendación sería eliminar todas estas cuentas, o desactivarlas si el borrado no es una opción. En el caso de que la cuenta no pueda desactivarse ni eliminarse, podemos modificar nuestros datos con información falsa, que no permita que sea vinculada con nosotros. Es importante eliminar también las cuentas de correo que ya no utilizamos, pues un acceso no autorizado puede revelar cuantiosa información sobre nuestras compras, servicios a los que estamos suscritos, listas de contactos, etc. Así mismo, en cuentas de correo que no deseemos eliminar, puede ser un buen momento para cancelar aquellas subscripciones a listas de correo que ya no nos interesan. Esta cancelación conviene realizarla de manera manual, aunque algunos servicios como unroll.me nos faciliten la tarea, ya que dar acceso a una aplicación de terceros a nuestro correo es una actividad de alto riesgo,

como mínimo de que nuestros datos sean vendidos a terceros[72]. En estas cuentas de correo también debemos eliminar *emails* cuyo contenido pueda ser sensible, como contraseñas o números de cuenta.

En nuestras redes sociales, debemos revisar nuestra información de perfil, nuestras fotos y todas nuestras interacciones, eliminando todo aquel contenido de escaso valor o con connotaciones negativas. También debemos revisar la lista de contactos, aquellas personas a las que damos permiso para acceder a nuestros contenidos con un nivel mayor de privilegios, y eliminar a aquellas personas con las que ya no tenemos relación.

En relación con nuestro historial de actividad, podemos empezar por limpiar el historial de navegación en todos los *browsers* y dispositivos que utilicemos. Si utilizamos Google, este también guarda su propio registro de actividad, así que podemos proceder a su borrado[73], y aprovechar para eliminar nuestro historial de ubicaciones, información sobre nuestros dispositivos, historial de pedidos[74], etc., así como revisar a qué aplicaciones les hemos concedido permisos para acceder a nuestra cuenta[75], siendo conveniente revocar el permiso si ya no utilizamos el servicio. Obviamente esto aplica también a otros servicios como Facebook, Instagram, Twitter… Aunque la eliminación de nuestro historial de actividad no es un proceso sencillo en algunas de estas plataformas y requiere la eliminación manual individual de cada elemento, conviene perder un poco de tiempo borrando al menos aquel contenido que no encaja con la imagen que queremos proyectar. En estas redes, además, es posible que hayamos sido etiquetados o enlazados en publicaciones o imágenes que no nos gustan, es el momento de eliminar estas etiquetas y borrar dicho rastro.

Finalmente, debemos realizar un repaso de las aplicaciones instaladas en nuestro *smartphone*. Es más que probable que hayamos instalado algunas aplicaciones para realizar una tarea concreta, no habiéndolas utilizado nunca más, o que hayamos usado una *app* durante cierto periodo de tiempo, pero ya no lo hagamos. También se puede dar el caso de que hayamos instalado una aplicación solamente para probarla y, al no habernos gustado, la hayamos dejado abandonada en la memoria de nuestro teléfono. En todos estos casos conviene desinstalarlas, ya que muchas de ellas tienen acceso a gran cantidad de información personal a través de permisos excesivos y mal configurados. Si se da la circunstancia de que la aplicación a borrar exigía nuestro registro o alta en el servicio, es recomendable darnos de baja previamente, solicitando la eliminación de nuestros datos si es posible.

No debemos olvidar que en plataformas como Google Play Store tenemos acceso no solamente a las aplicaciones instaladas en nuestro dispositivo, sino a un historial completo de *apps* que hemos tenido instaladas en algún momento en cualquier dispositivo asociado a nuestra cuenta, y que quizás queramos eliminar también.

6.2.2. ELEMENTOS BAJO EL CONTROL DE TERCEROS

Al realizar la búsqueda sobre nuestra huella digital, puede que hayamos encontrado información sobre nosotros publicada por entidades ajenas y que no deseamos que esté disponible en la red. Bajo esta circunstancia se agrupan multitud de casos, desde noticias de prensa hasta información oficial en el Boletín Oficial del Estado, pasando por fotos o comentarios sobre nosotros publicadas por amigos o conocidos nuestros.

Ante esta variedad de casos, se abren también diferentes posibilidades. Si el contenido está en la página de otra persona en una red social, podemos pedirle que lo elimine o denunciarlo en la plataforma, si es que existe esta opción. Si estuviera publicado en un sitio web, podríamos pedirle al administrador que lo quite o lo reemplace. Si no conseguimos que lo borren, podemos pedirle por ejemplo a Google, a través de una solicitud de "derecho al olvido", que omita el contenido en sus resultados de búsqueda[76]. El derecho a solicitar que se eliminen ciertos tipos de contenidos personales de los resultados de búsquedas en línea está restringido a determinados lugares, por ejemplo se puede solicitar en la Unión Europea. Una solicitud de "derecho al olvido" puede incluir información personal confidencial y datos sensibles como nuestro número de cuenta bancaria, una imagen de nuestra firma manuscrita, una foto de un desnudo, o una imagen o vídeo con contenido sexual explícito que está siendo compartido sin nuestro consentimiento. También puede aplicarse a contenido desactualizado que ya se ha borrado de un sitio web, pero que sigue apareciendo en los resultados de búsquedas.

Sin embargo, como casi siempre sucede en el contexto legal, a veces unos derechos pueden entrar en conflicto con otros, como la libertad de prensa. Difícilmente podremos solicitar la retirada de una información periodística veraz, por mucho que dañe a nuestra imagen, y menos si ha habido alguna sentencia condenatoria en nuestra contra. Lo mismo aplica en caso información de deudas que no hayan prescrito, o con datos publicados en boletines oficiales como el BOE relativos a procedimientos públicos en los que nos encontremos implicados.

6.3. LIMITACIÓN

Aunque debemos repetir los procesos de investigación y saneamiento en el futuro, es perentorio limitar la cantidad de información que volcamos en nuestra huella digital, puesto que es mucho más complicado mantener el control de una identidad amplia y dispersa que de una huella acotada y perfectamente definida. Los mecanismos son varios, y están íntimamente relacionados con las maneras en que la huella digital se constituye, y que ya hemos estudiado. A continuación veremos algunas estrategias para reducir la cantidad de información que filtramos a la red.

6.3.1. ANÁLISIS DE CONSECUENCIAS

Quizás el gesto más importante y básico, pero al que prestamos menos atención, consista en evaluar las consecuencias de cada acción que realizamos en la red. La recomendación, obvia por otra parte, sería tomarnos un tiempo antes de publicar un comentario, subir una fotografía, o retuitear un mensaje. Si podemos arrepentirnos en el futuro, si no quisiérmos que algunas personas tuvieran acceso a la información, o si simplemente estamos compartiendo más información de la necesaria, lo más sensato siempre es no publicar.

Otra cosa que debemos hacer es convertirnos en ciudadanos informados, al día de las consecuencias y de los riesgos de nuestras decisiones en línea. Por desgracia, no siempre tenemos tiempo para conocer el alcance de nuestras acciones. Por ejemplo, leer los términos y condiciones de un servicio para decidir si merece la pena su uso puede llevarnos horas[77], sin tan siquiera garantía alguna de comprenderlos por completo. Así pues, ante la duda se debe imponer una postura conservadora, y no dar un dato por mucho que se nos ofrezca una y otra vez la posibilidad de hacerlo.

6.3.2. DIVERSIFICACIÓN

Una segunda estrategia debería consistir en diversificar nuestra presencia *online*, de manera que no utilicemos las mismas cuentas y servicios para realizar distintas actividades. En resumen trataremos de compartimentar, utilizando distintos perfiles para afrontar distintas facetas de nuestra vida, aislando especialmente nuestra esfera íntima y personal cuya proyección al exterior deberíamos reducir a su mínima expresión.

Un ejemplo de esto sería no utilizar nuestra dirección de correo del trabajo para actividades de índole personal, donde además amigos y conocidos con escasa formación tecnológica pueden filtrar nuestro *email* a través de listas de reenvío incontroladas. Por supuesto si realizamos compras *online*, es conveniente contratar una tarjeta de débito con este propósito, a ser posible de tipo recarga, de forma que minimicemos el riesgo de que se filtren los datos de ésta en algún comercio poco seguro. Si nos gusta bromear y hacer chistes en Internet, no es necesario que nuestra cuenta de Twitter esté asociada a nuestro nombre ni a un correo que se pueda relacionar con nosotros fácilmente, podemos actuar bajo un pseudónimo.

Otra buena práctica, como aconseja el dicho popular, es no poner todos los huevos en la misma cesta. Google es un gigante tecnológico omnipresente en la vida de casi todos nosotros, pero no es necesario utilizar todos sus servicios. Existen alternativas de correo enfocadas a la privacidad como Protonmail[78], buscadores alternativos que no rastrean nuestra actividad como DuckDuckGo[79], etc.

6.3.3. CONFIGURACIÓN DE SERVICIOS Y DISPOSITIVOS

Aceptar la configuración por defecto de las aplicaciones, productos y servicios que empleamos es garantía de poner más información de la estrictamente necesaria en manos de terceros. El nivel de configuración de nuestra privacidad en estas aplicaciones varía enormemente de unas a otras, pero en cualquier caso debemos dedicar algunos minutos a explorar nuestras opciones.

Los controles de actividad de Google[80], por ejemplo, permiten decidir si deseamos que la empresa almacene el historial de navegación y actividad, el historial de ubicaciones, información sobre contactos, calendario, aplicaciones; también si deseamos permitir la grabación de voz y audio, o que se almacene el historial de búsquedas y reproducción de servicios como Youtube.

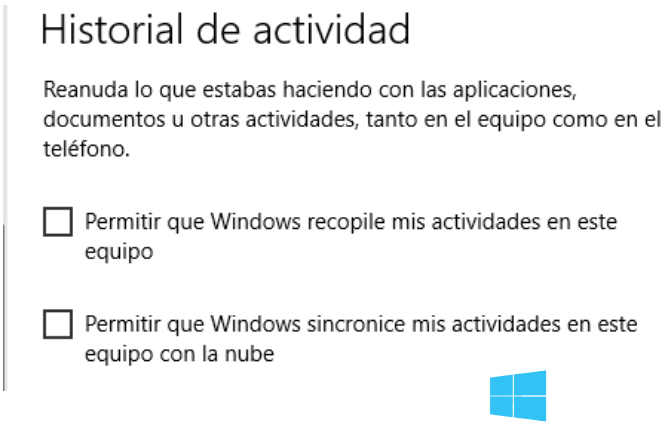
Redes sociales como Facebook permiten configurar a quién damos acceso a qué contenidos de los que compartimos. Sin embargo, a veces esto puede volverse en nuestra contra, creando un efecto de falsa seguridad que nos haga creer que nos encontramos en un entorno privado e íntimo, y que lo que compartimos solo podrá ser visto por un reducido grupo de personas cercanas.

La realidad es que normalmente en nuestros contactos en redes sociales también hay personas no tan íntimas, y que siempre hemos de seguir la máxima de que, una vez subido un contenido a la red, nuestra acción no tiene marcha atrás. Un conocido puede descargar una foto nuestra, subirla a otro

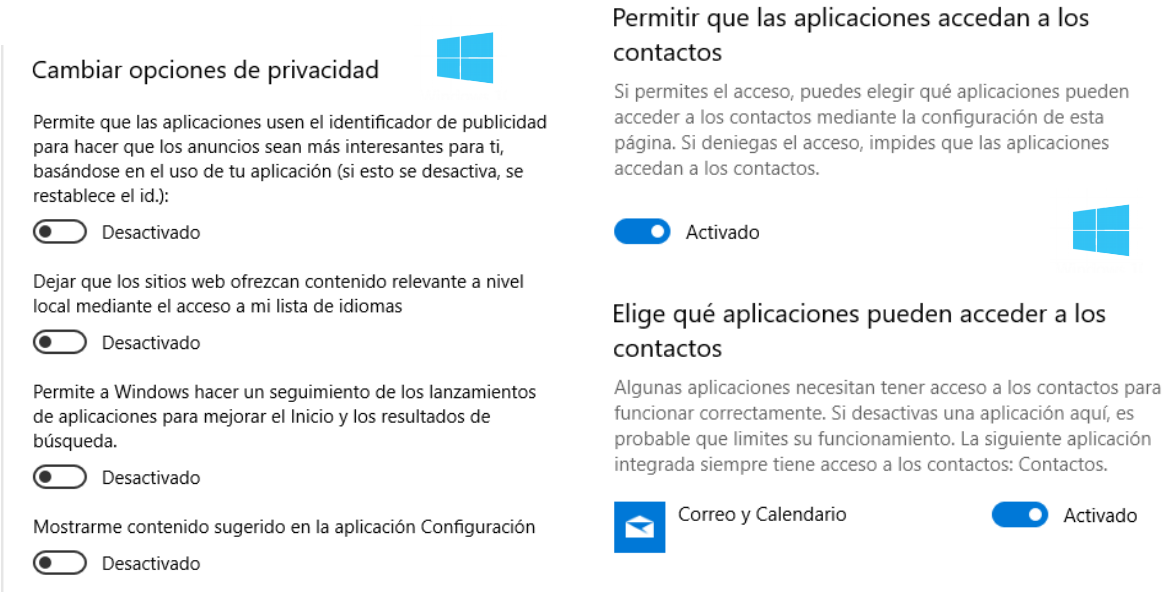
servicio, hacerla pública… Así pues, incluso configurando correctamente la privacidad de nuestras cuentas, debe imperar el principio de precaución y sentido común, y no compartir más de lo estrictamente necesario.

Otra opción interesante en la mayor parte de estas redes es limitar el contenido que nuestros contactos pueden publicar en nuestro perfil, o si pueden etiquetarnos en sus fotografías. Por muy apetecible que sea la interacción social resultante de estas acciones, lo cierto es que si abrimos la publicación de contenidos a terceros perdemos el control de nuestra huella digital, y por consiguiente damos opción a que se pueda dañar nuestra reputación.

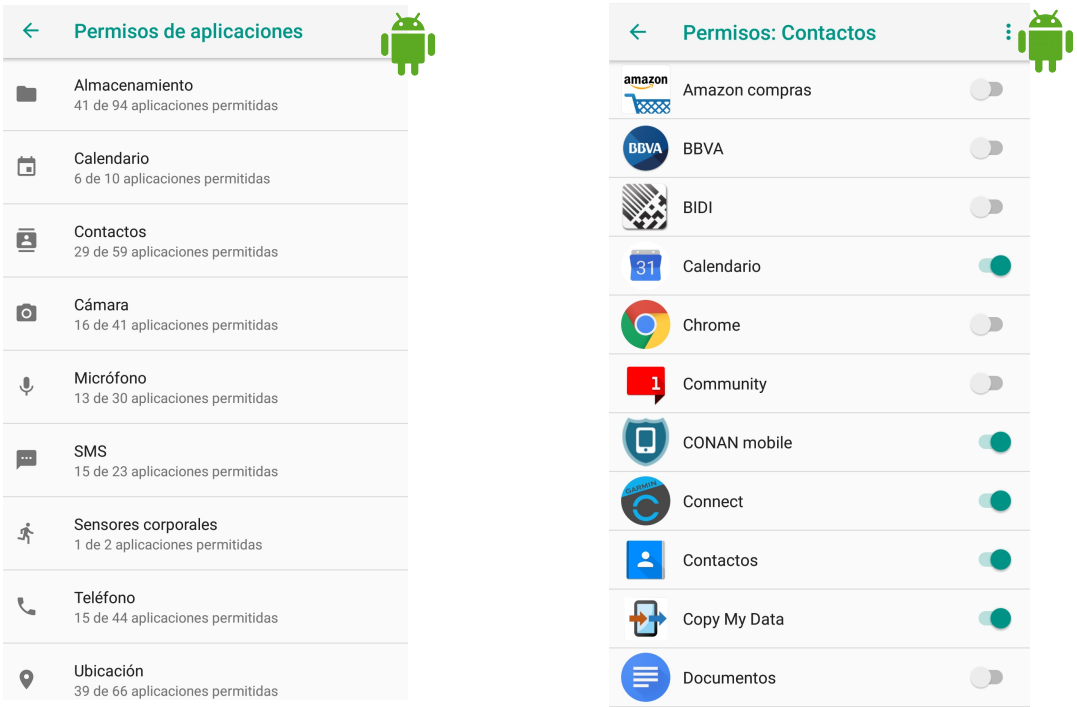
La máxima nuevamente debería ser la precaución, no permitir la publicación en nuestro perfil a terceros, y deshabilitar el etiquetado o al menos programar la aplicación para que nos mande un aviso en caso de que esto suceda, para que podamos revisar si el contenido es apropiado.



La configuración no debe limitarse a servicios en línea y redes sociales. Sistemas operativos como Windows 10 son grandes recolectores de información personal que, si no especificamos lo contrario, termina siendo subida a la nube, con la correspondiente pérdida de privacidad e incrementando nuestra huella digital de manera incontrolada. Algunos ejemplos de esto son el historial de actividades, el registro de aplicaciones utilizadas, el acceso a nuestra lista de contactos, etc.



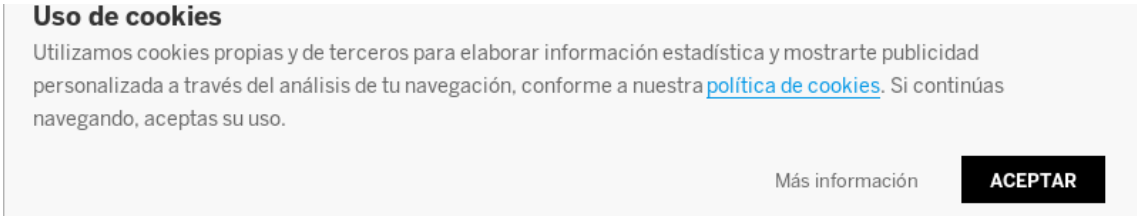
También nuestros dispositivos deben ser configurados. No es necesario que el Bluetooth de nuestro teléfono dé información sobre nosotros o sobre nuestro modelo de *smartphone*, al igual que nuestro punto de acceso WiFi no debería tener nuestro nombre o ninguna otra información personal. Y hablando de nuestro teléfono móvil, debemos repasar los permisos que concedemos a las aplicaciones. En el sistema operativo Android, por ejemplo, podemos hacerlo desde el menú de Configuración, opción Aplicaciones.



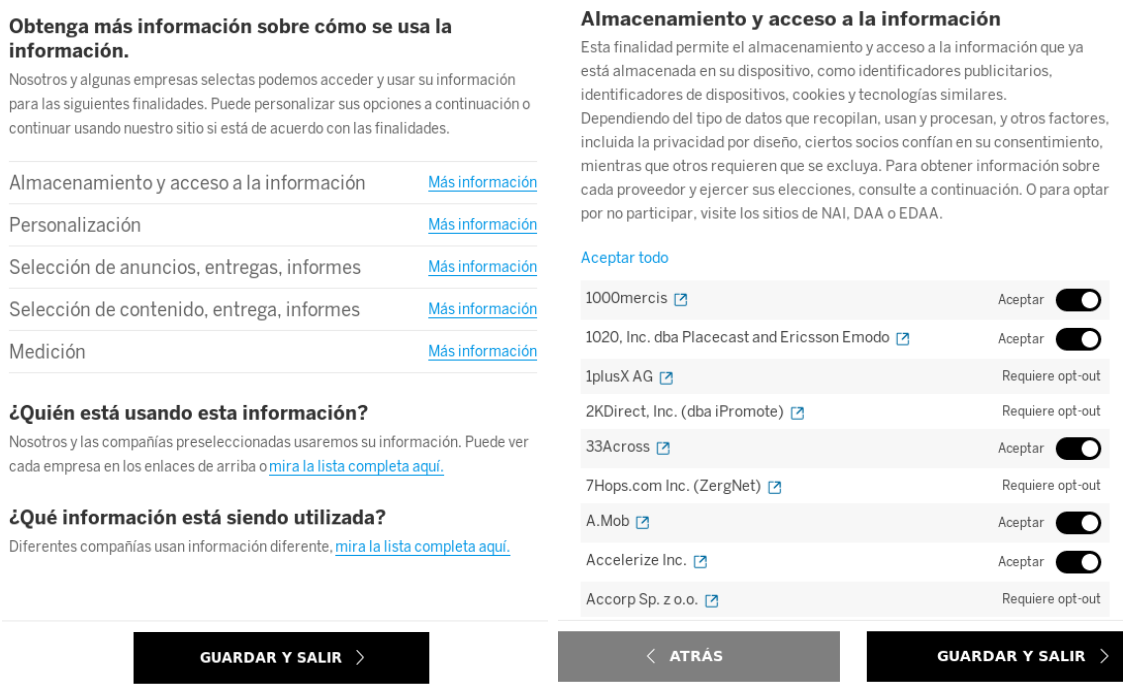
Para cada permiso podemos desplegar la lista de aplicaciones que lo han solicitado, y modificar la opción por defecto, denegándolo o concediéndolo, teniendo en cuenta eso sí que algunas aplicaciones pueden no funcionar si no se les da acceso a determinados recursos.

6.3.4. NAVEGACIÓN SEGURA

Como ya hemos visto, el control de *cookies*, especialmente las de terceros, forma parte de nuestra estrategia para prevenir el rastreo y disminuir nuestra huella digital, manteniendo el control de nuestra identidad en línea. La regulación europea obliga a los prestadores de servicios de la sociedad de la información a comunicar al usuario la utilización de *cookies*, su finalidad, así como a dar la opción de rechazarlas. Sin embargo, analicemos un caso concreto, por ejemplo la web del periódico El País, del grupo Prisa.



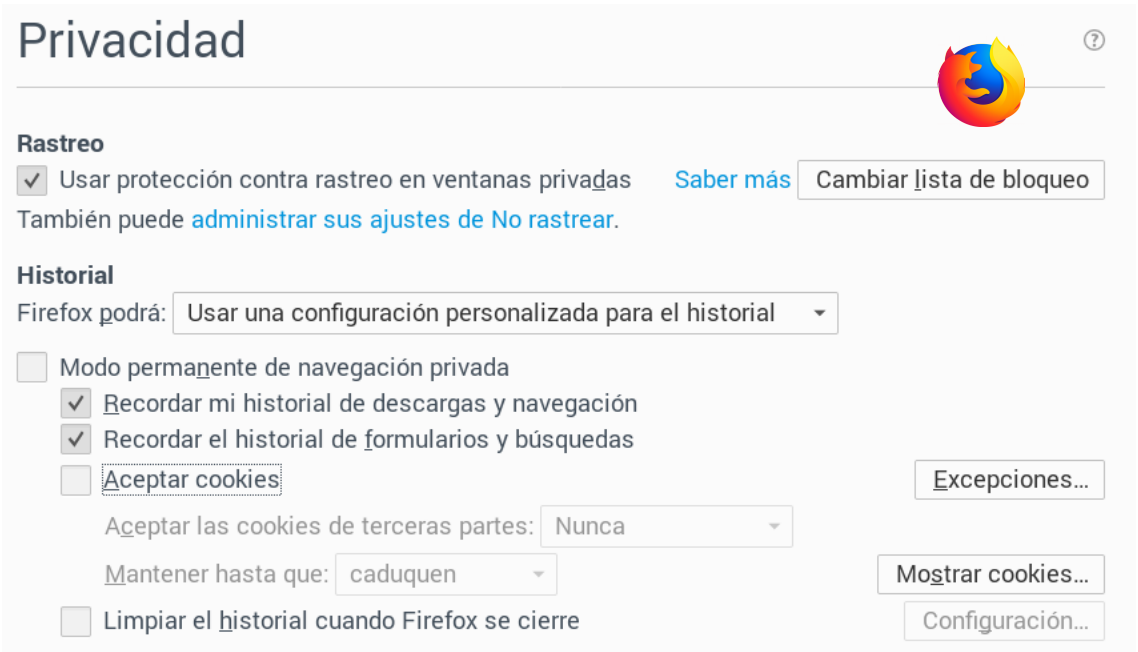
Lo primero que observamos es lo sencillo que resulta aceptar las *cookies* y continuar navegando por la página, no siendo igual de claro averiguar qué debemos hacer para rechazarlas. Si accedemos a la política de *cookies*[81] se nos informa de varios mecanismos para rechazarlas, como usar el servicio Your Online Choices[82] u otros tres similares para realizar un proceso conocido como *opt-out*. Algunos *trackers* requieren un *opt-out* individual; y si acudimos a su página web, como en el caso de AppNexus[83], nos encontramos con interminables y densos términos legales en inglés donde no resulta nada fácil encontrar dónde debemos hacer clic. En definitiva, el objetivo es desincentivar al usuario, pues si multiplicamos este caso por más de 100 sitios que recolectan *cookies*, y por varias páginas visitadas al día, la energía y tiempo requeridos hacen inviable la solución al problema.



Si decidimos escoger la opción *learn more*, se nos presenta otra vez la opción sencilla de aceptar las *cookies* y continuar navegando, con simplemente pulsar el botón colocado al pie de la ventana flotante, o por el contrario la posibilidad de consumir nuestro tiempo y energías en establecer nuestras preferencias para cada grupo de *cookies* definido por la web en cuestión. En ningún momento se nos permite rechazar todas las *cookies* de un solo clic, e incluso en este complejo menú se nos remite a páginas externas de *opt-out* para poder rechazar algunas.

Las empresas saben por supuesto que el usuario promedio no se va a tomar tantas molestias, de manera que cumplen con el requisito legal y consiguen su objetivo último de hacernos aceptar las *cookies*. Sin embargo, vale la pena buscar estrategias alternativas que no requieran tanto esfuerzo de nuestra parte y nos permitan evitar el rastreo.

Los navegadores web tienen la opción de configurar la privacidad, rechazando todas las *cookies* o las de determinados sitios, como por ejemplo en el navegador Firefox.



Podemos aprovechar para desactivar el almacenamiento del historial, descargas, búsquedas y formularios, teniendo siempre en mente eso sí que la página donde realicemos una búsqueda o en la que rellenemos un formulario, siempre tendrá acceso a dicha información. Otra posibilidad es navegar siempre en modo incógnito o privado, que no almacena el historial ni las *cookies*. En el caso de nuestro *smartphone*, también podemos emplear el modo incógnito o utilizar un navegador que ya esté pensado para proteger nuestra privacidad, como Firefox Focus.

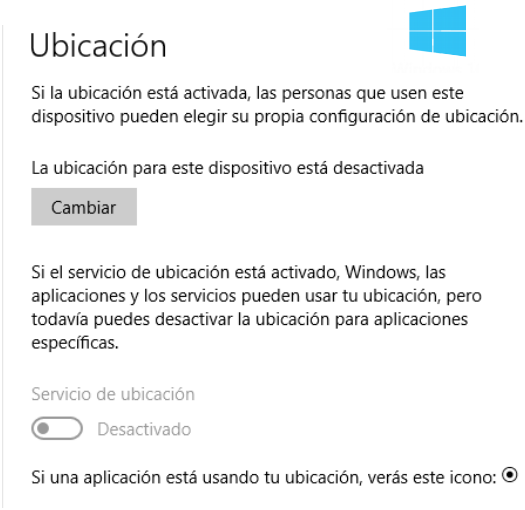
Sin embargo, también hemos visto que las *cookies* no son el único mecanismo de rastreo, siendo el *browser fingerprinting* un método común y difícil de detectar. Una opción altamente recomendable es instalar

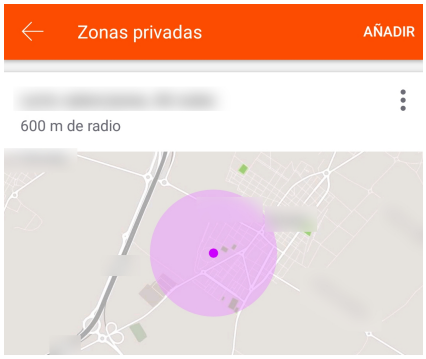
complementos a nuestro navegador que nos protejan contra el rastreo, como Privacy Badger[84], Ghostery[85], que además eliminará parte de la publicidad aligerando la carga de las páginas, o Disconnect[86]. Dado que JavaScript es una de las principales fuentes de información para el *browser fingerprinting*, además de un vector para ataques de *Cross Site Scripting*, *Clickjacking*, u otros *exploits*, también deberíamos instalar complementos que bloqueen esta tecnología, como NoScript[87] para Firefox o ScriptSafe[88] para Chrome, si bien cabe destacar que el bloqueo de *scripts* puede deshabilitar determinadas funciones de un sitio web, lo cual hace que estos complementos a veces no sean sencillos de utilizar para un usuario promedio.

6.3.5. CONTROL DE UBICACIÓN

La mejor manera de no filtrar datos sobre nuestra ubicación, obviando por supuesto no publicarla de manera consciente e intencionada, es desactivarla en el eficiente dispositivo de rastreo que todos llevamos siempre encima, nuestro *smartphone*. Para ello, debemos desactivar el GPS del móvil y también el WiFi, ya que la lectura de tramas *beacon* es suficiente para algunos servicios para aproximar nuestra ubicación, incluso con la geolocalización apagada.

De todas formas, conviene configurar todas las aplicaciones, productos y servicios que no necesiten conocer nuestra ubicación, para que no tengan acceso a estos datos, como por ejemplo en Twitter o Windows 10. Nótese como en muchas ocasiones los servicios se esfuerzan en lograr que activemos estas opciones, prometiendo mejoras en la experiencia de usuario con ánimo de captar nuestros datos.





Aplicaciones deportivas como Strava permiten definir zonas privadas, de manera que incluso cuando queremos registrar nuestra actividad y utilizar el GPS, podemos al menos definir regiones que no queremos que se almacenen, manteniendo una privacidad parcial.

Finalmente, debemos recordar que al publicar nuestras fotografías en la red, muchas veces incluyen metadatos de ubicación. En el caso de *smartphones* y cámaras que permitan configurar esta opción, lo mejor es desactivar el registro de la ubicación en las fotografías. También es interesante desactivar el GPS antes de captar la imagen. Y en caso de fotografías cuyo origen no podemos controlar, existen herramientas gratuitas para borrar los datos EXIF, como GIMP para ordenadores de escritorio o Exif Eraser para el sistema Android.

6.3.6. DATOS DE CONTACTOS Y COMUNICACIONES

Aparte de no compartir nuestra agenda de contactos alegremente con las aplicaciones y servicios que soliciten acceso, una buena recomendación consiste en usar el campo CCO (BCC en inglés) cuando se va a enviar un correo electrónico a una lista de personas, para evitar que sus direcciones caigan en manos de *spammers* o empresas de mercadotecnia.

A la hora de establecer conversaciones *online*, deberíamos priorizar servicios que garanticen que la comunicación será cifrada extremo a extremo utilizando algoritmos criptográficos. De esta manera, ni siquiera el proveedor del servicio debería ser capaz de conocer el contenido de la conversación, aunque generalmente se reservan la posibilidad de revocar la confidencialidad en el caso de que tengan que intervenir las fuerzas del orden por orden judicial, por ejemplo.

Por supuesto, nunca debemos conectarnos a una red WiFi pública abierta para usar servicios sensibles, que trabajen con datos personales, y en cualquier caso extremar en este tipo de redes la precaución de usar solamente conexiones cifradas, pues cualquiera que se conecte a la misma red puede monitorizar nuestra actividad en caso contrario.

Así mismo, si queremos ocultar nuestra dirección IP, podemos contratar algún servicio de VPN, aunque todavía sería visible para el proveedor.

Finalmente, y para evitar que nuestros datos se filtren a través de una conexión no segura, deberíamos comunicarnos solamente a través del protocolo HTTPS. Aunque la tendencia es que la mayoría de sitios web ya están implementando esta tecnología, añadir a nuestro navegador un complemento como HTTPS Everywhere[89] fuerza a que la conexión se haga de modo seguro siempre que dicha opción esté disponible, protegiendo nuestras comunicaciones contra configuraciones pobres o directamente incorrectas.

6.3.7. CONTRASEÑAS

Aunque pueda parecer no estar relacionado, mantener una política de contraseñas adecuada es vital para evitar la filtración *online* de más información de la necesaria. Si nuestros *passwords* son frágiles o se ven

comprometidos de alguna manera, alguien podría acceder a nuestras cuentas personales y obtener, vender o publicar información personal.

Así pues, recomendaciones básicas como no repetir la misma contraseña en varios productos o servicios, cambiar las contraseñas con regularidad, no utilizar *passwords* cortos o combinaciones comunes, activar la verificación en dos pasos en aquellos sitios que lo permitan, etc. pueden evitar que, ante un ataque a nuestras credenciales o ante un intento de robo de información a un servicio en línea, datos de nuestra huella digital acaben en malas manos.

6.4. PROYECCIÓN

Una vez comprobada nuestra huella digital, eliminada la información no deseada, y restringidos al máximo los mecanismos por los que se fugan nuestros datos sin nuestro control, es el momento adecuado para realizarnos las últimas preguntas: ¿qué imagen queremos proyectar en la red? ¿Qué podemos hacer para mejorar nuestra reputación *online*?

La respuesta a estas preguntas será muy personal, pero podemos considerar algunas recomendaciones generales:

- Mantener un perfil académico o profesional, a modo de currículum vitae, en páginas destinadas a este fin, como LinkedIn, destacando logros como premios, calificaciones, proyectos, cargos…
- Abrir un espacio personal como un blog, un portfolio, etc. donde mostrar nuestro trabajo, si es posible.
- Interactuar en redes sociales con perfiles profesionales, rigurosos y respetados, relacionados con nuestras áreas de interés. Nuestras aportaciones deben ser siempre sosegadas, reflexivas y añadir valor. Lo mismo aplica en foros, publicaciones diversas…
- Mantener nuestra vida personal en perfiles separados, preferiblemente bajo pseudónimos que no puedan ser asociados con nosotros.

En resumen, no se trata de suprimir nuestra huella digital, lo que además de casi imposible puede resultar contraproducente, sino de adoptar un papel proactivo, dando visibilidad a aquel contenido que contribuya a una reputación digital positiva y útil.

7. CONCLUSIONES Y TRABAJO FUTURO

El objetivo de este trabajo consistía en estudiar qué es y cómo se genera la identidad o huella digital, cómo se recopilan los datos de los usuarios, para qué pueden ser utilizados, y cómo puede influir la identidad digital en una persona. Como hemos visto, la huella digital de un individuo está constituida por el rastro que éste deja al utilizar las redes de comunicaciones. Mientras la huella digital hace referencia a los datos sin juicio de valor alguno, la reputación en línea se podría definir como el prestigio o la imagen que tenemos en Internet, proyectada a partir de nuestra huella.

En un mundo cada vez más digitalizado e interconectado, este concepto es de vital importancia. La capacidad de identificar a un individuo es clave en la sociedad de la información y del conocimiento, bien para ofrecer servicios a los ciudadanos, bien para obtener un beneficio económico o social. Conocer estos usos, así como estrategias de gestión de la reputación en línea, se presenta como una habilidad valiosa para los ciudadanos del futuro. Sin embargo, la preocupación de los ciudadanos por proteger su intimidad contrasta con su desconocimiento manifiesto acerca de los mecanismos de protección a su alcance. En una encuesta realizada, el 72% de los encuestados se mostraron muy o bastante preocupados por la privacidad de sus datos en Internet, pero el 45% desconocían las medidas de protección a su alcance o les resultaban difíciles de aplicar, considerándose solo un 10% de los encuestados como personas bien formadas en materia de seguridad.

Los motivos para este abismo entre la preocupación de los encuestados y su capacidad de respuesta ante lo que consideran una amenaza son variados. Para empezar, es habitual que llegado el momento entreguemos alegremente datos que, en frío, no desearíamos compartir. Intervienen aquí múltiples estrategias de captación por parte de las empresas, principalmente basadas en un sistema de recompensas para el usuario o en el uso de mecanismos tecnológicos como las *cookies*. De nuestra investigación se deduce que empresas y corporaciones invierten gran cantidad de tiempo y recursos para obtener datos de los ciudadanos, mientras que nosotros tenemos formación, energía y tiempo limitados. Se han propuesto para finalizar el trabajo diversas estrategias de protección de nuestros datos, pero entendemos que es un proceso que requiere formación continua por parte de la ciudadanía, y no una lección puntual que pueda impartirse y aplicarse una vez, para ser olvidada a continuación.

En general, podemos considerar que hemos alcanzado los objetivos que nos habíamos propuesto al inicio del trabajo. Al tratarse de un proyecto de investigación, la metodología era sencilla y la planificación inicial ha resultado de gran ayuda para no perder de vista las diferentes metas parciales a alcanzar. Por supuesto, muchos aspectos se podrían haber tratado con mayor

profundidad, sin embargo las limitaciones temporales y de espacio lo han hecho inviable. No obstante, pensamos que se han definido claramente el concepto y las partes de la huella digital, así como los mecanismos de generación de ésta y sus posibles usos. Si bien es imposible no haber dejado nada en el tintero, los elementos más relevantes han sido identificados y puestos sobre la mesa. En la parte final del trabajo se ha abordado el concepto de reputación en línea, ofreciendo una estrategia en cuatro fases para controlar y mejorar el contenido de la huella digital, con ejemplos y casos prácticos de las aplicaciones y servicios más comunes en la actualidad.

Es seguro que en el futuro surgirán nuevos métodos de recogida y tratamiento de los datos de la ciudadanía. En este sentido, este trabajo podría no terminar nunca, pues se trata de un tema vivo y en constante evolución. Por consiguiente, queda pendiente una revisión de nuevos datos personales a recolectar, nuevos mecanismos y nuevos tratamientos por parte de terceros.

Finalmente, la información y formación al ciudadano se ha revelado como clave para devolverle el control de su identidad digital. Una línea de trabajo futuro a seguir sería la generación de guías didácticas claras y prácticas para los usuarios, estableciendo previamente diversos perfiles demográficos para mejorar su impacto y utilidad.

8. GLOSARIO

APEC (Asia-Pacific Economic Cooperation): Foro multilateral creado en 1989, con el fin de consolidar el crecimiento y la prosperidad de los países del Pacífico, que trata temas relacionados con el intercambio comercial, coordinación económica y cooperación entre sus integrantes.

API (Application Programming Interface): Conjunto de subrutinas, funciones y procedimientos o métodos que ofrece cierta biblioteca para ser utilizado por otro software como una capa de abstracción.

APN (Access Point Name): Punto de acceso que permite el acceso a una red de datos de comunicación inalámbrica externa, por ejemplo la de un proveedor de telefonía móvil.

Beacon, trama: Trama de datos que contiene toda la información sobre una red inalámbrica. Son transmitidos periódicamente para anunciar la presencia de una red WLAN.

Big data: Concepto relativo a conjuntos de datos tan grandes y complejos como para que hagan falta aplicaciones informáticas de procesamiento de datos para tratarlos adecuadamente. El uso moderno del término tiende a referirse al análisis del comportamiento del usuario, extrayendo valor de los datos almacenados, y formulando predicciones a través de los patrones observados.

Bluetooth: Especificación industrial para Redes Inalámbricas de Área Personal que posibilita la transmisión de voz y datos entre diferentes dispositivos mediante un enlace por radiofrecuencia en la banda de los 2.4 GHz.

Branding: Proceso de hacer y construir una marca mediante la administración estratégica del conjunto de activos vinculados al nombre y/o símbolo que identifican a la marca influyendo en su valor, tanto para el cliente como para la empresa propietaria.

Broadcast: Forma de transmisión de información donde un nodo emisor envía información a una multitud de nodos receptores de manera simultánea, sin necesidad de reproducir la misma transmisión nodo por nodo.

Browser fingerprinting: Técnica de rastreo con la finalidad de identificar usuarios o dispositivos, y rastrearlos a través de diferentes sitios web, a partir de los datos que logre averiguar del navegador y del dispositivo que está realizando la navegación.

Canvas fingerprinting: Técnica de browser fingerprinting basada en la utilización del elemento canvas de HTML 5.

CCO (Copia De Carbón Oculta o Con Copia Oculta): Campo del encabezado de un mensaje de correo electrónico. A diferencia del campo Para y la casilla CC, las direcciones de correo electrónico añadidas a CCO permanecen invisibles a los destinatarios del mensaje.

Cifrado: Método que permite aumentar la seguridad de un mensaje o de un archivo mediante la codificación del contenido, de manera que sólo pueda leerlo la persona que cuente con la clave del algoritmo para descodificarlo.

Clickjacking: Técnica maliciosa para engañar a usuarios de Internet con el fin de que revelen información confidencial forzándoles a hacer clic en un elemento distinto al que el usuario percibe, por ejemplo en elementos invisibles.

Cookie: Pequeña información enviada por un sitio web y almacenada en el navegador del usuario, de manera que el sitio web puede consultar la actividad previa del navegador.

CPU (Central Processing Unit): Es el hardware dentro de un ordenador u otros dispositivos programables, que interpreta las instrucciones de un programa informático mediante la realización de las operaciones básicas aritméticas, lógicas y de entrada/salida del sistema.

Cross-site scripting: Es un tipo de vulnerabilidad informática de las aplicaciones web que puede permitir a una tercera persona inyectar en páginas web visitadas por el usuario código JavaScript o en otro lenguaje similar. Puede ser utilizado para robar información delicada, secuestrar sesiones de usuario, y comprometer el navegador, subyugando la integridad del sistema.

Cross-site tracking: Práctica consistente en recolectar información del usuario a través de la navegación en distintas páginas web.

CSS (Cascading Style Sheets): Lenguaje de diseño gráfico para definir y crear la presentación de un documento estructurado escrito en un lenguaje de marcado. Es muy usado para establecer el diseño visual de los documentos web.

EXIF (Exchangeable Image File Format): Especificación para formatos de archivos de imagen a los que se agregan tags específicos de metadatos que cumplen con un amplio espectro de información que permite enriquecer la imagen.

Exploit: Fragmento de software, de datos o secuencia de comandos o acciones utilizada con el fin de aprovechar una vulnerabilidad de seguridad de un sistema de información para conseguir un comportamiento no deseado del mismo.

Flag: Uno o más bits que se utilizan para almacenar un valor binario o código que tiene asignado un significado. Las banderas normalmente forman parte de una determinada estructura de datos, y el significado de su valor se definirá en relación a la estructura de la que forma parte.

Geolocalización: Capacidad para obtener la ubicación geográfica real de un objeto, como un radar, un teléfono móvil o un ordenador conectado a Internet. El término está estrechamente relacionado con el uso de sistemas de posicionamiento.

Geotagging: Proceso de agregar información geográfica en los metadatos de archivos de imágenes, vídeos, sonido, sitios web, etc. que sirva para su georreferenciación. Por

lo general estos datos suelen ser coordenadas que definen la longitud y latitud donde el archivo multimedia ha sido creado.

GDRP (General Data Protection Regulation): Reglamento europeo relativo a la protección de las personas físicas en lo que respecta al tratamiento de sus datos personales y a la libre circulación de estos datos.

GPS (Global Positioning System): Sistema que permite determinar en toda la Tierra la posición de cualquier objeto con una precisión de hasta centímetros si se utiliza GPS diferencial, aunque lo habitual son unos pocos metros de precisión.

GPU (Graphics Processing Unit): Coprocesador dedicado al procesamiento de gráficos u operaciones de coma flotante, para aligerar la carga de trabajo del procesador central en aplicaciones como los videojuegos o aplicaciones 3D interactivas.

Hash, función: Función computable mediante un algoritmo que habitualmente toma como entrada una cadena y la mapea en un rango de salida finito, como una cadena de longitud fija, Se utiliza para producir un resumen criptográfico.

HSTS (HTTP Strict Transport Security): Política de seguridad web establecida para evitar ataques que puedan interceptar comunicaciones, cookies, etc. Según este mecanismo un servidor web declara que los navegadores solamente pueden interactuar con ellos mediante conexiones HTTP seguras.

HTTPS (Hypertext Transfer Protocol Secure): Protocolo de aplicación basado en el protocolo HTTP, destinado a la transferencia segura de datos de Hipertexto, es decir, es la versión segura de HTTP.

IMEI (International Mobile Station Equipment Identity): Código pregrabado en los teléfonos móviles. Este código identifica al aparato de forma exclusiva a nivel mundial, y es transmitido por el aparato a la red al conectarse a esta.

IMSI (International Mobile Subscriber Identity): Código de identificación único para cada dispositivo de telefonía móvil, integrado en la tarjeta SIM, que permite su identificación a través de las redes GSM y UMTS.

Ingeniería social: Es una técnica que pueden usar ciertas personas para obtener, a través de la manipulación, información confidencial, acceso o privilegios en sistemas de información que les permitan realizar algún acto que perjudique o exponga la persona u organismo comprometido a riesgo o abusos.

Internet de las cosas: Interconexión digital de objetos cotidianos con Internet.

IP (Internet Protocol), dirección: Número que identifica, de manera lógica y jerárquica, a una Interfaz en red de un dispositivo que utilice el protocolo IP, que corresponde al nivel de red del modelo TCP/IP.

ISP (Internet Service Provider): Empresa que brinda conexión a Internet a sus clientes.

Log: Grabación secuencial en un archivo o en una base de datos de todos los acontecimientos (eventos o acciones) que afectan a un proceso particular (aplicación, actividad de una red informática, etc.). De esta forma constituye una evidencia del comportamiento del sistema.

LSO (Local Shared Object): Colección de datos que se almacenan en el computador del usuario al visitar un sitio web que hace uso de Adobe Flash. También se llaman Flash cookies.

MAC (Media Access Control), dirección: Identificador de 48 bits que corresponde de forma única a una tarjeta o dispositivo de red. Se la conoce también como dirección física.

Metadatos: Son datos que describen otros datos. En informática cualquier recurso tiene, cuando está almacenado conjuntamente con otros, la necesidad de ser descrito para facilitar las búsquedas que pudieran tratar de encontrarlo a partir de sus características distintivas.

Microtargeting: Metodología vinculada a la mercadotecnia que tiene como objetivo influir en las decisiones de los clientes, consumidores o el público en general. Consiste en la gestión de enormes cantidades de datos, en los cuales se buscan patrones comunes mediante criterios de selección que consideran inclinaciones, intereses, preocupaciones, situación socio económica, nivel educacional, franja de edad, etc. con el objeto de crear segmentaciones del conjunto total. Las personas de cada uno de estos subconjuntos serán las destinatarias de mensajes diseñados a su medida, con un alto grado de personalización, logrando así un incremento en el impacto y la respuesta esperada.

NFC (Near Field Communication): Tecnología de comunicación inalámbrica, de corto alcance y alta frecuencia que permite el intercambio de datos entre dispositivos.

OCDE (Organización para la Cooperación y el Desarrollo Económicos): Organismo de cooperación internacional cuyo objetivo es coordinar políticas económicas y sociales. Los representantes de los países miembros se reúnen para intercambiar información y armonizar políticas con el objetivo de maximizar su crecimiento económico y colaborar a su desarrollo y al de los países no miembros.

Opt-out: Hace referencia a varios métodos por los que los usuarios pueden evitar recibir productos o publicidad no deseados. Estos métodos son normalmente asociados a campañas de marketing directo como marketing de email o correo.

Plug-in: Aplicación informática que se relaciona con otra para agregarle una función nueva y generalmente muy específica. Esta aplicación adicional es ejecutada por la aplicación principal e interactúan por medio de la interfaz de programación de aplicaciones.

PNG (Portable Network Graphics): Formato gráfico de imágenes basado en un algoritmo de compresión sin pérdida para bitmaps no sujeto a patentes.

Podcast: distribución de archivos multimedia (normalmente audio o vídeo que suelen ser de larga duración, que pueden incluir texto como subtítulos y notas) mediante un sistema de redifusión (RSS) que permite opcionalmente suscribirse y usar un programa que lo descarga para que el usuario lo escuche.

RGB (Red, Green, Blue): Composición del color en términos de la intensidad de los colores primarios de la luz, considerando la mezcla aditiva.

Scam: Término anglosajón que se emplea familiarmente para referirse a las estafas por medios electrónicos. Se usa para definir los intentos de estafa a través de un correo electrónico o página web fraudulenta.

Script: Programa usualmente simple, que por lo regular se almacena en un archivo de texto plano. El uso habitual de los guiones es realizar diversas tareas como combinar componentes, interactuar con el sistema operativo o con el usuario.

Session replay: Es el acto de reproducir los pasos que un visitante ha dado en un sitio o aplicación web. Puede incluir la pantalla, interacciones de teclado y ratón, eventos de red y logs de consola.

SIM (Subscriber Identity Module): Tarjeta inteligente desmontable usada en teléfonos móviles y módems HSPA o LTE. Las tarjetas SIM almacenan de forma segura la clave de servicio del suscriptor usada para identificarse ante la red.

SSID (Service Set Identifier): Secuencia incluida en todos los paquetes de una red inalámbrica para identificarlos como parte de esa red. El código consiste en un máximo de 32 caracteres. Todos los dispositivos inalámbricos que intentan comunicarse entre sí deben compartir el mismo SSID.

Streaming: Distribución digital de contenido multimedia a través de una red de computadoras, de manera que el usuario utiliza el producto a la vez que se descarga. Se usa habitualmente para la difusión de audio o video.

Target comercial: En el ámbito de la publicidad, destinatario ideal de una determinada campaña, producto o servicio.

URL (Uniform Resource Locator): Cadena de caracteres con la cual se asigna una dirección única a cada uno de los recursos de información disponibles en Internet.

VPN (Virtual Private Network): Tecnología de red de computadoras que permite una extensión segura de la red de área local sobre una red pública o no controlada como Internet. Permite que un dispositivo envíe y reciba datos sobre redes compartidas o públicas con toda la funcionalidad, seguridad y políticas de gestión de una red privada. Esto se realiza estableciendo una conexión virtual punto a punto mediante el uso de conexiones dedicadas, cifrado o la combinación de ambos métodos.

Wearables: Complementos inteligentes que incorporan sensores y conectividad, permitiendo que puedan compartir información a través de Internet con el fabricante, operador u otros dispositivos conectados, sin necesidad de intervención humana. Relojes, pulseras y gafas con los más relevantes actualmente.

Web bug: Suelen ser pequeñas imágenes de un píxel por un píxel, visibles o invisibles colocadas dentro del código fuente de las páginas web de un sitio o en un mensaje de correo electrónico que se diseña para controlar quién lo lee. Suelen utilizarse para medir el tráfico de usuarios que visitan una página web y poder sacar un patrón de los usuarios de un sitio. Normalmente se utilizan para realizar análisis web.

Zombie cookie: Cookie que se regenera tras ser borrada, de manera que aunque el usuario desee borrar el rastro de su navegación, esto no es posible.

9. BIBLIOGRAFÍA

[1] "More Info About You". *WhatsMyIP*. [Fecha de consulta: 20 de octubre de 2018].
<http://www.whatsmyip.org/more-info-about-you/>

[2] "Configuración de anuncios". *Google*. [Fecha de consulta: 20 de octubre de 2018].
<https://adssettings.google.com/authenticated>

[3] "Tus preferencias de anuncios". *Facebook*. [Fecha de consulta: 20 de octubre de 2018].
<https://www.facebook.com/ads/preferences>

[4] "Cronología". *Google*. [Fecha de consulta: 20 de octubre de 2018].
<https://www.google.com/maps/timeline?pb>

[5] "Mi actividad". *Google*. [Fecha de consulta: 20 de octubre de 2018].
<https://myactivity.google.com/myactivity>

[6] "Dispositivos utilizados recientemente". *Google*. [Fecha de consulta: 20 de octubre de 2018].
<https://myaccount.google.com/device-activity>

[7] **Diuk, Carlos G.** (2014, 14 de febrero). "The Formation of Love". *Facebook*. [Fecha de consulta: 21 de octubre de 2018].
<https://www.facebook.com/notes/facebook-data-science/the-formation-of-love/10152064609253859>

[8] **Friggeri, Adrien** (2014, 15 de febrero). "When Love Goes Awry". *Facebook*. [Fecha de consulta: 21 de octubre de 2018].
<https://www.facebook.com/notes/facebook-data-science/when-love-goes-awry/10152066701893859>

[9] "Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal". *Boletín Oficial del Estado número 298, de 14 de diciembre de 1999, páginas 43088 a 43099.* [Fecha de consulta: 22 de octubre de 2018]
<https://www.boe.es/buscar/doc.php?id=BOE-A-1999-23750>

[10] **Dans, Enrique** (2018, 21 de septiembre). "Insurance, Wearables And The Future Of Healthcare". *Forbes*. [Fecha de consulta: 22 de octubre de 2018].
<https://www.forbes.com/sites/enriquedans/2018/09/21/insurance-wearables-and-the-future-of-healthcare>

[11] **Hern, Alex** (2018, 28 de enero). "Fitness tracking app Strava gives away location of secret US army bases". *The Guardian*. [Fecha de consulta: 22 de octubre de 2018].
<https://www.theguardian.com/world/2018/jan/28/fitness-tracking-app-gives-away-location-of-secret-us-army-bases>

[12] **Marr, Bernard** (2015, 17 de diciembre). "Barbie Wants To Chat With Your Child -- But Is Big Data Listening In?". *Forbes*. [Fecha de consulta: 22 de octubre de 2018].
<https://www.forbes.com/sites/bernardmarr/2015/12/17/barbie-wants-to-chat-with-your-child-but-is-big-data-listening-in>

[13] Redacción La Sexta (2018, 24 de abril). "Después queremos violar todos: los chats de 'La Manada' que son clave para la sentencia por violación múltiple". *La Sexta*. [Fecha de consulta: 22 de octubre de 2018].
<https://www.lasexta.com/noticias/sociedad/despues-queremos-violar-todos-chats-manada-que-son-claves-sentencia-violacion-multiple_201804245adf6e0f0cf220db8622e09f.html>

[14] **Schroeder, Stan** (2018, 4 de julio). "Google says no one is reading your emails, except...". *Mashable*. [Fecha de consulta: 23 de octubre de 2018].
<https://mashable.com/article/google-reading-your-emails-response>

[15] **Pot, Justin** (2017, 20 de abril). "How to Stop Gmail From Adding Events to Google Calendar". *How-to-Geek*. [Fecha de consulta: 23 de octubre de 2018].
<https://www.howtogeek.com/303056/how-to-stop-gmail-from-adding-events-to-google-calendar/>

[16] "Cookies. The EU Internet Handbook". *ec.europa.eu*. [Fecha de consulta: 24 de octubre de 2018].

<http://ec.europa.eu/ipg/basics/legal/cookies/index_en.htm>

[17] "Evercookie". *Wikipedia*. [Fecha de consulta: 25 de octubre de 2018].

<https://en.wikipedia.org/wiki/Evercookie>

[18] **Davidov, Mikhail** (2012, 4 de abril). "The double-edged sword of HSTS persistence and privacy". *Leviathan Security Group*. [Fecha de consulta: 25 de octubre de 2018].

<https://www.leviathansecurity.com/blog/the-double-edged-sword-of-hsts-persistence-and-privacy/>

[19] **Alonso, Chema** (2015, 6 de enero). "HSTS Super Cookies: Cómo te pueden espiar la navegación". *Un informático en el lado del mal*. [Fecha de consulta: 25 de octubre de 2018].

<http://www.elladodelmal.com/2015/01/hsts-super-cookies-como-te-pueden.html>

[20] "Cookie (informática). Privacidad y cookies de terceros". *Wikipedia*. [Fecha de consulta: 25 de octubre de 2018].

<https://es.wikipedia.org/wiki/Cookie_(inform%C3%A1tica)#Privacidad_y_cookies_de_terceros>

[21] "Trackers rank". *whotracks.me*. [Fecha de consulta: 28 de octubre de 2018].

<https://whotracks.me/trackers.html>

[22] **Briz, Nick** (2018, 26 de julio). "This is Your Digital Fingerprint". *Internet Citizen*. [Fecha de consulta: 28 de octubre de 2018].

<https://blog.mozilla.org/internetcitizen/2018/07/26/this-is-your-digital-fingerprint/>

[23] "What every Browser knows about you". *webkay.robinlinus.com*. [Fecha de consulta: 28 de octubre de 2018].

<http://webkay.robinlinus.com/>

[24] "Is your browser safe against tracking?". *panopticlick.eff.org*. [Fecha de consulta: 28 de octubre de 2018].

<https://panopticlick.eff.org/>

[25] "User-Agent". *developer.mozilla.org*. [Fecha de consulta: 30 de octubre de 2018].

<https://developer.mozilla.org/es/docs/Web/HTTP/Headers/User-Agent>

[26] **Englehardt, Steven y Narayanan, Arvind** (2016). "Online Tracking: A 1-million-site Measurement and Analysis". *Princeton University*. [Fecha de consulta: 2 de noviembre de 2018].

<https://webtransparency.cs.princeton.edu/webcensus/#canvas-fp>

[27] "WebGL Browser Report". *browserleaks.com*. [Fecha de consulta: 2 de noviembre de 2018].

<https://browserleaks.com/webgl>

[28] **jkula** (2017, 1 de septiembre). "AudioContext Fingerprinting". *Darkwave Technologies*. [Fecha de consulta: 2 de noviembre de 2018].

<https://www.darkwavetech.com/index.php/device-fingerprint-blog/audiocontext-fingerprinting>

[29] "Session replay". *Wikipedia*. [Fecha de consulta: 3 de noviembre de 2018].

<https://en.wikipedia.org/wiki/Session_replay>

[30] **Englehardt, Steven** (2017, 15 de noviembre). "No boundaries: Exfiltration of personal data by session-replay scripts". *Freedom to tinker*. [Fecha de consulta: 3 de noviembre de 2018].

<https://freedom-to-tinker.com/2017/11/15/no-boundaries-exfiltration-of-personal-data-by-session-replay-scripts/>

[31] "Data release: list of websites that have third-party session replay scripts". *webtransparency.cs.princeton.edu*.[Fecha de consulta: 3 de noviembre de 2018].

<https://webtransparency.cs.princeton.edu/no_boundaries/session_replay_sites.html>

[32] "Location tracking". *Me and my shadow*. [Fecha de consulta: 5 de noviembre de 2018].

<https://myshadow.org/location-tracking>

[33] **Kiss, Jemima** (2010, 15 de mayo). "Google admits collecting Wi-Fi data through Street View cars". *The Guardian*. [Fecha de consulta: 5 de noviembre de 2018].

<https://www.theguardian.com/technology/2010/may/15/google-admits-storing-private-data>

[34] "FOCA. Fingerprinting Organizations with Collected Archives". *Eleven Paths*. [Fecha de consulta: 5 de noviembre de 2018].

<https://www.elevenpaths.com/es/labstools/foca-2/index.html>

[35] **Edwards, Jim** (2012, 11 de octubre). "Apple Has Quietly Started Tracking iPhone Users Again, And It's Tricky To Opt Out". *Business Insider*. [Fecha de consulta: 6 de noviembre de 2018].

<https://www.businessinsider.com/ifa-apples-iphone-tracking-in-ios-6-2012-10?IR=T>

[36] "El esquema de permisos en Android". *Máster en Desarrollo de Aplicaciones Android. Universidad Politécnica de Valencia*. [Fecha de consulta: 6 de noviembre de 2018].

<http://www.androidcurso.com/index.php/recursos/41-unidad-7-seguridad-y-posicionamiento/282-el-esquema-de-permisos-en-android>

[37] "Privacy and the :visited selector". *developer.mozilla.org*. [Fecha de consulta: 7 de noviembre de 2018].

<https://developer.mozilla.org/en-US/docs/Web/CSS/Privacy_and_the_:visited_selector>

[38] **Munroe, Randall** (2018, 13 de junio). "Customer rewards". *Xkcd*. [Fecha de consulta: 8 de noviembre de 2018].

<https://xkcd.com/2006/>

[39] **Hunt, Troy** (2017, 5 de mayo). "Password reuse, credential stuffing and another billion records in Have I been pwned". *troyhunt.com*. [Fecha de consulta: 10 de noviembre de 2018].

<https://www.troyhunt.com/password-reuse-credential-stuffing-and-another-1-billion-records-in-have-i-been-pwned/>

[40] Garrigues comunica (2018, 24 de mayo). "¿Cómo se regula la protección de datos en Latinoamérica y cómo influye el RGPD?". *garrigues.com*. [Fecha de consulta: 11 de noviembre de 2018].

<http://www.garrigues.com/es_ES/noticia/como-se-regula-la-proteccion-de-datos-en-latinoamerica-y-como-influye-el-rgpd>

[41] "How Does Legislation Affect Digital Footprints?". *Internet Society*. [Fecha de consulta: 11 de noviembre de 2018].

<https://www.internetsociety.org/tutorials/digital-footprint-matters/module-9-legislation-affect-digital-footprints>

[42] "Data protection and online privacy". *europa.eu*. [Fecha de consulta: 11 de noviembre de 2018].

<https://europa.eu/youreurope/citizens/consumers/internet-telecoms/data-protection-online-privacy/index_en.htm>

[43] "Pwned websites". *haveibeenpwned.com*. [Fecha de consulta: 11 de noviembre de 2018].

<https://haveibeenpwned.com/PwnedWebsites#QatarNationalBank>

[44] "Pwned websites". *haveibeenpwned.com*. [Fecha de consulta: 11 de noviembre de 2018].

<https://haveibeenpwned.com/PwnedWebsites#Vodafone>

[45] **Jolly, Jasper** (2018, 25 de octubre). "British Airways: 185,000 more passengers may have had details stolen". *The Guardian*. [Fecha de consulta: 11 de noviembre de 2018].

<https://www.theguardian.com/business/2018/oct/25/british-airways-data-breach-185000-more-passengers-may-have-had-details-stolen>

[46] **Pérez, Enrique** (2018, 2 de noviembre). "Ahora podrás pedirle por voz a tu Roomba que limpie tu cocina gracias al acuerdo entre iRobot y Google". *Xataka*. [Fecha de consulta: 11 de noviembre de 2018].

<https://www.xataka.com/domotica-1/google-conocera-mejor-interior-nuestra-casa-gracias-a-su-nuevo-acuerdo-irobot-para-utilizar-mapeado-roomba>

[47] **Stokel-Walker, Chris** (2018, 10 de noviembre). "Will you be getting a smart home spy for Christmas?". *The Guardian*. [Fecha de consulta: 11 de noviembre de 2018].

<https://www.theguardian.com/technology/2018/nov/10/spy-christmas-smart-home-facebook-portal-google-home-hub-amazon-show-alexa>

[48] **Vives, Judith** (2018, 7 de noviembre) "Google advertirá de restaurantes en los que es posible intoxicarnos". *La Vanguardia*. [Fecha de consulta: 11 de noviembre de 2018].

<https://www.lavanguardia.com/tecnologia/20181107/452786563793/google-advertira-restaurantes-posible.html>

[49] **Ruddick, Graham** (2017, 2 de noviembre). "Admiral to price car insurance based on Facebook posts". *The Guardian*. [Fecha de consulta: 11 de noviembre de 2018].

<https://www.theguardian.com/technology/2016/nov/02/admiral-to-price-car-insurance-based-on-facebook-posts>

[50] **Timberg, Craig** (2013, 12 de marzo). "Personal data usage: what your car really says about you". *The Guardian*. [Fecha de consulta: 11 de noviembre de 2018].

<https://www.theguardian.com/technology/2013/mar/12/cars-internet-data-privacy-debate>

[51] **Karp, Paul** (2018, 14 de agosto). "Coalition's surveillance laws give police power to access electronic devices". *The Guardian*. [Fecha de consulta: 11 de noviembre de 2018].

<https://www.theguardian.com/australia-news/2018/aug/14/coalitions-surveillance-laws-give-police-power-to-access-electronic-devices>

[52] BBC News (2018, 12 de noviembre). "Amazon asked to share Echo data in US murder case". *BBC*. [Fecha de consulta: 14 de noviembre de 2018].

<https://www.bbc.com/news/technology-46181800>

[53] **Piña, Raúl** (2018, 20 de noviembre). "Los partidos políticos 'espiarán' los datos personales de los ciudadanos para captar votos". *El mundo*. [Fecha de consulta: 22 de noviembre de 2018].

<https://www.elmundo.es/espana/2018/11/20/5bf31b2d468aeb5e1e8b4648.html>

[54] **Angwin, Julia y Parris, Terry** (2016, 28 de octubre). "Facebook Lets Advertisers Exclude Users by Race". *ProPublica*. [Fecha de consulta: 12 de noviembre de 2018].

<https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race>

[55] **Statt, Nick** (2018, 21 de agosto). "Facebook will remove 5,000 ad targeting categories to prevent discrimination". *The Verge*. [Fecha de consulta: 12 de noviembre de 2018].

<https://www.theverge.com/2018/8/21/17764480/facebook-ad-targeting-options-removal-housing-racial-discrimination>

[56] **Sánchez, J.M.** (2018, 10 de junio). "Polémica en la app de La Liga: puede acceder al micrófono para evitar fraudes en las emisiones de los bares". *ABC*. [Fecha de consulta: 12 de noviembre de 2018].

<https://www.abc.es/tecnologia/moviles/aplicaciones/abci-polemica-laliga-puede-acceder-microfono-para-evitar-fraudes-emisiones-bares-201806101632_noticia.html>

[57] **FM, Yúbal** (2018, 11 de junio). "LaLiga reconoce que su app usa tu micrófono y ubicación para espiar qué bares emiten fútbol sin licencia". *Xataka*. [Fecha de consulta: 12 de noviembre de 2018].

<https://www.xataka.com/privacidad/laliga-reconoce-que-su-app-usa-tu-microfono-ubicacion-para-espiar-que-bares-emiten-futbol-licencia>

[58] **Wiener-Bronner, Danielle** (2018, 22 de marzo). "¿Qué es Cambridge Analytica? Guía para entender el polémico caso del que todo el mundo habla". *CNN Español*. [Fecha de consulta: 12 de noviembre de 2018].

<https://cnnespanol.cnn.com/2018/03/22/que-es-cambridge-analytica-guia-para-entender-el-polemico-caso-del-que-todo-el-mundo-habla/>

[59] **Biosca, P.** (2018, 24 de mayo). "Facebook y Cambridge Analytica: 10 claves para entender el escándalo del robo de datos". *ABC*. [Fecha de consulta: 12 de noviembre de 2018].

<https://www.abc.es/tecnologia/redes/abci-facebook-y-cambridge-analytica-10-claves-para-entender-escandalo-robo-datos-201803202237_noticia.html>

[60] **Hill, Kashmir** (2018, 26 de julio). "When a Stranger Decides to Destroy Your Life". *Gizmondo*. [Fecha de consulta: 13 de noviembre de 2018].

<https://gizmodo.com/when-a-stranger-decides-to-destroy-your-life-1827546385>

[61] **Zimmerman, Kaytie** (2017, 25 de junio). "What Millennial Job Seekers Need To Know About Their Online Presence". *Forbes*. [Fecha de consulta: 13 de noviembre de 2018].

<https://www.forbes.com/sites/kaytiezimmerman/2017/06/25/what-millennial-job-seekers-need-to-know-about-their-online-presence/#7682f7ad97f9>

[62] **Hinduja, Sameer**. "The Importance of your Digital Reputation". *Cyberbullying Research Center*. [Fecha de consulta: 15 de noviembre de 2018].

<https://cyberbullying.org/the-importance-of-your-digital-reputation>

[63] "35 Percent of Employers Less Likely to Interview Applicants They Can't Find Online". *CareerBuilder*. [Fecha de consulta: 15 de noviembre de 2018].

<https://www.careerbuilder.com/share/aboutus/pressreleasesdetail.aspx?sd=5%2F14%2F2015&id=pr893&ed=12%2F31%2F2015>

[64] **Riveiro, Aitor** (2014, 12 de marzo). "VÍDEO: El Movimiento por la Democracia presenta su hoja de ruta para un proceso constituyente". *eldiario.es*. [Fecha de consulta: 18 de noviembre de 2018].

<https://www.eldiario.es/politica/Movimiento-Democracia-presenta-proceso-constituyente_0_237977208.html>

[65] **Olaya, Vicente G.** (2015, 14 de junio). "Un edil de Ahora Madrid se burla en Twitter de los judíos y de Irene Villa". *El País*. [Fecha de consulta: 18 de noviembre de 2018].

<https://elpais.com/ccaa/2015/06/13/madrid/1434219265_951793.html>

[66] **Recuero, Marisa** (2016, 15 de noviembre). "La Audiencia Nacional absuelve a Zapata por el 'tuit' sobre Irene Villa". *El Mundo*. [Fecha de consulta: 18 de noviembre de 2018].

<https://www.elmundo.es/madrid/2016/11/15/582afd37468aeba3328b4579.html>

[67] @eldiarioCultura (2018, 9 de octubre). "James Gunn, despedido de Disney por sus bromas pederastas, se pasa a DC con 'Suicide Squad'". *eldiario.es*. [Fecha de consulta: 18 de noviembre de 2018].

<https://www.eldiario.es/cultura/James-Disney-Warner-Suicide-Squad_0_823118871.html>

[68] "Sobre el Data Detox Kit". *datadetox.myshadow.org*. [Fecha de consulta: 25 de noviembre de 2018].

<https://datadetox.myshadow.org/es/about>

[69] "Reverse Image Search". *TinEye*. [Fecha de consulta: 26 de noviembre de 2018].

<https://tineye.com/>

[70] "Google Imágenes". *images.google.com*. [Fecha de consulta: 26 de noviembre de 2018].

<https://images.google.com/>

[71] "Alertas". *Google*. [Fecha de consulta: 26 de noviembre de 2018].

<https://www.google.com/alerts>

[72] **Pot, Justin** (2017, 24 de abril). "Unroll.me Is Selling Your Information, Here's an Alternative". *How-To Geek*. [Fecha de consulta: 27 de noviembre de 2018].

<https://www.howtogeek.com/304373/unrollme-is-selling-your-information-heres-an-alternative/>

[73] "Mi Actividad". *Google*. [Fecha de consulta: 27 de noviembre de 2018].

<https://myactivity.google.com/delete-activity>

[74] "Otra actividad de Google". *Google*. [Fecha de consulta: 27 de noviembre de 2018].

<https://myactivity.google.com/more-activity>

[75] "Aplicaciones con acceso a tu cuenta". *Google*. [Fecha de consulta: 27 de noviembre de 2018].

<https://myaccount.google.com/permissions>

[76] "Retirada en virtud de la ley de privacidad de la UE". *Google*. [Fecha de consulta: 27 de noviembre de 2018].

<https://www.google.com/webmasters/tools/legal-removal-request?complaint_type=rtbf&pli=1>

[77] CHOICE Australia (2017, 13 de marzo). "How long does it take to read Amazon Kindle's terms and conditions?". *Youtube*. [Fecha de consulta: 28 de noviembre de 2018].

<https://www.youtube.com/watch?v=sxygkyskucA>

[78] "ProtonMail". *mail.protonmail.com*. [Fecha de consulta: 29 de noviembre de 2018].

<https://mail.protonmail.com/inbox>

[79] "DuckDuckGo". *duckduckgo.com*. [Fecha de consulta: 29 de noviembre de 2018].

<https://duckduckgo.com/>

[80] "Controles de la actividad de tu cuenta". *Google*. [Fecha de consulta: 30 de noviembre de 2018].

<https://myaccount.google.com/activitycontrols>

[81] "Política de cookies". *Prisa*. [Fecha de consulta: 30 de noviembre de 2018].

<https://www.prisa.com/es/info/politica-de-cookies>

[82] "Your Online Choices". *youronlinechoices.com*. [Fecha de consulta: 30 de noviembre de 2018].

<http://www.youronlinechoices.com/es/preferencias/>

[83] "AppNexus Platform Privacy Policy". *AppNexus*. [Fecha de consulta: 30 de noviembre de 2018].

<https://www.appnexus.com/platform-privacy-policy>

[84] "Privacy Badger". *Electronic Frontier Foundation*. [Fecha de consulta: 30 de noviembre de 2018].

<https://www.eff.org/privacybadger>

[85] "Ghostery". *ghostery.com*. [Fecha de consulta: 30 de noviembre de 2018].

<https://www.ghostery.com/>

[86] "Disconnect". *disconnect.me*. [Fecha de consulta: 30 de noviembre de 2018].

<https://disconnect.me/>

[87] "NoScript Security Suite". *Firefox Add-ons*. [Fecha de consulta: 30 de noviembre de 2018].

<https://addons.mozilla.org/en-US/firefox/addon/noscript/>

[88] "ScriptSafe". *chrome web store*. [Fecha de consulta: 30 de noviembre de 2018].

<https://chrome.google.com/webstore/detail/scriptsafe/oiigbmnaadbkfbmpbfijlflahbdbdgdf>

[89] "HTTPS Everywhere". *Electronic Frontier Foundation*. [Fecha de consulta: 30 de noviembre de 2018].

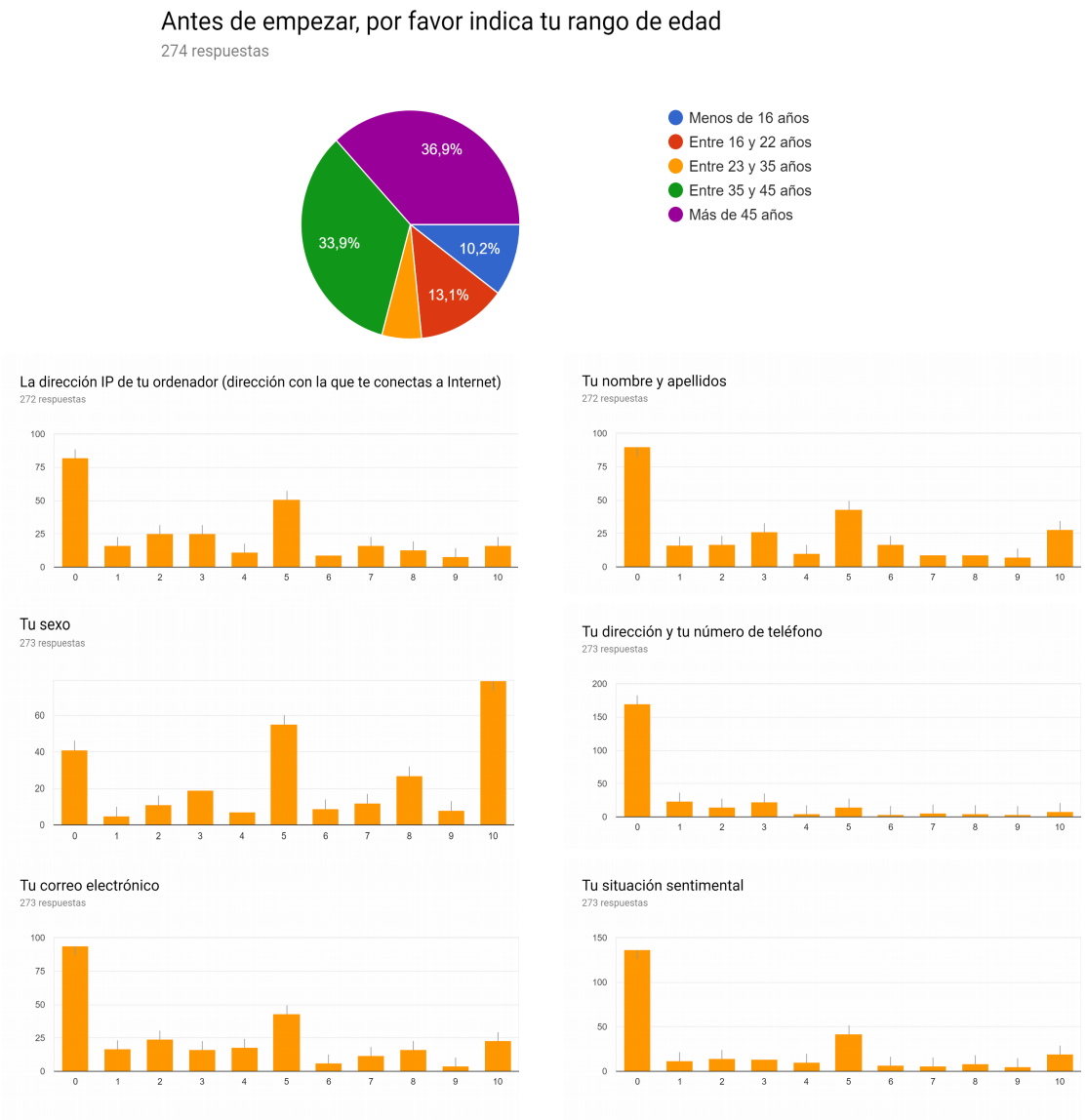
<https://www.eff.org/https-everywhere>

10. ANEXOS

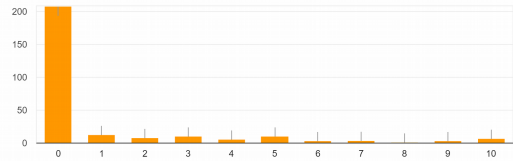
10.1. ANEXO 1

Resultados de una encuesta realizada a 274 participantes en los que se les preguntaba, de 0 a 10, su nivel de comodidad con que determinada información sobre ellos estuviera disponible en Internet. La encuesta se ha distribuido a través de redes de contactos personales y profesionales a través de la plataforma WhatsApp, así como se ha realizado a alumnos de instituto durante mi ejercicio de la docencia, y las respuestas se han recogido *online* a través de Google Forms.

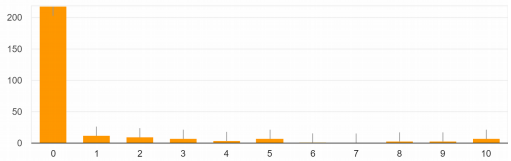
VALORA, DE 0 A 10 TU NIVEL DE COMODIDAD CON QUE LA SIGUIENTE INFORMACIÓN SOBRE TI SE ALMACENE EN INTERNET, DE MANERA PERMANENTE Y ACCESIBLE AL MENOS POR CIERTAS PERSONAS, EMPRESAS O GOBIERNOS, SIENDO 0 (NO ME SIENTO NADA CÓMODO) Y 10 (NO TENGO NINGÚN PROBLEMA CON QUE INTERNET SEPA DICHA INFORMACIÓN SOBRE MÍ)



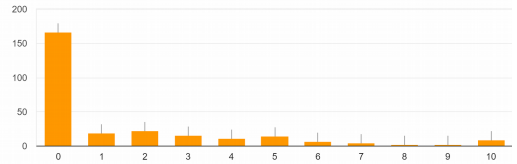
Tus contraseñas
273 respuestas



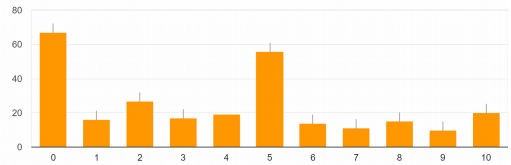
Datos financieros (préstamos, hipotecas, deudas...)
272 respuestas



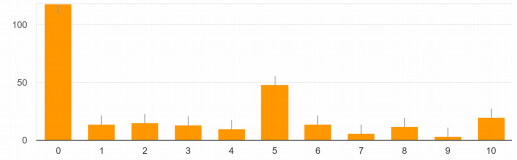
Datos de salud (enfermedades, ingresos hospitalarios...)
273 respuestas



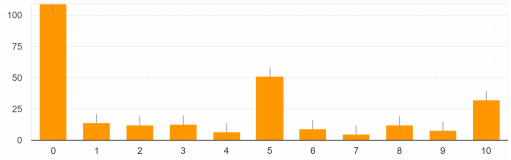
Datos académicos (formación, títulos, expediente académico, notas...)
272 respuestas



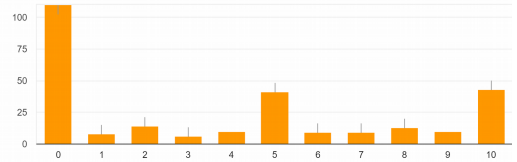
Orientación política
273 respuestas



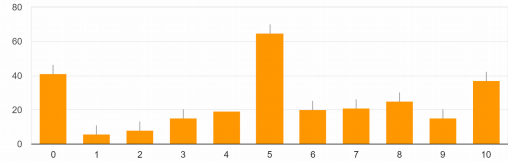
Orientación religiosa
272 respuestas



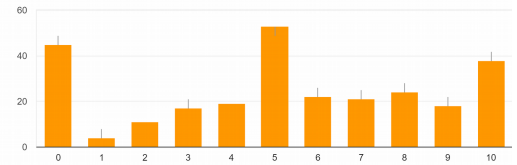
Orientación sexual
273 respuestas



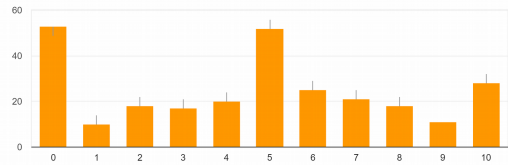
Gustos musicales, cinematográficos, literarios, artísticos...
272 respuestas



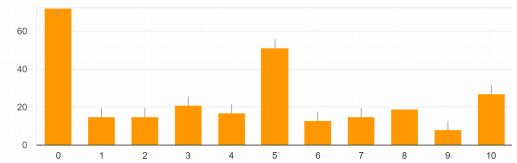
Restaurantes favoritos
272 respuestas



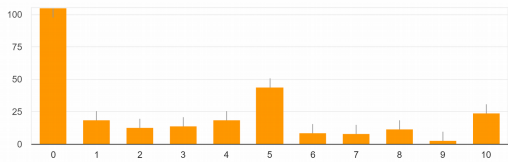
Actividades de ocio favoritas
273 respuestas



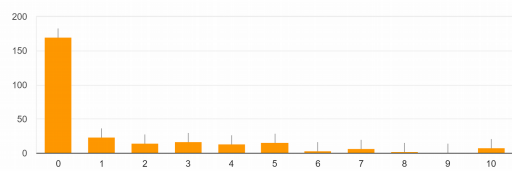
Viajes que has realizado
273 respuestas



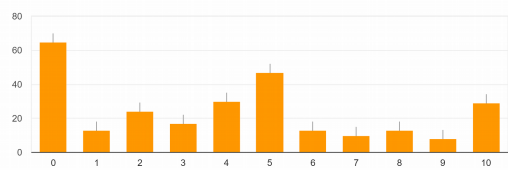
Las fotos que publicas en una red social como Facebook o Instagram
270 respuestas



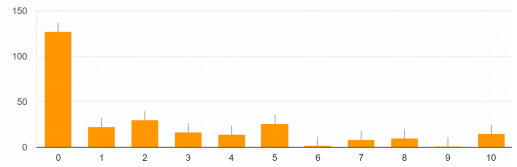
Registro completo de dónde hemos estado (dónde, cuándo, cuánto tiempo...)
273 respuestas



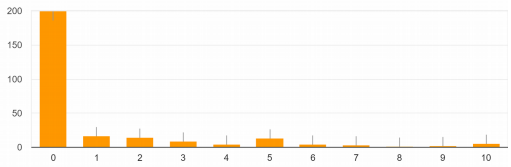
Registro de actividad de una pulsera inteligente (número de pasos, calorías quemadas...)
269 respuestas



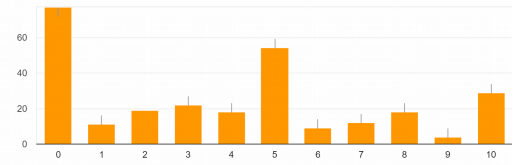
Historial de búsquedas en Internet
273 respuestas



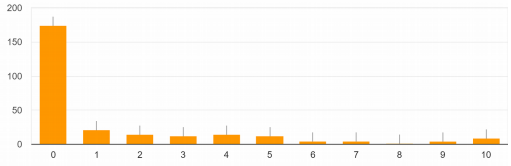
La posibilidad de que un dispositivo grabe, analice y almacene lo que dices
273 respuestas



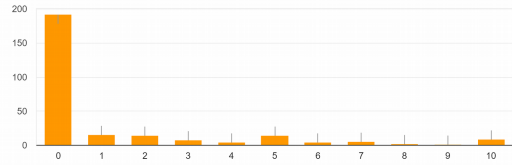
Qué teléfono móvil posees (marca y modelo)
273 respuestas



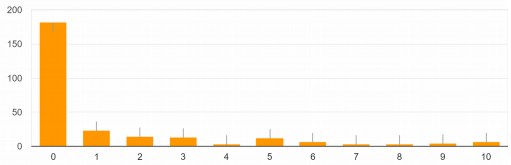
Agenda de contactos
272 respuestas



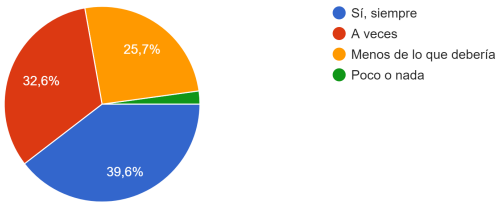
Contenido de los correos electrónicos que envías y recibes
273 respuestas



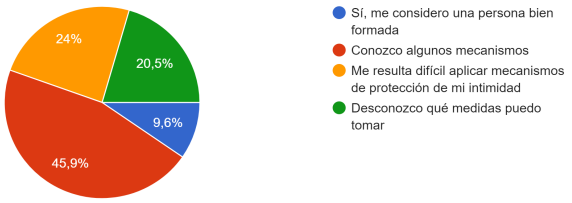
Tus mensajes en tus grupos de WhatsApp, Telegram...
273 respuestas



¿Te preocupas por tu privacidad, generalmente, cuando utilizas Internet?
230 respuestas



¿Crees que estás bien informado de los mecanismos de protección de tu privacidad que puedes utilizar?
229 respuestas



¿Crees que tu reputación digital en Internet puede tener repercusiones en tu vida personal?
228 respuestas