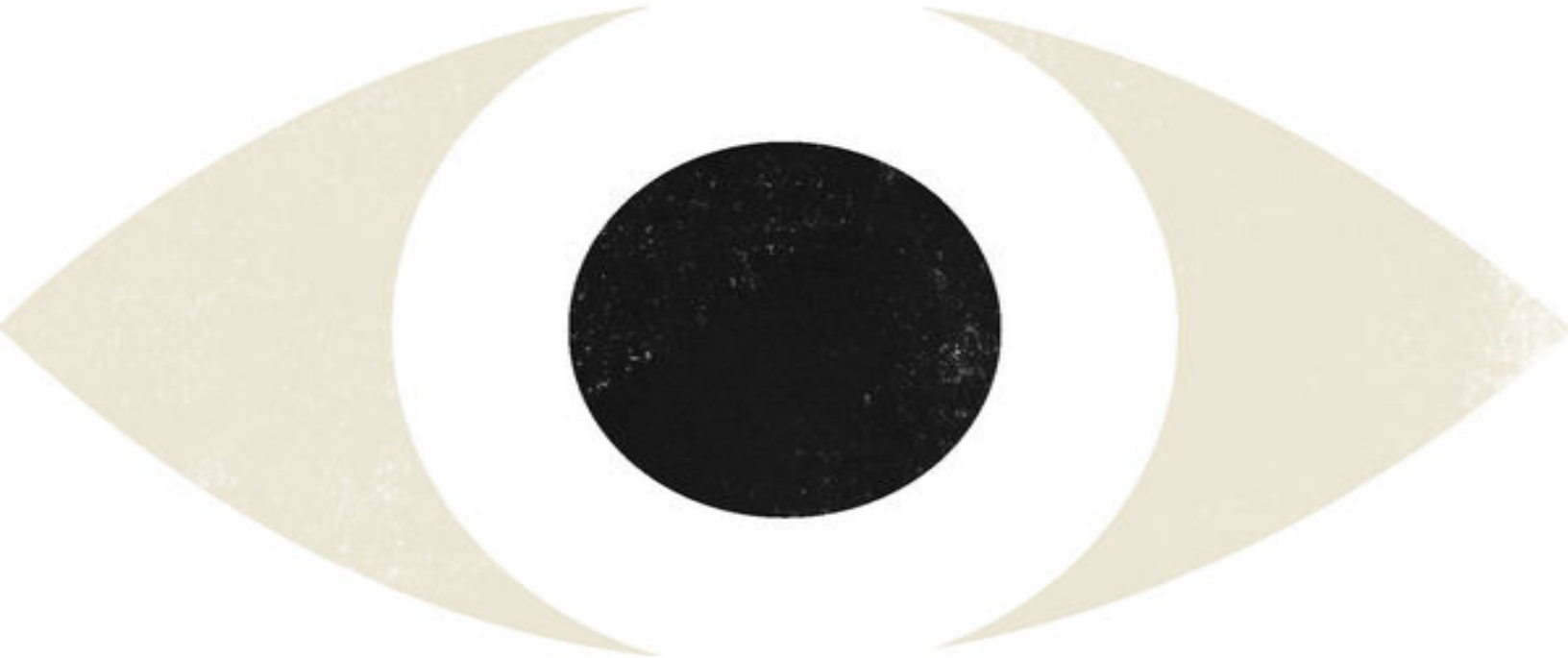
Paloma Llaneza

DATA NOMICS



TODOS LOS DATOS PERSONALES QUE DAS SIN DARTE CUENTA Y TODO LO QUE LAS EMPRESAS HACEN CON ELLOS

Claves, consejos y herramientas para proteger tu privacidad

DEUSTO

Índice

Gracias por adquirir este eBook

Visita **Planetadelibros.com** y descubre una
nueva forma de disfrutar de la lectura

¡Regístrate y accede a contenidos exclusivos!

Primeros capítulos
Fragmentos de próximas publicaciones
Clubs de lectura con los autores
Concursos, sorteos y promociones
Participa en presentaciones de libros

PlanetadeLibros

Comparte tu opinión en la ficha del libro
y en nuestras redes sociales:

Explora **Descubre** **Comparte**

SINOPSIS

¿Qué tienen en común tu cuenta de Instagram y la pulsera de actividad que llevas en tu muñeca? ¿Por qué te sientes intimidado cuando la policía te identifica por la calle, pero no te preocupa lo que las aplicaciones de tu móvil hacen con tus datos de geoposicionamiento?

Estas preguntas tienen una respuesta en común: cómo se usan los datos que damos, que generamos con nuestra vida, los que facilitamos voluntariamente en algún momento, los que dejamos sin saberlo, y los que se infieren de ellos.

Datanomics

Todos los datos personales que das sin darte cuenta y todo lo que las empresas hacen con ellos

PALOMA LLANEZA

EDICIONES DEUSTO

Todos entendemos las alegrías de nuestro mundo siempre conectado: las conexiones, que los demás nos valoren, las risas, la información. Sin embargo, apenas estamos empezando a considerar su coste.

ANDREW SULLIVAN (2016)
Actual presidente y director ejecutivo (CEO)
de Internet Society (ISOC)

Aquellos que renunciarían a una libertad esencial para comprar un poco de seguridad momentánea, no merecen ninguna de las dos.

BENJAMIN FRANKLIN

Un mundo construido a partir de lo familiar es un mundo en el que no hay nada que aprender. Es un mundo en el que existe autopropaganda invisible que nos adoctrina con nuestras propias ideas.

ELI PARISER en *The Economist*, 2011

Un Estado injusto es más peligroso que el terrorismo, y demasiada vigilancia fomenta un Estado injusto.

ROBERT STALLMAN

A Chiqui, por su fe inquebrantable.
Y por su paciencia también.

Introducción

¿Qué es más peligroso, una Roomba que barre tu casa o dejar el coche abierto? ¿Qué tienen en común tu cuenta de Instagram y la pulsera de actividad que llevas en la muñeca? ¿Por qué, a pesar de no haber impagado nunca una deuda, te deniegan un crédito por haber comprado en una determinada tienda? ¿Por qué te sientes intimidado cuando la policía te identifica por la calle pero no te preocupa lo más mínimo lo que las empresas de aplicaciones de tu móvil hacen con tus datos de geoposicionamiento?

Todo ello está relacionado con cómo se usan los datos que nos rodean: los que damos, los que generamos con nuestra vida, los que facilitamos voluntariamente en algún momento, los que dejamos sin saberlo y los que se infieren de ellos. Estos datos reflejan comportamientos y pensamientos profundos perfectamente identificados e individualizados, que facilitan a las empresas y a los Estados la toma de decisiones sobre nosotros, pues así saben lo que vamos a hacer en cada momento.

En algunas culturas primitivas las personas huían de las cámaras fotográficas con el temor de que éstas les robara el alma. Pero, en el mundo digital, las técnicas de análisis permiten a grandes compañías o instituciones analizar nuestras interacciones digitales y obtener de nosotros una fotografía interna con un enorme grado de precisión. Es una instantánea de nuestro pensamiento político, orientación sexual, filias y fobias, carácter, debilidades y fortalezas...

A finales de octubre de 2018, Tim Cook, CEO de Apple, la empresa estrella de Silicon Valley con unos beneficios anuales difíciles de asimilar por un cerebro humano, alabó con vehemencia

las bondades de la regulación europea de protección de datos personales. Cook pidió nuevas leyes de privacidad en Estados Unidos en línea con la norma europea, advirtiendo que la actual recopilación y tratamiento de cantidades ingentes de datos personales por parte de determinadas empresas perjudica gravemente a la sociedad.

Aunque en los últimos años haya sido difícil ver modelos de negocios tecnológicos que no se hayan basado en la prestación de servicios sin precio a cambio de una extracción de datos personales masiva, hay muchas empresas innovadoras que no centran sus beneficios en el mercadeo de los datos. Apple vende hardware caro y aspiracional, y sus beneficios no dependen, en modo alguno, de cuánto y cómo se entromete en la vida de sus clientes. No es de extrañar, por tanto, que Cook cargase contra las empresas extractivas de datos, a las que bautizó como el «complejo industrial de datos que se arma contra nosotros con eficiencia militar». El CEO de Apple fue un fiel defensor de la vida privada y, como alguien que sabe de lo que habla, dejó claro que tenemos un problema real, no imaginado ni exagerado.

No puedo negar que me sentí reforzada cuando dejó claro que los que compartimos su opinión no necesitamos una camisa de fuerza o un sombrero de papel de aluminio. Para Cook, «las plataformas y los algoritmos que prometieron mejorar nuestras vidas están magnificando nuestras peores tendencias... los actores malintencionados y los gobiernos han aprovechado la confianza de los usuarios para profundizar las divisiones, incitar a la violencia e, incluso, socavar nuestro sentido compartido de lo que es verdad y de lo que es falso... En Apple respaldamos plenamente una ley federal de privacidad integral en Estados Unidos. Es hora de que el resto del mundo siga el ejemplo de la UE». Para Tim Cook, regular cómo se recaban y usan nuestros datos no es una barrera para la innovación: «Esta noción no sólo es errónea, sino que es destructiva... El potencial de la tecnología se basa en la fe que la gente tiene en ella».

Tras más de veinticinco años dedicándome al Derecho TI, no puedo estar más de acuerdo con el diagnóstico de Cook. En este tiempo, hemos pasado de tener un contrato con Telefónica, monopolista en la prestación de los servicios de telecomunicaciones, a un entorno competitivo en donde el límite de los servicios de transporte de datos era la imaginación de los prestadores y del regulador. Con la apertura de los portadores de los datos y la generalización del ADSL como alternativa al cable, pudimos dejar de vivir en vilo cada vez que la música de nuestro router no seguía las notas necesarias para engancharse a la red. La conectividad trajo mejoras de las interfaces web y una cierta despreocupación en cómo escribíamos el código. El almacenamiento no era barato, pero los PC venían ya con inteligencia y capacidad mejoradas, lo que no permitía ser menos cuidadoso con lo que se transmitía. El negocio aún no dependía enteramente de los datos aunque las empresas de contenidos empezaban a sufrir la dictadura del clic, tan transparente en cuanto al impacto real de la publicidad.

Eran los años finales de la década de los noventa y la primera mitad de dos mil. Los portales de internet morían ante el imparable avance de Google y su *Don't be evil*, y todo el mundo hablaba de comercio electrónico aunque nadie compraba online. La protección de datos se refería a unos ficheros poco dinámicos de empresas que se dedicaban a los negocios tradicionales. El dato personal no era el objeto del negocio sino parte de los elementos necesarios para facturar los servicios. Como abogado asesoré a mis clientes sobre estos cambios con un código civil del siglo XIX, algunas directivas comunitarias, la normativa de telecomunicaciones y un par de leyes de protección de datos: la vieja LORTAD derogada por la LOPD en 1999.

La combinación mortal vino con un inesperado cambio tecnológico: el iPhone. La movilidad absoluta que trajo Apple requirió un almacenamiento en la nube propio de un entorno cerrado de aplicaciones y archivos en remoto. Empezamos a movernos con

un espía en el bolsillo permanentemente conectado que nos permitía acceder a las recién nacidas redes sociales. Los datos empezaron a fluir sin control. Contábamos nuestra vida y nuestros dispositivos se chivaban del resto.

A partir de 2008 empezamos a asesorar de manera intensiva a organizaciones que no cobraban en dinero por sus servicios pero que, a cambio, se quedaban con todo tipo de datos de sus usuarios; empresas que miraban al otro lado del Atlántico donde los campos de datos eran más verdes y no había un regulador dispuesto a sancionar por el uso irregular de los mismos. Mientras en Estados Unidos se hablaba sin complejo de la recogida masiva de datos, con una despreocupación típicamente protestante, en Europa no se contaban los pecados de las empresas que los usaban, más allá de lo que el consentimiento del usuario les permitía, lo que no impedía la queja persistente contra una regulación que limitaba la sacrosanta y siempre justificada innovación.

Los datos están en nosotros. Somos productores de datos. Una mezcla de percepción equivocada y conocimiento de nuestros sesgos y adicciones nos ha convertido en productores intensivos de datos, recursos necesarios para que siga funcionando una maquinaria de servicios por los que no queremos pagar con dinero. Los datos cuentan nuestro pasado con una precisión nunca vista hasta ahora, y son los posos sobre los que se lee nuestro futuro con los algoritmos que, en muchos casos, lo convierten en una profecía autocumplida. Frente a negocios extractivos de este nuevo petróleo, se alzan cuestiones sobre quién ha de recibir el dividendo de los beneficios que nuestros dobles, nuestros doppelgänger digitales, generan a las empresas que los usan. Y entre tanto ciberoptimismo, el control de los datos se usa como herramienta de control social en un mundo en el que todo lo que se almacena en un sistema informático es susceptible de ser robado. La cuestión, al final, es que ceder nuestros datos excede de la elección individual,

cuestionable, como veremos, al afectar a la convivencia y a los fundamentos de las reglas sociales que nos hemos dado para convivir.

No puedo negar que han sido y siguen siendo años apasionantes para ayudar a las empresas a usar los datos dentro de su negocio, y para tener una postura personal como ciudadana y usuaria crítica con alguno de estos usos... Todo ello me ha creado una cierta «bipolaridad». Por eso he intentado ser objetiva en el relato de cómo los datos afectan a nuestra vida. En *Datanomics* hay datos, sólo datos, de cómo se recaba y se usa nuestra información personal, y de cuáles han sido las consecuencias indeseadas de estos usos. De cómo hemos sido capaces de pasar de una economía productiva a una economía del dato, y cómo, para mantenerla, la sociedad que conocemos ha mutado con una fe casi religiosa en la que los datos son la solución y no el problema.

He dicho que sólo hay datos, y tal vez he faltado a la verdad. No he podido remediar dar mi opinión final sobre lo que pienso de quien no tiene nada que ocultar. Espero que el lector me disculpe por ello.

1

Dataísmo

La nueva religión

En el futuro cercano de 2020 y dos décadas después de la primera explosión de la burbuja puntocom, el modelo de negocio impulsado por la publicidad de las principales compañías de internet se desmorona. A medida que las empresas web sobrevaloradas se hunden, todos —grandes y pequeños, delincuentes y compañías decentes— compiten por quedarse con la propiedad de activos de datos subvalorados —y quién sabe si potencialmente valiosos— tras el hundimiento. Estamos ante un futuro distópico, una guerra por los datos con un trasfondo de estrés financiero, pánico, derechos de propiedad ambiguos, mercados opacos y trolls de datos en todas partes. En este mundo, la ciberseguridad y la seguridad de los datos se entrelazan de forma inextricable. Hay dos activos clave que los delincuentes explotan: los propios conjuntos de datos, que se convierten en los principales objetivos de ataque; y los humanos que trabajan en ellos, ya que el colapso de la industria deja a los científicos de datos desempleados en busca de empleo.

Los modelos de negocio basados en datos de la primera mitad de la década de 2010 desaparecerán. El ambiente, sombrío, anuncia el fin de la tercera era de las compañías de internet que tendrán que buscar nuevas fuentes de ingresos. Algunas compañías de hardware comenzarán a cobrar el precio completo por sus

dispositivos (por ejemplo, Amazon podría revocar todas las ofertas de precios especiales en su Kindle). Para reducir su dependencia de los datos, las empresas de servicios empezarán a cobrar. El «Freemium» será una palabra del pasado, y la mayoría de las aplicaciones «gratuitas» que han sido símbolos icónicos de la web 2.0, dejarán de serlo.

Este escenario es uno de los planteados por el Centro de Ciberseguridad a Largo Plazo de la Universidad de Berkeley[1] en 2016. Mirando a un futuro que casi se nos viene encima, parece poco probable que la economía basada en datos o con alta dependencia de éstos vaya a desaparecer de un plumazo.

Siempre se han usado datos para dibujar la realidad o tomar decisiones para mantenerla o cambiarla. La información siempre ha sido poder y no debemos caer en el adanismo de considerar que hemos inventado el análisis de datos y los algoritmos. Lo que es cierto es que nunca antes habíamos tenido de manera conjunta la capacidad de recabar datos con enorme granularidad (hasta el punto de identificar a un individuo por cómo interactúa con un teléfono), almacenarlos durante largo tiempo y sin límite de capacidad, y de analizarlos con rapidez y precisión. Por tanto, no es de extrañar que algunos piensen que los datos han venido a redimirnos.

Harari y los humanismos liberales

Nuestras religiones sostienen que existe un orden sobrehumano que no es producto de caprichos o convenios humanos, sino la base y los principios sobre los que establecemos normas y valores a los que dotamos de valor obligatorio y que, por su origen divino, no admiten crítica. Esta estructura de pensamiento se extiende a ideologías que participan de los mismos elementos constituyentes que le hemos otorgado a la religión, como el humanismo liberal o el

dataísmo, otro fenómeno que, junto con el entusiasmo tecnológico, son percibidos en las últimas décadas como un acontecimiento religioso que no admite crítica u oposición.

Los dataístas han encontrado sentido a la sinrazón evolutiva a través de la atractiva idea de que los datos lo explican todo, le dan significado a nuestra existencia. Como resalta Harari,[2] «el dataísmo sostiene que el universo consiste en flujos de datos, y que el valor de cualquier fenómeno o entidad está determinado por su contribución al procesamiento de los mismos [...]. El dataísmo une ambos[3] y señala que las mismas leyes matemáticas se aplican tanto a los algoritmos bioquímicos como a los electrónicos. De esta manera, el dataísmo hace que la barrera entre animales y máquinas se desplome, y espera que los algoritmos electrónicos acaben por descifrar los algoritmos bioquímicos y los superen». Tan potente es la propaganda a favor de los datos y en defensa de la desaparición de la privacidad (como si su reivindicación fuera pecaminosa, propia de personas con hábitos dudosos que prefirieran ocultar) que Harari coloca el dataísmo dentro de los movimientos humanistas de corte religioso. Y es verdad que si uno se atreve a oponerse al tratamiento de los grandes datos es señalado como alguien contrario, en su totalidad, al avance y a la innovación. Continuando con Harari, «para los políticos, los empresarios y los consumidores corrientes, el dataísmo ofrece tecnologías innovadoras y poderes inmensos y nuevos. Para los estudiosos e intelectuales promete asimismo el santo grial científico que ha estado eludiéndonos durante siglos: una única teoría global que unifique todas las disciplinas científicas».

Esta visión elevada de los datos como el GATACA que todo lo explica, aterriza en el frío suelo de los números cuando nos enfrentamos al uso que las empresas dan a los datos que les facilitamos de un modo u otro. Las dotamos del poder de los augures, y les dejamos leer nuestro futuro en los posos de nuestros hábitos. Gracias a nuestros datos, siguen comportamientos y sacan conclusiones que les permiten predecir lo que va a suceder, no el porqué, que hoy ha pasado a ser irrelevante.

Esta lectura sólo se puede hacer usando grandes cantidades de datos. Los usuarios suelen proporcionar información de contacto fidedigna para poder disfrutar del servicio o producto que se oferta en la plataforma, y añaden luego más datos con el uso recurrente del servicio.

Porque, como veremos a lo largo de esta obra, somos datos. Datos que nos definen con una precisión que supera la maternal y que, llevado al extremo y sin control, nos puede colocar, como individuos y sociedad, en una precaria situación. El propio Harari lo reconocía en una entrevista al diario *El País*:[4] «El mayor problema político, legal y filosófico de nuestra época es cómo regular la propiedad de los datos. En el pasado, delimitar la propiedad de la tierra fue fácil: se ponía una valla y se escribía en un papel el nombre del dueño. Cuando surgió la industria moderna, hubo que regular la propiedad de las máquinas. Y se consiguió. Pero ¿los datos? Están en todas partes y en ninguna. Puedo tener una copia de mi historial médico, pero eso no significa que yo sea el propietario de esos datos, porque puede haber millones de copias. Necesitamos un sistema diferente. ¿Cuál? No lo sé. Otra pregunta clave es cómo conseguir una mayor cooperación internacional».

Harari coloca la propiedad de los datos entre los tres principales problemas humanos de carácter global, lo que compromete y dificulta su resolución al no depender de las soberanías nacionales sino de un consenso internacional imposible en un mundo de intereses locales egoístas: «Nuestros tres principales problemas son globales. Un sólo país no puede arreglarlos. Hablo de la amenaza de una guerra nuclear, del cambio climático y de la disrupción tecnológica, en especial del auge de la inteligencia artificial y de la bioingeniería. Por ejemplo, ¿qué podría hacer el Gobierno español contra el cambio climático? Aunque España se convirtiera en el país más sostenible y redujera sus emisiones a cero, sin la cooperación de China o Estados Unidos no serviría de mucho. En cuanto a la tecnología, aunque la UE prohíba experimentar con los genes de una persona para diseñar superhumanos, si Corea o China lo

realizan, ¿qué haces? Es probable que Europa acabará creando seres superinteligentes para no quedarse atrás. Es difícil ir en la dirección contraria».

Los datos como subproducto

La RAE define los datos como «información sobre algo concreto que permite su conocimiento exacto o sirve para deducir las consecuencias derivadas de un hecho». Ésta y otras definiciones conexas inciden en la idea de que la transmisión de la información reflejada en los datos causará efectos, inmediatos o diferidos, en el comportamiento del receptor de la transmisión.

Como indica Pérez-Chirinos,[5] hasta la llegada de los artificios técnicos, los seres vivos hemos transmitido la información por medios bioquímicos: muchos animales son capaces de comunicarse a grandes distancias acústicamente y algunas especies cuentan con lo que pueden considerarse propiamente lenguajes con propósitos específicos, capaces de causar efectos predecibles en sus congéneres receptores.

Además de la capacidad de transmitir información verbalmente de efecto inmediato, los humanos disponemos desde hace milenios de instrumentos que nos permiten almacenar los datos que queremos comunicar diferidamente. Desde la representación primitiva de nuestras ideas en obras de arte hasta las últimas técnicas de almacenamiento digital, tenemos la posibilidad de comunicarnos a través del tiempo, incluso con personas que nacerán cuando ya no estemos vivos.

Para ello recurrimos a codificar el mensaje que queremos transmitir siguiendo las reglas apropiadas al soporte material en el que queremos conservarlo. Este proceso de codificación es conceptualmente idéntico tanto si se trata de escribir un documento, transcribir una canción en su partitura o generar una factura

electrónica: tan sólo cambia el algoritmo de codificación utilizado para obtener el resultado deseado, que es registrar el mensaje de forma duradera para que pueda ser decodificado posteriormente.

Gracias a esta posibilidad de registrar mensajes de forma persistente, los humanos podemos construir sistemas de información capaces de dirigir la conducta de otros humanos en un futuro distante del momento en que fueron construidos. Esta capacidad de dirigir conductas futuras opera sobre los receptores de estos mensajes del pasado cuando decodifican los mensajes persistentes.

Los sistemas informáticos producen datos de manera constante. Está en su naturaleza que sea así. Los alimentamos con datos para que nos den respuestas y, a cambio, nos vomitan datos estructurados, nos contestan a las preguntas que les hacemos o actúan como les hemos pedido que hagan. En este proceso, además, generan más datos, como un subproducto necesario para que las máquinas, sin que las entendamos, se hablen entre sí. En algunos casos, y por una mala programación, los sistemas, aplicaciones y redes generan más datos, más ruido de fondo del que sería deseable. El número y variedad de datos accesorios producidos por un sistema tenía, hasta hace unos años, una limitación basada en la eficiencia y la eficacia: se generaban los datos que la capa de transporte, la de aplicaciones o los sistemas de almacenamiento pudieran soportar. Con sistemas limitados en alguno de estos aspectos se tenía especial cuidado en no generar datos que no se podían procesar, transportar o almacenar. Eran los tiempos felices de los disquetes y los Amstrad en los que tener 100 megas de memoria era todo un lujo. Sobre esos ordenadores corrían procesadores de texto que eran capaces de guardar versiones de los documentos y los metadatos necesarios para que no se desconfiguraran. Ahora, cualquier procesador guarda una cantidad ingente de información y todas las versiones anteriores de los documentos.

Todo esto ocurría en un entorno cerrado y controlado: el de los grandes centros de datos, conectados en redes propietarias o en nuestro PC de sobremesa. Los datos se gestionaban de manera interna por principios de necesidad y eficiencia. Y llegó internet y conectamos nuestras máquinas entre sí. Transmitiendo, al principio, la información estrictamente necesaria —porque el ancho de banda no daba para más— y tras una década de movilidad a conexiones 5G que permiten que todas las cosas se conecten en una red infinita en la que se habla de nosotros a nuestras espaldas.

Una vez que conectas un sistema a una red abierta, los datos que produce y los que producimos nosotros al llevarlo en el bolsillo se multiplican y van más allá de las páginas webs visitadas, los anuncios clicados —aunque sea por error— o las palabras tecleadas. Todas las máquinas que intervienen en el proceso (nuestro PC, los routers, los servidores de contenidos y comunicaciones, las antenas de comunicaciones...), todos ellos, sin excepción, generan datos pasivos: los navegadores facilitan a las webs qué modelo son, desde qué ordenador se conectan, cuál es su sistema operativo, qué permisos se han dado y cuáles no, la IP y, por tanto, la compañía de telecomunicaciones con la que tenemos contratado el servicio. En muchos casos, sólo con estos datos es posible identificar de manera única el ordenador y, con un poco más, a los usuarios de éste.

Con la movilidad inteligente hemos dado un paso más allá: ha traído la individualidad, la identificación unívoca de todos y cada uno de nosotros. Es un sistema, pues un ordenador puede ser usado por cualquiera o tener usuarios intercambiables, igual que el teléfono fijo familiar. Antes, para hablar con el novio había que pasar por el filtro del padre malhumorado, la hermana chinchona o la madre indiferente. Ahora las relaciones se ahogan en la hiperconectividad y el hipercontrol del doble check de WhatsApp. Al ser el móvil un identificador personal más potente que el DNI, todo lo que se instale, todo lo que se almacene, todo lo que se diga o haga nos

identifica de manera unívoca. Controlar los móviles es controlar la identidad. Por eso, no es casual que Google regale el sistema operativo a los fabricantes de móviles que compiten con Apple.

Pero la movilidad no sería nada sin la hiperconectividad. Los teléfonos son ordenadores que, además, hacen llamadas. Los coches ya son gigantescos ordenadores que, además, tienen ruedas y motor para llevarte de un punto a otro. Los inofensivos electrodomésticos son ordenadores que le dicen a la compañía de seguros si bebemos demasiada cerveza o abusamos de los embutidos, o les cuentan a los ladrones cómo son nuestras viviendas y a qué hora no estamos en casa. De hecho, con los vestibles o los ingeribles, nuestros cuerpos pasan a ser un repositorio de datos que se transmiten y analizan para tomar decisiones sobre nosotros. Está cercano el día en que se penalice a los enfermos coronarios con dejarles sin tratamiento por su indolente alimentación y falta de ejercicio. Un adecuado análisis de varias fuentes de datos, por ejemplo, Instagram y una pulsera de ejercicio, junto con la edad y el peso que le habremos facilitado a la aplicación de la pulsera al darnos de alta en el servicio, permitirá saber si hemos comido de manera inadecuada y no hemos hecho el ejercicio necesario. Con estos datos se nos podrá penalizar por nuestro pasado y predecir nuestro futuro estado, invirtiendo en nuestra recuperación en una ponderación coste-beneficio-castigo.

Y todo eso sin haber introducido un único dato, sin haber tecleado más que unos pocos datos de registro y haber sido algo indolentes con los permisos otorgados a las aplicaciones o servicios. En muchos casos, esas aplicaciones o servicios están diseñados para que la búsqueda de los parámetros de configuración de privacidad sean más sencillos que la octava de las doce pruebas de Astérix (sí, la del formulario A-38).

Datos por servicios. Pago por datos

La publicidad lo cambió todo. En 1998, Sergey Brin y Larry Page presentaron un artículo en una conferencia en Australia en el que mantenían que la publicidad corrompía la tecnología del buscador. En un tiempo en que el contenido se indexaba mal y que incluso se vendían agendas para anotar las URL más usadas como si de listines telefónicos se tratasen, los buscadores como Yahoo! funcionaban bajo el concepto de portal de entrada a internet: no tenías que recordar o anotar todas las webs sino sólo recordar terra.com o yahoo.com para acceder a la red comercial. Todos luchaban por ser la entrada a la ciudad medieval de la web y, como tal, presentaban una home recargada y llena de banners publicitarios. Google.com fue una completa revolución: se presentó como una página en blanco, sin publicidad ni distracciones, tan sólo había una casilla central de búsqueda, y dos botones: «enter» y «voy a tener suerte». Todo ello bajo el eslogan *Don't be evil*. Tras esa puerta inmaculada, que ya no preside ninguna declaración de intenciones, se alza toda una maquinaria orientada a concentrar los servicios más útiles, mejor diseñados y sin intercambio monetario que se hayan visto en los últimos veinte años.

Google, y después Facebook, gracias a una política de competencia pensada en la era Reagan, han crecido en su posición de mercado hasta constituirse en un duopolio en varias áreas, una de ellas indiscutible: la publicidad online. En 2017, ambas ingresaron por este concepto 135 millones de dólares controlando, según las cifras, entre el 60 y el 75 por ciento del mercado global publicitario.

Veremos cómo muchos de estos servicios, los no necesariamente adictivos pero sí prácticos y útiles, representan una concentración del mercado indeseada e indeseable que hace que el idílico internet descentralizado sea un recuerdo romántico, casi tanto como los pitidos agónicos de los primeros módems. Un ejemplo: Facebook estuvo caído durante 45 minutos el 3 de agosto de 2018. Durante ese tiempo, el tráfico web global subió un 2,3 por ciento, lo

que supuso un aumento el 11 por ciento del tráfico directo a las páginas web y un 22 por ciento en otras aplicaciones, ampliándose en un 8 por ciento el tráfico de buscadores.

En este contexto, son varios los debates que se abren: el problema de la seguridad de los datos concentrados en los repositorios de unas pocas empresas; la obtención de los datos sin un consentimiento real para un tratamiento poco transparente que ponga en riesgo el derecho fundamental a la intimidad, para empezar, pero también el de la libre expresión, el derecho a la participación política y la libertad de movimientos; y el carácter ético —o no— de que los titulares de los datos comercien con los mismos, de que obtengan un pago por ellos más allá de recibir servicios útiles pero poco transparentes, en muchos casos diseñados para ser adictivos, y con consecuencias impensadas en la propia democracia representativa.

Monetización, dividendo y sindicatos de los trabajadores del dato

La monetización de los datos es un argumento casi tan recurrente y falaz como poner la responsabilidad del mal uso de los datos en el lado del usuario. Los datos son imprescindibles para establecer tendencias de consumo o estados de ánimo, para detectar tempranamente una enfermedad, o para individualizar y personalizar lo que se ve y lo que se ofrece a un usuario. Sin embargo, el paquete de datos individualizado de un ser humano no tiene gran valor monetario —o no al menos como para vivir únicamente de él —. Tomemos como ejemplo a Jennifer Lyn Morone. La artista estadounidense se registró como una empresa en Delaware para explotar sus datos personales con la finalidad de obtener un pago por los mismos y, de paso, hacer una acción artística. Hizo varios paquetes de datos, subconjuntos que exhibió en una galería de Londres en 2016 ofreciéndolos para su compra. Los paquetes

tenían distintos precios, desde 100 libras por unos datos básicos, hasta 700 libras por todos ellos, incluidos los de salud y su número de seguridad social.

Tal vez sus datos eran valiosos, o lo eran en una primera venta (los datos reutilizados se deprecian en la Deep web), pero la media está muy por debajo de las 100 libras de Morone. Haciendo un sencillo cálculo, si Facebook compartiera sus beneficios entre todos sus usuarios, cada uno obtendría sólo nueve dólares al año, cuantía que se acerca bastante a los precios que se manejan en otros entornos. Tal vez los precios de los datos frescos de tarjetas de crédito o de cámaras hackeadas de menores sean superiores en el internet profundo, pero sigue estando muy lejos de lo que un ser humano necesita anualmente para sobrevivir.

Los motivos son numerosos, pero el primero que se nos viene a la cabeza, por evidente, es que los datos valen si se analizan de manera combinada con otros datos. Un voto suelto no hace ganar unas elecciones (por mucho que Delibes y *El disputado voto del señor Cayo* se empeñen), pero lo cierto es que sin cada uno de ellos no se llega al Congreso de los Diputados.

De adoptarse esta aproximación —que los ciudadanos controlasen sus datos y comerciasen con ellos—, ¿cómo sería una economía de datos en la que los gigantes tecnológicos tuvieran que pagar por el acceso a éstos?

Como resalta *The Economist*,[6] no sería la primera vez que un importante recurso económico haya pasado de ser usado a ser poseído y negociado, como ocurrió con la tierra, el agua o, añado, la propia imagen. Sin embargo, este mismo medio considera que la información digital es un candidato improbable para vivir esta transición. A diferencia de los recursos físicos, los datos personales son un ejemplo de los bienes que pueden usarse más de una vez. De hecho, cuanto más se utilizan correctamente y con todas las garantías, mejor es para la sociedad, que dirían los dataístas.

Precisamente porque se pueden reutilizar, nada impide que una empresa los compre una vez y los use cientos de veces. Un sistema de monetización justa requeriría una transparencia total de las compañías que, desde el más profundo conocimiento de cómo funcionan, es improbable. No mencionaremos los problemas que causan las frecuentes filtraciones —que hacen que los datos fluyan como por una cañería averiada— o el hecho de que los sistemas de información estén diseñados desde la vagancia, generando un número desproporcionado de datos que podrían evitarse o minimizarse, y que, en la mayor parte de los casos, son proporcionados por el usuario, que no tiene la más remota idea de que están ahí y que están siendo recogidos, agregados, tratados y analizados.

Pero no dejemos que una evidencia desanime a una teoría florida. E. Glen Weyl, autor de *Radical Markets*, escrito en colaboración con Eric Posner, de la Universidad de Chicago,[7] considera que en el futuro, como ocurrió con cualquier avance de la Revolución industrial, el uso será tan intensivo que el precio subirá. Tan convencido está que milita por el estudio de un mecanismo para distribuir la riqueza creada por la IA en el futuro, a la vista de la alta concentración actual de los tratamientos de grandes datos. Weyl argumenta que las habilidades necesarias para generar datos valiosos pueden estar más ampliamente diseminadas de lo que se podría pensar, por lo que el trabajo de datos podría afectar la jerarquía estándar del capital humano. En una toma de los cuarteles de invierno digitales, Weyl avisa de una unión de los trabajadores de datos del mundo en una internacional dataísta que acabará con la desigualdad y traerá un mundo soleado, sin lluvia ni replicantes.

Hasta que el sindicato universal de los datos llegue, la BBC, en un reportaje de finales de 2018,[8] evidenció la realidad de las granjas de humanos que entrenan a las IA por cantidades de miseria. Toda la magia que hay tras estas maravillas pensantes se llama Brenda, una madre soltera de veintiséis años que vive en Kibera (Nairobi), el

barrio pobre más grande de África, y quizás el vecindario más duro del mundo, donde cientos de miles de personas viven en un espacio no mucho mayor que el Hyde Park de Londres.

Brenda trabaja para Samasource, una compañía con sede en San Francisco que cuenta entre sus clientes a Google, Microsoft, Salesforce y Yahoo. A la mayoría de estas empresas no les gusta hablar sobre la naturaleza exacta de su trabajo con Samasource, pero cabe afirmar que la información en la que trabajan Brenda y sus compañeros en jornadas de ocho horas es una parte crucial de una parte de sus procesos con la IA. La joven carga una imagen en la que busca todo lo que se encuentre en ella y lo marca para que los coches sin conductor sepan lo que están viendo. Etiqueta a las personas, los automóviles, las señales de tráfico, las líneas en el pavimento, hasta el cielo, especificando si está nublado o soleado. Se asegura de que no haya un sólo píxel etiquetado incorrectamente. Su trabajo será revisado por un superior, quien lo devolverá si no está a la altura. Los entrenadores más rápidos y precisos son premiados con el honor de tener su nombre en una de las muchas pantallas de televisión de la oficina y con, lo que es más importante, cupones para la compra. Samasource paga a Brenda y a sus compañeros un salario de unos nueve dólares diarios, digno para Nairobi, miserable para Silicon Valley.

Llegados a este punto, incluso Angela Merkel, la canciller de Alemania —uno de los países más estrictos con la protección de datos personales, pues llega a exigir que se almacenen dentro del territorio nacional—, está considerando que se ponga o se pague un precio por el uso de los datos personales.

Frente a esta opción se alza la cuestión ética y legal de si se puede comerciar con derechos tales como la dignidad o la intimidad, y si el derecho de cancelación de los datos, que no es más que una revocación de uso de los mismos, iría en contra del concepto de venta, que se basa en el carácter irrevocable cuando es legal. Es decir, si vendo una casa no puedo exigir que me la devuelvan por igual precio. En cuanto a los datos, sí puedo venderlos y pedir que

se borren al día siguiente. Además, está la discusión de si es ética su venta y, si se me permite, si es práctica. La UE está trabajando en un nuevo concepto, el del «dividendo» de los datos: un reparto de los beneficios del tratamiento entre la empresa que los usa y quien los provee.

Otra opción que está encima de la mesa es la de los trabajadores de los datos, la del pago por entrenar a las Inteligencias Artificiales (IA).

En 1770, el inventor húngaro Wolfgang von Kempelen construyó una máquina de ajedrez conocida como Mechanical Turk. Su objetivo, en parte, era impresionar a la emperatriz María Teresa de Austria. Este dispositivo era capaz de jugar ajedrez contra un oponente humano, y tuvo un éxito espectacular al ganar la mayoría de los juegos en sus demostraciones en Europa y América durante casi nueve décadas. Pero Mechanical Turk era sólo una ilusión que permitía a un maestro de ajedrez humano esconderse dentro de la máquina y operarla.

Unos 160 años después, Amazon.com llamó a su plataforma de crowdsourcing basada en micropago con igual nombre. Según Ayhan Aytes, la motivación inicial de Amazon para construir Mechanical Turk surgió después del fracaso de sus programas de inteligencia artificial en la tarea de encontrar páginas duplicadas de productos en su web.[9] Después de una serie de intentos inútiles y caros, los ingenieros del proyecto recurrieron a los humanos para trabajar detrás de los sistemas en red.[10] Amazon Mechanical Turk simulaba los sistemas de inteligencia artificial mediante el control, la evaluación y la corrección de los procesos de aprendizaje automático usando la capacidad intelectual humana: los usuarios tenían la impresión de que era la inteligencia artificial avanzada la que realizaba tareas cuando, en realidad, estaba más cerca de una forma de «inteligencia artificial-artificial», impulsada por una fuerza de trabajo remota, dispersa y mal remunerada que ayudaba al cliente a lograr sus objetivos comerciales. Como observó Aytes, «en ambos casos —tanto el Mechanical Turk de 1770 como la versión

contemporánea del servicio de Amazon— la actuación de los trabajadores que animan el artificio queda oscurecida por el espectáculo de la máquina».[11]

Este tipo de trabajo invisible y oculto, externalizado o subcontratado detrás de interfaces, y camuflado dentro de procesos algorítmicos es común, sobre todo en los procesos de marcar y etiquetar miles de horas de archivos digitales con el fin de alimentar las redes neuronales. A veces, este trabajo no se paga en absoluto, como en el caso del reCAPTCHA de Google. En una paradoja que muchos de nosotros hemos experimentado: para demostrar que no somos un agente artificial, nos vemos forzados a entrenar el sistema de IA de reconocimiento de imagen de Google de forma gratuita, al seleccionar múltiples cajas que contienen números de calles, automóviles o casas. En otras ocasiones, nos hacemos selfies para ver si nos parecemos más a la Mona Lisa o a la Chica de la Perla, como la aplicación de Google «Arts & Culture», que es un ejemplo perfecto de cómo regalamos un dato biométrico y, al tiempo, entrenamos una herramienta de reconocimiento visual.

Frente a esta situación tan habitual, los defensores de «datos como mano de obra» creen que a los usuarios se les debería pagar por usar servicios en línea.

El trabajo, como los datos, es un recurso difícil de precisar que no ha sido compensado adecuadamente durante la mayor parte de la historia humana. Incluso cuando los humanos pasamos a ser empleados aparentemente libres para decidir, tuvimos que sobrevivir a unas cuantas revoluciones antes de que el pago por nuestra labor alcanzara una cifra digna.

Weyl considera que los datos son como era el carbón, y que las materias primas o los recursos humanos son un elemento productivo más con el que alimentar a lo que ellos prefieren llamar «inteligencia colectiva». La mayoría de los algoritmos de IA necesitan ser entrenados con montones de ejemplos generados por humanos en un proceso llamado «aprendizaje automático». A menos que sepan cuáles son las respuestas correctas

(proporcionadas por humanos), los algoritmos no pueden traducir idiomas, entender el habla ni reconocer objetos o imágenes. Desde esa perspectiva economicista y nada filosófica, los datos serían una forma de trabajo con la que alimentar la IA. Sería un trabajo inconsciente o pasivo en el que se generan los datos —aun de manera voluntaria y activa por actividades como publicaciones en medios sociales, escuchar música o recomendar restaurantes, que generan los datos necesarios para impulsar nuevos servicios—; o un trabajo consciente y activo —en el que se toman decisiones sobre los datos, como etiquetar imágenes o conducir un automóvil a través de una ciudad, que pueden usarse como base para el entrenamiento de sistemas de inteligencia artificial.

Sin embargo, e independientemente de que estos datos se generen de manera activa o pasiva, pocas personas tendrán el tiempo necesario para realizar un seguimiento de toda la información que generan o para estimar su valor. Incluso aquellos que lo hagan carecerán del poder de negociación para obtener un buen trato de las empresas de IA.

Weyl entiende que la historia del trabajo ofrece una pista sobre cómo podrían evolucionar las cosas: gracias a la negociación colectiva a través de la fuerza sindical. De manera similar, espera ver el surgimiento de lo que llama «sindicatos de datos y trabajo», organizaciones que servirían como guardianes de los datos. Al igual que sus predecesores en el mundo laboral, Weyl imagina una Arcadia en la que estos «sindicatos del dato» negocian las tarifas, supervisan el trabajo de datos de los miembros y garantizan la calidad de su producción digital. Los sindicatos podrían, en su visión, canalizar el trabajo de datos especializados a sus miembros e incluso organizar huelgas, por ejemplo, bloqueando el acceso a los datos de sus miembros para ejercer presión sobre una empresa que los emplea sin su permiso o sin respetar sus derechos. Del mismo modo, estos sindicatos podrían canalizar las contribuciones de los datos de sus miembros —al más puro estilo de una

cooperativa— al tiempo que hacen labores de seguimiento, identificación, persecución y facturación a las empresas de IA que se beneficien de ellos.

La cuestión es cómo vamos a conseguir que Google o Facebook renuncien a su actual modelo de negocio de datos gratis para vender publicidad (y alguna cosa más). Como veremos cuando hablemos de las GAFA, ambas se encuentran en una situación de dominio en el mercado porque ninguna autoridad de competencia ha movido un dedo para evitar que se posicionaran ni para que abusaran de esa situación de monopolística. La interferencia en las elecciones presidenciales de 2016 en Estados Unidos ha motivado que se hayan comenzado a tomar tímidas medidas para paliar ese poder. Medidas desde el completo desconocimiento del funcionamiento interno de estas compañías y de cómo usan los datos que recogen a paladas.

Un par de datos. Como hemos dicho, en 2017 ambas compañías ganaron 135 mil millones en dólares en publicidad. Al respecto, la Comisión Europea sancionó a Google con dos multas que suman algo más de 7.500 millones de euros (por abuso de posición de dominio en el mercado de los buscadores y por instalar su navegador por defecto en su sistema operativo móvil Android); y la Agencia de Protección de Datos española hizo lo mismo con Facebook, a la que multó con 1,2 millones de euros por la filtración de los datos en el escándalo de Cambridge Analytica (septiembre de 2017).

Las multas en el caso de Google son recurribles, molestas pero no irremediables. Sobre todo, no animan a cambiar un modelo de negocio altamente rentable. En el caso de Facebook, la sanción es tan ridícula comparada con su volumen de negocio que seguro que gasta más en vasos de papel en una semana que en lo que ha de pagar por saltarse la normativa europea. Aunque el nuevo reglamento europeo de protección de datos aumenta significativamente la cuantía de las sanciones, acercándolas a las

que se imponen en materia de competencia, no parece que éste sea el camino para incentivar a estas compañías en el camino del uso ético y del pago de los datos.

Es más, si tuvieran que compensar a las personas por sus datos, éstos serían mucho menos rentables y, además, seguro que las empresas encontrarían maneras de inducir a la entrega gratuita de los mismos ya que, debido a la concentración de los servicios online, los usuarios (que no clientes) están cautivos.

Hay ya ejemplos de trabajos intensivos en la categorización de datos y plataformas de crowdworking, como la mencionada Mechanical Turk de Amazon (etiquetado de imágenes) o Mighty AI, una startup con sede en Seattle que contrata a miles de trabajadores en remoto para que etiqueten imágenes de escenas callejeras usadas para entrenar los algoritmos que se utilizan en los coches sin conductor.

A medida que los servicios de IA se vuelven más sofisticados, los algoritmos necesitarán alimentarse con una «dieta de información digital» de mayor calidad. *The Economist* mantiene que esto sólo sería posible si los datos son proporcionados por los usuarios previo pago. Me temo que con un internet de cosas conectadas que facilita toda una panoplia de datos en contexto, esta afirmación se tambalea.

En este mundo distópico en el que los humanos —como si de un episodio de *Black Mirror* se tratase—, los trabajadores, comenzarían su jornada laboral comprobando sus tareas para el día, facilitadas por su sindicato de datos, en la que podrían elegir desde mirar publicidad —mientras la cámara del ordenador recogería las reacciones faciales— hasta traducir un texto a un idioma raro, pasando por comprobar lo fácil o difícil que es explorar un edificio virtual, el trabajador tendría un panel de control en el que vería sus ganancias, un histórico de su comportamiento, su ranking e, incluso, sugerencias para adquirir nuevas habilidades.

Para poder comerciar con los datos, los titulares de los mismos habrían de ser conscientes de su valor y de su importancia, algo que, de ocurrir, será más tarde que pronto. Incluso los que dicen ser conscientes de su importancia, los regalan con total inconsciencia, convirtiéndose así en socios fundadores de la «paradoja de la privacidad».

Uno se pregunta si con este panorama distópico quedará alguien a quien venderle los servicios y los productos obtenidos con estos avanzados análisis conductuales.

2

Doppelgänger

Somos datos

Dorian, un joven y rico aristócrata del Londres victoriano, lleva una vida disoluta. Los años pasan crueles para sus amigos y coetáneos mientras él mantiene la misma tersura, agilidad y hambre que en sus años de juventud. En una habitación de su casa, un retrato que en otro tiempo presidió el salón, enmohece bajo siete llaves. Una manta tapa la fealdad del otro Dorian, su doppelgänger, que absorbe en lugar del original el transcurso de los años, la vileza de su carácter, el peso de sus vicios.

Si Oscar Wilde hubiera escrito *El retrato de Dorian Grey* ahora, éste se habría sustituido por todos los perfiles de Dorian en redes sociales, juegos online, aplicaciones de ligue... y Dorian ya hasta dudaría de si él es el original o lo es su doppelgänger digital.

Es una pesadilla recurrente en la literatura del XIX la figura del doble malvado, ése que, siendo nosotros, teniendo nuestro yo y sirviendo a nuestra naturaleza, actúa fuera de nuestro control, en contra de nuestro bienestar, y es una fuente infinita de dolores de cabeza. Nuestro doppelgänger somos nosotros, pero, al mismo tiempo, no nos obedece.

Doppelgänger[12] es el vocablo alemán para definir el doble fantasmagórico de una persona viva. La palabra proviene de doppel, que significa «doble» y de gänger, que es «andante». Su forma más

antigua, acuñada por el novelista Jean Paul en 1796, es «Doppeltgänger», que significa «el que camina al lado». El término se utiliza para designar a cualquier doble de una persona, comúnmente en referencia al «gemelo malvado» o al fenómeno de la bilocación, un extraño fenómeno que aparece tanto en numerosas corrientes filosóficas y misticistas como en supuestos teóricos de la física cuántica y mecánica. La bilocación o multilocación se menciona como la aparición de un mismo individuo o un objeto en dos o más lugares y tiempos diferentes en un mismo instante.

Así pues, la figura del doppelgänger no podría ajustarse más a nuestra dicotomía yo físico-yo digital. Como el retrato, nuestra identidad digital es más completa, rugosa y precisa que el yo que creemos que somos. Incluye todos aquellos elementos de nuestro carácter que, aunque deseemos no tenerlos, existen. En el mundo de los datos, somos más nuestros dobles digitales que nosotros mismos. Incluso cuando jugamos a ser alguien distinto, no hacemos más que mostrar nuestros anhelos más profundos, proyectarnos en el yo que querríamos haber sido. Nuestros doppelgänger, además, están en más de un sitio a la vez, como los datos que una vez creados se expanden como una mala gripe imposible de parar, y están definitivamente fuera de nuestro control. Por mucho que la legislación consagre un derecho a perseguir a nuestros dobles[13] allí donde se hallen, lo cierto es que se encuentran en tantos lugares y bajo tantas formas que la cacería parece imposible.

Nos prefieren de perfil

Uno de los mayores repositorios de doppelgänger es Facebook. Un reciente estudio de la Universidad Carlos III concluía que la red social[14] maneja para usos publicitarios datos sensibles del 25 por ciento de los ciudadanos europeos, que son etiquetados en la red social en función de asuntos tan privados como su ideología política, orientación sexual, religión, etnia o salud. Esto da una idea de la

profundidad de conocimiento que un prestador de servicios online sin pago tiene de nosotros. En ocasiones me imagino a estas entidades como factorías de clones nuestros, colgados en inmensos almacenes a la espera de sacarlos al mercado.

Estas plataformas, desde Netflix hasta Amazon pasando por Twitter o Facebook, elaboran perfiles de nosotros a partir de los datos que les facilitamos, dibujando con una enorme precisión a nuestros *alter ego* digitales.

Esa labor se lleva a cabo gracias a datos facilitados tanto directa y conscientemente por nosotros (datos activos) como indirecta e inconscientemente —con el mero uso de nuestros equipos y con una débil protección de los perfiles de uso de los servicios y de los propios equipos—. No hablaremos de datos entregados voluntaria o involuntariamente desde el momento en que todos hemos aceptado, sin ser coaccionados, unas condiciones generales de uso y de privacidad en las que se nos informa de la recogida de esos datos silenciosos tan reveladores.

Es difícil hacer esta distinción cuando uno está llevando a cabo una acción determinada. Me explico: mientras uno, por ejemplo, contesta un test, un concurso o una encuesta en Facebook —generalmente por divertimento adolescente—, no sólo proporciona las contestaciones sino que además da información adicional.[15] Por ejemplo, un test puede revelar que queremos mejorar nuestra autoestima, para sentirnos especiales al acertar todas las respuestas o para descubrirnos. Además de eso, y como pasó en el escándalo de Cambridge Analytica, la encuesta puede ser de un tercero que deja los datos obtenidos a la red social y a cambio extrae los perfiles del que contesta y de toda su red. Así, con una media de 200.000 personas contestando una encuesta, Cambridge Analytica se hizo con más de 85 millones de perfiles, la inmensa mayoría de usuarios que jamás completaron la encuesta.

Aquí podemos hacer la distinción entre datos activos y pasivos. Sin entrar en los metadatos que se transmiten con cualquier entrega activa de datos, los típicos de este tipo serían nombre y apellidos,

correo electrónico y una contraseña para crear la cuenta, ya sea para darse de alta en Facebook, Twitter, Instagram, WhatsApp o YouTube. Facebook, además, solicita información extra sobre si eres hombre o mujer, la fecha de nacimiento y el número de teléfono. Luego, subimos pensamientos, fotos, entramos en discusiones, apoyamos causas... y todo lo hacemos de manera activa y voluntaria. Además, se nos olvida lo que ya hemos hecho durante los últimos años en esa red o lo que buscamos hace dos años en el buscador. Es el agregado de todo un comportamiento consolidado en el tiempo lo que les permite hacer un doble casi perfecto de nosotros mismos.

No es un secreto que las cookies permiten a los anunciantes y a las páginas en las que entramos rastrear nuestra navegación general y específica. Los botones sociales activados en las webs son trackers que le chivan a las redes sociales lo que hemos estado haciendo fuera de ellas. Tampoco lo es que las elecciones de nuestros amigos o a quienes seguimos identifican nuestro perfil ideológico, religioso, nuestra profesión o nuestra orientación sexual —aunque aún no hayamos salido del armario—. Precisamente por lo sencillo que es ahora hacer una foto o grabar un vídeo, exponemos mucho de nuestra vida diaria. Cada vez que las compartimos facilitamos mucha información: a dónde vamos de vacaciones, con quién, cuántas veces salimos a comer fuera, qué comemos, qué hacemos y con quién. Los enlaces que compartimos revelan también grupos, relaciones, intereses y afiliaciones políticas, sociales, religiosas y comerciales.

Hay organizaciones que convierten la creación de doppelgänger en algo virtuoso, pues lo hacen de manera personalizada y cuidando cada detalle al milímetro. Y dado que nuestros dobles son suyos (al menos bajo la legislación estadounidense), ¿por qué no venderlos enteros o por piezas a empresas que los necesitan para redondear sus propios doppelgänger? En Estados Unidos existe y se permite la figura de los data bróker,[16] que recogen y venden listas de perfiles

no sólo para «mailings», sino también para aseguradoras, farmacéuticas y organizaciones políticas, sobre datos considerados sensibles.[17]

Tras el escándalo del uso de datos personales de los usuarios de Facebook por Cambridge Analytica, una infinidad de medios de comunicación dedicaron especiales informativos para educarnos sobre la información que facilitamos a las organizaciones y administraciones de manera consciente y, en mayor medida, inconsciente.

Noche de peli y mantita: datos activos y pasivos

The Wall Street Journal se ha propuesto analizar los datos facilitados a los distintos proveedores de servicios en una anodina noche cualquiera de «peli y mantita», así como el uso que las empresas involucradas hacían de los mismos. El periódico llegó a la conclusión de que el precio pagado en datos superaba con mucho lo abonado por una pizza. La diferencia entre pedir una pizza desde un teléfono fijo y ver una película de la televisión o hacerlo desde un móvil o televisión inteligente, y elegir una de las ofertas de cualquiera de las plataformas de entretenimiento disponibles, se reveló enorme.[18]

Incluso una noche tranquila en casa supone entregar información personal muy valiosa a compañías de alta tecnología. Es habitual que el usuario piense en los datos sueltos que facilita pero no en que esos datos, aparentemente sin importancia o desconectados, se puedan convertir en información relevante que, de ser consciente de ello, habría preferido no compartir.

Los teléfonos inteligentes, aparatos aparentemente inofensivos, prácticos y amables, junto con las aplicaciones y las cuentas en redes sociales que tenemos permanentemente instaladas y abiertas reúnen de nosotros más información de lo esperable. Sorprende la cantidad de datos que somos capaces de facilitar sólo por evitarnos

hacer tareas sencillas pero que, cada vez más, nos parecen tediosas. Dejaremos de usar los teclados y hablaremos con nuestros dispositivos, cambiando comodidad por datos biométricos de enorme valor.

The Wall Street Journal dibujó los gestos sencillos de una noche de peli y mantita, analizó qué datos recogía cada actor, determinó cuáles se facilitaron de manera consciente y cuáles inconscientemente, y reveló qué uso se iba a hacer de cada uno de ellos conforme a sus declaraciones de privacidad.

El plan propuesto por el diario resume las siguientes acciones de las amigas Sally y Kristen, con el consiguiente vaivén de datos.

Acción	**Datos activos (datos proporcionados a...)**	**Datos pasivos (adicionales recogidos por...)**
Sally saca su iPhone X e intercambia algunos mensajes de texto con su amiga Kristen. Sally y Kristen usan para comunicarse Apple iMessage. Los mensajes están cifrados de modo que Apple no podrá leer, en principio, los mensajes intercambiados.	Apple:[19] texto cifrado de extremo a extremo, información de la dirección de iMessage.	Apple: marcas de tiempo anonimizadas e información de enrutamiento de mensaje anonimizada.
Mientras Kristen limpia su apartamento, recurre a su Amazon Echo: «Alexa, abre Domino's y haz un pedido».La aplicación de Domino's instalada en el Echo extrae la información almacenada de la tarjeta de crédito de	ALEXA:[20] características de voz, contenido de la solicitud. DOMINO'S:[21] información de pago y facturación, tipo de pizza pedida, cantidad.	ALEXA: historial de interacciones, tipo de dispositivo Echo, ubicación, últimos cuatro dígitos de la tarjeta de crédito. DOMINO'S: transcripción de lo que dijo, configuraciones de hardware, sistema

Kristen. «¿Quieres usar tu Visa que termina en 1234?», pregunta Alexa. La información almacenada de la tarjeta de crédito se usa para completar la compra de la pizza. Alexa también registra la interacción, y Domino's crea una transcripción de lo que dijo.		operativo y estadísticas de rendimiento.
Sally se monta en su coche y conecta Google Maps en su iPhone para obtener indicaciones para ir a casa de Kristen. La aplicación utiliza sensores de iPhone para determinar su ubicación a medida que viaja, tocando el acelerómetro para obtener velocidad y el giroscopio para la dirección. Google recopila datos anónimos sobre su velocidad y ubicación, así como los de los conductores cercanos, para detectar si hay mucho tráfico.	GOOGLE:[22] dirección de su destino y ubicación.	GOOGLE: velocidad, punto cardinal hacia el que se dirige, tipo de dispositivo (iPhone X), dirección IP asignada al dispositivo, los enrutadores wifi más cercanos, y torres de telefonía más cercanas.
Sally y Kristen se ven muy de vez en cuando, así que aprovechan para sacarse un selfie. Sally sube la foto a Facebook, y la	FACEBOOK: foto cargada, texto enviado con foto, reconocimiento facial.	FACEBOOK: análisis de la foto, ubicación de la foto (si está incluida en los metadatos), fecha, tipo de dispositivo (iPhone X),

aplicación le sugiere que etiquete a Kristen basándose en su sistema de reconocimiento facial, que Kristen le ha dado permiso para usar. Facebook podría recopilar la ubicación de Sally en función de la dirección IP utilizada para cargar la foto, que podría utilizar para sugerir eventos locales que podrían interesarle o mostrar sus anuncios dirigidos a personas cercanas a un lugar específico. Su sistema también analiza la foto como lo hace con todas las imágenes para asegurarse de que no haya contenido inapropiado.		ID del dispositivo, sistema operativo del dispositivo, nivel de batería, intensidad de señal, señal Bluetooth, velocidad de conexión, espacio libre, nombres y tipos de aplicaciones y sus archivos, balizas de wifi cercanas y torres de telefonía, dispositivos cercanos, como un televisor para la transmisión de teléfono a TV, zona horaria, operador móvil o proveedor de servicios de internet, dirección IP, tiempo, frecuencia y duración de las actividades, versión del hardware y versión del software.
Kristen enciende su Apple TV y busca *Wonder Woman*. Compran la película. Apple podría luego sugerir otras películas que Kristen debería comprar, como *Batman v Superman: Dawn of Justice*. Por defecto, Apple ofrece recomendaciones personalizadas, pero los usuarios pueden desactivar esa configuración. En el proceso, Apple	Apple: película seleccionada, ID de Apple, información de tarjeta de crédito.	Apple: información de ancho de banda de internet, historial de compras, el coste de la noche en datos.

comprueba el ID de Kristen y le carga la información de su tarjeta de crédito almacenada. También utiliza información de ancho de banda de internet para asegurarse de que la película se descargue a la velocidad adecuada.		

Las amigas que optaron por una noche sin pretensiones, que quedaron a cenar una inofensiva pizza delante del televisor, habrían facilitado 53 campos de información. La tipología de datos que el WSJ manejó, recoge toda la información que las empresas involucradas en el proceso de proporcionar una noche anodina podrían haber recopilado conforme a sus declaraciones de privacidad, términos de servicio y documentos relacionados.

Los datos recabados responderían al siguiente reparto: 15 datos o campos de datos son proporcionados conscientemente por el usuario (lo que supone un 28 por ciento del total de la información facilitada), mientras 23 campos de datos provendrían de Facebook y 38 habrían sido recopilados por las distintas empresas intervinientes. Un 72 por ciento de los datos habrían sido facilitados por los usuarios sin ser conscientes de ello y, en la práctica, sin su consentimiento. Según el análisis del WSJ, las políticas de privacidad de Apple, Amazon, Google, Facebook y Domino's, en conjunto, suponen 76.069 palabras que, a una velocidad de lectura media de 250 palabras por minuto, requerirían al menos cinco horas de lectura continua. Ello sin contar con que cambian cada cierto tiempo, lo que obligaría a releerlas cada vez que se hace uso del servicio. Parece poco viable.

No todas las empresas hacen igual uso de los datos. Apple, por ejemplo, suele disociar la información de los usuarios y la usa principalmente para mejorar los dispositivos. Facebook y Google usan principalmente datos para mejorar los servicios y respaldar sus negocios publicitarios.

La información entregada por Sally y Kristen esa noche casera representa una mínima parte de los datos que las compañías intervinientes tienen de ellas. No sólo cuentan con sus datos de registro, sino con toda su actividad anterior, el resultado de recopilar ciento de noches de pizza u otras actividades.

Del ejemplo de la noche de «mantita y peli» llama poderosamente la atención el número y la granularidad de los datos que Facebook recaba para subir una simple foto. En el estudio del WSJ se amplió el alcance de búsqueda y se identificaron los datos que recababa Facebook en la prestación de sus servicios a partir de la lectura de su política de privacidad. Su lectura es todo un reto para los que afirman no tener nada que ocultar.

Sólo para el uso del servicio, Facebook recabaría los siguientes datos:

– Nombre.
– Dirección de correo electrónico.
– Contenido compartido.
– Contenido visto.
– Tipo de contenido dedicado a cada contenido.
– Si se comentó o no sobre mensajes y comunicaciones con otros.
– Conexiones con amigos, grupos, cuentas y hashtags.
– Eventos de la vida.
– Puntos de vista religiosos.
– Opiniones políticas.
– Intereses.
– Salud.
– Raza u origen étnico.

– Creencias filosóficas.
– Pertenencia a sindicatos.
– Registro de llamadas.
– Historial de registro de SMS.
– Detalles de contacto.
– Información de pago.
– Información de envío.
– Número de teléfono móvil.
– Ubicación precisa del dispositivo.
– Fotos y vídeos subidos.
– Configuración de dispositivo de reconocimiento facial.
– Acciones de comunicación de Messenger en Facebook.
– Interacciones con amigos, grupos, cuentas y características de hashtags.
– Ubicación de una foto (como metadatos).
– Fecha, hora, frecuencia y duración de las actividades.
– Hardware del sistema operativo.
– Versión de software.
– Nivel de batería.
– Potencia de señal disponible.
– Almacenamiento del navegador.
– Tipo de aplicación y nombres de archivo.
– Complementos.
– Comportamiento del dispositivo (movimientos del mouse, Windows en primer plano o en segundo plano).
– Identificador del dispositivo.
– Si el dispositivo usa Bluetooth.
– Señal wifi, beacons y torres de telefonía cercanas.
– Operador móvil.
– Proveedor del servicio de internet.
– Idioma.
– Zona horaria.
– Dirección IP.
– Velocidad de conexión.

– Dispositivos cercanos, como el televisor, para realizar compras de transmisión de TV a teléfono.
– Donaciones.
– Servicios.
– Actividad fuera de Facebook, incluidos sitios web visitados, compras realizadas, anuncios vistos y servicios usados en línea y acciones offline de proveedores terceros de datos.
– Actividad de Instagram: dónde vives, lugares a los que vas, empresas, cercanía con personas, información de eventos, de empresas, asistencia a conciertos, citas, comentarios de amigos, acerca de ti, mensajes de amigos, información de contacto de tus amigos, fotos de tus amigos, preguntas de búsqueda de Facebook.

En su política de privacidad, Facebook repite su mantra de haber escuchado alto y claro las peticiones de sus usuarios para gestionar la privacidad, que son conscientes de su complejidad y que van a trabajar en ello; un mensaje muy parecido al que dio Mark Zuckerberg en su declaración ante el Congreso de Estados Unidos a principios de abril de 2018. En esa comparecencia respondió unas 500 preguntas de un modo bastante vago y se quedó en blanco cuando le preguntaron cuánta información acumulaba, si se borraba alguna vez y qué hacía con los datos que recababa. Es cierto que Zuckerberg no estuvo muy brillante, pero los congresistas tampoco. El nivel de complejidad del funcionamiento de las plataformas de redes sociales y de prestación de otros servicios es tan abrumador que requiere algo más de tiempo de estudio que el que los congresistas dedicaron a esta comparecencia histórica, que marcará el principio de una nueva regulación en Estados Unidos. David Baser, Director de Gestión de Producto, publicó en el blog corporativo de Facebook algunas respuestas a las preguntas que quedaron sin contestar por parte de su CEO (unas cuarenta),

empezando por contestar a cómo y cuándo Facebook accede a información de personas cuando los servicios son prestados por un tercero.

Facebook recibe información de esos servicios incluso si el usuarios está desconectado a través de los siguientes medios:

– Uso de complementos sociales (como los botones de «Me gusta» o «Compartir»).
– Inicio de sesión con Facebook connect, que permite usar el usuario y contraseña de este servicio para iniciar sesión en otro sitio web o aplicación.
– Uso de Facebook Analytics.
– Cuando se sirven anuncios de Facebook.
– Uso de herramientas de medición, que permite que los sitios web y las aplicaciones muestren anuncios de los anunciantes de Facebook, publiquen sus propios anuncios en Facebook o en cualquier otro lugar y comprendan la efectividad de sus campañas.

Baser concluye, rotundo, que no venden los datos personales, aunque nada dice sobre que no permita a los terceros usar sus herramientas de segmentación para buscar usuarios que coincidan con determinados parámetros.

Esta abrumadora lista de datos se usa para construir a nuestros dobles digitales. Pero para esa tarea se necesita la argamasa que una las piezas y dé sentido al conjunto, en definitiva, para evitar un Frankenstein de datos confusos, inútiles y sin sentido. Para eso sirven los algoritmos, que no sólo hacen un análisis de quiénes somos, sino de quiénes vamos a ser.

Ya que parece que nos hemos rendido a que nuestros dobles pueblen la galaxia digital, al menos exijamos que la creación de nuestros doppelgängers sea cuidadosa, correcta y refleje con exactitud lo que somos. Porque sobre él se echarán las cartas en ese proceso de adivinar cómo seremos y cómo seremos tratados.

3

Todo es percepción

La paradoja de la privacidad

En 1673, el polímata jesuita Atanasio Kircher inventó la estatua citofónica, la «estatua que habla». Kircher, un extraordinario erudito, publicó más de cuarenta obras en los campos más diversos, como la medicina, la geología, la religión comparada o la música. Inventó el primer reloj magnético, muchos autómatas tempranos y el megáfono.

Su estatua parlante era, en realidad, un sistema de escucha; esencialmente un micrófono formado por un enorme tubo en espiral capaz de transmitir las conversaciones desde la plaza en la que estaba situada hasta la boca de una estatua. El sistema de escucha podía espiar conversaciones cotidianas en la plaza y transmitirlas a los oligarcas italianos del siglo XVII.

La estatua parlante de Kircher fue una forma temprana de extracción de información para las élites: las personas que hablaban en la calle no tenían ninguna indicación de que sus conversaciones se canalizaban a aquellos que instrumentarían ese conocimiento en beneficio de su propio poder, entretenimiento y riqueza. Los habitantes de las casas de los aristócratas tampoco tenían idea de cómo una estatua mágica podía hablar y transmitir información, pues, tras ella, se escondía, con toda su opacidad, un sistema de escucha.

Nadie se preocupaba, por tanto, de una estatua porque, simplemente, no la percibían como un instrumento de escucha sino como un elemento decorativo. En los tiempos del storytelling, las fake news y las realidades alternativas, es más cierto que nunca que todo es percepción.

Los papeles

Pensemos en un ejemplo de cómo percibimos una situación de riesgo. En una manifestación pacífica a la que hemos acudido por un impulso tuitero, un policía se nos acerca y nos pide de manera seca que le mostremos nuestro DNI. La primera reacción sería temer las consecuencias, desde una multa hasta una detención. Algunos más concienciados —o con más ganas de discutir que nosotros— se indignarían por lo que considerarían un atentado a su libertad de manifestarse pacíficamente. Recordemos la oposición pública que la Ley de Seguridad Ciudadana, llamada por sus opositores Ley Mordaza, tuvo en un momento de gran conflictividad social. Con su aprobación se limitaba el poder de los manifestantes y se ampliaba la capacidad de la policía para sancionar conductas que o bien eran impunes o bien estaban mejor ubicadas, desde un punto de vista de las garantías legales, en el Código Penal.

Volvamos a la manifestación, pacífica pero sin permiso, en la que estamos participando. La policía tiene recursos limitados y no pueden identificar a todos y cada uno de los manifestantes. Aprovechándonos de esta limitación de recursos, conseguimos llegar al final sin ser identificados. Nadie nos ha pedido el DNI, así que nos volvemos a casa tranquilos con la satisfacción del deber democrático cumplido. Craso error.

De vuelta a casa, en el metro, no somos conscientes de que hemos sido identificados probablemente de manera más eficaz que a aquellos a los que la Policía ha parado delante de nosotros. Durante toda la manifestación, hemos llevado un identificador en el

bolsillo que puede ser rastreado eficazmente usando una tecnología disponible y opaca conocida como buscador de IMSI,[23] también llamado Stingray[24] por uno de sus modelos más comunes. Esta tecnología permite convertir un teléfono en un identificativo más preciso que un carnet de identidad con la ventaja de que su uso no es percibido por su propietario. Incluso alguno de los modelos más evolucionados permiten interceptar datos como el contenido de las llamadas, de los mensajes de texto o del tráfico de internet, e incluso editar las comunicaciones o bloquear el servicio a determinados usuarios o a todos los participantes en una manifestación, lo que puede ser muy útil si se quiere evitar que los grupos se comuniquen y organicen entre sí.[25] Estos dispositivos de vigilancia, que no dejan rastro en los teléfonos intervenidos, se han utilizado con poca o ninguna supervisión judicial[26] en el Reino Unido, Estados Unidos[27] y China, pero también en diversos países europeos como Alemania, Francia y, parece que también España.

Así que mientras seguimos con nuestra vida anodina alguien ha documentado, con todo detalle, nuestra participación en la manifestación sin entrar en nuestro espacio físico, sin ofendernos o indignarnos y, sobre todo, sin nuestro conocimiento. Desde un punto de eficiencia, resulta más útil identificar a miles de personas con reducidos medios, que hacerlo mediante la filiación en persona, que proporciona menos información con un coste social y económico muy elevado. Sin mencionar la riqueza de los datos que se recaban cuando la vigilancia es electrónica.

GenZ y la intimidad cero

Esta ausencia de percepción de peligro o de riesgo se aprecia en todo nuestro comportamiento online. Creemos que cuando compartimos nuestros datos en una red social o en un servicio, el problema está en configurar adecuadamente las preferencias de privacidad que el propio prestador nos proporciona. Según un

estudio de la Pew Research Center,[28] la generación X (la nacida después de 1995) ha aprendido de sus hermanos millennials a gestionar su reputación online, limitando el acceso a sus perfiles y evitando, sobre todas las cosas, tener en el perfil «bueno» a sus padres como amigos. De este estudio,[29] resulta sorprendente la candidez de unos y otros, y la confianza en la buena fe de las empresas que les prestan el servicio.

La Generación Z está, en general, más preocupada por la privacidad que los millennials, pero menos que la Generación X o los Baby Boomers. Sin embargo, esta preocupación no se traslada a los pagos con aplicaciones móviles o cuando se usan las redes sociales. La Generación Z trabaja en la nube y desde la movilidad, y confiaría en sus teléfonos inteligentes y en las organizaciones que les facilitan, gratuitamente, servicios en la nube y aplicaciones móviles, sin limitación. El uso intensivo de aplicaciones como Snapchat, Secret y Whisper por parte de la Generación Z se usa como ejemplo de que esta generación daría una mayor importancia a la privacidad, tras haber aprendido de los efectos que una exposición incontrolada habría tenido para sus hermanos mayores, los millennials. Según esta apreciación, los riesgos e inconvenientes que implican compartir toda su información en internet habrían calado en ellos, pues han pasado a usar plataformas que les permitirían gestionar su información de manera efímera y segura.

Esta afirmación no está carente de aristas, ya que esas aplicaciones, si bien evitan la sobreexposición frente a terceros de la intimidad de sus usuarios, permiten a los proveedores de esos servicios acceder a información sensible y a datos de sus usuarios que, tras ser tratados, pueden ser cedidos a organizaciones que toman decisiones con base a esos datos, decisiones con un indudable impacto en su vida, patrimonio, economía y perspectivas laborales. Esto es especialmente importante si se considera que la Generación Z es la primera generación que vive totalmente en la nube: allí guardan sus fotos, su información sensible, allí trabajan de manera colaborativa usando herramientas facilitadas por los OTT.

Esta actitud tiene un importante impacto en cómo la Generación Z percibe la privacidad, entendida como limitación del acceso de la información que publican en redes, pero con una total despreocupación de los efectos que tiene para su intimidad y desarrollo personal futuro el ceder toda su información a corporaciones que regalan servicios a cambio de datos. Es mayor su conciencia sobre la reputación online y la gestión de su identidad digital, pero exhiben, como todos nosotros, una despreocupada ignorancia sobre la trascendencia del tratamiento de sus datos personales por las organizaciones y redes sociales

La Generación Z es la primera que carece de intimidad, la primera que tiene huella digital desde el día de su nacimiento, incluso antes, si una ecografía compartida por WhatsApp cuenta. Desde que nacen sus fotos, su estado de ánimo, su estado físico se comparte entre familiares y amigos bajo la percepción de que es una comunicación privada que sólo ven los destinatarios. Sin embargo, como ya sabemos, esos datos compartidos alimentan la máquina de dibujar a su gemelo digital: tanto las publicaciones de padres orgullosos de las fotos de sus hijos, como el comportamiento que del menor se deduce a través del uso de aplicaciones móviles sin supervisión, alimentan el doble virtual de una personalidad que se encuentra en formación. Su habilidad espacial jugando, o su pasividad ante canales de YouTube o Disney, se usan para determinar su capacidad cognitiva y el éxito o fracaso futuro en los estudios. No estamos lejos de valorar si vale la pena gastar recursos públicos en niños que prevemos no van a ser muy brillantes en los estudios. Todo ello, sin entrar a valorar las demandas colectivas contra YouTube o Disney por vulnerar la privacidad de menores[30] o las aplicaciones móviles espías dirigidas, específicamente, para rastrear el comportamiento de los niños y adolescentes.[31]

El juego de las percepciones

En esa extraña percepción del riesgo que tenemos con nuestros dispositivos y nuestros perfiles sociales, no percibimos cuánto de nuestra vida privada exponemos. Mantener fuera del alcance de otros no deseados aquellas partes más sensibles de nosotros es cada vez más ilusorio. Todos hemos escuchado, leído o comentado cómo la cesión de nuestros datos puede ser peligrosa o nos puede perjudicar. Sólo puedo imaginar la cara de los usuarios españoles que se vieron expuestos tras el ataque sufrido por la web de ligue para infieles Ashley Madison.[32] En todos los medios de comunicación se señalaron las instituciones y empresas en las que muchos trabajaban e, incluso, se representaron en un mapa visual por toda la geografía nacional. Un momento de bochorno personal e intransferible que, sin embargo, dudo que haya influido en las altas en otros servicios de citas como Adopta a un tío, Tinder o Grindr, orientado éste a la comunidad LGTBI.

En este juego de las percepciones, creemos que un ataque informático es algo excepcional que no tiene por qué afectar el uso de nuestros dispositivos y servicios. También creemos que nuestros datos están seguros, guardados en una caja fuerte virtual, inaccesibles para los piratas o terceros indeseables.

De nuevo, la percepción es errónea. Tomemos como ejemplo Grindr.

Grindr, que cuenta con más de 3,6 millones de usuarios activos diarios en todo el mundo, ha estado facilitando información sobre si sus usuarios eran seropositivos o no a las entidades que les ayudaban, al parecer, en el desarrollo de la aplicación. Apptimize y Localytics tuvieron acceso a la información que los usuarios de Grindr incluyeron en sus perfiles, entre la que se contaba su estado de VIH y la «última fecha de prueba».[33] Debido a que la información sobre el VIH se envía junto con los datos de GPS de los usuarios, ID de teléfono y correo electrónico, estas entidades podrían identificar a usuarios específicos y su estado de contagio. El problema, al parecer, es que la información sensible no estaba desagregada y

cualquier entidad que prestase servicios a Grindr y que tuviera acceso al set de datos también tenía acceso a los de salud de sus usuarios.

Ninguno de los usuarios de Grindr esperaría que su estado de contagio de VIH se compartiese más que con aquellos contactos elegidos por ellos dentro de la comunidad. Ni que ese dato estuviese disponible junto con su sexualidad, estado de relación, etnia e ID de teléfono.

Cuando damos nuestros datos, de manera consciente o inconsciente, confiamos en una empresa de la que no sabemos absolutamente nada más que su nombre. No sabemos si tiene medios para mantener la información segura, ni si tiene una sede física con trabajadores, o si, por el contrario, es una creación hecha desde un garaje y compartida con desarrolladores diseminados por todo el orbe. Ni lo sabemos, ni nos importa, ni llevamos a cabo la más mínima investigación. Nos fiamos porque nos lo dice nuestra red de amigos y familiares, porque lo dice nuestra comunidad, que sabe tan poco de lo que recomienda como nosotros mismos.

Porque, en realidad, tenemos una ausencia de percepción de peligro o, dicho de otro modo, percibimos el riesgo de que ocurra algo negativo como un evento lejano y poco probable. Es lo que llamamos la «paradoja de la privacidad».

La paradoja de la privacidad

La paradoja de la privacidad describe la discrepancia entre la actitud del usuario y su comportamiento real en relación con la privacidad online, lo que finalmente resulta en una dicotomía entre las actitudes de privacidad y el comportamiento real:[34] si bien los usuarios afirman estar muy preocupados por su privacidad, no hacen nada para protegerla.[35]

Así, aunque muchos usuarios muestran un interés teórico en su privacidad y mantienen una actitud positiva hacia el «comportamiento de protección de la privacidad», rara vez se traduce en una conducta protectora real.[36] Además, si bien existe la intención de limitar la divulgación de datos, la divulgación real a menudo excede significativamente la intención.[37] Diversas investigaciones —que el lector encontrará en las notas— demuestran que las decisiones de privacidad concretas y la conciencia del riesgo abstracto no son intercambiables. Las decisiones de privacidad no varían de acuerdo con las preferencias modificadas, lo que podría explicar la disparidad entre las preferencias establecidas para la privacidad y el comportamiento real.[38]

Aunque los usuarios son conscientes de los riesgos de privacidad en internet, tienden a compartir información privada a cambio de servicios personalizados.[39] Un cierto grado de percepción de riesgo[40] implica un mayor conocimiento de las estrategias de protección de la privacidad, pero parece ser un motivador insuficiente para aplicar tales estrategias.[41]

En el contexto de las actividades de las redes sociales, se observa un patrón similar. La utilización de varias estrategias de protección de la privacidad, como limitar el acceso al muro, restringir funciones tales como etiquetar fotografías o enviar mensajes privados en lugar de publicar contenido abierto que hemos visto que aplica la Generación Z, está diseñado para controlar el flujo de información entre amigos y compañeros. Sin embargo, tales estrategias muestran poca preocupación por la recopilación de datos por parte de la propia entidad que presta el servicio.[42] De ser cierta la preocupación que se predica, debería conducir a la provisión restringida de información en las redes sociales; sin embargo, se puede observar el efecto inverso, ya que muchos usuarios brindan información personal aparentemente sin ninguna vacilación.[43] Por otro lado, a la hora de descargarnos una aplicación en el móvil, pesa más la información que nos facilitan nuestros

amigos y familiares y la de la propia tienda de aplicaciones que la información real sobre la explotación de datos personales por parte de terceros.[44]

Parece que somos capaces de articular de manera teórica nuestras necesidades de privacidad, pero que la decisión real de usar una aplicación o darnos de alta en un servicio depende de otros factores como el deseo de pertenencia, de estar conectados, usar algo útil o estar en la onda.[45] La privacidad no cuenta en realidad entre los elementos a considerar a la hora de tomar decisiones sobre teléfonos inteligentes, aplicaciones o servicios, dando lugar a una conciencia de privacidad fallida. El usuario hace una evaluación de riesgo-beneficio en la que la evaluación del riesgo da un resultado nulo o insignificante, esto es, no hay un análisis de riesgos o, en los pocos casos en los que se da, se considera el riesgo como despreciable al analizarlo desde un punto de vista irracional y poco informado.

En definitiva, somos conscientes de los riesgos pero no queremos que la fiesta se acabe.

4

Google, el memorioso

Memoria y pasado

Durante muchos años de mi vida era capaz de parar, mirar la fecha y saber qué había hecho ese mismo día un año antes. Me retaba a mí misma haciendo este ejercicio en fechas anodinas y sin ningún valor ni social ni personal. Siempre lo recordaba todo a la perfección. Cualquiera podría cuestionar esta habilidad circense diciendo que no había prueba objetiva de mi acierto, que podría hacerme trampas al solitario.

Las geishas llevan un diario para saber qué kimono usan con cada cliente o grupo de clientes para no repetir el mismo *uchikake* en cada encuentro. De no ser así, podría dar la impresión de que su *okiya* no tiene medios y siempre viste el mismo atuendo. Si hace tiempo que no ves a alguien y le ves vestido igual que la última vez, es más que probable que tu primera reacción sea pensar que no tiene nada más en el armario, ignorando el hecho de que no le has visto en los meses intermedios. De aquí surgió mi costumbre de almacenar en la memoria, como una geisha, la ropa que llevaba en cada ocasión y con cada persona para nunca repetirme y dar la impresión de contar con un fondo de armario eterno. Esta memoria visual tan intensa no es habitual y tampoco es eterna. Un día ya no fui capaz de recordar lo que había hecho el año anterior ni cómo había ido vestida con tal cliente o cual amigo. Lo viví como una

pérdida irreparable producto de los años, aunque bien podría deberse a un funcionamiento más eficiente de mi memoria. O a que no me pasaba nada relevante o que, por el contrario, era en mi juventud cuando lo relevante se daba tan poco que era siempre memorable.

Sea lo que sea el motivo lo cierto es que ya no recuerdo ni cómo salí ayer a la calle... pero Google sí. El 28 de septiembre de 2018, Google cumplió veinte años. En este tiempo, corto o eterno, la compañía estrella de Alphabet se ha convertido en nuestra memoria externa, en nuestro segundo cerebro como si del propio Ireneo Funes[46] retratado por Borges se tratase.

Nuestro Ireneo, como Google, no puede olvidar. Cuenta con un extraordinario don: una memoria del todo. Sufre de hipermnesia, un síntoma del síndrome del sabio y, para Borges, esta pieza literaria es «una larga metáfora del insomnio» pero también de la carencia de inteligencia.[47] Borges considera que la memoria nada tiene que ver con aquélla, y en el caso de Funes resulta incluso un obstáculo para su habilidad de pensar.[48]

Google no duerme y almacena hasta la basura de apariencia más inútil, pero a diferencia de Funes cuenta con una inteligencia capaz de hacer un uso eficiente de todo lo que recuerda.

Dicen que la distancia es el olvido, pero en internet, por muy lejano que sea el servidor en el que alojemos nuestros datos o la red social en la que expresemos nuestro yo más carnavalero, el olvido no llega nunca a producirse.

Desde que la información de clientes y usuarios se ha convertido en una materia prima para las empresas, los riesgos para la intimidad que el tratamiento de sus datos conlleva van en aumento. La recogida y tratamiento de los datos personales de clientes y usuarios de los servicios web y de comercio electrónico, de los servicios en la nube (cloud computing), en movilidad (smartphone) o de las cada vez más extendidas redes sociales, se está revelando como un aspecto especialmente sensible. A ello se

une el carácter transnacional de internet y la complejidad de aplicar normas homogéneas en territorios con tradiciones jurídicas y niveles de protección legales diferentes.

«Odia el delito y compadece al delincuente», decía Concepción Arenal, inaugurando una nueva etapa del tratamiento criminal y penitenciario en España, pero, sobre todo, una manera de ver al individuo: como un ser cambiante, resultado de sus circunstancias, que no sólo podía evolucionar, sino que debía ser ayudado en la vía de la redención.

Producto de esta manera de pensar son nuestras leyes penales o el hecho, que pudiera ser poco importante pero que es trascendente, del borrado de antecedentes penales. Así evolucionamos desde la sociedad que unía a la pena que debía cumplirse el escarnio público (de ahí la expresión «poner en la picota») y el estigma para un individuo y sus descendientes, a una sociedad moderna en la que se cree en la evolución del individuo, la redención y las segundas oportunidades. Todos estábamos contentos, ya nadie tendría que llevar una letra escarlata en el pecho, pero llegó internet, los algoritmos de buscador, las redes sociales y la hiperconexión móvil.

Todas las entidades, desde los boletines oficiales hasta la prensa escrita, se lanzaron a publicar sus ediciones digitales y a cargar su hemeroteca en internet, permitiendo que cualquier buscador indexara hasta la noticia más irrelevante. En internet todo permanece, por siempre, fuera del contexto en que se hizo y sin que seamos conscientes de ello. No recuerdo qué busqué en Google hace una semana, pero Google sí. Así, nuestra vida aparece documentada por nosotros o por terceros hasta el último detalle en una línea infinita de código con todos nuestros datos, secretos y preocupaciones.

En este panorama, ni redención ni arrepentimientos parecen posibles. Tan es así que si se mira desde la óptica de la autocensura que una buena práctica aconseja, esta situación supone, *de facto*,

una inaceptable limitación a nuestra libertad de expresión, movimiento o pensamiento. O te autocensuras o asumes las consecuencias.

Pero aún nada dicen de los datos que Google o Facebook guardan de todos para siempre, de si serán borrados algún día, de a quién los venden y por cuánto, de qué minería de datos les aplican.

Por eso el olvido se ha convertido en un derecho y por eso Funes vivía la memoria como una maldición.

Toda la memoria del mundo

Google usa el tremendo alcance de sus productos para recopilar información detallada sobre los comportamientos online y en el mundo real de todos sus usuarios y de quien no lo son directamente, y los utiliza para dirigirse a ellos con publicidad pagada por terceros.[49] Y que los usuarios aceptan tácitamente dar sus datos a cambio de unos servicios extraordinariamente fiables, útiles y de enorme calidad; además, que recibir algo de publicidad no constituye tampoco nada noticiable. Lo que, tal vez, no es tan obvio y no entra dentro de este pacto tácito, que parece ser pacíficamente aceptado por todo el mundo, es la profundidad y la extensión de la vigilancia ejercida por Google y la persistencia en el tiempo de dicha vigilancia. Lo que Google sabe, lo que recuerda de nosotros, tiene una extensión tal que nuestro querido y amargado Funes quedaría reducido a un mero charlatán de feria.

No deja de sorprender cuánta información es capaz de recopilar de un solo usuario. El investigador Douglas Schmidt de la Universidad de Vanderbilt[50] ha publicado un estudio en el que demuestra cómo y cuántos datos recopila cada dispositivo en un uso normal[51] a través de cada uno de sus servicios. Google no es sólo propietario del primer navegador web del mundo,[52] sino que proporciona la aplicación y la plataforma móvil más instalada del universo conocido.[53] Junto con los servicios de vídeo, correo

electrónico y la aplicación de mapas de Google, cuenta con más de mil millones de usuarios activos mensuales[54] sin mencionar todo el catálogo de otros servicios también prestados por esta compañía.

Google recopila datos de usuarios de varias maneras: de manera consciente o activa, esto es, cuando el usuario comunica directa y conscientemente información a Google, como, por ejemplo, al iniciar sesión en cualquiera de sus aplicaciones como YouTube, Gmail, buscador, etc., y también, como todos los servicios de prestados sin pago, de manera pasiva o inconsciente. Este último supuesto se refiere a las aplicaciones diseñadas a propósito para recopilar información mientras se están ejecutando, casi siempre sin el conocimiento del usuario. Google recopila datos pasivos desde cualquiera de sus plataformas (como Android y Chrome), aplicaciones (como el buscador, YouTube o Maps), herramientas de edición y herramientas del anunciante (por ejemplo, AdMob, AdWords, Google Analytics, AdSense). Tanto los usuarios como los comentaristas pasan por alto el alcance y la magnitud de la recopilación de datos pasivos que Google realiza. De ahí la sensación de seguridad en el uso que se aprecia en los usuarios, que no ven un peligro específico al no ser conscientes del trasiego de datos que se produce.

El estudio de Schmidt, publicado en agosto de 2018, analiza los datos que Google recaba de sus usuarios y el impacto de los mismos. Usa varias fuentes, algunas propietarias de Google como «Mi actividad»[55] o «Takeout tools»[56] —que describen la información recopilada durante el uso de sus productos—, y otras no. Schmidt revisa los datos que se envían desde los dispositivos del usuario a los servidores de Google durante el uso de sus productos o de terceros, así como las políticas de privacidad de la compañía, tanto generales como específicas para cada producto. Mediante el uso combinado de estos recursos, proporciona una visión única e integral de los enfoques de recopilación de datos de Google y profundiza en los tipos específicos de información que recopila de los usuarios.

Un día de tu vida con Google

Para ilustrar la multitud de puntos de contacto entre Google y un individuo, así como el alcance de la información recopilada durante estas interacciones, se diseñó un experimento en el que un investigador llevó un dispositivo de teléfono móvil Android durante sus actividades cotidianas en una jornada cualquiera. El teléfono móvil se borró realizando un restablecimiento de datos a los valores de fábrica y se configuró como un nuevo dispositivo para evitar que la información previa del usuario asociada con el dispositivo distorsionara el resultado. Se creó una nueva cuenta, por lo que Google no tenía conocimiento previo del usuario y no tenía intereses publicitarios asociados con la cuenta. La SIM también era nueva. El investigador pasó un día normal utilizando el teléfono móvil asociado con la nueva cuenta.

Los datos recopilados por Google se verificaron utilizando dos herramientas provistas por Google: My Activity y Takeout. La herramienta My Activity muestra datos recopilados por Google de cualquier actividad relacionada con el buscador, uso de aplicaciones de Google (por ejemplo, reproducciones de vídeos de YouTube, buscador de mapas, asistente, etc.), visitas a páginas web de terceros (mientras está conectado a Chrome) y clics en los anuncios. La herramienta Google Takeout proporciona una información más completa sobre todos los datos históricos del usuario, recopilados a través de las aplicaciones de Google (por ejemplo, contiene todos los mensajes de correo electrónico anteriores en Gmail, consultas de buscador, colección de ubicaciones y vídeos de YouTube).

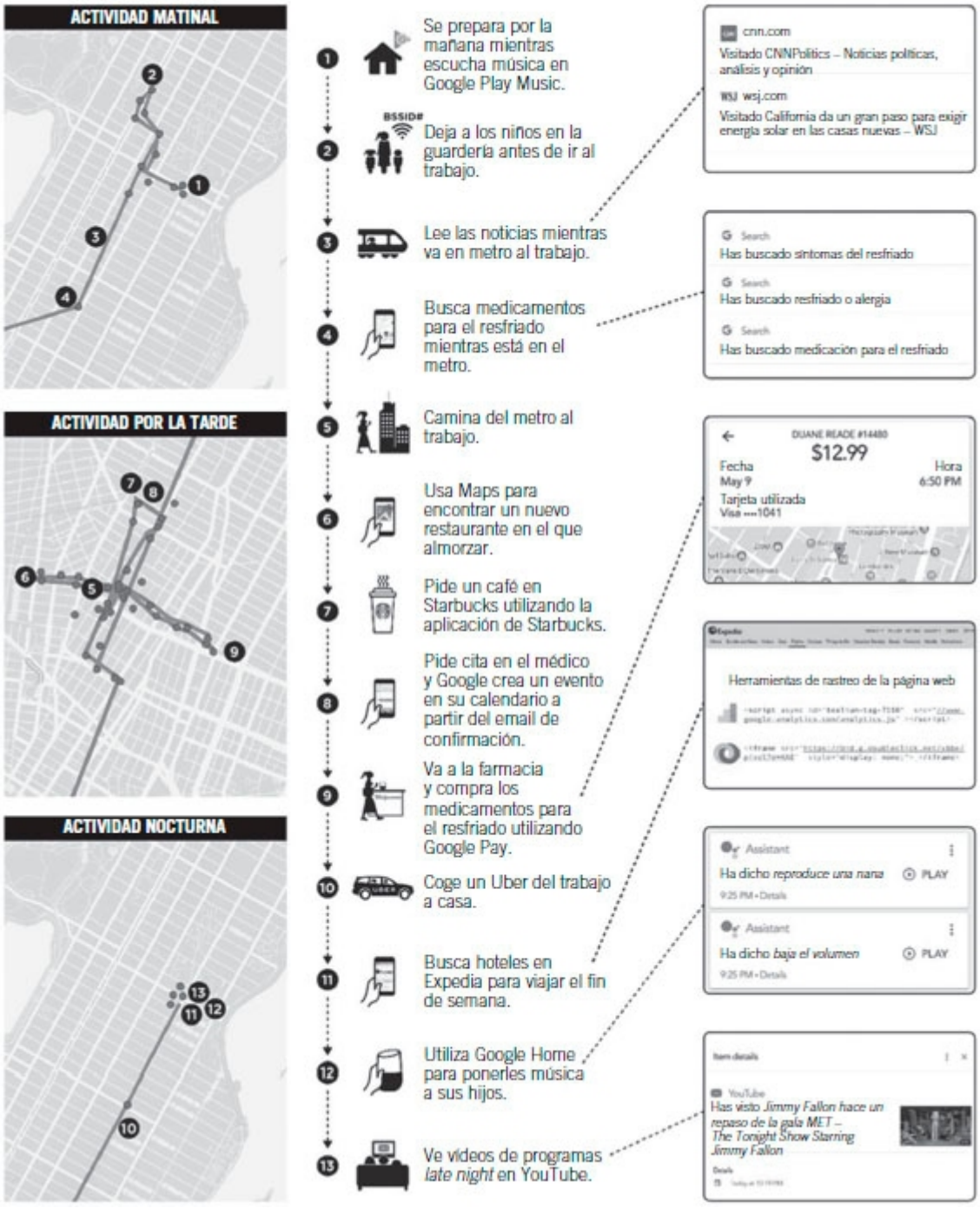
Recopilaron e identificaron dos tipos de datos: los activos y los pasivos. Los activos serían la información intercambiada directamente entre el usuario y un producto de Google, mientras que los pasivos serían toda la información intercambiada en segundo plano por el dispositivo y las aplicaciones sin ninguna notificación obvia al usuario.

Por ejemplo, mientras el usuario se prepara para salir de casa, escucha Google Play Music. Los datos activos serían sus preferencias musicales y los pasivos su geolocalización. A continuación, lleva a los niños al colegio y va al trabajo en metro, y le cuenta a Google las noticias que lee en el móvil, y de modo pasivo, la ruta escogida, el medio de transporte y el tiempo de desplazamiento.

Durante ese día, Google recopiló numerosos datos relacionados con la actividad del usuario, como su ubicación, sus rutas, productos adquiridos o la música que escuchó a lo largo del día. Lo que resultó más sorprendente es que Google recopiló o infirió más de dos tercios de la información a través de medios pasivos. Al final del día, identificó los intereses del usuario con una precisión notable.[57] Todo esto es posible gracias al sistema operativo Android utilizado por fabricantes de teléfonos inteligentes del mundo y su estrecha conexión con el ecosistema de Google a través de Google Play Services. Android ayuda a Google a recopilar información personal del usuario (por ejemplo, nombre, número de teléfono móvil, fecha de nacimiento, código postal y, en muchos casos, número de tarjeta de crédito), actividad en el teléfono móvil (por ejemplo, aplicaciones utilizadas, sitios web visitados) y coordenadas de ubicación. En un segundo plano, Android envía con frecuencia la ubicación del usuario de Google y la información relacionada con el dispositivo, como el uso de aplicaciones, los informes de fallos, la configuración del dispositivo, las copias de seguridad y varios identificadores relacionados con el dispositivo.

FIGURA 1

Un día en la vida con Google



Fuente: Schmidt, Douglas C. Google Data Collection. Vanderbilt University, agosto de 2018.

Al final de la jornada, Google le atribuyó al usuario varios «intereses», ocho de los cuales coincidían de manera estrecha con sus actividades.

Este mismo ejercicio fue realizado por un colaborador del medio online Xataka[58] y, el resultado, como si de un clickbait se tratara, sin duda le sorprendió. En su análisis usó las mismas herramientas Mi actividad y Takeout que ofrece Google y llegó a la conclusión de que «Google es el *Gran Hermano* definitivo».[59]

Aunque las herramientas de Mi actividad y Takeout tools son útiles para evaluar la cantidad de datos activos recopilados después de que un usuario interactúe con los productos de Google, no dibujan una imagen completa del tamaño y la escala de la recopilación de datos que la compañía de Alphabet realiza. Una comprensión exhaustiva de su amplitud requiere una revisión de las políticas de privacidad específicas de productos de Google, así como análisis del tráfico de datos reales pasados a los servidores de Google durante las instancias de interacción del usuario con sus servicios.

El efecto Martini: donde estés, a la hora que estés

Las plataformas Android y Chrome recopilan meticulosamente la información de ubicación y movimiento de los usuarios utilizando una variedad de fuentes. Por ejemplo, una evaluación de «ubicación aproximada» se hace mediante el uso de coordenadas GPS. La precisión de ubicación del usuario puede mejorarse aún más («ubicación precisa») mediante el uso de identificadores de la torre más cercana y los puntos de acceso wifi.[60]

Google puede determinar con un alto grado de confianza si un usuario está quieto, caminando, corriendo, montando en bicicleta o viajando en un tren o en un automóvil. Esto se consigue mediante el seguimiento de las coordenadas de ubicación de un usuario de un

dispositivo móvil Android a intervalos de tiempo frecuentes en combinación con los datos de los sensores integrados (como un acelerómetro) en teléfonos móviles.

Los teléfonos Android también pueden usar información de las balizas Bluetooth registradas con la Proximity Beacon API de Google. Estas balizas no sólo proporcionan las coordenadas de geolocalización del usuario, sino que también pueden identificar los niveles exactos del piso en los edificios. Es difícil para un usuario móvil de Android deshabilitar el seguimiento de su ubicación.[61] Una investigación de Associated Press descubrió que muchos servicios de Google en dispositivos Android e iPhone almacenan tus datos de ubicación incluso si ha utilizado la configuración de privacidad que, en teoría, impediría que Google te ubicase, lo que confirmaron diversos investigadores de la Universidad de Princeton. Google pediría permiso para la información de geoposicionamiento en Google Maps. Si se acepta, Google Maps lo almacenará en una «línea de tiempo» que describe con total precisión los movimientos diarios. La compañía permite «pausar» el «historial de ubicación» con lo que, en teoría, Google dejaría de almacenar esa información.

Lo revelado por AP es que, incluso en pausa, determinadas aplicaciones de Google almacenarían sin preguntar y sin permiso de manera automática los datos de ubicación con su marca de tiempo. Algunos servicios en Android y iPhone almacenan automáticamente nuestros movimientos incluso después de detener la configuración del «historial de ubicaciones». La filial de Alphabet acecha silenciosamente detrás de algunas de las características más innovadoras de su sistema operativo móvil Android: los teléfonos que usan Android transmiten silenciosamente datos a los servidores de Google, que incluyen desde coordenadas GPS hasta redes wifi cercanas, presión barométrica —que pueden usar la presión del aire para calcular dónde se encuentra una persona en un edificio de varios pisos— e incluso una estimación de la actividad actual del titular del teléfono. Aunque el producto detrás de esas transmisiones es opcional, para los usuarios de Android puede ser difícil de evitar

e incluso más difícil de entender. El «historial de ubicaciones» es un producto de Google heredado de Google Latitude, ahora desaparecido. Lanzada en 2009, esa aplicación permitía a los usuarios transmitir constantemente su ubicación a sus amigos. En la actualidad, dicho historial se utiliza para potenciar funciones como predicciones de tráfico o recomendaciones de restaurantes. Si bien no está habilitado en un teléfono Android de forma predeterminada, la activación del historial de ubicaciones está sutilmente configurada para aplicaciones como Google Maps, Fotos, o el asistente personal de Google. Al probar varios teléfonos, Quartz[62] descubrió que ninguna de esas aplicaciones usa el mismo lenguaje para describir lo que ocurre cuando el historial está habilitado, y ninguna indica explícitamente que la activación permita que todas las aplicaciones de Google, no sólo la que solicita permiso, accedan a los datos del historial de ubicaciones.

El organismo de control, la Comisión Australiana de Competencia y Consumo (ACCC), está investigando las alegaciones de la compañía tecnológica Oracle: aseguran que los dispositivos Android envían a Google información sobre búsquedas, lo que se está viendo, y sobre datos de ubicación, incluso si los servicios de localización están apagados. Oracle también afirma que los millones de usuarios de Android del país gastan sus datos y, por tanto, subvencionan a Google al pagar la conexión que la compañía utiliza.

El 4 de diciembre de 2017, el Tesorero, el MP Hon Scott Morrison, se dirigió la ACCC para realizar una investigación sobre plataformas digitales. La investigación analiza el efecto que los motores de buscador digital, las plataformas de redes sociales y otras plataformas de agregación de contenido digital tienen sobre la competencia en los mercados de servicios de publicidad y medios. En particular, la investigación analiza el impacto de las plataformas digitales en el suministro de noticias y contenido periodístico y sus implicaciones para los creadores de contenido de medios, anunciantes y consumidores. La investigación se inició antes de que se revelara que Facebook había permitido que Cambridge Analytica

recolectara los datos de 50 millones de usuarios de Facebook sin su permiso, y encendió un debate global sobre la privacidad de éstos. Las prácticas de datos de Facebook están siendo investigadas en otras partes del mundo.[63]

Oracle se dirigió a la ACCC como parte de esta investigación, y reveló que Google puede rastrear dispositivos Android sin servicios de ubicación activados, tarjetas Sim o aplicaciones instaladas, a través de direcciones IP asignadas, puntos de acceso wifi y torres móviles que muestran dónde se conectan los dispositivos o intentan conectarse. Los teléfonos Android también tienen dispositivos barométricos que pueden usar presión de aire para ubicar a una persona dentro de un edificio. La ACCC investigará los informes sobre que Google recolecta enormes cantidades de datos de los teléfonos Android, incluida información detallada de la ubicación.[64]

Junto con las autoridades australianas, Arizona investigará a Google por almacenar las ubicaciones de sus usuarios que no lo desean. El descubrimiento de hace unas semanas comienza a tener consecuencias, con la fiscalía de este estado siendo la primera en abrir dirigencias.[65]

Lo que Google sabe aunque no lo uses

Android y Chrome son las plataformas clave de Google que ayudan en la recopilación de datos de usuarios importantes debido a su amplio alcance y frecuencia de uso. En enero de 2018, Android tenía el 53 por ciento del mercado total de sistemas operativos móviles de Estados Unidos (Apple iOS tenía un 45 por ciento). Desde mayo de 2017, se han venido activando más de 2 mil millones de dispositivos Android al mes en todo el mundo. El 60 por ciento de los navegadores instalados en el mundo son Chrome, que ya cuenta con más de mil millones de usuarios activos mensuales.

Ambas plataformas facilitan el uso de Google y contenido de terceros y, por tanto, proporcionan acceso a Google a una amplia gama de información personal, de actividad web y de ubicación.

El navegador Chrome permite a Google recopilar datos de usuarios desde dispositivos móviles y de escritorio, con más de dos mil millones de instalaciones activas en todo el mundo.[66] Recopila información personal (por ejemplo, cuando un usuario completa formularios en línea) y la envía a Google como parte del proceso de sincronización de datos. También realiza un seguimiento de las visitas a las páginas web y envía las coordenadas de ubicación del usuario. Si bien Chrome no exige compartir información personal adicional recopilada de los usuarios, sí tiene la capacidad de capturar dicha información. Por ejemplo, Chrome recopila un rango de información personal a través de su función de «autocompletar», y tales campos de formulario generalmente incluyen nombre de usuario, dirección, número de teléfono, nombre de usuario y contraseñas.

Si el usuario inicia sesión en Chrome utilizando la cuenta de Google y habilita su función «sincronizar», esta información se envía y se almacena en los servidores de Google. Recientemente se ha comprobado que Google Chrome fuerza a los usuarios a estar conectados a Gmail. La versión 69 del navegador[67] lleva la integración de las cuentas de Google activando automáticamente la conexión a Gmail si se está conectado al navegador. Antes de esta versión, era posible consultar el correo sin identificarse en el navegador y, así, personalizar aún más la navegación.

Chrome también podría aprender sobre los idiomas que habla una persona durante sus interacciones con su función de traducción, que está habilitada de forma predeterminada. Además de los datos personales, tanto Chrome como Android envían información a Google sobre la navegación web de un usuario y las actividades de la aplicación móvil, respectivamente. Cualquier visita a la página web se rastrea automáticamente y recopila credenciales de usuario de Google si el usuario ha iniciado sesión en Chrome.

Chrome también recopila información sobre el historial de navegación de un usuario, contraseñas, permisos específicos del sitio web, cookies, historial de descargas y datos agregados.

Android envía actualizaciones periódicas a los servidores de Google, incluidos el tipo de dispositivo, el nombre del proveedor del servicio móvil, los informes de fallos e información sobre las aplicaciones instaladas en el teléfono. También notifica a Google cada vez que se accede a una aplicación en el teléfono (por ejemplo, Google sabe cuándo un usuario de Android accede a su aplicación de Uber).

De interés potencialmente mayor, sin embargo, es la información pasiva que estas plataformas recopilan, que va más allá de los datos de ubicación y que permanece relativamente desconocida por los usuarios.

Para evaluar el tipo y la frecuencia de ocurrencia de dicha recopilación con mayor detalle, Douglas Schmidt hizo un experimento en el que monitoreó los datos de tráfico enviados a Google desde teléfonos móviles (tanto Android como iPhone).

En aras de la comparación, este experimento también incluyó el análisis de datos enviados a Apple a través de un dispositivo de iPhone.

Para simplificar, los teléfonos se mantuvieron estacionarios, sin interacción del usuario. En el teléfono Android, una sola sesión del navegador Chrome permaneció activa en segundo plano, mientras que en el iPhone se utilizó el navegador Safari. Esta configuración proporcionó una oportunidad para el análisis sistemático de la recogida de datos de Google desde Android y Chrome, así como la recopilación en ausencia de éstos, esto es, desde un dispositivo iPhone, sin solicitudes de recopilación adicionales generadas por otros productos o aplicaciones, es decir, cuando el iPhone no tiene instaladas apps o se usan servicios de la familia Google como YouTube o Gmail. Resultó que el teléfono Android comunicó un número mayor de datos a las plataforma de Google que el iPhone sin aplicaciones de Google instaladas. Este resultado destaca el

hecho de que las plataformas Android y Chrome juegan un papel importante en la recopilación de datos de Google. Además, la comunicación del dispositivo iPhone con los servidores de Apple fue diez veces menos frecuente que las comunicaciones del dispositivo Android con Google.[68]

Tanto Android como Chrome envían datos a Google incluso en ausencia de interacción del usuario. El estudio de Douglas Schmidt demostró que un teléfono Android inactivo (con Chrome activado en segundo plano) comunica información de ubicación a Google 340 veces durante un período de 24 horas, o a un promedio de 14 comunicaciones de datos por hora. De hecho, la información de ubicación constituyó el 35 por ciento de todas las muestras de datos enviadas a Google. En contraste, al hacer la misma prueba con un dispositivo Apple iOS con Safari (en donde no se usa ni el sistema operativo Android ni el navegador Chrome), Google no pudo recopilar ningún dato apreciable (ubicación o de otro tipo) en ausencia de una interacción del usuario con el dispositivo.

Después de que un usuario comienza a interactuar con un teléfono Android (por ejemplo, se desplaza, visita páginas web, usa aplicaciones), las comunicaciones pasivas a los dominios de los servidores de Google aumentan significativamente, incluso en casos en los que el usuario no usa ninguna aplicación destacada de Google (es decir, sin usar Google Search, YouTube, Gmail o Google Maps). Este aumento se debe principalmente a la actividad de datos del editor y los productos de anunciantes de Google (por ejemplo, Google Analytics, DoubleClick —ahora Google Ad Manager—, AdWords —ahora Google Ads—). Dichos datos constituyeron el 46 por ciento de todas las solicitudes a los servidores de Google desde el teléfono Android.[69] Al usar un dispositivo iOS, si un usuario decide renunciar al uso de cualquier producto de Google (es decir, sin Android, Chrome, o aplicaciones de Google) y visita sólo páginas web que no sean de Google, la cantidad de veces que se comunican los datos a los servidores de Google sigue siendo sorprendentemente alta. Esta comunicación es impulsada por la

publicidad servida al usuario. La cantidad de veces que llaman esos servicios de Google desde un dispositivo iOS es similar a un dispositivo Android. El informe de Schmidt comprobó que la magnitud total de los datos comunicados a los servidores de Google desde un dispositivo iOS es aproximadamente la mitad que la de un Android.

Los identificadores de publicidad (que supuestamente son «anónimos para el usuario» y recopilan datos de actividad en aplicaciones y visitas a sitios web de terceros) pueden conectarse con la identidad de Google de un usuario. Esto es debido a que Android identifica al usuario del móvil y pasa a los servidores de Google la información de publicidad junto con su identidad. Del mismo modo, el ID de cookie de DoubleClick (que rastrea la actividad de un usuario en páginas web de terceros) es otro identificador «anónimo» que Google puede conectar con la cuenta de un usuario si accede a una aplicación de Google desde el mismo navegador

Todo ello demostraría que Google tiene la capacidad de conectar los datos anónimos recopilados a través de medios pasivos con la información personal del usuario.

Y llegó la publicidad

Una fuente importante para la recopilación de datos de actividad del usuario de Google proviene de sus herramientas centradas en anunciantes, como Google Analytics, DoubleClick,[70] AdSense, AdWords y AdMob.[71] Estas herramientas tienen un gran alcance: más de un millón de aplicaciones móviles usan AdMob; más de un millón de anunciantes usan AdWords; más de 15 millones de web usan AdSense; y, por último, más de 30 millones de webs utilizan Google Analytics.

Existen dos grupos principales de usuarios de herramientas centradas en anunciantes de Google:

- **Editores de sitios web y aplicaciones**: organizaciones que poseen webs y crean aplicaciones móviles. Estas organizaciones utilizan las herramientas de Google para:
  1. Ganar dinero, al permitir la visualización de anuncios a los visitantes en sus sitios web o aplicaciones, y
  2. hacer un mejor seguimiento y comprender mejor quién visita sus webs y usan sus aplicaciones. Este seguimiento se realiza sirviendo cookies, ejecutando scripts en los navegadores de los visitantes de la web que ayudan a determinar la identidad de un usuario, rastrear su interés en el contenido y seguir su comportamiento en línea. Las bibliotecas de aplicaciones móviles de Google rastrean el uso de aplicaciones en teléfonos móviles.
- **Anunciantes que usan las herramientas de Google para dirigirse a perfiles específicos para aumentar el rendimiento de sus inversiones en marketing** (los anuncios mejor orientados generan mayores porcentajes de clics y conversiones). Dichas herramientas también les permiten analizar sus audiencias y medir la eficacia de su publicidad digital al rastrear en qué anuncios se hicieron clic, con qué frecuencia, y al proporcionar información sobre los perfiles de las personas que hicieron clic en los anuncios.

Estas herramientas, juntas, recopilan información sobre las actividades de los usuarios en las webs y en las aplicaciones, como el contenido visitado y los anuncios en los que se hace clic. Funcionan en segundo plano, en gran medida imperceptibles para los usuarios.

La información recopilada por dichas herramientas incluye un identificador no personal que Google puede usar para enviar anuncios dirigidos sin identificar la información personal del

individuo. Los identificadores pueden ser específicos del dispositivo o sesión, así como permanentes o semipermanentes. Para proporcionar a los usuarios un mayor anonimato durante la recopilación de información para la orientación de anuncios, Google ha cambiado recientemente hacia el uso de identificadores únicos de dispositivos semipermanentes (por ejemplo, GAID). En otras secciones se detalla cómo estas herramientas recopilan datos de usuarios y el uso de dichos identificadores durante el proceso de recopilación de datos.

No hay datos anónimos

Google recopila datos a través de editores y productos publicitarios y los asocia a una variedad de identificadores semipermanentes y anónimos. Además, tiene la capacidad de asociar estos identificadores con la información personal de un usuario en concreto. Así se deduce de la lectura de la política de privacidad de Google que indica claramente que la compañía puede asociar datos de servicios publicitarios y herramientas analíticas con la información personal del usuario, dependiendo de la configuración de la cuenta de éste. Esta opción está habilitada de manera predeterminada. Además, un análisis del tráfico de datos intercambiado con los servidores de Google identificó dos ejemplos clave (uno en Android y el otro en Chrome) que apuntan a la capacidad de Google de relacionar los datos recopilados anónimamente con la información personal de los usuarios.

El identificador de publicidad móvil puede anonimizarse a través de datos enviados a Google por Android.

Los análisis del tráfico de datos que se comunican entre un teléfono con Android y los dominios del servidor de Google sugieren una forma posible a través de la cual los identificadores anónimos (GAID en este caso) pueden asociarse con la cuenta de Google de un usuario. Esta comunicación particular proporciona una

sincronización de datos de Android a los servidores de Google y contiene información de registro de Android (por ejemplo, registro de recuperación), mensajes del kernel, volcados de memoria y otros identificadores relacionados con el dispositivo.

A través del proceso de registro, Android envía a Google una variedad de identificadores permanentes e importantes relacionados con el dispositivo, incluida la dirección MAC del dispositivo, IMEI/MEID y el número de serie del aparato. Además, estos envíos también contienen la ID de Gmail del usuario de Android. Estos datos permiten a Google conectar la información personal de un usuario con los identificadores permanentes del dispositivo Android. Estos intercambios de datos con los servidores de Google desde un teléfono Android, muestran cómo Google puede conectar información anónima recopilada en un dispositivo móvil Android a través de las herramientas de DoubleClick, Analytics o AdMob con la identidad personal del usuario.[72]

Hasta ahora se ha establecido que Google recopila una amplia variedad de datos de usuarios a través de sus herramientas de publicación y anunciantes, sin un conocimiento directo del usuario. Si bien estos datos se recolectan con identificadores anónimos del usuario, Google tiene la capacidad de conectar esta información con las credenciales personales de un usuario almacenadas en su cuenta de Google.

Vale la pena señalar que la recopilación pasiva de datos de usuario de Google desde páginas web de terceros no se puede evitar utilizando herramientas populares de bloqueo de anuncios, ya que están diseñadas principalmente para evitar la aparición de anuncios mientras los usuarios navegan por páginas web de terceros, no para evitar saber quién navega.

Las muchas caras de Google

Google tiene docenas de productos y servicios en constante evolución. A menudo se accede a estos productos a través de una cuenta de Google, o se la asocia, lo que permite a la empresa vincular directamente los datos de la actividad del usuario de unos productos con otros. Además de los datos de uso del producto, Google también recaba identificadores relacionados con el dispositivo y datos de ubicación cuando se accede a los productos y servicios de Google. Algunas de las aplicaciones de Google (por ejemplo, YouTube, buscador, Gmail y Mapas) son fundamentales para las tareas básicas que muchos usuarios realizan a diario.

1. **Buscador**.[73] Google utiliza sus productos de buscador para recopilar datos relacionados con consultas, historial de navegación y actividad de clic/compra de anuncios. Por ejemplo, Google Finance recopila información sobre el tipo de acciones que los usuarios pueden rastrear, mientras que Google Flights rastrea las reservas de viajes de los usuarios y las solicitudes de búsqueda. Cuando se utiliza el buscador, Google recopila datos de ubicación y registra toda la actividad de búsqueda que realiza un usuario y la vincula a su cuenta de Google si el usuario ha iniciado sesión.

   Además de ser el motor de búsqueda predeterminado en los dispositivos Chrome y Google, el buscador de Google también es la opción predeterminada en otros navegadores y aplicaciones de terceros a través de varios acuerdos de distribución. Por ejemplo, Google se convirtió recientemente en el motor buscador predeterminado en el navegador Firefox de Mozilla en ubicaciones geográficas clave (incluidos Estados Unidos y Canadá). Del mismo modo, Apple cambió de Bing de Microsoft a Google para los resultados de buscador web de Siri en dispositivos iOS y Mac.

2. **YouTube**.[74] La cantidad de contenido subido y visto en YouTube es sustancial: unas 400 horas de vídeo cargadas cada minuto y unos mil millones de horas de vídeo descargadas diariamente. Los usuarios pueden acceder a YouTube a través de ordenadores de escritorio (navegador web), dispositivos móviles (aplicación y/o navegador web) y Google Home (a través de un servicio de suscripción de pago llamado YouTube Red). Google recopila y almacena el historial del buscador, el historial de reproducciones, listas de reproducción, suscripciones y comentarios en vídeos. Toda esta información está marcada con un sello de fecha y hora de cuando tuvo lugar la actividad. Si un usuario entra a su cuenta de Google en cualquier aplicación de la compañía dentro de un navegador (por ejemplo, Chrome, Firefox, Safari), Google reconocerá la identidad del usuario, incluso si se accede a ella a través de un sitio web que no es de Google (por ejemplo, vídeos de YouTube reproducidos por CNN. com). Esta característica permite a Google rastrear el uso que hace de YouTube un usuario a través de múltiples plataformas de terceros. Google también ofrece un producto de YouTube para niños, conocido como YouTube Kids, que pretende ser una versión «amigable para la familia», con funciones de control parental y filtros de vídeo. Google recopila información de YouTube Kids, que incluye el tipo de dispositivo, el sistema operativo, los identificadores únicos del dispositivo, la información de registro y detalles sobre cómo se usó el servicio. Google utiliza esta información para ofrecer anuncios limitados a los que no se pueden hacer clic y que tienen restricciones de formato, duración y servicio del sitio.
3. **Maps** es la aplicación de navegación insignia de Google. Google Maps puede determinar las rutas, la velocidad y los lugares que un usuario visita con frecuencia (por ejemplo,

su casa, el trabajo, restaurantes...). Esta información proporciona a Google una ventana a los intereses de un usuario (por ejemplo, preferencias de alimentos y compras), movimiento y comportamiento. Maps usa la dirección IP, el GPS, la señal del móvil y los datos del punto de acceso wifi para calcular la ubicación de un dispositivo. Los dos últimos se recopilan desde el dispositivo a través del cual se usa Maps y se envía a Google para su evaluación de ubicación a través de su API de ubicación. Esta API proporciona datos detallados sobre un usuario, incluidas las coordenadas geográficas, si el usuario está parado o en movimiento, la velocidad y la determinación probabilística del modo de transporte del usuario (por ejemplo, bicicleta, automóvil, tren, etc.). Maps almacena una línea de tiempo histórica de los lugares visitados por un usuario que inició sesión en Maps utilizando su cuenta de Google. La precisión de la información de ubicación capturada por las aplicaciones de navegación permite a Google no sólo dirigirse a públicos publicitarios, sino también ayudar a entregar anuncios a los usuarios a medida que se acercan a las tiendas. Además, Google Maps usa esta información para generar actualizaciones de tráfico en tiempo real.

4. **Gmail** almacena todos los mensajes (enviados/recibidos), el nombre del remitente, la dirección de correo electrónico y la fecha/hora de los mensajes. Dado que Gmail actúa como depósito de correo central para muchas personas, puede determinar sus intereses escaneando el contenido del correo electrónico, identificando direcciones comerciales a través de sus correos electrónicos promocionales o recibos de ventas enviados a correos electrónicos, aprendiendo sobre los planes de un usuario (por ejemplo, reservas de cena, citas con el médico...). Dado que los usuarios pueden usar su ID de Gmail para otras plataformas de terceros (por

ejemplo, Facebook, LinkedIn), Gmail puede escanear cualquier contenido que provenga de ellos en forma de un correo electrónico (por ejemplo, notificaciones, mensajes).[75]

Desde su inicio en 2004 hasta al menos finales de 2017, Google ha escaneado los contenidos de los correos electrónicos de Gmail para mejorar la orientación de anuncios y los resultados del buscador, así como para filtrar el correo no deseado. En el verano de 2016, Google dio un paso más y cambió su política de privacidad para permitirle combinar datos de navegación web anónimos anteriormente de su filial DoubleClick (que sirve anuncios personalizados en internet) con los datos de identificación personal que Google tiene a través de sus otros productos, incluido Gmail. El resultado fue que los anuncios de DoubleClick que siguen a las personas en la web ahora se pueden personalizar según las palabras clave que usaron en Gmail. También significa que Google ahora puede crear un retrato completo de un usuario por nombre, basado en todo lo que escribe en el correo electrónico, cada sitio web que visita y las búsquedas que realiza. A finales de 2017, Google anunció que interrumpiría la práctica de la personalización de anuncios basada en mensajes de Gmail. Recientemente, sin embargo, la empresa aclaró que todavía estaría escaneando mensajes de Gmail para algunos propósitos.

5. **Google Play Music y Play Movies and TV** son servicios bajo demanda que ofrecen transmisión de música, podcasts, programas de TV y películas. Estos servicios se pueden considerar como el equivalente de iTunes de Apple. Al igual que YouTube, recopilan información sobre búsquedas de usuarios, contenido

comprado/alquilado/reproducido, información sobre la geografía de los usuarios (a través de la dirección IP) e información del dispositivo.

6. **Waze** fue adquirido por Google en 2013 y opera como una subsidiaria de Google. A diferencia de Maps, Waze es una aplicación de fuentes múltiples en la que los datos proporcionados por el usuario (como coordenadas de ubicación GPS, tiempos de viaje, información de tráfico, accidentes, monitoreo policial, carreteras bloqueadas y en construcción) se analizan para proporcionar en tiempo real actualizaciones sobre el estado del tráfico o las condiciones del viaje. Además de la información de ubicación recopilada a través del dispositivo móvil, Waze recopila información sobre el uso de sus servicios desde el dispositivo en el que está instalado, incluido el nombre del móvil, sistema operativo, visitas a páginas web, información visualizada en la aplicación, uso/creación de contenido de la aplicación, y anuncios vistos y clics. Waze funciona como una red social y ofrece a los usuarios la posibilidad de hacerse amigo de otros usuarios y crear una comunidad en línea de conductores locales.[76] Los usuarios pueden vincular la lista de contactos de su teléfono, con Facebook o Twitter. En general, Waze le da a Google la capacidad de acceder a más datos de usuario en tiempo real, así como también información sobre conocidos que pueden o no estar usando una cuenta de Google.
7. **Google Docs y Drive**. Las herramientas de productividad de Google (documentos, hojas de cálculo, presentaciones, formularios y Drive, que son parte del «G Suite» de productos más amplio) son aplicaciones basadas en la nube que utilizan tanto las personas como las empresas. Las políticas de privacidad de Google para G Suite[77] prohíben a Google escanear archivos almacenados con fines publicitarios en las versiones empresariales. Por el

contrario, las versiones gratuitas de estas herramientas se rigen por la política de privacidad general de Google, por lo que la empresa puede acceder a la información de los documentos en Drive para la orientación de anuncios.[78]

8. **Vídeo chat y aplicaciones de mensajería social**. Google Hangouts es una plataforma de comunicación similar a Skype y una parte de las aplicaciones de G Suite conectadas a la nube de Google. Los usuarios pueden comenzar y unirse a videoconferencias o conversaciones grupales desde la aplicación Hangout disponible en: Android e iOS, desde un navegador web, desde la aplicación de escritorio de Hangouts o la extensión de Chrome, y desde otros productos de Google (por ejemplo, Gmail y Calendar).[79] Google almacena detalles de estos intercambios (incluidas las marcas de tiempo de conferencias y llamadas), información de los participantes y contenido de los mensajes. Estos detalles están disponibles para los usuarios y se pueden descargar a través de la herramienta Google Takeout. Algunos datos grabados en una conversación de Google Hangouts incluyen nombres de participantes, ID de los mismos y el contenido de la conversación.[80]
9. **Google+** es una red social que fue lanzada en 2011 para competir con Facebook. Aunque Google no publica estadísticas sobre los usuarios activos de Google+, ahora es principalmente un lugar para comunidades de nicho.[81] El servicio fue un fracaso y objeto de numerosas bromas sobre el silencio sepulcral que reinaba en Google+. En octubre de 2018, sufrió un ataque y expuso las cuentas de 500.000 usuarios del servicio (muchos dicen que parece increíble que hubiera tanta gente). Google ha decidido cerrar el servicio a causa de este grave incidente.

10. **Google Translate** es un servicio gratuito de traducción automática compatible con más de 100 idiomas que está disponible en la web y a través de aplicaciones para Android e iOS. También está integrado en Google Assistant y Google Chrome, además de estar disponible para desarrolladores de terceros a través de una API de pago.[82] En total, atiende a más de 500 millones de usuarios mensuales. Si los usuarios tienen la aplicación Google Translate en su teléfono, pueden usarla para traducir idiomas dentro de otras aplicaciones, como WhatsApp. Aunque Google afirma que no rastrea las consultas web de Google Translate, y estas consultas no aparecen en el historial del buscador de un usuario, ni en ningún otro lugar de Google Takeout, la política de privacidad de Google le permite almacenarlas durante un período breve y ocasionalmente durante uno más prolongado para depuración y otras pruebas.
11. **Google Calendar** es una herramienta de programación que permite a los usuarios realizar un seguimiento de sus actividades diarias y semanales. Se usa ampliamente en dispositivos de escritorio y dispositivos móviles, con más de 500 millones de descargas de Google Play Store. La información personal, como el nombre y la información de contacto de una persona, a menudo se asocia con un usuario de Calendar. Mediante el uso de calendarios, los usuarios proporcionan a Google los detalles (como la hora, la ubicación y la información de contacto) de todos los participantes de un evento. Como parte de esta recopilación de información, Google lee y almacena las direcciones de correo electrónico de todos los miembros incluidos en un evento del calendario, independientemente de su afiliación a Google. Siempre que una persona en una

invitación de calendario esté usando Calendar de Google, Google registra la información de contacto de cada una de las personas en la invitación.

12. **Google Keep** es una herramienta que permite al usuario tomar notas que se sincronizan entre dispositivos asociados con su cuenta de Google. Keep se ha descargado más de 100 millones de veces desde Google Play Store. Google recopila y almacena todo el contenido de las notas, así como la hora en que fueron creadas. Google escanea y clasifica las notas que se crean en función de sus contenidos. Ejemplos de categorías de clasificación son comida, lugares y viajes.
13. **Chromecast**. Al igual que Apple TV, Google Chromecast es un dispositivo que actúa como interfaz para ver vídeos de una variedad de aplicaciones (por ejemplo, Netflix, YouTube, Hulu, Play Store), y para proyectar vídeo desde teléfonos inteligentes y ordenadores en televisores y monitores más grandes. Cada dispositivo Cast tiene un identificador único asociado a la cuenta de Google de un usuario durante el registro. Google usa Chromecast para recopilar la actividad del sistema, informes de fallos y datos de uso, como, por ejemplo, detalles sobre el uso de la funcionalidad de transmisión de dispositivos, incluidas las aplicaciones y los dominios que se emiten. Chromecast utiliza Google Cast, que es una plataforma de software que permite la transmisión continua de audio/vídeo entre dispositivos en la misma red. Además de una multitud de otros productos de Google (por ejemplo, Google Home) que usan la funcionalidad Cast, una plataforma también utilizada por dispositivos de terceros que usan «Chromecast incorporado» (por ejemplo, televisores y consolas de videojuegos), que ofrecen una funcionalidad similar. Google también recopila datos de uso e informes de fallos de dispositivos Cast de terceros.

14. **Google DNS**. Google lanzó un servicio gratuito de sistema de nombres de dominio (DNS), llamado Google Public DNS, en diciembre de 2009. Google DNS está destinado a mejorar la experiencia de navegación web mejorando su velocidad, seguridad y precisión. En diciembre de 2014, Google Public DNS informó que había servido 400 mil millones de respuestas. Para detectar el abuso (como ataques DDoS) y solucionar problemas, Google Public DNS mantiene un registro temporal de las direcciones IP completas que elimina en 24-48 horas. Para esfuerzos a más largo plazo para depurar y prevenir el abuso, Google mantiene información de ubicación de ciudad no identificable personalmente durante dos semanas, y toma muestras al azar de un pequeño subconjunto para almacenamiento permanente.
15. **Wifi Google router**. Google comenzó a desplegar un enrutador wifi de malla, Google Wifi, en diciembre de 2016. Los enrutadores de malla permiten a los usuarios ampliar el acceso a una red wifi en un área más grande con enrutadores conectados adicionales. En diciembre de 2017, Google Wifi se convirtió en el mejor sistema de wifi de malla en Estados Unidos. La información que recopila se clasifica en las siguientes tres categorías:
   – Servicios en la nube, que incluyen información transmitida desde dispositivos conectados, información inferida de dispositivos conectados (como el nombre del fabricante; ejemplo: Samsung), estado de conexión, velocidad de transferencia de datos y consumo histórico, configuración de red e información sobre el entorno inalámbrico (por ejemplo, otros enrutadores en el área). Recopila información de dispositivos conectados y uso de datos.
   – Estadísticas wifi, que incluyen el uso de datos anónimos, informes de fallos e información de rendimiento del dispositivo.

– Estadísticas de la aplicación wifi, que incluye informes de uso y bloqueo.

Según Google, el enrutador wifi no realiza un seguimiento de los sitios web visitados ni recopila el contenido del tráfico de la red.

16. **Nest**. En enero de 2014, Google adquirió Nest, una empresa de automatización del hogar. La línea de productos de Nest incluye muchos dispositivos domésticos inteligentes, incluidos termostatos, cámaras, timbres, sistemas de alarma, cerraduras y detectores de humo/CO_2. Además, Nest puede integrar más de 100 productos de terceros, desde refrigeradores hasta camas. Los dispositivos Nest recopilan una variedad de información, que incluye no sólo los ajustes directos de los usuarios en los dispositivos, sino también los datos sobre el entorno dentro del hogar. Por ejemplo, Nest Learning Thermostat recopila datos tales como temperatura, humedad, luz ambiental y movimiento y, por tanto, sabe cuándo hay personas en casa e incluso en qué habitaciones. Cuando un usuario conecta productos de terceros que se integran con Nest, éste comparte información con esas partes, pero informa al usuario sobre qué información se está compartiendo. Nest no comparte datos con otros terceros, como compañías de energía o compañías asociadas, sin obtener primero el consentimiento del usuario. En su sitio web, Nest declara que sus cuentas y las de Google no están vinculadas (a menos que un usuario opte por integrarse con productos y servicios de Google) y que Google no vende datos de Nest.

    Recientemente, sin embargo, surgió la preocupación acerca del intercambio de datos a partir de un anuncio de Google sobre la fusión de los equipos de hardware de las

mencionadas compañías. Además, ya han saltado las alarmas sobre las relaciones futuras entre Google, Nest y las compañías eléctricas y de seguros del hogar.

17. **Google Fiber**. Es un servicio telefónico de banda ancha, protocolo de internet (IP) y voz por IP (VOIP) que conecta a los usuarios a través de redes de fibra óptica de ultra alta velocidad que se extienden hasta sus residencias, conocidas como Fiber Internet, Fiber TV y Fiber Phone, respectivamente.[83] Fiber Internet recopila información técnica conectada a la cuenta de Google de un usuario, pero no comparte detalles de la cuenta con otros servicios de Google sin el consentimiento adicional del usuario. Fiber Internet requiere el consentimiento del usuario antes de asociar la cuenta de Google de un usuario con otra información, como los sitios visitados o el contenido de las comunicaciones. Fiber TV recopila información sobre programas y aplicaciones utilizadas, y la asocia con la cuenta de Google del usuario. Del mismo modo, Fiber Phone recopila información de su uso (por ejemplo, registros del historial de llamadas, correos de voz, mensajes SMS y conversaciones grabadas) y la asocia con la cuenta de Google de un usuario. Aunque esta información no se comparte con terceros, a menos que el usuario brinde su consentimiento, la información se puede compartir con terceros para el procesamiento externo y por razones legales. Si bien Google no menciona explícitamente el uso de Fiber TV/teléfono para la orientación de anuncios, su política de privacidad general permite el uso de la información personal recopilada para fines de orientación. Además, Google afirma que compartirá públicamente la información identificable no personal del usuario de Google Fiber con proveedores de contenido, editores, anunciantes y/o sitios conectados.

El Google del futuro

Google tiene productos adicionales que muestran el potencial futuro de la compañía en la ardua tarea de recopilar todos los datos del mundo.

1. **Accelerated Mobile Pages (AMP)** es una iniciativa de código abierto encabezada por Google para permitir tiempos de carga más rápidos para sitios web y anuncios.[84] Los usuarios pueden acceder al contenido de varios editores cuyos artículos aparecen en los resultados de buscador mientras navegan por el carrusel de AMP, todo mientras permanecen dentro del dominio de Google. En efecto, el caché de AMP funciona como una red de entrega de contenido (CDN) de propiedad y operada por Google. Al crear una herramienta de código abierto con una CDN, Google ha atraído a una gran base de usuarios que sirven webs móviles y anuncios, lo que constituye una cantidad significativa de información (por ejemplo, el contenido en sí, páginas vistas, anuncios publicados e información sobre quién está entregando el contenido). Toda esta información está disponible para Google, alojada en sus servidores CDN, lo que proporciona a la compañía estadounidense muchos más datos de los que podría acceder de otra manera.[85]
2. **Google Assistant** es un asistente personal virtual al que se accede a través de teléfonos móviles y dispositivos inteligentes. Es uno de los asistentes virtuales más populares, junto con Siri de Apple, Alexa de Amazon y Cortana de Microsoft. Se puede acceder al Asistente de Google a través del botón de inicio de los dispositivos móviles con Android 6.0 o superior. También se puede acceder a través de una aplicación dedicada en dispositivos con iOS, así como a través de altavoces

inteligentes, como Google Home. El asistente de Google realiza numerosas funciones, como enviar mensajes de texto, buscar correos electrónicos, controlar música, buscar fotos, obtener respuestas a preguntas sobre el clima o el tráfico y controlar dispositivos domésticos inteligentes. Google recopila todas las consultas del asistente, ya sean de audio o mecanografiadas. También recopila la ubicación donde se produjo la consulta. Además de su uso en Google, el asistente se puede habilitar en otros altavoces producidos por terceros (por ejemplo, auriculares inalámbricos Bose). En general, Google Assistant está disponible en más de 400 millones de dispositivos. Google puede recopilar datos a través de todos estos dispositivos, ya que las consultas al asistente pasan por los servidores de Google.

3. **Google Photos**. Este servicio es utilizado por más de 500 millones de personas en todo el mundo y almacena más de 1,2 mil millones de fotos y vídeos. Google registra las coordenadas de tiempo y GPS para cada foto tomada. Google carga imágenes a su nube y realiza el análisis de éstas para identificar un amplio conjunto de objetos, como modos de transporte, animales, logotipos, puntos de referencia, texto y caras. Las capacidades de detección de rostros de Google incluso permiten la detección de estados emocionales asociados con caras en fotos cargadas y almacenadas en su nube. Google Photos realiza este análisis de imágenes de forma predeterminada cuando se utiliza el producto. Si un usuario autoriza a Google a agrupar rostros similares, Google identifica a diferentes personas mediante la tecnología de reconocimiento facial y permite a los usuarios compartir fotografías en función de su tecnología de «agrupamiento por rostros». Google utiliza

este servicio para hacerse con una base de datos valiosa de identificación facial lo que ha provocado una fuerte oposición judicial en distintos estados de Estados Unidos.

4. **Chromebook** es la tableta de Google que se ejecuta en el sistema operativo Chrome (Chrome OS) que permite a los usuarios acceder a las aplicaciones en la nube.[86] Se puede entrar a Chromebook a través de una cuenta de Google, así que el usuario accede de manera automática a las aplicaciones de Google que use a través de Chromebook. La recopilación de datos de Chromebook es similar al navegador Google Chrome.[87]
5. **Google Pay** es un servicio de pago digital que permite a los usuarios almacenar información de tarjetas de crédito, cuentas bancarias e información de PayPal para realizar pagos en tiendas, sitios web o aplicaciones usando Google Chrome o un dispositivo Android conectado. Gracias al pago, Google recoge la dirección de usuario y los números de teléfono verificados, ya que están asociados con cuentas de cargo. Además de información personal, Pay también recopila información de transacciones, como la fecha y el monto de la transacción, ubicación y descripción del comerciante, tipo de pago utilizado, descripciones de los artículos comprados, cualquier foto que el usuario decida asociar con la transacción, nombres y las direcciones de correo electrónico del vendedor y del comprador, la descripción del usuario del motivo de la transacción y cualquier oferta asociada con la transacción. Google trata esta información como información personal bajo su política de privacidad general. Por tanto, puede utilizar esta información en todos sus productos y servicios para publicidad enriquecida. La configuración de privacidad de Google permite dicho uso de los datos recopilados por defecto.

Aunque no lo uses

Google recopila datos de terceros además de la información que recopilan directamente de sus servicios y aplicaciones. Por ejemplo, en 2014, Google anunció que comenzaría a rastrear las ventas en tiendas físicas comprando datos de transacciones de tarjetas de crédito y débito. Dichos datos cubrieron el 70 por ciento de todas las transacciones de crédito y débito en Estados Unidos. Contenía el nombre del individuo, así como la hora, la ubicación y el monto de su compra.

Estos datos de terceros también se utilizan como respaldo de Google Pay, incluidos los servicios de verificación, información proveniente de transacciones de Google Pay en establecimientos comerciales, métodos de pago, identidad de emisores de tarjetas, información sobre acceso a saldos en la cuenta de pago de Google, información de facturación del operador e informes del consumidor. Aunque la información de usuario de terceros que Google recibe actualmente es de alcance limitado, ya ha atraído la atención de las autoridades gubernamentales. De hecho, Google llegó a un acuerdo con Mastercard con la finalidad de tener información offline de operaciones. Google pagó a Mastercard millones de dólares por los datos, según dos personas que trabajaron en el acuerdo, y las compañías discutieron compartir una parte de los ingresos publicitarios, de acuerdo con una de las personas. Los anunciantes gastan generosamente en Google para obtener información valiosa sobre el vínculo entre los anuncios digitales, la visita a un sitio web o una compra online.

Desde 2014, Google ya venía facilitando esa información a los anunciantes, relacionando los clics del usuario con la visita a una tienda física, utilizando la función «historial de ubicación» en Google Maps. Aun así, el anunciante sólo sabía que había entrado en la tienda pero no que hubiera completado la compra.

En 2017, y gracias al acuerdo con Mastercard, Google anunció que su servicio «Store Sales Measurement» tenía acceso a aproximadamente el 70 por ciento de las tarjetas de crédito y débito de Estados Unidos. A través de este programa de prueba, Google puede hacer coincidir los perfiles de usuario existentes con compras realizadas en tiendas físicas. El resultado es poderoso: Google, que sabe quién hizo clic en un anuncio, ahora le puede dar la tasa de conversión en ventas reales en las tiendas.

Funciona así: una persona busca «lápiz de labios rojo» en Google, hace clic en un anuncio, navega por la web pero no compra nada. Más tarde, entra en una tienda y compra lápiz labial rojo con su Mastercard. El anunciante que sirvió el anuncio recibe un informe de Google que enumera la venta junto con otras transacciones en una columna llamada «Ingresos sin conexión», pero sólo si el internauta se conectó a una cuenta de Google en línea y realizó la compra en 30 días. A los anunciantes se les proporciona un informe general con el porcentaje de compradores que hicieron clic o vieron un anuncio y luego realizaron una compra relevante.

No es una coincidencia exacta, pero es la herramienta más poderosa que Google, el vendedor de anuncios más grande del mundo, ha ofrecido para comprar en el mundo real.

Por ejemplo, la FTC anunció una medida cautelar contra Google en julio de 2017 al entender que la recopilación de datos de compras de consumidores de Google infringe la privacidad electrónica.[88] La medida cautelar cuestiona la afirmación de Google de que pueden proteger la privacidad del consumidor durante todo el proceso utilizando su algoritmo. Aunque aún no se han tomado medidas adicionales, la orden de la FTC es un ejemplo de preocupación pública con la cantidad de datos de consumidores que recoge Google.[89]

5

Casandra

Predicción y toma de decisiones

Casandra, sacerdotisa de Apolo, pactó con éste, sexo mediante, la concesión del don de la profecía. Tan pronto adquirió su superpoder, Casandra perdió interés por él, quien no se tomó bien el rechazo y le mantuvo el don de ver el futuro pero la maldijo con que nadie creería en sus pronósticos. Se hartó de repetir que desconfiaran de los griegos aunque trajeran regalos, pero los troyanos dejaron entrar el caballo. El resto es historia.

No niego que cuando me pongo pesada advirtiendo sobre los riesgos que tiene para la privacidad el uso de determinadas aplicaciones, noto una mirada de conmiseración en mis interlocutores que me imaginan durmiendo con papel de plata en la cabeza para que no lean mis pensamientos. Personalmente, me veo más como Casandra pontificando en el desierto.

Lo curioso es que, sin necesidad de acceso carnal con Apolo, las empresas de tecnología que basan su modelo de negocio en la explotación de los datos son unas verdaderas Casandras dotadas del don de la adivinación.

Omega

Imaginemos el año que viene, 2020. Los científicos de datos habrán desarrollado modelos capaces de predecir y manipular el comportamiento de individuos con un alto grado de precisión. La capacidad de los algoritmos para adivinar cuándo y dónde una persona específica llevará a cabo acciones particulares es considerada por algunos como una señal del último algoritmo, el Omega, el paso final en la transferencia de poder de la humanidad a tecnologías ubicuas. Para los responsables de la ciberseguridad, los riesgos nunca han sido más altos.

Con el desarrollo acelerado del aprendizaje automático, los algoritmos y los sensores que rastrean la acción humana y permiten que los conjuntos de datos se alimenten mutuamente, llegará el internet de 2020, que incorporará modelos capaces de predecir y manipular un sorprendente rango de comportamientos humanos. En lugar de inferir tendencias individuales a partir de tendencias y grupos con características similares, estos nuevos modelos harán predicciones verdaderamente individualizadas granulares, discriminatorias y precisas sobre comportamientos complejos.

El poder de la ciencia de datos para predecir el comportamiento individual con tal nivel de precisión se convertirá en el debate más polarizador de la década: ¿es un indicador de que la humanidad ha rendido el poder, las libertades y los misterios insondables de su alma a las máquinas? ¿O es un indicador de un progreso sorprendente, que permite a las sociedades resolver más eficazmente algunos de sus problemas más recalcitrantes? Si bien este debate continuará en abstracto, los análisis predictivos generarán nuevas vulnerabilidades de seguridad que superarán los conceptos y prácticas de defensa existentes, que se dirigirán más a las personas en lugar de a la infraestructura.

En este escenario, la disponibilidad de cantidad y variedad de datos de alta calidad, junto con algoritmos y análisis avanzados capaces de preguntar a esos datos, permitirá predicciones altamente precisas e individualizadas del comportamiento humano. Si bien hoy es posible predecir los comportamientos agregados de

grupos y poblaciones, en 2020 tales predicciones serán órdenes de magnitud más precisas y, lo que es más importante, mucho más personalizadas, hasta el punto de predecir el comportamiento de una sola persona.

En 2020, los algoritmos de la próxima generación podrán saltarse los análisis demográficos y limitarse a las preferencias específicas de una sola persona (ejemplo: Mari Carmen prefiere ver *Juego de Tronos*). Más importante aún, las predicciones probabilísticas se convertirán en contingentes, y contarán con afirmaciones muy precisas sobre las condiciones exactas en las que la persona X tomará la acción Y. Sabremos exactamente en qué condiciones (hora, lugar, coste, etc.) Mari Carmen verá *Juego de Tronos*.

En este mundo, los modelos predictivos desempeñarán un papel cada vez más importante en la vida cotidiana, ya sea para organizar mejor el tráfico aéreo global, mostrar productos o calcular cuándo y dónde desplegar tropas. Las empresas que viven de recomendar productos para adelgazar, podrán hacer recomendaciones sobre la dieta más adecuada a partir de un análisis que predecirá cuándo los clientes tendrán antojos.

A medida que mejoremos nuestra capacidad de predecir elecciones y comportamientos individuales, la capacidad de adivinar el comportamiento grupal será menos útil y menos precisa: la agregación de predicciones individuales en predicciones de grupo es difícil y, curiosamente, menos precisa que el antiguo enfoque de desagregar hacia abajo de grupos a individuos.

Ésta es la visión que desde 2016 nos hacía para el año próximo la Universidad de Berkeley.[90] Y lo cierto es que no va del todo desencaminada.

Otra vez el big data

Como agentes humanos, somos visibles en casi todas las interacciones con plataformas tecnológicas. Siempre estamos siendo rastreados, cuantificados, analizados y mercantilizados. El nuevo horizonte infinito es la extracción de datos, el aprendizaje automático y la reorganización de la información a través de sistemas de inteligencia artificial de procesamiento humano y mecánico combinados. Los territorios están dominados por pocas megaempresas globales, que están creando nuevas infraestructuras y mecanismos para la acumulación de capital y la explotación de recursos humanos y planetarios.

El big data, el uso de inteligencia artificial (IA) y el «machine learning» están en boca de todos. Llevamos años hablando del diluvio de datos, de las grandes oportunidades que para la ciencia y los negocios tiene acceder a una enorme cantidad de información y extraer de ella patrones y hacer prospectiva. Es necesario separar lo que es la acumulación de grandes cantidades de datos que se almacenan y las técnicas de análisis que se usan, que incluyen el data mining o el uso de la de la IA en sus múltiples variables.

Una definición popular de «grandes datos» o «big data», proporcionada por el glosario de la consultora Gartner IT[91] sería: «Son los activos de información de alto volumen, alta velocidad y alta variedad que exigen formas innovadoras y rentables de procesamiento de la información para una visión mejorada y la toma de decisiones». Así pues, el concepto big data se suele definir por sus tres uves: el volumen se refiere a conjuntos de datos masivos a analizar, la velocidad de los datos en tiempo real y la variedad de las diferentes fuentes de datos, si bien hay quien considera que no sería necesario que se dieran estos tres requisitos.[92]

Por su parte, ICO,[93] el supervisor de protección de datos del Reino Unido, propone definir los grandes datos por sus siguientes aspectos distintivos de otro tipo de análisis: el uso de algoritmos, la opacidad del procesamiento, la tendencia a recoger «todos los datos», la reutilización de los datos y el uso de nuevos tipos de datos.

Recetas de cocina e inteligencia artificial

Hasta ahora hemos venido hablando de datos. Dediquemos un momento a la IA y a los algoritmos de los que todo el mundo habla y a los que culpamos de nuestras penas.

Pero ¿qué es un algoritmo? Wikipedia define un algoritmo como «un conjunto de reglas que define con precisión una secuencia de operaciones» y la RAE como «conjunto ordenado y finito de operaciones que permite hallar la solución de un problema». Aunque es un concepto que se asocia habitualmente a las matemáticas, en realidad está muy próximo a la quinta acepción de la palabra receta: «Nota que comprende aquello de que debe componerse algo, y el modo de hacerlo. Receta de cocina».

De hecho, el uso ha cambiado de maneras interesantes desde el surgimiento de internet, y en particular de los motores de búsqueda, a mediados de la década de 1990. En la raíz, un algoritmo es algo pequeño y simple; una regla utilizada para automatizar el tratamiento de una pieza de datos. Si sucede A, entonces haz B; si no, entonces haz C. Ésta es la lógica «*if/then/else*» de la informática clásica. Si un usuario dice tener dieciocho años, déjale entrar en la web; si no, carga el mensaje «lo siento, debes tener dieciocho años para entrar». En palabras de PérezChirinos:

> Algoritmos y recetas de cocina comparten la idea de transformar ordenadamente unos ingredientes para obtener un resultado más apetecible que los ingredientes a transformar. Además, en ambos casos el resultado puede no ser evidente a partir de los ingredientes; de aquí obtienen su valor los algoritmos: tanto ellos como las recetas son a veces celosamente guardados por sus descubridores por las ventajas económicas, reputacionales o de otro tipo que puede suponer su conocimiento.
>
> Los algoritmos, como las recetas, pueden ser triviales o muy complicados. En el ámbito contable, el algoritmo que permite obtener el saldo de una cuenta tras un abono o un adeudo es trivial teniendo en cuenta que los ingredientes de un algoritmo se denominan datos.

> Este ejemplo no presupone que el algoritmo vaya a ser automatizado en un sistema de información: podría ser parte de un manual o un curso de contabilidad dirigido a personas, aunque tiene alguna peculiaridad: contempla expresamente el caso de que pueda intentar utilizarse con tipos de operación incorrectos.
>
> A diferencia de lo que ocurre con la documentación informal, en la que se cuenta con la inteligencia del lector para cubrir posibles omisiones, la descripción de un algoritmo de ejecución automatizada debería ser exhaustiva y cubrir todas las opciones posibles si queremos evitar sorpresas inesperadas como resultado de la ejecución automática de las instrucciones del algoritmo.[94]

En esencia, los programas son paquetes de algoritmos, recetas para tratar datos. En el nivel micro, nada podría ser más simple. Si los ordenadores parecen mágicos es porque son rápidos, no inteligentes.

Hay algoritmos más veniales y otros más peligrosos. El algoritmo determinista es casi de la familia, un algoritmo que nos acompaña y nos persigue por las pantallas de nuestros navegadores y nuestras cuentas de correo. Un algoritmo determinista es aquel que, en términos informales, es completamente predictivo si se conocen sus entradas. Es casi un algoritmo de Perogrullo, si conoces los datos con los que se ha alimentado el algoritmo, sabes por adelantado los resultados. Un modelo simple de algoritmo determinista es la función matemática, pues ésta extrae siempre la misma salida para una entrada dada. Este tipo de algoritmo es reversible, ya que conociendo el resultado es bastante sencillo saber qué datos se introdujeron al inicio.

Ésta es la razón por la que la página de reserva de hoteles nos sigue bombardeando con mensajes de hoteles en Cuenca después de haber ido hace un año, una página de moda nos ofrece un pantalón que ya tenemos o nos siguen insistiendo con un vuelo a Las Maldivas después de haber hecho una búsqueda, de haberlo hablado con alguien por WhatsApp o de haber escrito el hashtag #MaldivasForEver en tu cuenta de Instagram.

Por el contrario, un algoritmo no determinista es un algoritmo que con la misma entrada ofrece muchos posibles resultados y, por tanto, no ofrece una solución única. No se puede saber de antemano cuál será el resultado de la ejecución de un algoritmo no determinista.

En los últimos años ha surgido un significado más ambiguo de la palabra «algoritmo»: que es cualquier sistema de software de toma de decisiones grande y complejo; cualquier medio de tomar una serie de datos de entrada y evaluarlos rápidamente, de acuerdo con un conjunto determinado de criterios (o «reglas»). Son estos algoritmos los que han revolucionado las áreas de la medicina, la ciencia, el transporte y la comunicación, por lo que es fácil comprender la visión utópica de la informática que dominó durante muchos años.

Pero ¿por qué no hacemos más que hablar de algoritmos y en qué se diferencian de la IA?

Inteligencia artificial y aprendizaje automático

La IA se define como la capacidad de una máquina para realizar funciones cognitivas que asociamos con la mente humana, como la percepción, el razonamiento, el aprendizaje y la resolución de problemas. Ejemplos de tecnologías que permiten a la IA resolver problemas comerciales son la robótica, los vehículos autónomos, la visión por ordenador, el lenguaje, los agentes virtuales y el aprendizaje automático.[95]

Por su parte, el aprendizaje automático o *machine learning* es una subdivisión de la IA, en concreto cuando se aplica el aprendizaje automático a conjuntos de grandes datos. Los algoritmos de aprendizaje automático detectan patrones y aprenden a hacer predicciones y recomendaciones mediante el procesamiento de datos y experiencias, en lugar de recibir instrucciones explícitas

de programación. Los algoritmos también se adaptan en respuesta a nuevos datos y experiencias para mejorar la eficacia a lo largo del tiempo.

Finalmente, el aprendizaje profundo o *deep learning* es, a su vez, un tipo de aprendizaje automático que puede procesar una amplia gama de recursos de datos, pero que requiere menor intervención humana en el procesamiento previo y, a menudo, puede producir resultados más precisos que los enfoques tradicionales de aprendizaje automático. En el aprendizaje profundo, las capas interconectadas de máquinas basadas en software, conocidas como «neuronas», forman una red neuronal. La red puede cargar grandes cantidades de datos de entrada y procesarlos a través de múltiples capas que aprenden características cada vez más complejas de los mismos en cada capa. A continuación, la red puede tomar una decisión sobre los datos, saber si su decisión es correcta y usar lo que ha aprendido para tomar otra resolución sobre otros nuevos. Por ejemplo, una vez que aprende cómo se ve un objeto, puede reconocer el objeto en una nueva imagen.[96]

Hay tres tipos de aprendizaje automático:

A. **El aprendizaje supervisado.** Un algoritmo utiliza los datos de entrenamiento y los comentarios de los humanos para conocer la relación de los inputs para un resultado dado (por ejemplo, cómo los inputs «época del año» y «tasas de interés» predicen los precios de la vivienda). Se clasifican los datos de entrada y el tipo de comportamiento que desea predecir, pero necesita el algoritmo para calcularlos en base a nuevos datos.

Cómo funciona:

1. Un ser humano etiqueta cada elemento de los datos de entrada (por ejemplo, para predecir los precios de la vivienda, se etiquetan los datos de entrada como «época del año», «tasas de interés», etc.) y se define la variable de salida (por ejemplo, «precios de la vivienda»).

2. El algoritmo se entrena con los datos para encontrar la conexión entre las variables tanto de entrada como de salida.
3. Una vez que se completa el entrenamiento, generalmente cuando el algoritmo es lo suficientemente preciso, éste se aplica a nuevos datos.

Dentro del aprendizaje supervisado, se usan diversos tipos de algoritmos para cada caso de uso empresarial:

Algoritmo	Definición	Casos de uso
Regresión lineal	Método estándar altamente interpretable para modelar la relación pasada entre variables de entrada independientes y variables de salida dependientes (que pueden tener un número infinito de valores) para ayudar a predecir valores reales de las variables de salida.	Entender los impulsores de ventas de productos, como precios de la competencia, distribución, publicidad, etc. Optimizar los puntos de precio y estimar las elasticidades precio-producto.
Regresión logística	Extensión de la regresión lineal que se usa para las tareas de clasificación, lo que significa que la variable de salida es binaria (por ejemplo, sólo blanco o negro) en lugar de continua (por ejemplo, una lista infinita de colores potenciales).	Clasificar a los clientes según la probabilidad de que devuelvan un préstamo. Predecir si una lesión cutánea es benigna o maligna según sus características (tamaño, forma, color, etc.).
Linear/quadratic discriminant analysis (Análisis discriminante lineal/cuadrático)	Actualiza una regresión logística para tratar problemas no lineales, aquéllos en los que los cambios en el valor de las variables de entrada no se resuelven en cambios	Predecir la rotación de clientes. Predecir la probabilidad de cierre de una ventaja de ventas.

	proporcionales a las variables de salida.	
Árbol de decisión	Modelo de clasificación o regresión altamente interpretable que divide los valores de las características de los datos en ramas y nodos de decisión (por ejemplo, si una entidad es un color, cada color posible se convierte en una nueva rama) hasta que se tome una decisión final.	Proporcionar un marco de decisión para la contratación de nuevos empleados. Comprender los atributos del producto que aumentan las posibilidades de compra.
Naive Bayes	Técnica de clasificación que aplica el teorema de Bayes, que permite calcular la probabilidad de que un evento se base en el conocimiento de los factores que podrían afectar ese evento (por ejemplo, si un correo electrónico contiene la palabra «dinero», entonces la probabilidad de que sea spam es alta).	Analizar el sentimiento para evaluar la percepción del producto en el mercado. Crear clasificadores para filtrar correos spam.
Support vector machine	Técnica que normalmente se usa para la clasificación, pero se puede transformar para realizar una regresión. Dibuja una división óptima entre clases (lo más amplia posible). También se puede generalizar rápidamente para resolver problemas no lineales.	Predecir a cuántos pacientes necesitará atender un hospital en un período de tiempo. Predecir la probabilidad de que alguien haga clic en un anuncio online.
Random forest	Modelo de clasificación o regresión que mejora la precisión de un árbol de decisión simple al generar múltiples árboles de decisión	Predecir el volumen de llamadas en los CAU. Predecir el uso de energía en una red de distribución eléctrica.

	y tomar un voto mayoritario para predecir el resultado, que es una variable continua (por ejemplo, la edad) para un problema de regresión y una variable discreta (por ejemplo, ya sea negro, blanco o rojo) para la clasificación.	
AdaBoost	Técnica de clasificación o regresión que utiliza una multitud de modelos para tomar una decisión, pero los sopesa en función de su precisión para predecir el resultado.	Detectar actividad fraudulenta en transacciones con tarjeta de crédito; y conseguir una menor precisión que el aprendizaje profundo. Una forma sencilla y económica de clasificar las imágenes (por ejemplo, reconocer el uso de la Tierra de las fotos de satélite para los modelos de cambio climático). Conseguir una menor precisión que el aprendizaje profundo.
Gradient-boosting trees	Técnica de clasificación o regresión que genera árboles de decisión de forma secuencial, en la que cada árbol se enfoca en corregir los errores provenientes del modelo anterior. El resultado final es una combinación de los resultados de todos los árboles.	Prevenir la demanda de productos y niveles de inventario. Predecir el precio de los automóviles según sus características (por ejemplo,
Red neuronal simple	Modelo en el que las neuronas artificiales (calculadoras basadas en software) forman tres capas (una de entrada, una oculta	Predecir la probabilidad de que un paciente se una a un programa de salud. Predecir si los usuarios

	en la que se realizan los cálculos y una de salida) que se pueden usar para clasificar datos o encontrar la relación entre variables en problemas de regresión.	registrados estarán dispuestos o no a pagar un precio particular por un producto.

B. **Aprendizaje sin supervisión**. Un algoritmo explora los datos de entrada sin recibir una variable de salida explícita (por ejemplo, explora los datos demográficos del cliente para identificar patrones). No sabe cómo clasificar los datos y quiere que el algoritmo encuentre los patrones y los datos.

Cómo funciona:

1. El algoritmo recibe datos sin etiquetar (por ejemplo, un conjunto de datos que describen los viajes de los clientes en un sitio web).
2. Infiere una estructura a partir de los datos.
3. El algoritmo identifica grupos de datos que muestran un comportamiento similar (por ejemplo, forman grupos de clientes que muestran comportamientos de compra similares).

Dentro del aprendizaje no supervisado, por su parte, se usan los siguientes tipos de algoritmos para cada caso de uso empresarial.

Algoritmo	Definición	Casos de uso
K-means clustering	Pone los datos en un número de grupos que contienen datos con características similares (según lo determinado por el modelo, no de antemano por los humanos).	Segmentar a los clientes en grupos por distintas características (por edad), para asignar mejor campañas de marketing o evitar la pérdida de clientes.

Gaussian mixture model	Una generalización de la agrupación de medidas que proporciona más flexibilidad en el tamaño y la forma de los grupos (agrupaciones).	Segmentar a los clientes para asignar mejor las campañas de marketing utilizando características de clientes menos distintas (por ejemplo, preferencias de productos). Segmentar empleados en función de la probabilidad de agotamiento.
Hierarchical clustering	Divide o agrega agrupamientos a lo largo de un árbol jerárquico para formar un sistema de clasificación.	Agrupa a los clientes de tarjetas de fidelización en grupos progresivamente más microsegmentados. Informa del uso/desarrollo del producto agrupando clientes que mencionan palabras clave en redes sociales.
Sistema de recomendación	A menudo utiliza la predicción de comportamiento de grupo para identificar los datos importantes necesarios para hacer una recomendación.	Recomendar qué películas deberían ver los consumidores según las preferencias de otros clientes con atributos similares. Recomendar noticias que un lector pueda querer leer en función del artículo que esté viendo.

C. **Aprendizaje reforzado**. Un algoritmo aprende a realizar una tarea simplemente intentando maximizar las recompensas que recibe por sus acciones (por ejemplo, maximiza los puntos que recibe para aumentar los rendimientos de una cartera de inversiones). No se cuenta con muchos datos de entrenamiento o no

se puede definir claramente el estado final ideal o la única forma de aprender en el terreno interactuando con él. Es similar al entrenamiento de una mascota, con castigos y recompensas. Cómo funciona:

1. El algoritmo realiza una acción sobre el terreno (por ejemplo, realiza un intercambio en una cartera financiera).
2. Recibe una recompensa si la acción lleva a la máquina un paso más cerca de maximizar las recompensas totales disponibles (por ejemplo, el rendimiento total más alto en la cartera).
3. El algoritmo optimiza para la mejor serie de acciones corrigiéndose a sí mismo con el tiempo.

Son casos típicos en los que se emplea el aprendizaje de refuerzo para: optimizar la estrategia de trading para una cartera de opciones; equilibrar la carga de las redes eléctricas en ciclos de demanda variables; realizar un inventario y selección de inventario utilizando robots; optimizar el comportamiento de conducción de los coches autoconducidos, u optimizar los precios en tiempo real para una subasta en línea de un producto con un suministro limitado.

Es interesante, ante el ruido y confusión existentes, hacer un pequeño resumen del funcionamiento de la IA y del aprendizaje automático y reflejar para qué tipo de decisiones se emplean.

De lo anterior podemos extraer una primera conclusión: que la IA usa algoritmos, pero que no todos los algoritmos se usan en la IA.

Que últimamente sólo se hable de IA y algoritmos es debido a que en 2009 se produjo una convergencia de avances algorítmicos, proliferación de datos y un aumento enorme en la capacidad de computación y del almacenamiento de aquéllos. Estos tres elementos (datos, modelos matemáticos y capacidad de cómputo y almacenamiento) han hecho que la IA salga de los laboratorios y que sea una realidad usable hoy día.

Aunque no todo comenzó en 2007, ahí fue cuando cambió el paradigma de la recogida de datos y de todo lo que ha venido después. En enero de ese año, Steve Jobs presentaba el iPhone que impulsó la revolución de los teléfonos inteligentes y amplificó de un modo inimaginable la generación de datos. El número total de teléfonos inteligentes vendidos en 2007 fue de, aproximadamente, 122 millones. Con aquella famosa keynote de Jobs, dio comienzo la era del consumo las 24 horas del día, los 7 días de la semana, y la generación de datos y creación de contenido por parte de los usuarios de teléfonos inteligentes en tiempo real.

En 2009, la Universidad de California, en Berkeley, presentó Spark para manejar modelos de big data de manera más eficiente. Desarrollado por el científico informático rumano-canadiense Matei Zaharia en AMPLab, Spark transmite enormes cantidades de datos que aprovechan la RAM, lo que lo hace mucho más rápido en el procesamiento de datos que el software que usa el disco duro para leer/escribir. Revolucionó la capacidad de actualizar big data y realizar análisis en tiempo real.

Ese mismo año, el científico informático estadounidense Andrew Ng y su equipo en la Universidad de Stanford usaron GPU para entrenar modelos de aprendizaje profundo de manera más efectiva, y demostraron que el entrenamiento en redes de creencias profundas —con 100 millones de parámetros en las GPU— es 70 veces más rápido que hacerlo en las CPU, un hallazgo que reduciría el entrenamiento de las IA de semanas a unas pocas horas.

El equipo de Geoffrey Hinton consiguió en 2012 clasificar imágenes de ImageNet con una tasa de error del 15,3 por ciento en comparación con la tasa de error del 26,2 por ciento, que usaba una red neuronal convolucional (CNN). El equipo de Hinton entrenó a su CNN con 1,2 millones de imágenes usando dos tarjetas de GPU. Así es cómo el sistema de aprendizaje profundo ganó por primera vez el famoso concurso de clasificación de imágenes.

Fue también en 2012 cuando Google demostró la efectividad del aprendizaje profundo en el reconocimiento de imágenes. Usó 16.000 procesadores para entrenar una red neuronal artificial profunda con mil millones de conexiones en diez millones de miniaturas de vídeos de YouTube, seleccionados al azar, en el transcurso de tres días. Sin recibir ninguna información sobre las imágenes, la red reconoció imágenes de gatos, marcando el inicio de avances significativos en el reconocimiento de imágenes.

DeepMind lanzó en 2013 un algoritmo para jugar con Atari utilizando el aprendizaje por refuerzo y el aprendizaje profundo. Mientras que el primero se remonta a finales de la década de 1950, no fue hasta que Vlad Mnih, de DeepMind (antes de ser comprada por Google), lo aplicó junto con una red neuronal para jugar con los videojuegos Atari a niveles sobrehumanos.

En 2014, el número de dispositivos móviles superó a toda la población mundial.

Fue ya en 2016 cuando comenzó a tomar forma una consideración más matizada de nuestra nueva realidad algorítmica. Si entendemos los algoritmos como entidades independientes con vida propia es porque se nos ha animado a pensar en ellos de esta manera. Facebook o Google han vendido defendiendo sus algoritmos con la promesa de objetividad, de su capacidad de sopesar un conjunto de condiciones con frialdad matemática y ausencia de emoción. No es de extrañar que la toma de decisiones algorítmica se haya extendido a la concesión de préstamos, fianzas, ayudas, plazas universitarias, entrevistas de trabajo y a casi cualquier cosa que requiera una toma de decisión.[97]

Tal sed irresistible de nuevos recursos y campos de explotación cognitiva ha llevado a la búsqueda de capas cada vez más profundas de datos que puedan utilizarse para cuantificar la psique humana, consciente e inconsciente, privada y pública, idiosincrásica y general. De esta manera, hemos visto el surgimiento de múltiples

economías cognitivas de la economía de la atención,[98] la de la vigilancia, la economía de la reputación,[99] y la economía emocional, así como la cuantificación y mercantilización de la confianza

Cada vez más, el proceso de cuantificación está llegando al mundo humano afectivo, cognitivo y físico. Existen conjuntos de entrenamiento para la detección de emociones, el parecido familiar, el seguimiento de un individuo a medida que envejece, y para acciones humanas como sentarse, saludar con la mano, levantar un vaso o llorar. Todas las formas de datos biológicos, incluidos el forense, biométrico, sociométrico y psicométrico, se están capturando y se registran en las bases de datos para la formación en IA.

Las salidas de los sistemas de aprendizaje automático son predominantemente inexplicables y no controladas, mientras que las entradas son enigmáticas. Para el observador casual, parece que nunca ha sido tan fácil construir sistemas de IA o de aprendizaje automático de lo que es hoy. La disponibilidad de herramientas de código abierto para hacerlo en combinación con poder de cálculo rentable a través de superpotencias de la nube, como Amazon (AWS), Microsoft (Azure) o Google (Google Cloud), está dando lugar a una falsa idea de la «democratización» de la IA. Mientras que las herramientas de aprendizaje automático, como TensorFlow, se vuelven más accesibles desde el punto de vista de la configuración de su propio sistema; las lógicas subyacentes de esos sistemas y los conjuntos de datos para su capacitación son inaccesibles y están controlados por muy pocas entidades. En la dinámica de recopilación de conjuntos de datos a través de plataformas como Facebook, los usuarios están alimentando y entrenando a las redes neuronales con datos de comportamiento, voz, imágenes etiquetadas y vídeos o datos médicos. En una era de extractivismo, el valor real de esos datos es controlado y explotado por unos pocos en la parte superior de la pirámide.

Toma de decisiones: sesgo y transparencia

Tradicionalmente, el análisis de un conjunto de datos implica, en términos generales, decidir lo que se quiere averiguar a través de ellos y la construcción de una consulta para encontrarlo mediante la identificación de las entradas pertinentes.[100] Los conjuntos de entrenamiento para los sistemas de IA pretenden llegar a la naturaleza refinada de la vida cotidiana, pero repiten los patrones sociales más estereotípicos y restringidos, reinscribiendo una visión normativa del pasado humano y proyectándolo en el futuro humano. En este momento del siglo XXI, vemos una nueva forma de extractivismo que está en marcha: una que llega a los rincones más alejados de la biosfera y las capas más profundas del ser humano cognitivo y afectivo. Muchas de las suposiciones sobre la vida humana, hechas por los sistemas de aprendizaje automático, son estrechas, normativas y están cargadas de errores. Sin embargo, están inscribiendo y construyendo esas suposiciones en un mundo nuevo, y cada vez más desempeñarán un papel en cómo se distribuyen las oportunidades, la riqueza y el conocimiento.

El perfilado que nos prometen como parte de una experiencia de compra o de percepción de un servicio más agradable o ajustado a nuestras necesidades, puede convertirse, sin una aplicación escrupulosa de la ley y sin la transparencia debida a los clientes, en una fuente de perpetuación de la desigualdad y de los sesgos cognitivos. En Estados Unidos, la Comisión Federal de Comercio encontró la evidencia de que se estaba limitando el acceso al crédito a personas no con base en su propio historial sino usando los historiales crediticios de otras personas que compraban en las mismas tiendas que aquéllos. Éste es un perfecto ejemplo de cómo funciona el tratamiento de grandes datos: no busca la causalidad, sino la correlación. No hace falta que las personas que compren en una tienda sean «culpables» de tener un mal crédito, sino que, al comprar en la misma tienda que los que sí lo tienen, y por correlación, se les incluye en el mismo grupo que la gente que no

atiende sus pagos. En este supuesto, las personas afectadas no lo son por pertenecer a un determinado grupo étnico, religioso o sexual —como hasta ahora se ha percibido en este tipo de análisis—, sino que serán discriminadas por compartir determinados factores circunstanciales con personas no merecedoras de crédito.

Este sistema de toma de decisiones tiene, además, un efecto pernicioso: el sistema se retroalimenta y acaba haciendo que las predicciones se cumplan. Si excluyes a alguien de un trabajo porque le has sometido a un test que concluye que es problemático, tiene alguna enfermedad mental o problema de carácter, y ese test se convierte en un estándar en el mercado, aunque el resultado se equivoque, los numerosos rechazos que sufrirá esa persona la convertirán en alguien iracundo o deprimido, un clásico ejemplo de profecía autocumplida. Es el caso del test Kronos, que copian las preguntas de las pruebas médicas para enfermos mentales en los sistemas de acceso a puestos de trabajo. Obligar a pasar un examen médico para obtener un trabajo en Estados Unidos está prohibido, pero cuando las preguntas las realiza una máquina opaca, y la toma de decisiones no motivadas pero basadas en las matemáticas, convierte lo que era un problema legal en una oportunidad de negocio.

Como comentábamos, todo esto sucedió a costa de los sistemas de cálculo de rating y del nacimiento de FICO, cuando los banqueros eran personas que prestaban con interés en ciudades de tamaño medio y conocían a sus convecinos. Sus prejuicios pesaban sobre un análisis objetivo y sobre su capacidad de devolver los créditos: si eran hombres o mujeres, religiosos o no, blancos o negros. Los prejuicios y las injusticias no han venido ni con los datos ni con los algoritmos. De hecho, el uso de modelos matemáticos, como FICO, que hace una proyección del riesgo de impago de un particular o una empresa basado en su historial crediticio, democratizó el acceso al dinero y trasladó los criterios de concesión de manera transparente al solicitante del crédito. Si pagas tus deudas a tiempo, tienes crédito. Así de sencillo.

Sin embargo, las cosas se han complicado un tanto con la aparición del análisis de grandes datos y de los algoritmos opacos. Haber pagado las deudas es un elemento poco relevante si de lo que se trata es de ver el futuro con precisión. Ya sabemos por los folletos de los fondos de inversión que los rendimientos pasados no prometen rendimientos futuros. La honradez o el miedo reverencial, que tanto da a la hora de pagar las deudas, como todos sabemos, se pueden ven alterados con los cambios propios de estar vivo. Comprar en supermercados más baratos, comer menos fuera de casa, ingerir demasiados hidratos de carbono, volverse un runner obsesivo a los cincuenta años o hacer rutinas diarias distintas pueden ser indicativos de problemas económicos, haberse echado un amante o aumentar las probabilidades de tener un fallo cardiaco, elementos todos ellos que limitarían la capacidad de devolución de un crédito o minarían la salud de quien lo solicita, lo que supone, de nuevo, un empeoramiento probable de la capacidad de pago. Un divorcio a un año vista te aleja de una hipoteca tanto como una repentina afición al ejercicio físico aeróbico te puede impedir conseguir un seguro médico, o al menos cerrarlo con una prima aceptable.

Ya no se trata de aguantar, pues, la molestia de una publicidad que, como somos tan perspicaces, evitamos con bloqueadores, sino de dar información que de manera individual es irrelevante pero que, en conjunto, no sólo define quiénes somos sino, lo que es más importante, quiénes vamos a ser.

La aparente frialdad de los modelos matemáticos nos podría llevar a la conclusión de que son neutrales y que sus resultados son la mejor de las opciones posibles, pero no es así. Como muy inteligentemente señala Cathy O’Neil,[101] los modelos matemáticos no son neutros y son la representación simplificada de un proceso. A pesar de su reputación de imparcialidad, reflejan objetivos e ideologías, así como los prejuicios y prioridades de sus creadores. «Los modelos son opiniones embebidas en matemáticas», afirma O’Neil con rotundidad. Y no sólo debido al sesgo de quien los

programa, sino a errores en su diseño que los convierte, en palabras de la autora, en «matemáticas de destrucción masiva» (WMD, en sus siglas en inglés). Para O'Neil son típicos WMD:

- Los que definen su propia realidad con la finalidad de justificar sus resultados. Estos modelos, en su opinión, se autoperpetúan, son muy destructivos y demasiado comunes.
- Los que camuflan entre su código asunciones peligrosas no comprobadas.
- Los que tienden a penalizar al pobre y a perpetuar la desigualdad.
- Los que funcionan como una caja negra, en la que nadie está dispuesto a explicar el razonamiento detrás del resultado.

Los algoritmos que señala O'Neil en su obra son opacos pero predecibles: hacen lo que han sido programados para hacer. Un codificador experto puede, en principio, examinar y desafiar sus fundamentos. Algunos de nosotros soñamos con un ejército de ciudadanos para hacer este trabajo, similar a la red de astrónomos aficionados que apoyan a profesionales en ese campo. La legislación para permitir esto parece lejana pero inevitable.

Un ejemplo de algoritmo tóxico a los que se refiere O'Neil es el sistema COMPAS (Correctional Offender Management Profiling for Alternative Sanctions, o Gestión del perfilado de los internos en centros penitenciarios para sanciones alternativas), una herramienta usada por los jueces estadounidenses para leer el futuro: el sistema les da una tasa de reincidencia que emplean en la toma de sus decisiones. Una especie de precrimen que, mirando al pasado del penado, lee su futuro.

En todo Estados Unidos, los jueces y los agentes de libertad condicional utilizan COMPAS para evaluar la probabilidad de que un acusado sea reincidente, un término utilizado para describir a los

delincuentes que vuelven a delinquir. Hay docenas de estos algoritmos de evaluación de riesgos, desarrollados por varios estados de la unión, y dos herramientas líderes a nivel nacional a disposición de quienquiera pagar por ellas.

ProPublica[102] se propuso evaluar COMPAS, desarrollada por Northpointe, Inc., para descubrir la precisión subyacente de su algoritmo de reincidencia y para probar si el algoritmo estaba sesgado contra ciertos grupos.

El análisis descubrió que los acusados negros eran mucho más propensos que los acusados blancos a que se les marcara incorrectamente como de mayor riesgo de reincidencia, siendo muy bajo el porcentaje de error en los blancos.

ProPublica revisó los expedientes de más de 10.000 acusados en el condado de Broward, Florida, y comparó sus tasas de reincidencia pronosticadas con la tasa reincidencia real en un período de dos años.

El sistema funciona de la siguiente manera: al ingresar en prisión los internos completan un cuestionario COMPAS. Sus respuestas alimentan el software y generan una puntuación que incluye la predicciones de «riesgo de reincidencia» y «riesgo de reincidencia violenta».

ProPublica comparó las categorías de riesgo de reincidencia predichas por la herramienta COMPAS con las tasas reales de reincidencia de los acusados en los dos años posteriores a su calificación, y encontró que predecía correctamente la reincidencia de un delincuente en un 61 por ciento de los casos y que sólo acertaba en un 20 por ciento de los casos de reincidencia violenta.

Al pronosticar quién reincidiría, el algoritmo predijo correctamente la reincidencia para los acusados blancos y negros a aproximadamente la misma tasa (59 por ciento para los acusados blancos y 63 por ciento para los acusados negros), pero cometió errores de maneras muy diferentes. El análisis de ProPublica encontró que:

– Se pronosticaba que los acusados negros corrían un mayor riesgo de reincidencia muy superior al real. ProPublica averiguó que los internos negros que no reincidieron durante un período de dos años tenían casi el doble de probabilidades de clasificarse erróneamente como de mayor riesgo en comparación con sus compañeros de desdichas blancos (45 por ciento frente a 23 por ciento).
– Era habitual que se estableciese que los acusados blancos fueran puntuados con un porcentaje de riesgo menor de lo que realmente les hubiera correspondido. Conforme el análisis, los acusados blancos que volvieron a cometer delitos en los siguientes dos años fueron etiquetados erróneamente de bajo riesgo casi en el doble de los casos que los reincidentes negros (48 por ciento contra 28 por ciento).
– El análisis también mostró que incluso cuando se controlaba por delitos previos, reincidencia futura, edad y género, los presos de raza negra tenían un 45 por ciento más de probabilidades de obtener puntuación de riesgo más alta que los acusados blancos.
– COMPAS clasificaba erróneamente como reincidentes en delitos violentos a los negros en el doble de casos que los blancos, mientras que los reincidentes violentos blancos eran un 63 por ciento más propensos a ser clasificados erróneamente como de bajo riesgo de reincidencia violenta, en comparación con reincidentes violentos negros.
– El análisis de la reincidencia violenta también mostró que incluso cuando se controla por delitos previos, reincidencia futura, edad y género, los acusados de raza negra tenían un 77 por ciento más de probabilidades de recibir ratios de mayor riesgo que los acusados blancos.

O’Neil mantiene que la mayor parte de los modelos matemáticos son opacos. Muchos argumentan que si no lo fueran, las personas objeto de análisis intentarían hacer trampas y forzar una decisión favorable a sus intereses, lo que carece de sentido. Mantener las reglas por las que uno es juzgado en la oscuridad es casi tan terrible como ocultar el nombre de quien te acusa, porque podrías tomar medidas contra él. Éste era el argumento de la Santa Inquisición que, como todo el mundo sabe, se usó para inventar acusaciones contra quien más convenía sin dar al acusado el derecho de defenderse. Por ello, nuestro sistema legal está basado en el principio acusatorio y las decisiones tienen que estar profusamente motivadas. Algo que nos parece tan obvio en el mundo físico, nos resulta perfectamente asumible cuando interviene la tecnología.

El sistema COMPAS permite hacer a los penados preguntas en el cuestionario inicial que no se permiten en Estados Unidos en una sala de juicios; se les cuestiona por su pasado o situación personal, algo que es simplemente inadmisible ante un tribunal. Sin embargo, las personas las contestan en un cuestionario basado en un consentimiento que no es tal. Como tampoco lo es cuando aceptamos los términos y condiciones de un servicio online.

Puede parecer que el prejuicio es un fenómeno humano específico que requiere de la cognición humana para formarse una opinión o estereotipar a cierta persona o grupo. Aunque algunos tipos de algoritmos informáticos ya han exhibido prejuicios, como el racismo y el sexismo, basados en el aprendizaje de registros públicos y otros datos generados por humanos, un grupo de expertos en informática y psicología de la Universidad de Cardiff y el MIT han demostrado que incluso cuando las máquinas se autoprograman pueden desarrollar prejuicios por sí mismas.[103]

El Tribunal Nacional de No Discriminación e Igualdad de Finlandia ha determinado[104] que una persona fue discriminada en la concesión de crédito al consumo porque no se realizó una evaluación individual de la solvencia de la misma. El acreedor sólo

evaluó la solvencia sobre la base de información de antecedentes, como edad, sexo, lugar de residencia y lengua materna. El Defensor del pueblo presentó el caso ante el Tribunal Nacional de No Discriminación e Igualdad para su consideración: declaró que la conducta del acreedor supone una discriminación múltiple e impuso una multa de 100.000 euros. Ésta sería la primera decisión europea sobre la discriminación en la concesión automática de crédito.

Los algoritmos son, por tanto, operaciones y sistemas matemáticos más o menos complejos, mejor o peor diseñados, que pueden generar resultados beneficiosos, como evitar desconectar a un enfermo en coma con posibilidades ocultas de recuperación o, por el contrario, acentuar la pobreza de los pobres, la desigualdad, el machismo y la reincidencia criminal.

No todas las IA son HAL, la voz malvada y testaruda de *2001*, *una odisea en el espacio*, ni parece que haya llegado, aún, el momento de las IA multitarea, omnicomprensivas y capaces de la singularidad. Hay usos de la IA en campos como el de la medicina que son incontestables.

Una IA desarrollada tras de ocho años de investigación por la Academia de Ciencias de China y el Hospital General PLA en Pekín ha sido capaz, con una precisión de casi el 90 por ciento, de acertar el pronóstico en enfermos en coma. Como en la película *Despertares* (*Awakenings*, en inglés) de 1990, basada en la novela homónima publicada por el neurólogo Oliver Sacks, los médicos desahuciaron a pacientes que, sin embargo, se consideraron viables por una IA.

El sistema que usa la IA puede rastrear la actividad cerebral invisible para el ojo humano y así detectarla donde los médicos sólo ven quietud. Los trastornos de la conciencia son una mezcla heterogénea de diferentes enfermedades o lesiones. Aunque se han propuesto algunos indicadores y modelos para el pronóstico, cualquier método único cuando se usa de manera individual conlleva un alto riesgo de falsas predicciones. El objetivo del estudio fue desarrollar un modelo de pronóstico multidominio que combinase la resonancia magnética funcional en estado de reposo con tres características clínicas para predecir resultados a un año vista por cada enfermo. El modelo discriminó entre pacientes que recuperarían la consciencia más tarde y aquellos que no lo harían con una precisión de alrededor del 88 por ciento en tres conjuntos de datos de dos centros médicos. También fue capaz de identificar la importancia para el pronóstico

de los diferentes predictores, incluidas las funciones cerebrales y las características clínicas. Hasta donde sabemos, ésta es la primera implementación precisa, robusta e interpretable de un modelo de pronóstico multidominio que se basa en la resonancia magnética funcional en estado de reposo y las características clínicas en los trastornos crónicos de la consciencia.

Los pacientes eran evaluados mediante un escáner cerebral con imágenes de resonancia magnética funcional, una tecnología que registra la actividad cerebral midiendo pequeños cambios en el flujo sanguíneo.

Las actividades neuronales son demasiado numerosas y sofisticadas para ser directamente visibles para los médicos, pero el sistema de inteligencia artificial, equipado con algoritmos de aprendizaje automático, puede examinar estos detalles cambiantes y descubrir patrones previamente desconocidos de casos anteriores.

Por ejemplo, la máquina puede buscar en regiones que se encargan de diferentes funciones, como control de movimiento, capacidad verbal, audición y visión, para ver cómo interactúan entre sí después de sufrir daños físicos.

Algunos pacientes pueden tener una mente activa, pero su comunicación con el mundo exterior está temporalmente bloqueada. Su estado es más probable que sea juzgado erróneamente por métodos de evaluación convencionales, que generalmente se basan en respuestas conductuales definidas operativamente a estímulos sensoriales específicos.

Al menos siete pacientes sin esperanza de recuperar la consciencia fueron reevaluados por este sistema de inteligencia artificial que predijo que despertarían en un año. Un paciente de diecinueve años pasó seis meses con el síndrome de vigilia sin respuesta, antes conocido como estado vegetativo, después de que un accidente le causara una lesión grave en la sien izquierda. Algunos de los mejores neurólogos de China realizaron cuatro rondas de evaluaciones sobre su potencial de recuperación y le otorgaron 7 de 23 puntos en una escala de recuperación de coma, una calificación que significaba que su familia tenía el derecho legal de desconectar su soporte vital. Sin embargo, después de realizar sus escaneos cerebrales, el sistema le dio más de 20 puntos, cerca del rating completo.

Los médicos dieron a una víctima de accidente cerebrovascular de cuarenta y un años de edad que había estado vegetativo durante tres meses una puntuación potencial de recuperación de seis. Y despertó.

El joven, la mujer de mediana edad y otros cinco pacientes a quienes los médicos creían que nunca se recuperarían se despertaron dentro de los 12 meses posteriores al escáner cerebral, tal como la IA había predicho.

Para tranquilidad de aquellos que ven un futuro a lo *Matrix*, en donde lo seres humanos seremos pilas alcalinas de grandes máquinas de IA, el sistema no es perfecto: un hombre de treinta y seis años que sufrió daño bilateral del tallo cerebral después de un accidente cerebrovascular recibió una puntuación igual de baja tanto por parte del cuerpo médico como de la IA. Se recuperó completamente en menos de un año.

En definitiva, y para simplificar, la fórmula del milagro y de toda la prospectiva basada en datos, es un número suficientemente grande de datos (aunque puedan ser no relevantes en apariencia para el análisis), que sean de calidad, a los que se aplique un sistema de análisis.

Después de todo, el software está escrito por hombres blancos y asiáticos abrumadoramente ricos, e inevitablemente reflejará sus suposiciones. El sesgo no requiere de mala fe para convertirse en dañino y, a diferencia de un ser humano, no podemos pedirle fácilmente a un controlador algorítmico que explique su decisión. O'Neil pide «auditorías algorítmicas» de cualquier sistema que afecte directamente al público, una idea sensata a la que tanto la industria tecnológica como las empresas que los usan se oponen con fiereza alegando secreto empresarial: no van a facilitar la lógica de sus negocios públicamente para que cualquier competidor pueda copiarlo.

Por eso la transparencia algorítmica es tan importante. Los modelos matemáticos son diseñados para ser cajas negras inescrutables ni por quien los diseña, entrena o usa. Las empresas están muy poco por la labor de explicar su funcionamiento y las motivaciones de sus decisiones, como si de la fórmula de la Coca-Cola se tratase.

6

Adiction by design

Condicionamiento

Iván Petróvich Pávlov nació el 26 de septiembre de 1849 en Leningrado. Estudió Teología, Medicina y Química. Se especializó en fisiología del intestino y en el funcionamiento del sistema circulatorio en Alemania, donde dedicó más de diez años a aprender a hacer orificios en el tracto intestinal para la extracción de los jugos gástricos, trabajo por el que ganó el premio Nobel de Fisiología (Medicina) en 1904. Pero nadie lo recuerda por este premio sino por su perro. Como en el monstruo de Frankenstein, la criatura es más famosa que su creador, y Pávlov es más recordado por formular la ley del reflejo condicional que por los diez años de trabajosos agujeros. En su experimento inicial, Pávlov utilizó un metrónomo a 100 golpes por minuto para llamar a los perros a comer y, tras varias repeticiones, los perros comenzaron a salivar en respuesta al metrónomo sin necesidad de que la comida estuviera presente.

Pávlov y su perro

Todos hemos estudiado al perro de Pávlov en el colegio, pero pocos de nosotros sabemos que gracias a la hiperconectividad, las redes sociales y los servicios online de extracción de datos nos hemos convertido en el perro.

Sean Parker, el arrepentido expresidente de Facebook (compañía también dueña de WhatsApp e Instagram), confesó en 2017 que el botón de «like» surgió de las estrategias para tratar de «consumir el mayor tiempo posible de atención consciente de la gente». Su diseño se hizo con la intención de dar a los usuarios un pequeño golpe de dopamina, de explotar una vulnerabilidad de la psique humana: la necesidad de validación social. Añadió: «Sólo Dios sabe lo que le está haciendo a la mente de nuestros hijos».[105]

Somos como pequeñas mascotas a las que se las entrena con premios o a las que se las tiene en estado de ansiedad esperando que el premio suceda. Tal vez ese ejemplo es demasiado encantador. En realidad, nos han convertido en tristes perros de Pávlov conectados.

Las redes sociales copian los métodos del juego, crean ansiedad o necesidad incontrolada, activando los mismos mecanismos cerebrales que la cocaína.[106] Los expertos advierten de que las plataformas utilizan las mismas técnicas que las empresas de juego para crear dependencias psicológicas e integrar sus productos en la vida de sus usuarios. Estos métodos son tan efectivos que pueden activar mecanismos similares a la cocaína en el cerebro, crear necesidades psicológicas e incluso provocar «llamadas y notificaciones fantasma», en las que los usuarios perciben el zumbido de un teléfono inteligente, incluso cuando no está realmente allí.

Facebook, Twitter y otras compañías estarían utilizando métodos similares a la industria de los juegos de azar para mantener a los usuarios enganchados a sus aplicaciones. Natasha Schüll, en su obra *Addiction by Design*,[107] mantiene que se están usando en las empresas de tecnología las mismas técnicas para crear adicción que las empleadas, por ejemplo, en las máquinas tragaperras, usando el mismo mecanismo de creación de «bucles lúdicos». Ya sea en las Snapchat streaks, el scrolling de fotos de Facebook o en la reproducción de CandyCrush, en todo ellos se crean, a propósito, ciclos repetidos de incertidumbre, anticipación y

retroalimentación, dando de vez en cuando y de manera aleatoria unas recompensas, que siempre serán suficientes para que el usuario siga viniendo. Si uno osa desconectarse, el servicio o la aplicación se encargará de mandar mensajes, ofertas o recompensas para llamar la atención del usuario y atraerlo de nuevo al servicio.

Consideremos que el número de usuarios activos mensuales de Facebook era de 2.130.00 millones a principios de 2018, un 14 por ciento más que el año anterior. A pesar de todos los problemas que Facebook ha ido acumulando desde las elecciones presidenciales estadounidenses de 2016 y, específicamente, desde el escándalo de Cambridge Analytica de 2018, la Red Social —con mayúsculas— tuvo ingresos récord de 11,97 mil millones de dólares el primer trimestre de 2018, un 49 por ciento más que el año anterior.

Cuando culpamos a los adolescentes de estar todo el día pegados a la pantalla de sus móviles, nadie mira en dirección a los diseñadores de las aplicaciones. Hay cientos de estudios sobre la adicción de los menores, y no tan menores, en los que se establecen pautas para evitar la adicción del móvil o salir de ella sin tener en consideración que los servicios están diseñados para ser adictivos a propósito. Cuando uno mete una moneda en una máquina tragaperras o entra en un casino sabe que es un juego de azar pensado para sacarnos el dinero a base de hacernos dependientes. La banca siempre gana. Cuando uno accede a Snapchat o descarga WhatsApp cree que está haciendo uso de los servicios que te ofrecen: compartir fotos y momentos temporales con amigos o comunicarse con la familia, el grupo de pádel o el de padres del 5.º C. Nadie cree estar entrando en una casa de apuestas ni cree que ha de tomar medida alguna al respecto. La adicción es vista como un hecho culpable, una debilidad de carácter en la que sólo caen los jóvenes y los que no se saben controlar. La realidad, como veremos, es muy distinta.

Las razones por las que todos estos servicios tienen tanto éxito es simple y llanamente porque no podemos dejarlos.

El psicólogo del comportamiento, Nir Eyal, el autor de *Hooked: How to Build Habit-Forming Products*,[108] describe perfectamente cómo nos relacionamos con las redes sociales: «Todo comienza con un disparador, una acción, una recompensa y luego una inversión, y es a través de ciclos sucesivos, a través de estos ganchos, cuando se forman los hábitos. Los vemos en todo tipo de productos, ciertamente en las redes sociales y en los juegos de azar. Esto es una gran parte de cómo se cambian los hábitos». Una vez que se forma un hábito, ya no son necesarios esos disparadores, esos activadores externos (una notificación, un correo electrónico o cualquier tipo de timbre o llamada) porque se reemplaza con un desencadenante interno mediante la asociación mental entre querer usar este producto y tratar de satisfacer una necesidad emocional.

Los mecanismos *pull-to-refresh* o *infinite scrolling* de cualquier red social son desconcertantemente similares a una máquina tragaperras, en palabras de Tristan Harris, quien trabajó en Google como *design ethicist*, algo así como un encargado de que el diseño de los productos cumpla estándares éticos. Este mecanismo es como el de una tragaperras: tiras de una palanca y recibes inmediatamente una recompensa o nada. Como en las máquinas, no sabemos cuándo vamos a recibir la recompensa (la mayoría de las veces los algoritmos de juegos nos hacen perder, y la mayoría de las veces por mucho que bajemos en nuestro timeline no vamos a encontrar nada interesante o gratificante, como ocurre en el juego). Pero eso es precisamente lo que nos hace volver. El condicionamiento operativo es un proceso de aprendizaje en el que se adquieren y modifican nuevos comportamientos a través de su asociación con las consecuencias. Reforzar un comportamiento aumenta la probabilidad de que vuelva a ocurrir en el futuro, mientras que castigar uno disminuye la probabilidad de que se repita. En el condicionamiento, los horarios de refuerzo son un componente importante del proceso de aprendizaje. Cuándo y con qué frecuencia reforzamos un comportamiento puede tener un mayor impacto en la fuerza y la velocidad de la respuesta.[109]

Un programa de refuerzo es básicamente una regla que establece en qué casos un comportamiento se reforzará. A veces, un comportamiento puede ser reforzado cada vez que ocurre. En otras, un comportamiento puede no ser reforzado en absoluto. Se puede usar refuerzo positivo o negativo, dependiendo de la situación. En ambos casos, el objetivo del refuerzo es siempre fortalecer el comportamiento y aumentar la probabilidad de que vuelva a ocurrir en el futuro.

Hay dos tipos de programas de refuerzo:

– **Programas de refuerzo continuo**. En el refuerzo continuo, el comportamiento deseado se refuerza cada vez que ocurre. Este programa se utiliza mejor durante las etapas iniciales de aprendizaje para crear una asociación fuerte entre el comportamiento y la respuesta.
– **Programas de refuerzo parcial**. En el refuerzo parcial o intermitente, la respuesta se refuerza sólo una parte del tiempo. Las conductas aprendidas se adquieren más lentamente con refuerzo parcial, pero la respuesta es más resistente y es muy difícil deshabituarse. Hay cuatro tipos de refuerzo parcial:
    1. Los programas de proporción fija, en los que una respuesta se refuerza solamente después de un número específico de respuestas. Este programa produce una tasa alta y constante de respuesta con apenas una breve pausa después de la entrega del reforzador. Un ejemplo de un programa de proporciones fijas sería entregar alimento a una rata después de que presione una barra cinco veces.
    2. Los programas de proporción variable se dan cuando una respuesta se refuerza después de un número impredecible de respuestas. Este programa crea una alta tasa constante de respuesta. Los juegos de azar y juegos de lotería son ejemplos de recompensas basadas

en un programa de proporción variable. En el caso español, por ejemplo, las tragaperras tienen que ir diseñadas para dar un porcentaje específico de premios, porque si se dejasen a la mera aleatoriedad el número de premios sería menor. En un entorno de laboratorio, esto supondría alimentar a la rata después de un número variable de veces (tres, dos, siete).

3. Los programas de intervalo fijo son aquéllos en los que la primera respuesta se recompensa sólo después de que haya transcurrido un tiempo específico. Este programa provoca grandes cantidades de respuesta cerca del final del intervalo, pero una respuesta mucho más lenta inmediatamente después de la entrega del reforzador. Un ejemplo de esto sería alimentar a la rata tras un lapso de tiempo después de que presione la barra para obtener el alimento.
4. En los programas de intervalo variable, una respuesta es recompensada después de que haya transcurrido un tiempo impredecible. Este programa produce una tasa de respuesta lenta y constante. Esto sería entregar alimento a nuestra querida rata del ejemplo en intervalos variables tras presionar la palanca; la haríamos esperar, por ejemplo, primero un minuto, después un intervalo de cinco minutos, y tres minutos la tercera vez. Las recompensas de los programas de refuerzo variable son la clave para que los usuarios de las redes sociales estén consultando sus móviles de manera constante.[110]

Las redes sociales están llenas de recompensas impredecibles que captan, de manera permanente, la atención de sus usuarios, creando en ellos una rutina que les obliga a verificar sus pantallas de forma habitual porque nunca saben cuándo ni cómo se va a producir la recompensa.

Al igual que en los juegos de azar, que alteran físicamente la estructura del cerebro y hacen que las personas sean más susceptibles a la depresión y a la ansiedad, se está documentando que el uso de redes sociales provoca un aumento de la tristeza y el potencial de tener un impacto psicológico adverso en los usuarios.

Una tragaperras en el bolsillo

Tristan Harris[111] mantiene y demuestra —porque formó parte de uno— que hay departamentos enteros de grandes empresas tecnológicas que diseñan sus sistemas para que sean lo más adictivos posible, para que estemos permanentemente conectados, para obtener nuestra atención, nuestros datos y servirnos publicidad. Harris mantiene que todos nosotros estamos conectados a este sistema: «Todas nuestras mentes están secuestradas. Nuestras opciones no son tan libres como creemos que son». Harris, insiste en que miles de millones de personas tienen pocas opciones si usan estas tecnologías, ahora omnipresentes, y desconocen en gran medida las formas invisibles en que una pequeña cantidad de personas en Silicon Valley están moldeando sus vidas.

Harris, graduado de la Universidad de Stanford, estudió con B. J. Fogg, un psicólogo del comportamiento muy respetado en los círculos tecnológicos, con la finalidad de controlar las formas en que el diseño tecnológico puede usarse para persuadir a las personas. Su caída del caballo como Saulo comenzó en 2013, cuando trabajaba como gerente de producto en Google. Gracias a su experiencia en la empresa Alphabet elaboró un estudio orientado a minimizar la distracción de los usuarios. Remitió el memorándum a diez colegas cercanos y se extendió como la pólvora, llegando a más de 5.000 empleados de Google, incluidos ejecutivos de alto nivel que recompensaron tanta sinceridad de Harris con lo que vulgarmente se conoce con «una patada *p'arriba*» o, como prefiero

llamarlo yo, un ascenso que lo colocó allí donde acababa la moqueta. Con el rimbombante título de *Google's in-house design ethicist and product philosopher*, Harris acabó siendo ascendido a un rol marginal, sin equipo, sin apoyo, sin poder, encapsulado para evitar que el virus de la conciencia se extendiese por la compañía.

Tanto tiempo libre permitió a Harris comprobar cómo LinkedIn explota la necesidad de reciprocidad social para ampliar su red; cómo YouTube y Netflix reproducen automáticamente los vídeos y los próximos episodios, privando a los usuarios de una opción sobre si quieren o no seguir viendo; o cómo Snapchat creó su característica adictiva Snap streaks, fomentando la comunicación casi constante entre sus usuarios adolescentes.

Como veremos en el apartado siguiente, las técnicas que utilizan estas empresas no siempre son genéricas; gracias a la nanopersonalización pueden adaptarse algorítmicamente a cada persona. Un informe interno de Facebook, que se filtró en 2018 y que ya hemos comentado, reveló que la empresa era capaz de identificar cuándo los adolescentes se sienten «inseguros», «sin valor» y «necesitan un impulso de confianza». Tal información granular, mantiene Harris, permite mantener un modelo perfecto de botones que pulsar para cada persona en particular. Harris cree que las empresas tecnológicas nunca se propusieron deliberadamente hacer que sus productos fueran adictivos. Estaban respondiendo a los incentivos de una economía publicitaria, experimentando con técnicas que podrían captar la atención de las personas, incluso tropezando con un diseño altamente efectivo por accidente. Cuenta Harris que un amigo en Facebook le dijo que los diseñadores inicialmente decidieron que el icono de notificación, el que alerta a las personas sobre nuevas actividades como «solicitudes de amigos» o «me gusta», estuviese diseñado en el color azul corporativo, «pero nadie lo usó», cuenta Harris. «Lo cambiaron a rojo y, por supuesto, todos lo usaron.»

Ese icono rojo está ahora en todas partes. Cuando los usuarios de teléfonos inteligentes consultan sus teléfonos, docenas o cientos de veces al día, se enfrentan a pequeños puntos rojos al lado de sus aplicaciones, suplicando que los toquen. El rojo es un color disparador; éste es el motivo de que se utilice como señal de alarma.

Un diseño más seductor, explica Harris, explota la misma susceptibilidad psicológica que hace que el juego sea tan compulsivo: las recompensas variables que hemos mencionado. Cuando pulsamos las aplicaciones con iconos rojos, no sabemos si descubriremos un correo electrónico interesante, una avalancha de «me gusta» o nada en absoluto. Es la posibilidad de decepción lo que lo hace tan adictivo e incontrolable.

Es esto lo que explica cómo el mecanismo *pull-to-refresh* —en el que los usuarios deslizan hacia abajo, se detienen y esperan para ver qué contenido aparece— se convirtió rápidamente en una de las características de diseño más adictivas y ubicuas de la tecnología moderna. «Cada vez que se desliza hacia abajo, es como una máquina tragaperras —dice Harris—. No sabes lo que viene después. A veces es una hermosa foto. A veces es sólo un anuncio.»

Loren Brichter, la diseñadora que creó el mecanismo de *pull-to-refresh* para actualizar —usado por primera vez para ello en el timeline (TL) de Twitter— es admirada por la comunidad de desarrolladores de aplicaciones por sus diseños elegantes e intuitivos.

Ahora, con treinta y dos años, Brichter dice que nunca pretendió que el diseño fuera adictivo, pero que no cuestiona la comparación con las máquinas tragaperras. «Estoy de acuerdo al ciento por ciento. Tengo dos hijos ahora y lamento cada minuto que no les esté prestando atención porque mi teléfono inteligente me ha absorbido.»

Brichter creó la función en 2009 para Tweetie, principalmente porque no pudo encontrar ningún sitio en el que poner el botón «actualizar» en la aplicación. Sostener y arrastrar el TL para

actualizarlo parecía, en ese momento, nada más que una solución agradable e inteligente. Twitter adquirió Tweetie al año siguiente, integrando el *pull-to-refresh* en su propia aplicación. Desde entonces, este el diseño se ha convertido en una de las características más copiadas por todas las aplicaciones.

Brichter está desconcertado por la longevidad de la función. En una era de tecnología de notificación de inserción, las aplicaciones podrían actualizar automáticamente el contenido sin que el usuario tire de su TL. Podría retirarse sin problema, pero cumple una función psicológica adictiva muy interesante para mantener al usuario pendiente. Después de todo, las máquinas tragaperras serían mucho menos adictivas si los jugadores no pudieran tirar de la palanca por sí mismos. Brichter prefiere otra comparación: es como el botón redundante de «cerrar puerta» en algunos ascensores con puertas automáticas: «A la gente le gusta seguir presionándolo».

El resultado de su diseño preocupa a Brichter. De hecho, tiene bloqueados ciertos sitios web, ha desactivado las notificaciones automáticas, ha restringido el uso de la aplicación Telegram para enviar mensajes sólo a su esposa y a dos amigos cercanos, y ha tratado de retirarse de Twitter, sin mucho éxito. Carga su teléfono en la cocina cuando llega por la noche y no lo toca hasta la mañana siguiente.

No todos en su campo aparecen agobiados por la culpa. Los dos inventores enumerados en la patente de Apple para «administrar conexiones de notificación y mostrar iconos», Justin Santamaria y Chris Marcellino, no sufren la misma preocupación. Fueron contratados por Apple para trabajar en el iPhone cuando eran un par de veinteañeros. Como ingenieros, trabajaron en la parte más tecnológica del sistema de notificación automática lanzado en 2009 para habilitar alertas y actualizaciones en tiempo real. Fue un cambio revolucionario que proporcionó la infraestructura para tantas experiencias que ahora forman parte de la vida cotidiana, desde pedir un Uber, hacer una llamada por Skype o recibir actualizaciones de noticias de última hora.

Rosenstein, el cocreador del «me gusta» de Facebook, cree que puede haber un caso para litigar para el supuesto de la «publicidad psicológicamente manipuladora», ya que el ímpetu moral de los grandes de la tecnología sería comparable a tomar medidas contra las compañías de combustibles fósiles o las tabaqueras. «Si sólo nos importa la maximización de las ganancias, iremos rápidamente a la distopía.»

James Williams no cree que hablar de distopía sea inverosímil. El que fuera estratega de Google, que construyó el sistema de métricas para el negocio de publicidad del buscador global de la compañía, ha tenido una visión desde primera fila de una industria que describe como «la forma de control atencional más grande, más estandarizada y más centralizada en la historia humana». Williams, de treinta y cinco años, dejó Google en 2017 y está a punto de completar un doctorado en la Universidad de Oxford que explora la ética del diseño persuasivo, en un viaje que le ha llevado a cuestionar si la democracia puede sobrevivir a la nueva era tecnológica. Su epifanía llegó hace unos años, cuando notó que estaba rodeado de tecnología que le impedía concentrarse en las cosas que quería.

Hackeando mentes

Tristan Harris publicó su memorándum interno como uno de sus ensayos,[112] en el que compara el diseño de producto con los trucos de magia. Los magos buscan los puntos ciegos, las vulnerabilidades y los límites de la percepción del público para así influir en lo que hacen sin darse cuenta. Una vez que se aprende a hacerlo, es una cuestión de saber qué botones se han de presionar. Los diseñadores de producto juegan con las vulnerabilidades cognitivas y psicológicas de los usuarios (consciente e inconscientemente) para captar su atención en su perjuicio y, tal vez, en contra de su verdadera voluntad.

Harris pone varios ejemplos reales de las técnicas de adicción usadas por las empresas tecnológicas en los últimos años:

1. **Si controlas el menú, controlas las opciones**. La cultura occidental está construida alrededor de ideales de elección individual y libertad. Muchos de nosotros defendemos ferozmente nuestro derecho a tomar decisiones «libres», mientras que ignoramos la forma en que nos manipulan en sentido ascendente mediante menús limitados que no elegimos. Esto es exactamente lo que hacen los magos: mientras que hacen creer al público que son libres, es el mago el que establece y limita las opciones a elegir. Cuando a los usuarios se les facilita un menú de opciones, rara vez preguntan qué es lo que no está en el menú, cuáles son los motivos por los que le dan estas opciones y no otras, cuáles son los motivos de esta elección, y si la misma responde al interés del usuario o al de la empresa que facilita las opciones. Harris pone el ejemplo de un grupo de amigos que quedan para tomar algo. No saben a dónde ir y todos abren Yelp (o lo que sería para nosotros El Tenedor o TripAdvisor) para encontrar recomendaciones cercanas y ver una lista de bares. Ya no son un grupo de amigos de copas buscando dónde ir: se han convertido en un grupo de usuarios comparando bares. De usuarios que creen que la aplicación amplía su capacidad de elegir, de tomar decisiones.
   Mientras la pantalla azul se refleja en sus caras, han dejado de ser un grupo, de disfrutar de su entorno o de improvisar unas cervezas en los bares que tienen a su alrededor sólo con levantar la vista.
   La aplicación ha cambiado la dinámica del grupo, que era quedar para hablar y ponerse al día de sus vidas, para pasar a ser la búsqueda de un sitio bonito e instagrameable. La libertad de elección es ilusoria si no se

puede diseñar el menú, como ilusoria es la libertad de aceptar las condiciones de uso de un servicio cuya longitud requeriría tres días de lectura.

En palabras de Harris, el menú «más poderoso» es diferente al que tiene la mayoría de las opciones. Pero cuando nos rendimos ciegamente a los menús que nos dan, es fácil perder de vista la diferencia, pues todas las interfaces de usuario son menús.

Al dar forma a los menús que seleccionamos, la tecnología secuestra la forma en que percibimos nuestras elecciones y las reemplaza por otras nuevas. Pero cuanta más atención prestamos a las opciones que recibimos, menos tienen que ver con nuestras verdaderas necesidades.

2. **Poner una máquina tragaperras en un billón de bolsillos**. Si eres una aplicación, ¿cómo mantienes a la gente enganchada? Conviértete en una máquina tragaperras. El usuario promedio verifica su teléfono 150 veces al día, no de manera consciente ni porque tenga una verdadera necesidad de hacerlo, sino porque está condicionado. Ya hemos hablado antes de las recompensas variables intermitentes: sólo hay que vincular una acción del usuario con una recompensa variable. La incertidumbre sobre si va a haber recompensa o no, y cuál será, es un potenciador de la adicción.

   Hemos visto cómo gestos típicos de una máquina tragaperras (tirar de la palanca) están incorporados en prácticamente todos los servicios y aplicaciones que usamos. ¿Me enteraré de algo interesante, divertido, picante si actualizo mi TL de Twitter, si autorizo las notificaciones, si actualizo mi correo electrónico o mi perfil de Instagram? ¿No le estoy dando a la palanca para ver qué foto que viene a continuación, si encuentro mi próxima cita en Tinder?

No creo que los diseñadores de Apple o Google tuvieran como objetivo que sus teléfonos funcionasen como máquinas tragaperras. Surgió por accidente y resultó que usar recompensas variables intermitentes en todos sus productos era bueno para el negocio.

3. **Miedo a perderse algo importante o de quedarse fuera: FOMSI (Fear of Missing Something Important) o FOMO (Fear of missing out)**. Otra forma en que las aplicaciones y webs exacerban nuestra adicción es induciéndonos a pensar que hay un 1 por ciento de probabilidad de que te estés perdiendo algo importante. Si una aplicación o un servicio convence a los usuarios de que es un canal para obtener información importante, mensajes, amistades o parejas sexuales, será difícil desconectarlos. Debo de ser una de las pocas personas en España que no usa ni tiene instalado WhatsApp habiendo sido de las primeras en usar el servicio. Tomé la determinación cuando la compañía fue comprada por Facebook por un precio calculado sobre su número de usuarios. No tuve duda de que la compraban por los datos y que, antes o después, integrarían toda la poderosa información que se obtiene de esa aplicación (conversaciones que se creen privadas o apartadas de terceros, estrategias empresariales, rupturas matrimoniales, nacimientos, amenazas...) se incorporaría a la enorme información que ya tiene Facebook de todos, usuarios o no. Así ha sido. Cuando lo cuento la gente me pregunta que cómo puedo vivir así, desconectada de amigos, sin saber lo que está pasando. Muchos me dicen que lo necesitan para el trabajo, que les imponen su uso para estar informados. Ninguna de las razones que me dan me convencen: esa aplicación es un perfecto ejemplo de uso de las técnicas de adicción que estamos viendo y veremos. La realidad es que un porcentaje altísimo de las conversaciones son insustanciales y evitables, que es un

instrumento de control de la presencia y la actividad del otro, que establece dinámicas sociales tóxicas que impiden a la gente irse de grupos que sólo les generan notificaciones, y que hay alternativas menos intrusivas y más sanas. Empezando por pensar que siempre te estás perdiendo algo y que, como decían las abuelas, «si te quiere, llamará».

4. **Aprobación social**. Todos somos vulnerables a la aprobación social. La necesidad de pertenecer, ser aprobado o apreciado por nuestros compañeros es una de las motivaciones humanas más elevadas. Pero ahora nuestra aprobación social está en manos de compañías tecnológicas: cuando me etiqueta un amigo me lo imagino haciendo una elección consciente, no lo veo como una confabulación de Facebook para obligar a éste a etiquetarme o a Twitter para darme un Fav.

   Facebook, Instagram o SnapChat pueden manipular la frecuencia con la que las personas etiquetan las fotos sugiriendo automáticamente todas las caras que deberían etiquetar. O Twitter con la función «Por si te lo perdiste». Así que cuando alguien me etiqueta o me da un Fav a un tuit sin importancia que publiqué hace 18 horas en realidad no es tanto por afecto como que el cerebro de los miembros de mi red está reaccionando a la sugerencia de Facebook. Aunque lo parezca, no están haciendo una elección libre.

   Gracias a opciones de diseño como éstas, Facebook o Twitter controlan la aprobación social. Por ejemplo, cuando cambiamos nuestra foto de perfil Facebook sabe que es un momento en el que somos vulnerables a la aprobación social y lo coloca más arriba en el feed y por más tiempo, para que el usuario se sienta recompensado y reconfortado en su pertenencia a la comunidad de Facebook.

5. **Reciprocidad social (*Tit-for-tat*)**. Somos vulnerables a la necesidad de corresponder los gestos ajenos. Pero al igual que con la aprobación social, las compañías de tecnología manipulan la frecuencia con la que experimentamos esta obligación social.

   En algunos casos, es por accidente. Las aplicaciones de correo electrónico, mensajes de texto y mensajería son fábricas de reciprocidad social. Pero en otros casos, las empresas explotan esta vulnerabilidad a propósito. Si alguien me hace un favor, se lo debo; si alguien me da las gracias, me manda un correo electrónico o un mensaje de texto sería una grosería no contestarle; si alguien me sigue, ¿cómo no voy a seguirle?

   Harris dedicó largo tiempo de su estancia en el cuarto de las escobas en Google a analizar cómo LinkedIn hace uso de la reciprocidad social para mantener enganchados a sus usuarios.

   LinkedIn quiere que la mayor cantidad posible de personas se creen obligaciones sociales, ya que cada vez que se las devuelven (aceptando una conexión, respondiendo a un mensaje, realizando una recomendación o respaldando una habilidad), tienen que hacerlo a través de LinkedIn. Se trata de conseguir que pasen el mayor tiempo posible dentro del servicio.

   Al igual que Facebook, LinkedIn explota una asimetría en la percepción. Cuando uno recibe una invitación de alguien para conectarse, imagina que esa persona ha tomado una decisión consciente, cuando en realidad es probable que respondan inconscientemente a la lista de contactos sugeridos por la red social. En otras palabras, LinkedIn convierte nuestros impulsos inconscientes (para «agregar» a una persona) en nuevas obligaciones sociales que millones de personas se sienten obligadas a pagar.

6. **Cuencos sin fondo, alimentaciones infinitas y reproducción automática**. El profesor de Cornell Brian Wansink[113] demostró que puedes engañar a las personas para que sigan comiendo sopa al darles un recipiente sin fondo que se rellena automáticamente a medida que comen. Con los tazones sin fondo, las personas comen un 73 por ciento más de calorías que las que tienen tazones normales y creen haber consumido 140 calorías de las que realmente han tomado.

   YouTube, Netflix, Amazon Prime, HBO... todos reproducen automáticamente el siguiente episodio o vídeo después de una corta cuenta atrás, en lugar de esperar a que tomemos la decisión consciente de no continuar. Una gran parte del tráfico de estos sitios web es impulsado por la reproducción automática de lo que viene a continuación.

   Otra de las técnicas de adicción es hacer que la gente siga comiendo aunque no tenga hambre, cogiendo una experiencia limitada y finita, y convirtiéndola en un flujo sin fin que nos mantiene consumiendo sin sentido ni criterio. Las empresas tecnológicas explotan este hack: los feeds de noticias están diseñadas a propósito para que se recarguen automáticamente para hacer un scroll sin fin, eliminando a propósito cualquier motivo para hacer una pausa, reconsiderar o abandonar.

   Es muy habitual encontrar en el argumentario de las empresas tecnológicas frases como «sólo facilitamos que el usuario vea el vídeo que quiere ver», «sólo recabamos los datos que el usuario libremente nos da» o «queremos servir publicidad relevante o mejorar su experiencia de navegación», cuando en realidad el libre albedrío se ve limitado por técnicas aplicadas estrictamente por motivos comerciales. No podemos culparles de que quieran que demos más datos, que estemos más tiempo, que

consumamos más. Es su negocio. Deberíamos ser conscientes de que no nos están prestando un servicio público.

7. **Interrupción instantánea *versus* entrega «respetuosa»**. Las empresas saben que los mensajes que interrumpen en tiempo real son más provocadores que una reacción anterior y más eficientes que los mensajes que se envían de forma asíncrona, como el correo electrónico o cualquier bandeja de entrada diferida. Sólo hay que pensar cuántas veces hemos interrumpido a la persona que ha venido a vernos después de habernos pedido una cita por coger una llamada de teléfono que no sabemos ni si nos interesa.

   Por este motivo, Facebook Messenger, WhatsApp, WeChat o SnapChat han diseñado sus servicios y aplicaciones para interrumpir y llamar nuestra inmediata atención, en lugar de ayudar a los usuarios a respetar la atención de los demás. De nuevo, la interrupción es buena para el negocio.

   Con estas interrupciones, igual que con la llamada de teléfono inesperada, aumentamos el sentimiento de urgencia y la reciprocidad social. Por eso Facebook o WhatsApp chivan al remitente que su mensaje se ha visto y cuándo se ha hecho, lo que aumenta la urgencia y obligación de contestar. Sin embargo, Apple permite a los usuarios activar o desactivar las notificaciones de su sistema de mensajería.

   El problema es que, si bien las aplicaciones de mensajería maximizan las interrupciones en nombre del negocio, esto supone un desgaste de energías y recursos que perjudica la economía mundial con tantas interrupciones diarias. Si cuantificásemos el tiempo que empleamos en la labor de consultar y contestar mensajes irrelevantes o que podrían ser contestados al final del día, nos sorprenderíamos de los

millones de horas de trabajo y de descanso que se pierden, con el impacto negativo en la productividad y la salud de todos nosotros.

8. **Combinando tus razones con sus razones**. Otra forma en que las aplicaciones te enganchan es coger las razones del usuario para visitarlas (para realizar una tarea de su interés) y hacerlas inseparables de las razones comerciales de la aplicación (maximizar la cantidad que consumimos una vez que estamos allí).

   En los supermercados colocan las cosas de primera necesidad al fondo para obligarte a pasar por otros productos que no estaban en tu lista de la compra pero que acaban en ella sin saber muy bien por qué.

   Las empresas tecnológicas diseñan sus servicios y aplicaciones de la misma manera. Facebook no permite entrar a ver un evento sin pasar primero por el canal de noticias, y eso no es casual. Cualquier motivo que uno tenga para entrar en Facebook, esta red intenta maximizar el tiempo que el usuario pasa en la web o en la aplicación. En un mundo ideal, las aplicaciones y servicios te darían siempre una forma directa de obtener lo que quieres y lo separarían de su objetivo, pero sabemos que eso no es posible.

   Harris imagina una «carta de derechos» digitales que describiese los estándares de diseño en la que se forzara a las tecnológicas a diseñar los productos en beneficio de los clientes.

9. **Opciones incómodas**. Igual que los magos, las tecnológicas hacen más sencillo elegir lo que ellos quieren que elijamos, y más difícil lo que no. Las tecnológicas nos ponen ante dilemas de «susto o muerte» en las que la capacidad de elección está trucada. Si quieres seguir recibiendo el servicio, has de aceptar las condiciones leoninas de privacidad que hemos redactado en cuatro

tomos. Si no te gusta Facebook, te puedes ir a otra red social, que existiría si Facebook no la hubiese comprado o ahogado. Si no te gusta nuestro servicio de mensajería, siempre puedes darte de baja, pero, como eres un adolescente, pasarás a ser un apestado social. Si eres adicto a nuestra aplicación, la culpa es tuya, contrólate; acude a las miles de charlas sobre cómo contener tu fea compulsión, pero no nos pidas que rediseñemos nuestros productos. Por eso son gratis. ¿Qué te creías?

Así, en lugar de ver el mundo en términos de la elección de opciones, debemos ver el mundo en términos de la fricción requerida para elegir.

Harris imagina un mundo en el que las elecciones vengan con una escala de dificultad (como los coeficientes de fricción) y con una agencia gubernamental como la del medicamento que sería quien las estableciese y controlase. Sería muy útil no sólo para las opciones de internet, sino para la vida en general.

10. **Errores de pronóstico y estrategias de «pie en la puerta»**. La gente no sabe intuitivamente cuál es el coste de un clic o una visualización. Por qué no va a hacer clic en ese vídeo de cachorros tan mono. Es la estrategia de los antiguos vendedores que hacían lo que se llamaba «puerta fría»: se hace una pequeña e inocua petición para comenzar, le regalo este libro en la venta puerta a puerta, o «sólo un clic» en el mundo virtual y, una vez que se ha metido el pie entre la puerta y el marco, lo siguiente es entrar y vender la enciclopedia o decirle al usuario que ya que estás aquí: «¿Por qué no te quedas un rato más y ves nuestros vídeos de gatitos?». Prácticamente, todas las webs utilizan este truco.

    Harris plantea cómo sería si la web y los teléfonos inteligentes ayudaran a pronosticar las consecuencias de los clics. Algunas plataformas como Medium informan del

«tiempo de lectura estimada» al inicio de la publicación. Muchos usuarios valoran estas opciones que manifiestan un respeto de la plataforma a su tiempo. El trato digno al cliente puede ser un valor.

Tiempo bien empleado

Harris cree que «la libertad última es una mente libre, y necesitamos la tecnología para ayudarnos a vivir, sentir, pensar y actuar libremente. Necesitamos que nuestros teléfonos inteligentes, pantallas y navegadores sean exoesqueletos de nuestras mentes y relaciones interpersonales que pongan en primer lugar nuestros valores, no nuestros impulsos».[114]

Muchos de los que han contribuido con su creatividad e inteligencia a que la tecnología nos cree adicción, están pasando por los medios de comunicación entonando el *mea culpa*, advirtiendo de lo que sus creaciones nos están haciendo y contando las medidas que toman para evitar caer en las garras de la compulsión.

Otros, como Harris, están impulsado el movimiento Time well spent (timewellspent.io) como parte del Center for Humane Technology (humanetech.com). Este centro es una alianza sin precedentes de exempleados de las tecnológicas más potentes. A Harris y McNamee se unieron Justin Rosenstein, creador del botón de «Me gusta» de Facebook, Lynn Fox, anterior responsable de comunicación de Apple y Google, y Sandy Parakilas, director de operaciones en el departamento de privacidad de Facebook entre 2011 y 2012.

Sobre Facebook, el colectivo denuncia que su algoritmo está diseñado para maximizar la atención de los usuarios y las horas que dedican a la plataforma, tiempo que está directamente ligado a los beneficios que la tecnológica obtiene por publicidad. El modelo de

negocio se basa en hacer crecer el número de usuarios y las conexiones e interacciones entre ellos para, de esa forma, aumentar su base de datos.[115]

Siendo conscientes de que sólo la sociedad civil y la presión política puede cambiar las cosas, el Center for Humane Technology ha iniciado contactos en el Capitolio para establecer límites legales a las tecnológicas y así conseguir poner freno a esta epidemia de adicción.

Si el lector ha llegado hasta aquí y no está agotado de la lectura, le recomendamos que vea las propuestas que hay encima de la mesa en el capítulo sobre las GAFA.

7

Manipulation by default

Filtro burbuja

El primer artículo sobre Donald Trump que Boris publicó en su web de noticias del mundo, describía cómo, durante un mitin de campaña en Carolina del Norte, el candidato había abofeteado a un asistente que habría manifestado en voz alta su desacuerdo.

La verdad es que esta noticia nunca sucedió. Boris encontró la información en algún foro conspiranoico de internet o en el perfil social de alguien desinformado con mala fe o al servicio de Moscú. Pero a Boris poco le importaba que fuera cierto, falso o que, con su publicación, estuviera haciéndole el juego a alguna potencia extranjera interesada en dar un vuelco a las elecciones a la presidencia de Estados Unidos. Lo único que le preocupaba era alimentar su blog «cosas de interés diario», así que hizo suyo el texto, hasta su última coma y error tipográfico, y lo publicó en su blog. Luego, para llegar a una audiencia más amplia, colgó el enlace en Facebook, diseminándolo en varios grupos dedicados a la política estadounidense. Cuando volvió a consultarlo, se llevó una sorpresa: la noticia del bofetón se había compartido más de 800 veces.[116]

Boris, el macedonio

Ese mes, febrero de 2016, Boris obtuvo más de 150 euros por publicidad de Google en su blog. Cuanto más tráfico, más impresiones, y cuantas más impresiones, más ingresos. En un país, Macedonia, donde el salario mensual ronda esa cifra, parece que publicar noticias falsas pero populares tiene un mayor retorno de la inversión que los estudios o un trabajo de periodista contando la verdad. Así que nuestro Boris, residente en Veles, dejó la escuela y se dedicó a seguir la campaña presidencial desde su pequeño blog.

Considerando que éste era el mejor uso posible de su tiempo, dejó de ir a la escuela secundaria. El trabajo tampoco era agotador. Después de haber trabajado como negro haciendo contenidos a céntimo de euro o haciendo clic en vídeos de YouTube en MicroWorkers.com, sabía todo lo necesario para montar un medio de comunicación y llenarlo de noticias o contenidos copiados. Compró una serie de dominios baratos en GoDaddy —Gossip Knowledge.com y DailyInterestingThings.com— usó plantillas básicas gratuitas de WordPress y los llenó de deportes, cuestiones de salud «magufas», cotilleos y noticias políticas sacadas de cualquier parte.

Tras el éxito cosechado por el inexistente bofetón a Trump, fundó NewYorkTimesPolitics.com, una web que copiaba el aspecto de la página de inicio de *The New York Times* y se dedicaba a la política estadounidense. Tras recibir una orden de cese y desistimiento del periódico, fundó PoliticsHall.com, y un par de meses más tarde, agregó USAPolitics.co a su cartera.

Boris desarrolló una rutina. Varias veces al día, buscaba en internet artículos a favor de Trump, los copiaba en uno de sus dos blogs y luego los viralizaba en grupos de Facebook con nombres tales como «My America, My Home», «The Deplorables», o «Friends Who Support President Donald J. Trump». Los grupos de Trump superaban con mucho el número de los grupos de Clinton, lo que simplificó la viralización de sus artículos.

Publicó también con su propio nombre, pero a través de 200 perfiles de Facebook que había comprado para este fin. Un perfil falso con un nombre ruso le costó alrededor de 10 centavos. Para usar uno estadounidense, tuvo que pagar algo más, unos 50.

Boris aprendió trucos para monetizar mejor sus sitios web: grandes anuncios con enormes interstitials que invadían la página para que los visitantes, bien voluntariamente, bien por error, acabaran haciendo clic en el anuncio. Poca inversión y gran retorno.

Tanto trabajo duro fue recompensado generosamente por los motores de publicidad automatizados, como AdSense de Google, y Boris se pudo comprar un ordenador nuevo e irse de vacaciones.

No fue el único. En Veles, una ciudad de 55.000 habitantes, crecieron las webs en favor de Trump como setas, hasta 100, todas ellas plagadas de noticias falsas o sensacionalistas. Tuvieron especial éxito las noticias que resaltaban el apoyo del papa a Trump o el inminente encarcelamiento de Hillary Clinton por el tema de los correos electrónicos mandados en su época de secretaria de Estado desde su servidor personal.[117]

The New Yorker, en una historia de David Remnick,[118] desveló los motivos por los que el propio Obama, la semana final antes de la votación, se dedicó casi obsesivamente a hablar de la ciudad de Veles y de su «fiebre del oro digital». Acertó con el diagnóstico, pero llegó tarde para frenar un fenómeno montado sobre la imparable y poderosa ola de la tecnología de datos.

En España tampoco nos podemos quejar, y aparte de algunos foros que nutren las noticias más de lo necesario, hay ejemplos de Boris locales como el de Luis Domínguez Villadiego y su web Digital Sevilla en la que publica noticias como «La Mr. Patato andaluza: de Jennifer a Susana Díaz» o «Tres cosas que tienen en común Díaz y Kim Jong-un». Nuestro Boris alcanzó su récord de audiencia con 496.000 usuarios únicos, superando en casi 50.000 a la web del periódico *Diario de Sevilla,* uno de los medios líderes de la provincia.

Su éxito, como el de Boris, se debió no sólo al efecto multiplicador del clickbait y las redes sociales, sino a la publicidad que, como hemos visto, recompensa el contenido de odio o típicamente troll. Se trata de generar contenidos que se visualicen mucho por su tono sensacionalista o insultón, «atraer» la publicidad programática, esa publicidad automática que Google y otras compañías colocan en las páginas con más tráfico. Como en el caso de las granjas de noticias de Macedonia, se trata de que se coloque el mayor número de publicidad posible en nuestro blog, algo que se consigue alimentándolo a menudo a base de falacias y bulos.

Una vez que las compañías han enganchado a los usuarios a las aplicaciones y servicios que corren sobre sus terminales inteligentes, tienen que monetizar ese esfuerzo, y devolver a sus accionistas el dinero invertido en forma de beneficios.

En Macedonia, hacer trampas para sacar dinero de la publicidad es una actividad a la que se dedican desde hace tiempo junto con granjas chinas de bots, donde programan móviles para entrar en páginas que suman visitas sin intervención humana.

En 2017, Google inició el reembolso de importantes cantidades de dinero a sus anunciantes tras comprobarse que su publicidad fue alojada en sitios web que generaban tráfico falso.[119] Según la Association of National Advertisers, las pérdidas derivadas del fraude de la publicidad (uso de webs falsas o bots) habría sido de 7.200 mil millones de dólares en 2016, habiendo descendido un 10 por ciento en 2017[120] hasta los 6.500 mil millones de dólares en todo el mundo. Esto podría suponer casi la mitad del mercado mundial publicitario.

El filtro burbuja

El mal de la publicidad no afecta sólo a la juventud de Macedonia o a la sevillana: es una desgracia que ataca a los medios de comunicación tradicionales, algunos centenarios, que rellenan sus

webs y sus redes sociales de noticias intrascendentes y que cuentan como toda fuente con el tuit de un famoso, la foto de una celebrity de extrarradio en Instagram o el último flame del turno de mañana en Twitter. El clickbait, el anzuelo para pescar visitas y así monetizar con publicidad barata la edición digital del medio, es consecuencia de un sistema adictivo de caza de datos y de audiencias, en el que lo único importante es captar la atención por el método que sea. La perversión de un sistema que llega cuando todo el mecanismo que se usa para entontecer con publicidad de consumo se usa para alelar a un votante.

En el capítulo anterior hemos recogido las informaciones y declaraciones de antiguos trabajadores de Facebook y Google en las que manifiestan, no sin horror y preocupación, que el uso de los datos de los usuarios para conocerlos en profundidad se está empleando en mecanismos de manipulación emocional y de refuerzo de sus posiciones, haciendo un uso tan brillante como insensato de los sesgos cognitivos de los que acceden a su plataforma: sólo vemos aquello que nos agrada y lo que nos reafirma que nuestra visión de la realidad es la correcta, eliminando del timeline de los usuarios cualquier información, dato o imagen que muestre una visión distinta del mundo. El algoritmo que hace posible servir publicidad personalizada, hecha a la medida de cada individuo, es el mismo que lo retiene cautivo dándole lo que quiere.

Roger McNamee[121] se ha unido al grupo de críticos con las prácticas monopolísticas de las GAFA tras haber sido asesor de Facebook durante muchos años y conocer en profundidad el funcionamiento de su algoritmo. En su artículo «How to Fix Facebook – Before It Fixes Us» nos narra su creciente estupefacción y preocupación ante la deriva que el uso de la tecnología de Facebook estaba tomando, para acabar en el activismo junto a Barry Lyn. Hablaremos más de ellos en el capítulo siguiente.

McNamee observó durante las primarias demócratas a las presidenciales de 2016 una avalancha de memes antiClinton claramente misóginos que parecían tener su origen en los grupos de Facebook que apoyaban a Bernie Sanders. En opinión de McNamee, esos mensajes no estaban basados en la construcción de tráfico orgánico sino que había algo más: en el caso de las primarias demócratas y en la posterior votación en la que resultó ganador Trump contra todo pronóstico, ni el sentido común ni los datos de votación explicaron por completo un factor crucial: el nuevo poder de las plataformas sociales para amplificar los mensajes negativos.

Los académicos han estudiado durante mucho tiempo la interacción entre los usuarios y las interfaces que utilizan para producir y consumir información y el papel activo que desempeñan éstas en la configuración de nuestro comportamiento de búsqueda y producción. En la era de internet, un pequeño número de programadores en Silicon Valley ha ejercido cada vez más poder sobre nosotros mientras nos mantenía felizmente ignorantes de cómo nos hacía adictos para, a continuación, manipularnos.

Los algoritmos pensados para maximizar la atención tienen una consecuencia perversa e inesperada: dan ventaja a los mensajes negativos. Parece que los seres humanos reaccionan más a los estímulos que llaman a su cerebro profundo o reptil. Como cualquier usuario puede comprobar, el miedo y la ira producen mucha más participación y engagement que la alegría, por lo que los algoritmos favorecen los contenidos sensacionalistas, irrelevantes o groseros sobre contenidos de más peso intelectual.

Alguien podría decir que siempre ha habido periódicos amarillistas que basaban su línea editorial en los sucesos y los escándalos poco corroborados. O que las líneas editoriales de los medios siempre han venido marcados por un determinado sesgo político, religioso o, simplemente, por los intereses de los grandes anunciantes. La gran diferencia es que el que compraba un diario por ser de tendencia conservadora, por ejemplo, era consciente

tanto de lo que compraba como de que rechazaba activamente otras opciones disponibles en el quiosco. Además, cualquier medio de comunicación tradicional, a pesar de su línea editorial, producía un contenido único para una pluralidad de lectores o televidentes.

Ya no es así. Cada usuario tiene su propio canal personalizado y, lo que es peor, cree que es el mismo que ve el resto, creando la impresión ya comentada de que la realidad es tal como él o ella la concibe. Si lo combinamos con el carácter adictivo de las redes y del uso de técnicas de manipulación para exacerbar nuestro lado más irracional, la negatividad, el extremismo y la alienación están servidos. Es un hecho que las plataformas como Google o Facebook segregan a los individuos y les sirve la realidad a través de un filtro burbuja de ideas afines, reduciendo el riesgo de exposición a ideas desafiantes.

Es lo que Eli Pariser bautizó en 2011 cuando introdujo el término «filtro burbuja», que se ha convertido en un referente para hablar de la manipulación en internet.[122]

El filtro burbuja es el aislamiento intelectual que se produce como consecuencia del uso, por parte de las webs, aplicaciones o servicios, de algoritmos para adivinar, de forma selectiva, la información que un usuario desearía ver (en función de la información del usuario, como la ubicación, el comportamiento anterior y el historial de búsqueda) y proporcionársela.

El filtro burbuja es la lógica consecuencia de cómo funcionan las empresas de datos y publicidad. Primero recaban todos los datos posibles sobre nosotros, nuestro comportamiento, gustos reales, sentimientos y pensamientos. A continuación, los analizan y construyen a nuestro doppelgänger para ofrecérselos a los anunciantes como objetivo de sus campañas: «Tenemos tantos adolescentes con tendencia al suicidio que viven en ciudades costeras del sur de Europa» u «ofrecemos tantos hombres con complejo de impostor y unos ingresos por encima de los 200.000 euros residentes en las afueras». Esas audiencias intuitivas, que se usaban en la época de los grandes medios de comunicación, son

detalladas pero poco específicas. Gracias a los grandes datos y los algoritmos podemos hacer microtargeting, dirigir campañas tan personalizadas como se quiera.

Eso es exactamente lo que hace el algoritmo de Facebook cuando construye nuestros feeds, y el de Google cuando nos responde a una búsqueda. Desde 2009, nadie recibe el mismo resultado buscando lo mismo. Google lo personaliza tanto que las búsquedas han perdido la riqueza de antaño, siendo planas, previsibles y, en muchos casos, inútiles.

Sólo nos muestran lo que creen que queremos ver para así servirnos la publicidad que adivinan va a tener más impacto en nosotros. De paso, refuerzan nuestro sesgo de confirmación: si internet dice lo mismo que yo pienso es que tengo razón. Como resultado, no vemos la información con la que no estamos de acuerdo, aislándonos en nuestras propias burbujas culturales o ideológicas.

Es llamativo cómo la concentración de las fuentes que usamos para informarnos en unas pocas manos, unido al hecho de que consideramos que representan la totalidad de internet, esto es, del conocimiento humano, nos mantiene peor informados que nunca.

Así pues, la herramienta más importante utilizada por Facebook y Google para mantener la atención del usuario (el filtro burbuja, el uso de algoritmos para dar a los consumidores «lo que quieren») conduce a un flujo interminable de publicaciones que confirman las creencias existentes de cada usuario. En Facebook es su feed de noticias, mientras que en Google son los resultados de buscador personalizados individualmente. El resultado es que todos ven una versión diferente de internet diseñada para crear la ilusión de que todos los demás están de acuerdo con ellos. El refuerzo continuo de las creencias existentes tiende a afianzar esas creencias más profundamente, al tiempo que las hace más extremas y más resistentes a los hechos contrarios. Facebook lleva el concepto un paso más allá con su función de «grupos», que alienta a los usuarios de ideas afines a congregarse en torno a intereses o

creencias compartidas. Si bien esto aparentemente proporciona un beneficio para los usuarios, las mayores bonificaciones se otorgan a los anunciantes, que pueden llegar a las audiencias con mayor eficacia.

Los sorprendentes resultados de las elecciones presidenciales de Estados Unidos de 2016 pusieron de manifiesto cómo la manera en la que unas empresas privadas gestionan su negocio no sólo afecta a la intimidad de los ciudadanos (algo que parece no importarle a nadie porque beneficia a todos), sino a la propia democracia y la estabilidad económica y geopolítica mundial.

Mark Zuckerberg, director ejecutivo de Facebook, lleva dos años dando explicaciones de lo ocurrido y se ha comprometido en diversos foros a poner solución al problema de las fake news y a la interferencia de países extranjeros en las elecciones del mundo libre. Pero nada ha dicho de solucionar el filtro burbuja.

Las medidas son tan cosméticas y poco eficientes como se podía esperar. Facebook no puede reconocer que es un editor que gestiona los contenidos que ven sus usuarios. Al negar su naturaleza de mero intermediario, de plataforma neutra en la que la conversación es de sus usuarios, deviene, tanto por la ley estadounidense como por la europea, responsable de todo lo que digan cada uno de sus miles de millones de usuarios. Y eso acabaría con el negocio.

Para no ser acusado de censura previa, organizaciones exteriores verifican el contenido de las noticias, haciendo un fact checking no exento de sesgo, y se informa al usuario de que la noticia o el contenido que va a ver podría ser falso. Teniendo en cuenta la polarización política creada por estos medios, y animada por el propio Donald Trump, quien acusa de sesgo izquierdista a los resultados de las búsquedas de Google con su nombre, es difícil pensar que quien vea ese aviso lo va a leer de modo objetivo. Simplemente porque hemos sido manipulados para perder la objetividad. Nuestros sesgos son la nueva objetividad.

Una de las medidas que ha adoptado tímidamente Facebook es mejorar el sistema de publicidad política a través de su plataforma que no parece haber tenido un gran impacto. Un clásico ejemplo de lo ocurrido hasta el momento fue la campaña del *brexit*.

El mensaje crudo y emocional del «Leave» se compartió mucho más que el del «Remain». A eso se unió muy probablemente el hecho de que los mensajes del Leave fueran más baratos para los anunciantes: el precio de los anuncios de Facebook (y Google) se determina mediante una subasta y el coste de dirigirse a consumidores más sofisticados hace aumentar la puja. En sentido contrario, inflamar los ánimos de la clase menos favorecida resultó ser más barato y, por su situación, más efectivo. Como consecuencia de este sistema, Facebook se presentaba como una plataforma mucho más barata y efectiva para los agitadores del Leave en términos de coste por usuario alcanzado, asegurando, además, gracias al filtro burbuja, que no recibirían información alguna que cuestionara sus creencias. El modelo de negocio de Facebook puede haber tenido el poder de remodelar la política y el bienestar de todo un continente.

La última ocurrencia de Facebook, con tal de no contar cómo funciona el algoritmo, ha sido montar una «sala de guerra», una «*war room*» con vistas a las elecciones de «midterm» estadounidenses, en las que en noviembre de 2018 se renovó el Congreso, Senado y un número importante de gobernadores estatales.

Facebook invitó a dos reporteros de *The New York Times* a la sala de guerra, aunque no les dejaron ir más allá de la entrada. El control de la información y la opacidad de la compañía es de todos conocida. En la visita, Facebook informó a los periodistas de las medidas que estaba acometiendo:

– Construir defensas contra los spammers, hackers y agentes extranjeros.

– Nuevas contrataciones para ayudar a moderar el contenido y establecer un catálogo de todos los anuncios políticos.
– Desarrollo de un software personalizado para rastrear la información que fluye a través de la red social en tiempo real.
– Paneles de seguimiento en tiempo real que muestran de un vistazo cómo (y si) está cambiando la actividad en la plataforma.

Y mientras Facebook se enfrenta a la tormenta perfecta, con caídas significativas en su cotización y ataques que comprometen las cuentas de sus usuarios, Google pasa desapercibida.

Mientras se centra el debate en los bots rusos en Twitter y Facebook nadie se fija en lo importante: desmontar el filtro burbuja y rediseñar todo el modelo de la economía de los datos relacionada con la publicidad.

En el capítulo siguiente, veremos algunas propuestas, que pasan por empezar a aplicar la normativa antimonopolio a las empresas tecnológicas de datos.

8

GAFA

Fake news, democracia y monopolios

En la obra *Towards the Year 2018* que recogía las ponencias presentadas en el curso de una conferencia de tres días organizada en 1968 con motivo del 50 aniversario de la Foreign Policy Association, el fundador del departamento de Ciencia Política del Massachussets Institute of Technology (MIT), Ithiel de Sola Pool, predecía el futuro pronosticando que el auge de los ordenadores permitiría la acumulación de datos personales y la creación de archivos con un enorme poder de manipulación: «En 2018 un investigador sentado ante una consola personal será capaz de cruzar datos de compra (de los registros de las tiendas) con datos de cociente intelectual (de los registros escolares) y empleo (de los registros de la Seguridad Social). Tendrá esa capacidad tecnológica, pero ¿tendrá derecho a hacerlo?».[123]

Como pasa en todos los ejercicios de futurología, sólo nos fijamos en aquellos que aciertan el pronóstico y, sin duda, no podemos negar a Ithiel de Sola Pool su capacidad profética: sólo cinco décadas después, las GAFA acumulan, analizan y venden los datos de sus usuarios, deciden lo que ven y lo que no, a quién votan y a quién odian, qué piensan y pensarán, qué desean y desearán. La pregunta final del profesor del MIT sigue siendo tan relevante como acertada su profecía. Las empresas cuentan con una potente

tecnología para el procesamiento de datos que les permite diseccionarnos hasta nuestra última emoción, pero ¿tienen derecho a hacerlo?

El poder monopolístico de las GAFA

Las GAFA (acrónimo de Google, Apple, Facebook y Amazon) han aumentado su poder como monopolios convirtiéndose, para algunos, en obstáculos para la innovación al abusar de su posición de dominio en el mercado, bien comprando a la competencia, bien anulándola. No sólo eso: la concentración del mercado en pocos prestadores que recaban millones de datos, junto con el uso intensivo que de los mismos hacen, supone un riesgo para la intimidad de los ciudadanos y para su desarrollo personal pleno. El diseño de algoritmos para analizar a los usuarios y servirles la publicidad personalizada a su carácter, ha sido explotado con la finalidad de ofrecer informaciones, en muchos casos falsas, como parte de campañas de desinformación en procesos electorales tan trascendentes como el *brexit* o la elección del presidente de Estados Unidos.

Si bien la economía del dato figura en la agenda de todas las empresas embarcadas en la transformación digital, las GAFA son, con diferencia, las empresas que más eficientemente han usado el dato o han aprovechado la ausencia de limitaciones a la innovación. Al empuje del argumento de la innovación como ariete, le ha acompañado un crecimiento económico en Estados Unidos que justificaba que el mismo se produjese sin supervisión. El hecho de que los derechos de los ciudadanos estadounidenses no pareciesen afectados, abundaba en la falta de interés del regulador de este país en tomar medida alguna.

No todas las GAFA tienen el mismo modelo de negocio, ni todas ellas hacen un uso igual de los datos. Apple y Amazon, las dos aes del acrónimo, están del lado del sector productivo o del

retailer, y las quejas con respecto a las mismas vienen por su tamaño, su modelo fiscal y, en el caso de Amazon, por su transversalidad y afectación del comercio minorista.

Las consonantes del acrónimo, Google y Facebook, con modelos de negocio completamente basados en los datos, su extracción y explotación, han concentrado las críticas más virulentas y han protagonizado los escándalos más sonados en cuanto al uso descontrolado de los mismos, afectando a instituciones tan sagradas como el propio funcionamiento de la democracia a través de las urnas.

En 2017, Google fue sancionada por la Comisión Europea por abuso de posición de dominio en el mercado de las búsquedas, y Facebook empezó a recibir serias acusaciones sobre el uso de su plataforma para la difusión de campañas de desinformación y manipulación política, sospechas que se vieron confirmadas con el escándalo, en 2018, de Cambridge Analytica, consultora política británica contratada para la campaña del presidente Trump. Un antiguo trabajador de esta última desveló cómo habían extraído grandes cantidades de datos de Facebook (más de 85 millones en Estados Unidos según las últimas estimaciones) con una simple encuesta. La plataforma permitía a los encuestadores extraer el perfil completo de los que contestaban a la misma junto con el de todos sus contactos, lo que facilitaba hacer un análisis de comportamiento de millones de personas, pudiendo así lanzarles mensajes personalizados aprovechando el sesgo del algoritmo de Facebook. El escándalo, que ha llevado a Mark Zuckerberg a comparecer ante el Congreso y Senado de Estados Unidos y ante el Parlamento Europeo, no sólo se centraba en la fuga de datos, sino en el propio funcionamiento de una plataforma que se habría diseñado para mostrar realidades paralelas y personalizadas manteniendo a sus usuarios dentro de su zona de confort, de su visión sesgada del mundo. Es el ya comentado filtro burbuja. En este sentido, el uso y abuso del algoritmo basado en el sesgo cognitivo de los usuarios de Facebook y su sistema publicitario,

pone en cuestión que Facebook y otras redes sociales sean un mero intermediario sin control sobre el contenido y, por tanto, sin responsabilidad sobre los mismos.

Las GAFA no son ajenas al uso de las WMD que nos ha enseñado O'Neil, y usan las matemáticas para predecir el comportamiento de los usuarios, el cómo y el cuándo, diseñando sus productos de un modo tan atractivo, «jugable», adictivo y útil que les permite recabar la enormidad de datos que van a necesitar para alimentar el algoritmo.

Como hemos indicado, tanto Facebook como Google y otras plataformas sociales viven del mercado publicitario, lo que coloca a los anunciantes en la posición del verdadero cliente, mientras que los usuarios de las plataformas son el producto. Es especialmente interesante ver cómo ha evolucionado el mercado publicitario y cómo este cambio tiene su impacto en el modelo de negocio de los datos y, por lo que vemos, en el propio proceso democrático.

Fake news y bots rusos

Hasta la década pasada, las plataformas de medios basaban su modelo de negocio en la transmisión de contenidos «para todos los gustos». Así, la televisión o la radio buscaban contenidos transversales que atrajesen al mayor número de personas posibles, lo que hacía que la publicidad se orientara a los contenidos con mayor audiencia. Este tipo de «contenido convincente» era esencial, ya que permitía a las audiencias elegir entre una variedad de medios de distribución, ninguno de los cuales podía esperar atraer la atención de ningún consumidor individual durante más de unas pocas horas. Los televisores no eran transportables y, aunque algunos ordenadores sí lo eran, eran lo suficientemente incómodos como para no ser un canal atractivo para el consumo de contenidos de manera intensiva.

Cuando el negocio de la generación de contenido se limitaba al entorno del PC como sistema de distribución, las plataformas de internet estaban en desventaja. Su contenido no podía competir con los medios tradicionales, y su canal, el ordenador, era incómodo de utilizar fuera de un escritorio. Su única ventaja, los datos, no era suficiente para superar la desventaja en cuanto al contenido. Como consecuencia de esto y de que la conectividad permitía conocer de primera mano el impacto de la publicidad sin tener que acudir a mediciones basadas en aparatos que nadie tenía, las plataformas web bajaron de manera muy importante el precio de su publicidad.

La publicidad, que seguía siendo genérica a pesar de las cookies, no valía lo que costaba hasta que el usuario se instaló en la plataforma y el algoritmo fue capaz de conocerlo con una precisión inusitada. Los teléfonos inteligentes acabaron de cambiar el mercado publicitario y la batalla por mantener la atención de los usuarios el mayor tiempo posible la ganaron Facebook y Google con sus enormes repositorios de contenidos en tiempo real. ¿Por qué pagar un periódico con la esperanza de captar la atención de una parte determinada de su audiencia, cuando se puede pagar a Facebook para llegar exactamente a las personas que uno quiere y a nadie más?

En el caso de las elecciones presidenciales estadounidenses, se ha demostrado que la trama rusa habría señalado a un grupo de usuarios susceptibles a su mensaje, usando las herramientas de publicidad de Facebook para identificar a los usuarios con perfiles similares y publicando anuncios para persuadir a esas personas para que se unieran a grupos dedicados a temas controvertidos. Los algoritmos de Facebook habrían favorecido el crudo mensaje de Trump y las teorías de conspiración antiClinton que entusiasmaron a sus seguidores, con la posible consecuencia de que, como en el caso británico, Trump y sus patrocinadores hubieran pagado menos que Clinton por la publicidad en Facebook en términos de persona alcanzada.

Y, si esto no reviste la suficiente gravedad, una vez que los usuarios estaban en los grupos, los rusos habrían usado cuentas falsas que se harían pasar por trolls estadounidenses y «bots» para compartir mensajes incendiarios y organizar eventos. Los trolls y los bots, que se hacían pasar por ciudadanos estadounidenses, habrían creado la ilusión de un mayor apoyo a las ideas radicales de lo que realmente existía. Los contenidos con «me gusta» de usuarios reales se compartían por estas cuentas fake y por los bots en sus propios feeds de noticias, de tal modo que una pequeña inversión en publicidad y memes publicados en grupos de Facebook llegaría a decenas de millones de personas.

El Parlamento Europeo, en una resolución de 25 de octubre de 2018,[124] solicitó una auditoría completa de Facebook, así como nuevas medidas contra la intromisión de las elecciones a raíz del escándalo de FacebookCambridge Analytica. La resolución resume las conclusiones alcanzadas tras la reunión de mayo de 2018 entre los principales eurodiputados y el CEO de Facebook, Mark Zuckerberg, y las tres audiencias posteriores. También hace referencia a la violación de datos sufrida por Facebook el 28 de septiembre de ese año, a la que nos referimos en el capítulo «Data Breach» más adelante. Los eurodiputados señalaron en esta resolución que los datos obtenidos por Cambridge Analytica podrían haber sido utilizados con fines políticos, tanto en el referéndum del *brexit* como en las elecciones presidenciales estadounidenses de 2016. Asimismo, destacaron la urgencia de contrarrestar cualquier intento de manipular las elecciones de la UE y, a tal fin, entendieron que toca adaptar las leyes electorales para reflejar la nueva realidad digital. Con la finalidad de evitar la injerencia electoral a través de las redes sociales, los eurodiputados propusieron en su resolución las siguientes medidas:

– Aplicar las salvaguardas electorales convencionales del mundo offline al entorno online, y en concreto, el uso de normas sobre transparencia y límites del gasto, respeto por

las jornadas de reflexión e igualdad de trato de los candidatos.
- Identificar quién está detrás de la publicidad electoral pagada en redes sociales y otros entornos.
- Prohibir los perfiles para fines electorales, incluido el uso de comportamientos online que pueden revelar preferencias políticas.
- Obligar a las redes sociales a etiquetar el contenido compartido por los bots, a acelerar el proceso de eliminación de cuentas falsas y a trabajar con verificadores de hechos independientes, así como con la universidad para abordar el problema de la desinformación.

Las investigaciones sobre el uso indebido por parte de «fuerzas extranjeras» del espacio político online deberían ser llevadas a cabo, en opinión del Parlamento Europeo, por los Estados miembros con el apoyo de Eurojust.

Acabemos con los monopolios tecnológicos

Toda esta historia se lee como la trama de una novela de ciencia ficción distópica: una tecnología que se celebra para unir a las personas se explota con un poder hostil para separarlas, socavar la democracia y crear miseria. Esto es precisamente lo que sucedió en Estados Unidos durante las elecciones de 2016. Los estadounidenses dedicaron ingentes recursos en ciberseguridad y ciberguerra para evitar ataques desde el extranjero, sin imaginar que un enemigo pudiera infectar las mentes de sus ciudadanos mediante un invento de su propia creación y a un mínimo coste. Como el tiempo y las investigaciones han acabado demostrando, el ataque no sólo fue un éxito abrumador, sino que también fue persistente, ya que el partido político que se benefició se niega a

reconocer la realidad. Los ataques continúan todos los días, planteando una amenaza existencial para los procesos democráticos de toda Europa occidental.

Facebook y Google han respondido reiterando su oposición a la regulación gubernamental, insistiendo en que la misma acabaría con la innovación y perjudicaría la competitividad global del país, asegurando que la autorregulación, ésa que habrían aplicado hasta ahora, produciría mejores resultados.

Por tanto, nos encontramos con una explosiva combinación nunca vista antes: crecimiento ilimitado y ausencia casi total de control legal o regulatorio. En este último aspecto, las GAFA participan de una estructura impositiva extractiva, y, hasta la entrada en vigor del Reglamento General de Protección de Datos Europeo, de una exclusión de las normas de privacidad de otros países por extraterritoriales, así como de una casi total ausencia de control por parte de los órganos de competencia. Las GAFA contaban y cuentan, en definitiva, con todas las ventajas de operar extraterritorialmente y con ausencia de control en el propio territorio.

Diversos sectores sociales han comenzado a solicitar a la administración estadounidense un cambio de rumbo mediante la aplicación, entre otras, de estrictas reglas antimonopolio a estas entidades, de responsabilidades por los contenidos que ofrecen de manera individualizada, todo ello con la finalidad de preservar la competencia, la innovación y el acceso universal justo y abierto a los servicios de internet. La cuestión del excesivo poder de las GAFA y de la ausencia de control sobre las mismas centró el discurso del hombre que hundió la libra en 1992, George Soros, en el Foro Económico Mundial de Davos, Suiza, el pasado enero de 2018:[125] «A medida que Facebook y Google han aumentado su poder como monopolios se han convertido en obstáculos para la innovación, causando una variedad de problemas de los cuales nos estamos empezando a dar cuenta ahora [...]. Afirman que simplemente distribuyen información, pero el hecho de que sean distribuidores casi monopolísticos los convierte en servicios públicos, y deberían

someterse a regulaciones más estrictas dirigidas a preservar la competencia, la innovación y el acceso universal justo y abierto [...]. Los dueños de estas gigantescas plataformas se consideran los amos del universo, pero, de hecho, no son más que esclavos de la necesidad de preservar su posición dominante [...]. Es sólo cuestión de tiempo que se rompa el dominio global de los monopolios de TI en Estados Unidos. Las causas de ese final serían la regulación y los impuestos».

Citó Soros a la comisaria de Competencia de la Unión Europea, Margrethe Vestager —quien ha impuesto las sonadas multas a Google—, como un modelo que debería inspirar a otros reguladores del mundo. Soros explicó de qué modo el monopolio permite que las redes sociales tengan una rentabilidad excepcional, comenzando con el hecho de que las GAFA «eluden la responsabilidad de —y evitan pagar por— el contenido de sus plataformas». Sin el coste de generar contenidos, basan su modelo en la publicidad y en la venta de producto: «El modelo de negocio de las compañías de redes sociales se basa en los anuncios. Sus verdaderos clientes son sus anunciantes», no los usuarios. Señaló, además, algo que es sabido por todos: que las ganancias de estas empresas crecen y les permiten ampliarse porque «venden productos y servicios directamente a los usuarios» gracias a «explotar los datos que controlan» con fines comerciales, algo que, en su opinión, «socava la eficiencia de la economía de mercado».

El multimillonario señaló todos y cada uno de los problemas que están encima de la mesa en estos momentos. Acusó a las GAFA de ser una amenaza para la sociedad, debido a su conducta monopolista y su estímulo a la «adicción», que comparó con el de las compañías dedicadas a los juegos de azar. Soros advirtió que «las redes sociales influyen en el modo en que la gente piensa y actúa, sin que se dé cuenta», un hecho con «consecuencias negativas de largo alcance sobre el funcionamiento de la democracia y la integridad de las elecciones».

Facebook y Google son, junto con Apple y Amazon, dos de las compañías más poderosas en la economía global. Parte de su atractivo para los accionistas es que su gigantesco negocio publicitario opera, como hemos visto, casi sin intervención humana. Los algoritmos pueden ser hermosos en términos matemáticos, pero sólo arrastran todos los defectos de las personas que los crean. Como hemos venido señalando, los algoritmos tienen fallos cada vez más obvios y peligrosos.

Gracias al enfoque de *laissez-faire* del Gobierno de Estados Unidos en cuanto a regular su actividad o a aplicar las normas de competencia a las GAFA, las plataformas de internet han venido aplicando estrategias comerciales que no se habrían permitido en décadas anteriores. Nadie les ha impedido usar productos gratuitos para centralizar internet y luego reemplazar sus funciones principales. Nadie les ha impedido desviar las ganancias de los creadores de contenido ni tampoco recopilar datos sobre cada aspecto de la vida en internet de cada uno de sus usuarios de manera individual. Tampoco nadie ha hecho nada por evitar que acumulasen una cuota de mercado que no se veía desde los días de Standard Oil, ni les impidió realizar experimentos sociales y psicológicos masivos con la vida de sus usuarios. Nadie, en definitiva, ha exigido que se controlasen estas plataformas; la mera sugerencia de que así fuera te colocaba del lado de los luditas.

La realidad es que Facebook y Google son tan grandes que es probable que las herramientas tradicionales de regulación sean inefectivas para reparar el daño realizado. La Unión Europea lo ha intentado: la Comisión Europea sancionó a Alphabet, la matriz del buscador Google, con una multa de 2.424 millones de euros por abuso de posición dominante en el servicio de comparación de productos y ha iniciado un nuevo procedimiento a cuenta del sistema operativo Android, que ha resultado en una sanción cercana a los 5.000 millones de euros.

En el primer caso, se acusó a Google de aprovechar su abrumadora posición de dominio en las búsquedas en internet para promocionar artificialmente las apariciones de Google Shopping y eliminar competidores. El daño fue claro: la mayoría de los competidores europeos de Google en su misma categoría sufrieron pérdidas paralizantes. El superviviente más exitoso perdió el 80 por ciento de su participación en el mercado en un año.

La cuantía de la multa, que hubiera sido un varapalo insuperable para cualquier otra compañía, en el caso de Google no tuvo efecto ni en sus inversores ni en su conducta. La multa antimonopolio más grande en la historia de la UE rebotó en Google como una pelota de ping-pong en el casco de un acorazado. Lamentablemente, las decisiones y acciones comerciales de las GAFA se toman en Estados Unidos sin tener en cuenta legislaciones o reglamentos europeos, asiáticos o africanos.

En febrero de 2018, el regulador antimonopolio de la India impuso una multa de 1,36 mil millones de rupias (21,17 millones de dólares) a Google por «sesgo de buscador» y abuso de su posición dominante.[126] La Comisión de Competencia de la India (CCI) estableció que Google abusaba de su dominio en la búsqueda y en los mercados de publicidad online. Es de suponer que el impacto de una multa tan exigua no afectará en absoluto a la posición de la compañía ni a sus prácticas monopolísticas. El fallo indio puso fin a una investigación iniciada por el organismo de control en 2012 sobre las denuncias presentadas por el sitio web Matchmaking Bharat Matrimony y una organización sin fines de lucro, Consumer Unity and Trust Society (CUTS). El organismo de control indio también expresó su desilusión con Google porque no fue posible recopilar todos los datos de ingresos en el tiempo asignado.

La segunda sanción europea a Google lo fue por presuntas prácticas anticompetitivas en el mercado de sistemas operativos para móviles. Al parecer, la Comisión contaría con pruebas de que la compañía habría abusado de su dominio del mercado de smartphones (el 90 por ciento de los dispositivos funcionan con el

sistema operativo Android de Google) imponiendo condiciones leoninas a los fabricantes de teléfonos, en detrimento de la oferta para los consumidores europeos.[127]

En este contexto de creciente control del poder monopolístico de Google se produjo el despido de Barry Lynn del grupo de expertos de tendencia izquierdista New America después de lo que se supuso una presión política ejercida por Google, principal contribuyente de ese think tank. Al parecer, y según su versión, este despido se habría debido a las críticas persistentes de Lynn al gigante de las búsquedas, acusándole de ser un monopolio que abusa de su posición de dominio y subvierte el mercado.

El despido dejó en evidencia a New America, creando la apariencia de que sus informes políticos y artículos de opinión no eran producto de su esforzado pensamiento sino que, más bien, venían modelados por los intereses de sus donantes. Los comentarios aprobatorios de Lynn a la primera multa de la UE a Google estarían, al parecer, en el origen de su despido.

Lo cierto es que el escándalo habría podido ser una bendición para Lynn, quien se ha convertido en un cruzado, en un luchador de los principios al más puro estilo estadounidense, que ha creado su propio think tank el Open Markets Institute con el objeto de luchar contra los monopolios. En este sentido, Lynn argumenta que dado que los gigantes como Amazon y Google hacen que la economía se estanque y sea más frágil, es necesario controlarlos de forma más agresiva usando la ley antimonopolio existente.

Tanto Lynn como McNamee, que se ha incorporado al equipo de Open Markets Institute, iniciaron a finales de 2017 contactos con los congresistas y senadores estadounidenses con el objetivo de influir en los candidatos y los votantes antes de las elecciones legislativas de 2018.

El propio Lynn en *The Guardian* resumía sus planes sobre cómo parar a Google y Facebook.[128] Lynn comienza comentando su encuentro con el senador Kennedy «el martes, el senador John Kennedy de Louisiana dijo a los consejos generales de Facebook y

Google: “A veces su poder me asusta”. El problema, dijo Kennedy, es que las corporaciones saben mucho sobre nosotros y nosotros muy poco sobre ellos. Kennedy ilustró sus temores con dos preguntas retóricas. “Si su CEO viene y les pregunta quiero saber todo sobre el senador Graham... Podría hacerlo, ¿no es así?”».

Lynn considera que esta aparente paradoja apunta a un problema fundamental que Facebook y Google no pueden resolver solos: estas instituciones están diseñadas para reunir grandes cantidades de información, pero no están diseñadas para gestionar esa información en interés del público en general, ni para evitar ataques de propaganda que subviertan las elecciones de sociedades democráticas. «Si está claro que Facebook y Google no pueden administrar lo que ya controlan —continúa Lynn—, ¿por qué dejar que esas corporaciones posean más? Las autoridades antimonopolio de Estados Unidos pueden imponer esa norma casi de inmediato.»

El Open Markets Institute no duda de que estas corporaciones deberían estar sujetas a la legislación antimonopolio. Facebook, por ejemplo, contaría con un 77 por ciento del tráfico de redes sociales móviles en Estados Unidos; más de la mitad de todos los adultos estadounidenses usa Facebook todos los días. Facebook y Google gestionarían más del 75 por ciento de la publicidad online, y casi la totalidad de las nuevas inversiones en ese sector. Ambas concentran más de la mitad de todo el tráfico de los sitios web de noticias. En total, Facebook tiene más de 2 mil millones de usuarios en todo el mundo.

En opinión de Lynn, tanto el Departamento de Justicia como la Comisión Federal de Comercio estadounidense cuentan con competencias para limitar la actividad de abuso de posición de dominio que tendrían las GAFA. «De hecho, la FTC —continúa Lynn — habría creado parcialmente el problema de las “noticias falsas” al no usar su autoridad para bloquear las adquisiciones previas de estas plataformas, como las compras por parte de Facebook de WhatsApp e Instagram. Si a esas compañías se les hubiera

permitido crecer y competir con Facebook, hoy veríamos menos poder y control concentrados en esa única corporación [...]. Uno de los mitos de las plataformas tecnológicas es que son actores innovadores que inventan nuevas formas de hacer las cosas. Realmente no son conglomerados que crecen a través de las adquisiciones. Larry Page y Sergey Brin de Google, gracias a una subvención de la National Science Foundation, crearon una forma de clasificar las páginas web que les permitió construir un excelente motor de búsqueda. Más adelante, usaron el dinero obtenido gracias a este éxito temprano en comprar la mayoría del resto de productos clave de Google, incluidos YouTube, Android, Deep Mind, Waze y Doubleclick. Google llegó a comprar una compañía por semana.»

Como vemos, lo que eran comentarios marginales, se han convertido en un clamor. Como ya ocurrió con otras entidades como Microsoft, las GAFA han excedido con mucho los límites de complacencia que la administración estadounidense les dio para crecer e innovar. La cuestión no se limita sólo a las cuestiones de competencia clásica —si ahogan a la competencia y si se han convertido en una traba para la innovación— sino que afecta a derechos fundamentales (a la privacidad, la intimidad y el acceso a información veraz), poniendo en peligro la propia supervivencia del sistema democrático. Sin duda nunca antes habíamos visto una situación tan grave creada por un número tan reducido de entidades privadas.

Divide y vencerás

Son muchas las propuestas que hay sobre la mesa, pero traemos aquí las de McNamee (de su artículo citado y siguiendo a Lynn), quien propone las siguientes reformas legislativas, entre las que se encuentra la limitación de adquisiciones para incentivar la

innovación, algunas relativas a la limitaciones necesarias para evitar las fake news y a protección de la vida privada de los usuarios. Las medidas serían las siguientes:

1. Prohibir los bots. Para McNamee es esencial prohibir los bots que se hacen pasar por humanos. Distorsionan la política de una manera inimaginable y carece de precedente histórico: por muchos panfletos y libelos que se hayan impreso en el pasado, el impacto es insignificante al lado del efecto de estas falsas identidades. La norma debería, al menos, requerir el etiquetado explícito de todos los bots, la capacidad de los usuarios para bloquearlos y la responsabilidad de los proveedores de la plataforma por los daños que los mismos causen.
2. En segundo lugar, no se debería permitir a las plataformas que hicieran nuevas adquisiciones hasta que hayan abordado el daño causado hasta la fecha, hayan tomado medidas para evitar daños en el futuro, y hayan demostrado que tales adquisiciones no tendrían como resultado una competencia disminuida.

   Un aspecto poco apreciado del crecimiento de las plataformas es su patrón de engullir empresas más pequeñas: en el caso de Facebook, Instagram y WhatsApp; en el de Google, YouTube, Google Maps, AdSense y muchos otros, usando estas adquisiciones para aumentar su monopolio y frenar la innovación.

   Esta cuestión es relevante porque internet ha perdido algo muy valioso, la descentralización. Internet fue diseñado para tratar todos los contenidos y a todos los propietarios de contenido por igual. Esa igualdad tiene valor para la sociedad, ya que mantiene el campo de juego nivelado y anima la entrada de nuevos participantes. La descentralización tuvo un coste: nadie tenía un incentivo para hacer que las herramientas de internet fueran fáciles

de usar. Frustrados por esas herramientas, los usuarios adoptaron alternativas fáciles de usar como las que ofrecían Facebook y Google, lo que permitió a ambas empresas centralizar internet, colocándose entre los usuarios y el contenido, e imponiendo de manera efectiva un impuesto en ambos lados.

Este modelo de negocio es central para Facebook y Google, y extremadamente conveniente para los usuarios a corto plazo. Sin embargo, el coste social a medio plazo es inasumible.

3. En tercer lugar, las plataformas deberían ser transparentes sobre quién está detrás de la comunicación política basada en cuestiones problemáticas. La Honest Ads Act[129] es un buen comienzo, pero no va lo suficientemente lejos por dos razones: la publicidad fue una parte relativamente pequeña de la manipulación rusa; y los contenidos falsos generados a propósito jugaron un papel mucho más importante que los anuncios orientados de los candidatos. La transparencia con respecto a quienes patrocinan publicidad política de todo tipo es un paso hacia la reconstrucción de la confianza en las instituciones políticas.
4. En cuarto lugar, las plataformas deben ser más transparentes con respecto a sus algoritmos. Los usuarios merecen saber por qué ven lo que ven en sus feeds de noticias y resultados de buscador. Facebook y Google adoptan millones de decisiones editoriales cada hora y deben aceptar la responsabilidad de las consecuencias de esas elecciones. Los consumidores también deberían poder ver cuáles de sus atributos son tenidos en cuenta por los anunciantes.

   Si Facebook y Google tuvieran que ser directos y decir a los usuarios que la razón por la que ven bulos y fake news es simplemente porque es bueno para su negocio, sería

mucho menos probable que siguieran haciéndolo. Permitir que terceros auditen los algoritmos mantendría la transparencia.

5. En quinto lugar, se debería exigir a las plataformas que tengan una relación contractual más equitativa con los usuarios.

   Facebook, Google y otros han incluido en sus *end-user license agreements* (EULA) —los contratos que especifican la relación entre la plataforma y el usuario— condiciones y derechos a su favor sin precedentes.

   Cuando uno instala o actualiza un software, la aceptación del EULA es un requisito para completar la instalación. Si no se desea aceptar las nuevas condiciones que vienen con una actualización, el usuario puede seguir utilizando la versión anterior durante un tiempo, a menudo, años. No ocurre lo mismo con las plataformas de internet como Facebook o Google: el uso del producto viene con la aceptación implícita del último EULA, que puede cambiar en cualquier momento. Si hay términos que uno elige no aceptar, la única alternativa es abandonar el uso del producto. Para Facebook, donde los usuarios han contribuido con el ciento por ciento del contenido, esta opción es particularmente problemática.

   Se debe exigir a todas las plataformas de software que ofrezcan una opción de exclusión voluntaria que permita a los usuarios seguir con la versión anterior si no les gusta o no están conformes con las nuevas condiciones de contratación.

   Estas «forking platforms» entre versiones antiguas y nuevas tendrían varios beneficios: mayor elección del consumidor, mayor transparencia en el EULA y más cuidado en el despliegue de nuevas funcionalidades, entre otras. Limitaría el riesgo de que las plataformas realizasen experimentos sociales masivos sobre millones o miles de

millones de usuarios sin notificación previa. El mantenimiento de más de una versión de sus servicios sería costoso para Facebook, Google y el resto, pero nada que su cuenta de resultados no pudiera soportar.

6. En sexto lugar, se necesita un límite en la explotación comercial de los datos de los consumidores por parte de las plataformas de internet, algo que sucede en Europa desde hace años, pero que las empresas con sede en Estados Unidos no han respetado en absoluto. Los clientes entienden que el uso «gratuito» de plataformas como Facebook y Google concede a éstas una cierta licencia para explotar sus datos personales. El problema es que las plataformas están usando esa información de manera que ni los consumidores entienden ni aceptarían si lo hicieran. Por ejemplo, Google adquirió la base de datos Mastercard, que hemos comentado, a principios de 2018; Facebook usa software de reconocimiento de imágenes y etiquetas de terceros para identificar a los usuarios en contextos sin su participación, y en donde podrían preferir ser anónimos. Lo que demuestran ambos ejemplos es que las plataformas no sólo usan los datos que les facilitan voluntaria e involuntariamente sus usuarios, sino también los que compran a terceros. Y usarán esa información de manera ilimitada en el tiempo. En este sentido, McNamee propone una ley de prescripción sobre el uso de los datos de los consumidores por parte de una plataforma y sus usuarios, dotando a los mismos del derecho de renegociar los términos de cómo se utilizan sus datos, tal como ocurre en la UE.
7. Séptimo, los consumidores, no las plataformas, deben poseer sus propios datos. En el caso de Facebook, esto incluye publicaciones, amigos y eventos; en resumen, todo. Los usuarios que crearon esta información deberían tener derecho a exportarla a otras redes sociales.

Este derecho de portabilidad ha sido recogido en el RGPD pero no hay norma en Estados Unidos que ampare dicha portabilidad. Dada la inercia, conveniencia y ausencia de competencia no sería de esperar una migración masiva de usuarios pero, en cambio, se produciría una explosión de innovación y emprendimiento.

Facebook es tan poderoso que la mayoría de los nuevos participantes evitarían la competencia directa a favor de crear una diferenciación sostenible. Las empresas nuevas y los jugadores establecidos crearían nuevos productos que incorporarían los datos portados de los usuarios, forzando a Facebook a competir nuevamente.

8. En octavo lugar, McNamee considera que ha llegado el momento de revivir el enfoque tradicional sobre el tratamiento de los monopolios.

Desde la era Reagan, la ley antimonopolio ha operado bajo el principio de que el monopolio no es un problema, siempre que no genere precios más altos para los consumidores. Bajo ese marco, Facebook y Google han podido dominar varias industrias, no sólo las redes sociales y de búsquedas, sino también el correo electrónico, vídeos, fotos y ventas de anuncios digitales, entre otros, aumentando sus monopolios comprando rivales potenciales como YouTube e Instagram. Si bien es atractivo, este enfoque ignora los costes que no aparecen en una etiqueta con un precio. La adicción a Facebook, YouTube y otras plataformas tiene un coste; la manipulación electoral tiene un coste; el uso indiscriminado de los datos personales tiene un coste; y la reducción de la innovación y la contracción de la economía empresarial tienen un coste. Todos estos costes son evidentes ahora, y podemos cuantificarlos lo suficientemente bien como para apreciar que el coste de la concentración en internet podría ser inaceptable para los consumidores.

9

Familia

Datos genéticos

Hay personas que viven en la melancolía. Mi tía es una de ellas. No sólo mantiene contacto con todos los familiares, por muy improbable y lejana que sea la conexión, sino que es capaz de acordarse de cualquier ser humano que haya protagonizado una relación irrelevante con cualquiera de nosotros y, además, tener su teléfono.

La serie *Raíces* fue un bombazo en la televisión española. Su emisión en la única cadena de la época (a la otra disponible se la llamaba despectivamente UHF) mantuvo a la España de finales de los setenta enganchada a las peripecias de Kunta Kinte, un guerrero mandinga capturado por traficantes de esclavos y trasladado, obviamente en contra de su voluntad, en una nave negrera a las colonias inglesas en Norteamérica (futuro Estados Unidos). La serie trataba de su peripecia, de su deseo de regreso y de su necesidad de conocer su identidad.

Buscando mis raíces

Mi tía, inspirada por los pasos de Kunta Kinte junto con otros muchos españoles de la época, se entregó a la búsqueda de nuestros ancestros, lo que la llevó hasta Portugal. No fue muy lejos pero estuvo entretenida unos cuantos meses mientras intentaba

encontrarse a sí misma. Si esa necesidad de saber la hubiese atacado en estos días, habría acudido sin duda a uno de los múltiples servicios online que te permiten hacer esa búsqueda cómodamente desde el sillón de casa. Habría empezado en Ancestry. com, en el que no habría encontrado mucho consuelo: es un servicio para los estadounidenses y lo más cerca que te dejan de tus antepasados es en la versión mexicana de la web.

Hagamos un ejercicio de imaginación y consideremos la posibilidad de tener familia estadounidense o mexicana, algo que no es del todo tan improbable teniendo en cuenta la larga tradición de emigración española a aquel continente.

Ancestry empezó prestando sus servicios haciendo uso de árboles genealógicos en los que las personas buscaban sus orígenes, enriqueciendo el árbol con los datos que cada usuario proporcionaba de sí mismo y en conexión con otros miembros de la comunidad. Con la llegada de los kits caseros de recogida de muestras genéticas, esta y otras páginas han incorporado la coincidencia genética a través del análisis del ADN como método para hallar coincidencias.

Esto supone un salto sustancial. Al igual que ocurre con los datos biométricos, los datos genéticos no se pueden cambiar, lo que es una verdadera lata si la base de datos del prestador es pirateada. Una vez que el usuario remite el bastoncillo con la saliva a Ancestry, y ésta secuencia el ADN, el resultado junto al bastoncillo quedan depositados en una base de datos accesible tanto por Ancestry como por aquellas personas con las que puedas tener alguna coincidencia de ADN.

Ancestry recopila, según su propia declaración de prácticas de privacidad, la clásica información que creemos intrascendente como nombre, dirección de correo electrónico e información de facturación, contraseña o número de teléfono. Lo interesante es el resto de información que recopila y su valor esencial como identificador unívoco que no se puede cambiar y que queda a disposición de un tercero para siempre: nuestro código genético.

Ancestry guarda, entre otros, los datos de árboles genealógicos pero también los «árboles de ADN», que son aquellos que mantienen conectada la información genética de distintos usuarios relacionados de alguna manera entre sí. Como es obvio, el ADN permite analizar la salud futura de aquellos que se hacen las pruebas en una búsqueda romántica de su pasado glorioso. Como vamos viendo, pasado y futuro se funden a través de los datos como un continuo espacio tiempo.

Ancestry no entra en más detalles sobre qué otra información guarda de sus usuarios,[130] pero deja claro para qué la va a usar: aparte de para las finalidades obvias de la prestación del servicio y publicitarias, Ancestry empleará los datos, incluido el ADN, en proyectos de investigación genealógica o genómica, o verificación de identidad. También lo usará para hacer encuestas y para completar los datos del perfil de cada usuario, suponemos que con la finalidad de mejorar la información de las investigaciones médicas y farmacéuticas que realiza con sus «socios comerciales»: cuánto fumamos, si comemos tofú o qué opinamos de tal o cual participante de *Gran Hermano*, lo que dará una idea de cuánto tiempo pasamos sentados, sin hacer nada, delante de la televisión. Nuestro GATACA se usará también con fines de estudios científicos, estadísticos e históricos, así como para el desarrollo e investigación de productos.

Que Ancestry no se queda la información o la usa únicamente para encontrar a ese abuelo nuestro que fue príncipe italiano, es algo que queda claro en su política de privacidad: los datos se ceden a terceros que se categorizan en diversos tipos, como observador, colaborador y administrador, así como a la Policía y a los tribunales bajo orden judicial.[131] En el año 2017, Ancestry recibió peticiones de las autoridades de Estados Unidos, Reino Unido, Canadá y Alemania, todas relacionadas, según su informe de transparencia, con investigaciones respecto al uso indebido de tarjetas de crédito y al robo de identidad. En el mismo informe señalan que no recibieron solicitud alguna relacionada con la salud

ni la genética de ningún miembro de Ancestry, y no revelaron, según su información, ningún dato de ese tipo a ninguna agencia encargada del cumplimiento de la ley.

Como parte del uso recreativo de la genética, y siguiendo la estela de Ancestry, varias páginas nos leen el futuro a cambio de un poco de nuestra saliva, nuestro GATACA personal.[132] BabyGlimpse, a partir de un estudio completo del genoma, nos cuenta cómo va a ser nuestro hijo de mayor, si le gustarán más las comidas saladas o los dulces, si será intolerante a la lactosa o si el uso de antibióticos aumentará su riesgo de contraer una enfermedad cardiaca. 23andMe, competidor de Ancestry, anuncia kits para analizar las bacterias de nuestro intestino o conocer el ADN de nuestras mascotas, en un ejercicio de diversificación del mercado de los marcadores genéticos. Orig3n, por su parte, promete revelar a sus clientes «cuál es su superpoder» mientras que la española 24Genetics ofrece conocer si «te afecta más o menos la cafeína que a los demás» o «por qué eres más goloso».

Los resultados de ADN de cinco millones de clientes de 23andMe (la mayor empresa del sector) se cederán a GlaxoSmithKline para el diseño de nuevos fármacos. 23andMe solicitará el consentimiento a sus usuarios para participar en la investigación científica, pero no parece que les vayan a pagar por ello. Glaxo ha invertido 300 millones de dólares en 23andMe con quien ha suscrito un acuerdo de cuatro años que le otorga los derechos exclusivos para el desarrollo de fármacos usando el repositorio de muestras de ADN y los datos genéticos de todos sus clientes. El primer proyecto será sobre medicamentos para tratar la enfermedad de Parkinson a partir de la mutación del gen LRRK2 que tienen algunos pacientes de esta enfermedad.[133] La cuestión que queda en el aire es cómo de esforzada será la petición de consentimiento a los clientes para participar en el proyecto de investigación y cuántas cesiones y reutilizaciones posteriores de sus datos se producirán. Es entendible que Glaxo y otras compañías se interesen por un repositorio de datos de tal magnitud, pero preocupa

que unos datos facilitados, tal vez sin mucha reflexión, para otras finalidades se cedan de manera no transparente, sin control de las cesiones en cascada, reusos, y reempaquetamiento para otras entidades y otras finalidades. Aunque se anonimicen, está demostrado que con un número limitado de datos es posible desanonimizar un dato y saber a quién corresponde.

El ADN es el dato de salud con mayúsculas, que determina el riesgo de desarrollar determinadas enfermedades, las cuales, conocidas de antemano, pueden limitar el acceso de sus portadores a seguros privados de salud o al mundo laboral.

Pero eso no es todo.

CSI Sacramento

Como sociedad amedrentada que somos, cuando las libertades se topan con los razonamientos, generalmente interesados, a favor de la defensa de la seguridad nacional y el orden público, las cuestiones relativas a la intimidad ceden. No transmitimos tampoco mucha seriedad en nuestros postulados en defensa de la privacidad cuando vamos regalando datos esenciales a cambio de servicios, en muchos casos, prescindibles o sustituibles por servicios de pago más respetuosos con la privacidad de sus usuarios. Así las cosas, ni a los cuerpos policiales ni a la sociedad que, alentada por los medios de comunicación, demandan la justicia tumultuaria del Juez de la Horca, les incomoda lo más mínimo que se usen las bases de datos de empresas de genoma recreativo para la persecución del delito.

Joseph James DeAngelo, de setenta y dos años, vecino de Citrus Heights, Sacramento (California), fue detenido por culpa —o gracias— a la curiosidad de un familiar cuarenta años después de haber cometido sus crímenes. De hecho, sin esta colaboración inesperada habría continuado impune de doce asesinatos y al menos 50 violaciones que habría cometido en California entre 1976

y 1986. DeAngelo habría dejado una muestra de ADN que fue posteriormente secuenciado y cotejado con los perfiles genéticos disponibles online en varias empresas de servicios de genealogía genética. Cuando encontraban una posible coincidencia, los investigadores exploraban los árboles genealógicos de la persona en busca de posibles sospechosos. Entre ellos estaba el de DeAngelo, al que investigaron para asegurarse de qué otras circunstancias coincidían (fechas, lugares...). Cuando tuvieron un aceptable nivel de certeza, la Policía recogió una muestra de su ADN de su basura y confirmaron sus sospechas.

La base de datos usada por la fiscalía y la Policía en esta investigación, GEDmatch, ha dejado en evidencia a todas las empresas que recaban datos de ADN con uno u otro propósito. GEDmatch es una base de datos genómicos personales, abierta a la que cualquiera puede subir sus árboles genealógicos. Su página es conocida por ser «la base de datos de ADN y genealogía *de facto* de las fuerzas del orden», según Sarah Zhang, de *The Atlantic*.[134] La página es una especie de Facebook genético en el que los usuarios publicitan sus datos genéticos, de los que son dueños, así como todas sus relaciones, lo que permite trazar sin ninguna limitación las relaciones genéticas y familiares. Esto es lo que pasa cuando generas un dato y lo compartes. Resulta imposible aventurar en manos de quién va a caer.

GEDmatch reconoció que la detención del asesino se produjo gracias a los datos disponibles en su plataforma que, al ser abierta, no requería orden judicial alguna. GEDmatch advirtió a quienes envían su ADN que deben ser conscientes de esta posibilidad: «Es importante que los usuarios de GEDmatch entiendan los posibles usos de su ADN, incluida la identificación de familiares que hayan cometido delitos o hayan sido víctimas de ellos. Si te preocupan los usos no genealógicos de tu ADN, no deberíais subirlo o deberíais borrar lo que haya sido subido».

DNA Doe Project[135] —una organización de voluntarios en la que trabajan algunos de los mejores genealogistas genéticos de la industria alrededor de la iniciativa DoeFundMe y que utiliza la genealogía genética para identificar a personas desconocidas— descubrió la identidad de una mujer asesinada en 1981 secuenciando su genoma a partir de una muestra bastante deteriorada. Buscaron coincidencias con los perfiles archivados en Ancestry.com, donde localizaron a una prima de la víctima. La desconocida resultó ser una mujer llamada Marcia King.

Más allá de los peligros de la gestión de los propios datos biométricos, lo que el análisis de ADN plantea es la afectación de los datos de terceros, de su intimidad o, incluso, de su libertad. Pensemos en los datos de ADN de los menores facilitados por los padres sin su consentimiento, conocimiento, y excediendo, con mucho, las capacidades de tutela y cuidado que les corresponden. Una cosa es dar el consentimiento para pruebas médicas, incluido un análisis genómico, para curar la enfermedad de un menor, pues en este caso se actúa en su beneficio, en un entorno médico controlado y sujeto al secreto profesional; y otra muy distinta es hacerle un análisis de saliva a través de un servicio online con la finalidad lúdica o pretendidamente preocupada de saber si va a ser intolerante a las semillas de amapolas. Los datos de ese menor, que crecerá y tendrá una vida independiente de sus padres, quedan, para siempre, almacenados con finalidades cambiantes y bajo el control de entidades que viven de dar rendimiento a sus accionistas o sujetas a legislaciones de una panoplia de países con distintas visiones de lo que es delito o no, visiones que cambian con el tiempo y sus gobernantes.

Por lo demás, nadie puede evitar tener un familiar que decida que quiere secuenciar su genoma para saber si le engorda el tomate o si debería tomar más aceite de oliva para mejorar su rendimiento como deportista de élite ocasional.

Lo cierto es que, en ambos casos, hay terceros que pueden distribuir mi ADN o parte de él sin mi consentimiento, y con efectos en mi vida y en las decisiones futuras que se puedan tomar sobre mí. La tentación de hacer selecciones genéticas por raza, sexo o enfermedades no es nueva. Lo que sí lo es, es poderlo hacer de una manera tan exhaustiva, barata y masiva.

Y eso sin mencionar la persistencia de ese dato. Por mucho que se quiera, el ADN no muta ni cambia a lo largo de la vida.

Sólo puedo agradecerle a mi tía que en los tiempos de Kunta Kinte no hubiera internet.

10

Sensorium

Internet de las cosas

Hay un cilindro en el centro de la habitación, impasible, suave, simple y pequeño. Con una altura de 14,8 cm de altura y una única luz circular azul verdosa alrededor de su borde superior, el cilindro espera en silencio. Una mujer entra a la habitación con un niño dormido en brazos y se dirige al cilindro. «Alexa, enciende las luces del pasillo.» El cilindro vuelve a la vida, y la habitación se ilumina. La mujer hace un leve gesto con la cabeza y se lleva al niño al piso de arriba.

Esta escena no está sacada de *Ex Machina* sino del vídeo de 2017 de presentación del dispositivo Echo y del asistente Alexa de Amazon, en el que se publicita la última versión de Amazon Echo. El vídeo comienza diciendo: «Saluda al nuevo Echo» y explica que Echo se conectará con Alexa (el agente de inteligencia artificial) para «reproducir música, llamar a amigos y familiares, controlar dispositivos domésticos inteligentes y más». El dispositivo contiene siete micrófonos direccionales, por lo que se puede escuchar al usuario en todo momento, incluso cuando se reproduce música. El dispositivo viene en varios estilos, como gris metalizado o beis básico, diseñado para «mezclarse o destacarse». Pero incluso las opciones de diseño vienen con un gran vacío: nada alertará al propietario sobre la vasta red que subyace y conduce sus

capacidades interactivas. El vídeo promocional simplemente indica que el rango de cosas que puede pedirle a Alexa que haga siempre aumentará. «Como Alexa está en la nube, siempre se está volviendo más inteligente y agregando nuevas funciones.»[136]

Un comando breve y una acción o respuesta es la forma más común de interacción con este dispositivo con el que los consumidores, por un precio aceptable, se comunican con una inteligencia artificial usando el lenguaje natural.

Pero para que esta sencilla interacción y reacción fugaz e inane se produzca es necesario que hayan ocurrido muchas cosas: cadenas entrelazadas de extracción de recursos, trabajo humano y procesamiento algorítmico a través de redes de minería, logística, distribución, predicción y optimización.

Alexa

¿Cómo es posible? Alexa es una voz incorpórea que representa la interfaz de interacción humano-IA para un conjunto extraordinariamente complejo de capas de procesamiento de información. Estas capas están alimentadas por mareas constantes: los flujos de voces humanas se traducen en preguntas de texto, que se utilizan para consultar bases de datos de posibles respuestas, y el reflujo correspondiente de las respuestas de Alexa. Por cada respuesta que da Alexa, su efectividad se deduce de lo que sucede a continuación: ¿es la misma pregunta pronunciada de nuevo? (¿el usuario se sintió escuchado?), ¿la pregunta fue reformulada? (¿el usuario sintió que la pregunta fue entendida?), ¿hubo una acción después de la pregunta? (¿La interacción dio como resultado una respuesta rastreada: una luz encendida, un producto comprado, una pista reproducida?)

Con cada interacción, Alexa se está entrenando para escuchar mejor, para interpretar con mayor precisión, para desencadenar acciones que se correspondan con los comandos del usuario con

mayor precisión y para construir un modelo más completo de sus preferencias, hábitos y deseos.

¿Qué se necesita para llegar a este estado de perfección? Datos. Cada pequeño momento, ya sea para responder una pregunta, encender una luz o poner una canción, requiere una vasta red planetaria, impulsada por trabajo y uso intensivo de datos. La escala de recursos requeridos es mucho mayor que la energía o el trabajo que se necesitarían para hacer funcionar un electrodoméstico o presionar un interruptor con un simple gesto de la mano.

Cuando un ser humano interactúa con un Echo u otro dispositivo de inteligencia artificial activado por voz, actúa mucho más que con sólo un producto. Es difícil ubicar al usuario humano de un sistema de IA en una sola categoría: más bien, merecen ser considerados como un caso híbrido. Así como la quimera griega —animal mitológico parte león, parte cabra y parte serpiente—, el usuario de Echo es simultáneamente un consumidor, un recurso, un trabajador y un producto. Esta identidad múltiple de los usuarios es recurrente en buena parte de los sistemas basados en la extracción y análisis de datos que estamos comentando. En el caso específico de Amazon Echo, el usuario ha comprado un dispositivo que pretende hacer su vida más cómoda, pero, a la vez, es propietario del dispositivo, y también es un recurso, ya que sus comandos de voz se recopilan, analizan y conservan con el fin de construir un cuerpo cada vez más grande de voces e instrucciones humanas. Y proporciona mano de obra, ya que al usar el servicio contribuyen a entrenar a la IA que es el corazón del sistema, retroalimentándolo y mejorando su precisión, la utilidad y la calidad general de las respuestas de Alexa. En esencia, están ayudando a entrenar las redes neuronales de Amazon. Todo lo que esté más allá de las limitadas interfaces físicas y digitales del dispositivo está fuera del control del usuario.

Echo presenta una superficie lisa sin capacidad para abrirla, repararla o cambiar la forma en que funciona. El objeto en sí es una extrusión de plástico muy simple con una colección de sensores. Su poder y complejidad reales se encuentran en otro lugar, lejos de la vista. El Echo no es más que un «oído» en el hogar: un agente de escucha desencarnado que nunca muestra sus conexiones profundas con sistemas remotos.

La Roomba cotilla

A finales del julio de 2017, Colin Angle, fundador y consejero delegado de la empresa iRobot fabricante de la Roomba, en una entrevista con Reuters[137] reconoció sus planes para comercializar los datos recogidos por los dispositivos IoT de limpieza que fabrica: «Hay todo un ecosistema de aparatos y servicios que se pueden ofrecer a los hogares una vez que tienes un mapa de la casa bien hecho y cuyo dueño nos ha dado permiso para compartir». Estos dispositivos IoT de limpieza fueron los primeros en incluir la tecnología de navegación para almacenar un mapa y reconocer los espacios. Para ello cuenta con una cámara con una inclinación de 45 grados que registra el espacio en busca de cambios y obstáculos. Después, almacena los datos y los analiza para poder limpiar mejor, y ahora los comparte y los vende. El fundador de iRobot manifestó que los datos no se cederían sin permiso del consumidor, pero no quedó claro cómo va a recabar el consentimiento. iRobot es un perfecto ejemplo de cómo un dispositivo IoT puede afectar a la privacidad de sus usuarios, de los problemas jurisdiccionales de la protección de datos y de la recogida del consentimiento para tratar y ceder los mismos. Sin mencionar la posibilidad de poner al alcance de los ladrones mapas exactos de las casas, junto con otra información útil si uno pretende robar con el menor de los inconvenientes: si tiene mascotas o no, su tamaño o a qué hora están los propietarios en sus domicilios.

La idea maléfica de incorporar a las cosas cotidianas actuadores y sensores con la finalidad de que reconozcan su entorno, tomen decisiones, recaben datos y los compartan a través de una red de comunicaciones generalmente pública, ha abierto frente a nosotros un panorama extraño. Difícil resistirse a tantas ventajas cuando un número muy importante de usuarios y proveedores de servicios generan información en grandes cantidades, comportándose como sensores en movimiento susceptibles de recogerlo todo. Piénsese, por ejemplo, en los coches y motos compartidos que se mueven por nuestras ciudades recogiendo datos medioambientales, de comportamiento de los usuarios y de otros conductores, imágenes del entorno, siendo aptos para vigilar lo que ocurre en la calle en todo momento y en tiempo real. Y, además, trasladan gente de un punto a otro.

La propia naturaleza de las cosas conectadas (conocido de manera genérica como el internet de las cosas o Internet of Things —IoT—)[138] crea sofisticadas interdependencias entre los productores de los productos, los desarrolladores del software para las cosas y sus componentes, y los prestadores de servicios que se generan alrededor de su uso. Debido, precisamente, a la alta complejidad del ecosistema IoT (objetos físicos, software, infraestructura de internet, datos personales y no personales, comportamiento del usuario final, analítica de datos, etc.), y a la variedad de actores implicados (fabricantes de productos, fabricantes de sensores, productores de software, proveedores de infraestructura, otros actores involucrados en el suministro de diferentes servicios, usuarios finales, etc.), hace que cualquier actor que participe en este ecosistema tenga acceso a la multitud de datos que un dispositivo IoT genera.

El fenómeno es tan amplio como el número de cosas que son susceptibles de conexión. Hablamos de objetos cotidianos conectados a internet con capacidad de reconocer su entorno, de tomar decisiones más o menos inteligentes y de actuar con base a la información que reciben. Se prevé que la incorporación del IoT a

la vida cotidiana asistida sea de gran importancia, sobre todo si consideramos que nos dirigimos a una sociedad envejecida, con cada vez más ancianos y dependientes. Un ejemplo de ello es la ciudad de Songdo, en Corea del Sur, primer ejemplo de ciudad inteligente completamente equipada. Prácticamente todo en esa ciudad está cableado, conectado y convertido en una fuente continua de datos que son monitorizados y analizados por multitud de ordenadores, con escasa o nula intervención humana.

El concepto comprende, como vemos, desde juguetes, lavadoras, neveras, televisores, hasta marcapasos, bombas de insulina, ascensores, coches o aviones altamente autónomos. Casi cualquier cosa del mundo físico puede sentir, actuar y conectarse con las implicaciones legales de privacidad y de responsabilidad por daños que ello pueda suponer. Para optimizar el uso de la energía, los sistemas domésticos inteligentes necesitan hacer un seguimiento de cuándo los residentes están ausentes y cuándo no, lo que facilita una información valiosa a los ladrones pero también a los proveedores de contenidos digitales, que así pueden mejorar sus emisiones cuando los usuarios están en casa. Para su correcto funcionamiento, un robot de limpieza necesita reconocer su entorno y hacer un mapa del espacio por el que se mueve, para evitar obstáculos y no repetir la limpieza en áreas por las que ya ha pasado. Cuando ese robot de limpieza se conecta con la nube del proveedor con la finalidad de facilitar datos que, en teoría, mejorarán el funcionamiento de todas las unidades de la compañía, también facilita un plano preciso de la casa y de lo que hay dentro de ella. Puede que el dueño de esa empresa decida vender esos datos, convirtiéndose no sólo en un proveedor de robots IoT de limpieza sino en una empresa que monetiza información.

Vestibles e ingeribles

Lo que tal vez impresione más es que estos dispositivos no sólo nos rodean como elementos externos a nosotros (prendas, vestibles, electrodomésticos o medios de transporte), sino que, por primera vez, los tragamos o los llevamos adheridos, implantados o inyectados. Ejemplo de este último tipo de tecnología es el de la pastilla digital, aprobada en noviembre de 2017 por la Administración de Alimentos y Medicamentos de Estados Unidos (FDA).[139]

Esta pastilla, que permite a los facultativos saber si sus pacientes están tomando la medicación y cuándo lo hacen, cuenta con un sensor comestible que, al ingerirlo, envía un mensaje a un parche y éste, a su vez, a una aplicación móvil que se actualiza en una plataforma, lo que, como es obvio, hace viajar la información por diversas redes con las implicaciones de seguridad que ello supone.

Centrándonos en el caso concreto, las autoridades sanitarias estadounidenses sólo han permitido su uso en un fármaco antipsicótico, al entender que es de interés público hacer seguimiento del tratamiento de una persona que se puede hacer daño a sí misma o a los demás. No se nos debe escapar la trascendencia que este tipo de uso del IoT tiene para la intimidad del enfermo, la seguridad de sus datos y, a medio plazo, para la expropiación del cuidado de la salud como una cuestión de interés general.

Sensorium

Sobre esta cuestión trata el escenario 4 (internet intencional de las cosas) planteado por el Centro de ciberseguridad a largo plazo, de la Universidad de Berkeley:[140] en 2020, el internet de las cosas es capaz de abordar los problemas de la educación, el medio ambiente, la salud, la productividad laboral y el bienestar personal. California liderará el camino con su robusto sistema «inteligente»

para la gestión del agua, y las ciudades adoptarán sensores conectados en red para gestionar problemas sociales, económicos y ambientales complejos, como la atención médica y el cambio climático, que solían parecer imposibles de arreglar. Sin embargo, no todos son felices. Los críticos hacen valer sus derechos y autonomía a medida que estas tecnologías se afianzan, y las tensiones internacionales aumentan a medida que los países desconfían de integrar estándares y tecnologías. Los piratas informáticos encuentran innumerables nuevas oportunidades para manipular y reutilizar la vasta red de dispositivos, de diez de manera sutil e indetectable. Debido a que el IoT está en todas partes, la ciberseguridad se convierte en simplemente «seguridad», y es esencial para la vida diaria. Pero mientras este nuevo mundo ofrecerá beneficios para la vida pública y revitalizará el papel de los gobiernos, también generará mayores vulnerabilidades a medida que las tecnologías del IoT se vuelvan esenciales para las funciones de gobierno y el bien colectivo.

Los gobiernos identificarán enormes beneficios para los dispositivos de diseño inteligente. A través de actos de «paternalismo positivo» (acción gubernamental intencional diseñada para mejorar la vida pública), los gobiernos desplegarán e implementarán el IoT en innumerables aspectos de la vida humana.

En este mundo, el IoT no sólo significará frigoríficos que reemplazarán automáticamente la leche cuando se agote, o tarjetas de crédito que vibrarán cada vez que se carga un gasto. Significará pulseras inteligentes que diagnosticarán problemas de salud en tiempo real y que pedirán atención médica sin intervención humana. Significará la medición inteligente de petróleo, gas y electricidad; semáforos que cambiarán automáticamente en función de los patrones de congestión; y sensores portátiles, el sucesor de Google Glass, que ayudará a los maestros en el aula a realizar un seguimiento de si los estudiantes prestan atención. En este mundo,

los gobiernos se convertirán en proveedores de infraestructura pública, creando nuevos productos altamente técnicos que sirvan al interés público.

Las fuerzas impulsoras detrás de la aparición de este «IoT intencional» son claramente visibles desde 2016. Los sistemas integrados y los sensores se están generalizando. Las pulseras MagicBand de Disneyworld permiten a los visitantes del parque pagar por artículos, reservar paseos, pedir comida y obtener experiencias personalizadas; también que el parque siga los flujos de visitantes y optimice la distribución de empleados, alimentos y otros servicios. Las redes de iluminación inteligente en calles, estacionamientos y centros comerciales utilizan led, sensores y datos, para encender y apagar las luces automáticamente, controlar los contaminantes, escuchar disparos o rastrear el tráfico e incluso a los compradores. El movimiento del «yo cuantificado» una vez descartado como un pasatiempo geek, está demostrando que las personas pueden usar sensores para rastrear datos de salud significativos y procesables sobre sí mismos. Del lado del Gobierno, tanto en Estados Unidos como en Europa, se están desarrollando modelos para un ecosistema de carreteras y vehículos conectados a internet: una red de comunicaciones inalámbricas que conecta automóviles, autobuses, camiones, trenes, señales de tráfico, teléfonos inteligentes y otros dispositivos para mejorar la seguridad y los flujos de tráfico, creando opciones de transporte más respetuosas con el medio ambiente.

Las ciudades inteligentes en lugares de alta tecnología y control como Corea del Sur (la ciudad de Songdo) y los Emiratos Árabes Unidos (Mazdar) serán los primeros indicadores de este cambio, ya que implementan visiones más expansivas del IoT para combatir los problemas inherentes a la densa vida urbana. En los próximos años, los planificadores de estas ciudades argumentarán abiertamente que el comportamiento humano podría y debería ser «administrado» por las aplicaciones de IoT para lograr una mayor eficiencia. Tratar de manera efectiva los desafíos sociales, económicos y ambientales

de la ciudad. Es importante destacar que esas ciudades, por sus regímenes políticos, estarán en posición de presentar ese argumento sin la carga negativa de ser considerada vigilancia que acompañaría a argumentos similares en el mundo occidental.

El verdadero cambio hacia la adopción generalizada de IoT ocurriría cuando los gobiernos federales de Estados Unidos adopten este nuevo modelo de una manera más enfocada, probablemente como una respuesta a las necesidades públicas urgentes. Por ejemplo, el gobernador de California podría anunciar una inversión estatal masiva en tecnologías de IoT para responder a la sequía y la crisis del agua en el estado. Los sensores se instalarían en ríos, presas, granjas, aguas subterráneas, distritos de agua, alcantarillado, negocios y hogares, junto con instrumentos reguladores del agua y bajo demanda, y dispositivos de reciclaje de agua para crear una red IoW (Internet of Water) que proporcionaría datos precisos al estado y administraría de manera más efectiva los recursos escasos. La publicación de estos datos en el dominio público crearía un mercado dinámico en el que las empresas privadas desarrollarían nuevos servicios y dispositivos vinculados al sistema, asumiendo, por supuesto, que el estado de California estuviera dispuesto a abstenerse de reglamentarlo en exceso.

Una iniciativa de IoT masiva y de alto perfil como ésta podría obtener un amplio apoyo público como una acción de «paternalismo positivo», cuyos beneficios eclipsarían las preocupaciones, que se percibirían como vagas e hipotéticas sobre la privacidad. Los partidarios argumentarán que, durante una grave sequía que amenaza la sostenibilidad fundamental de California como sociedad, la cantidad de agua que utiliza un hogar o una empresa ya no puede considerarse un asunto privado, como tampoco puede serlo el estado de vacunación de un individuo.

Y así, con unos artefactos conectados generando miles de datos por hora, construimos nuevas realidades en las que el uso del dato se pone al servicio del interés general, un interés cambiante y opaco, sibilino en su adopción, que puede poner en peligro

derechos como el de propiedad, intimidad, el derecho de defensa (¿cómo discutir en un juicio un resultado si no se tienen los mismos medios técnicos que la administración que nos controla?) o, incluso, libertad pública de naturaleza política; porque recordemos que todos estos artefactos conectados generan datos de todo tipo que son recogidos para conocer el comportamiento del individuo que los usa, pero también, de manera acumulada, para definir patrones de comportamiento generando una suerte de «inteligencia colectiva» que se incorpora a las cosas conectadas mediante modificaciones de su software e instrucciones de funcionamiento. Hablamos, entonces, del «Internet of Everything» o del IoT cognitivo, que nos acerca más a los retos de la robótica y del control total.

Con estos dispositivos podemos encontrarnos con escenarios de riesgo sencillos, como, por ejemplo, el de un sensor de temperatura casero que, cuando la calefacción de la comunidad no funciona, toma la decisión de poner en funcionamiento la bomba de calor de nuestro equipo individual de aire acondicionado; o escenarios más complejos, como aquel en el que todos los dispositivos y los datos generados por un determinado usuario están sincronizados de tal manera que, en atención al tiempo atmosférico, la distancia al puesto de trabajo, la situación del tráfico analizada en tiempo real, el histórico de las actividades del individuo, y el mantenimiento del vehículo autónomo, su despertador «toma la decisión» de sonar diez minutos antes o después, como si de un capítulo de la serie británica *Black Mirror* se tratara, la cocina conectada nos dejaría preparado, conforme a nuestros hábitos, el desayuno, la ducha estaría a la temperatura adecuada en atención a nuestros gustos, nuestra temperatura corporal y la del entorno, y una aplicación nos recomendaría la ropa más adecuada según lo programado en la agenda para ese día. En este escenario, junto con nuestra agenda, estarían en los dispositivos portables la programación del ejercicio necesario y las opciones de dieta para mantenernos en un estado óptimo.

Parece evidente que nos encontramos, por primera vez en la historia, una situación en la que la combinación de objetos, conectividad y software pueden obtener datos de una granularidad nunca vista. Esta nueva inmersión no es ya de nosotros en la tecnología, sino de la tecnología en nosotros.

No hay que ser un lector de literatura distópica, por tanto, para imaginar escenarios en los que el mal funcionamiento de alguno o todos estos elementos puedan causar un daño a una persona o a sus bienes, como, por ejemplo, un exceso de calor que provoque un incendio en la vivienda, quemaduras graves por duchas con agua demasiado caliente, o la planificación de un accidente catastrófico basado en el conocimiento de todos los conductores de una ciudad y en el acceso no consentido a sus vehículos.

En definitiva, esta conectividad sin precedentes que constituye lo mejor del IoT es, a la vez, lo peor; no sólo crea grandes oportunidades a la investigación y al bienestar personal, sino que lo que hace del IoT una tecnología tan usable y competente es lo mismo que genera riesgos considerables, y nuestro sistema de protección de la intimidad no está preparado.

Los sistemas ciberfísicos conectados, el IoT vitaminado de los coches autónomos, por ejemplo, requieren de grandes cantidades de datos para operar de manera efectiva, lo que plantea una infinidad problemas relacionados son la privacidad. Y con la seguridad.

La noche del 18 de marzo de 2018 en una carretera de cuatro carriles mal iluminada en Tempe, Arizona, un Uber Volvo XC90 autoconducido y conectado detectó un objeto delante de sí. Llevaba un controlador por exigencia legal que en lugar de mirar hacia la carretera iba consultando su teléfono móvil durante 19 minutos. Los sensores Lidar de radar y emisión de luz permitieron a los algoritmos de a bordo calcular que, dada la velocidad constante de su vehículo de 43 mph, el objeto que tenía delante estaba a seis segundos de distancia, suponiendo que no se moviese. Pero un objeto en una carretera rara vez permanecen parado, así que los

algoritmos buscaron en una base de datos de entidades mecánicas y biológicas reconocibles en busca de un ajuste del que se pudiera inferir el comportamiento probable del bulto que tenía delante.

En un primer momento, el ordenador de a bordo dibujó un espacio en blanco; segundos más tarde, decidió que se trataba de otro automóvil, por lo que esperó que se alejara y no requiriera ninguna acción especial. Sólo en el último segundo encontró una identificación clara: era una mujer con una bicicleta, bolsas de la compra colgando confusamente del manillar que, ilusa, pensó que el Volvo la rodearía como lo haría cualquier vehículo conducido por un humano. Como el vehículo tenía prohibido por programación tomar acciones evasivas sin intervención humana, devolvió bruscamente el control al controlador humano que estaba leyendo algo en la pantalla de su móvil y fue incapaz de reaccionar. Elaine Herzberg, de cuarenta y nueve años, fue arrollada por un coche que ni intentó frenar. Murió en el acto, dejando a todos los tecnooptimistas con varias preguntas incómodas: ¿era inevitable esta tragedia algorítmica? ¿Y estamos preparados para aceptar estos errores? ¿Los fanáticos de la privacidad deberían dejarse de tonterías y permitir ser escaneados, sondados, vigilados y observados para que la IA de los coches no atropelle a tranquilas ciclistas a la vuelta de una tarde de compras?

Hace veinte años, George Dyson anticipó mucho de lo que está ocurriendo hoy en su clásico libro *Darwin Among the Machines*. El problema sería que estamos construyendo sistemas que están más allá de nuestras capacidades de control. Para Dyson, lo preocupante es que la sociedad móvil conectada funciona como un organismo digital multicelular, casi como la conciencia colectiva de los Borg de *Stark Trek*.

Los problemas graves que el IoT y la movilidad en general suponen para la privacidad no son nuevos. Ya la CCN, en su guía IoT, destacaba que: «las preocupaciones más obvias tienen que ver con la privacidad. La recolección masiva de nuestros datos y metadatos puede que sea lo mismo que instalar cámaras de

vigilancia, pero es una forma de vigilancia digital incluso más peligrosa para nuestra intimidad y libertad que la mera observación visual». En la misma guía al relacionar las superficies de ataque, se menciona la privacidad en relación con la divulgación de datos del usuario; la publicación de la ubicación del usuario a través del seguimiento de su dispositivo; y la existencia de sistemas con privacidad diferencial, en la que unos pocos monitorizan a todos y nadie los monitoriza a ellos.

Por si había alguna duda, nuestras bombillas IoT nos vigilan.

11

Nuestros cuerpos, nosotros

Datos biométricos: voz e imagen

Theodore Twombly es un hombre solitario, introvertido y deprimido con un bigote ridículo que escribe cartas personales para familiares o seres queridos de aquellos que, por algún motivo, son incapaces de hacerlo por sí mismos. Theodore lleva una vida anodina para anestesiar la pena causada por una larga relación que ha tocado a su fin. Intrigado y harto de tanta soledad, adquiere un nuevo y avanzado sistema operativo que aprende de la interacción con su usuario a través de un interfaz de voz con lenguaje natural. Cuanto más habla con él, mejor lo entiende, más natural es el lenguaje, mejor fluye la relación. Su sistema operativo es «Samantha», y cuenta con la voz aterciopelada de Scarlett Johansson. Imposible para Twombly no caer enamorado de esa IA a la que, sin saberlo, entrena y llena de curiosidades y que, por ello, al final de la cinta se ve obligada a seguir aprendiendo en otro lugar; tal vez en la nube, tal vez en la plataforma del prestador de servicios junto con otras Samanthas. Spike Jonze escribió, produjo y dirigió *Her* (2013), una de las mejores películas sobre la IA aplicada a las labores de acompañamiento, cuidado y asistencia, una evolución de Alexa o la propia Siri, todas ellas con nombres de mujer.

Cuando en 2011 Phil Schiller, vicepresidente de marketing de Apple, presentó al mundo a Siri, el primer asistente virtual activado por voz de la historia, nadie imaginó que el futuro dejaría de ser escrito y tecleado. Siete años después, 710 millones de personas utilizan y hablan con estos ayudantes 2.0 al menos una vez al mes, según la consultora Tractica, y se calcula que en 2018 los ayudantes virtuales alcanzarán los mil millones de usuarios. El habla es la manera más natural que tenemos los humanos de interactuar. Si aplicamos este modo de comunicarnos a nuestra relación con los dispositivos electrónicos, el paradigma de la relación entre los usuarios y las máquinas cambiará por completo. En palabras de Schiller en aquella ya lejana presentación: «Lo único que queríamos era hablar a nuestros dispositivos, hacerles preguntas sencillas y que nos dieran una respuesta. Queríamos poder hablar con ellos como con cualquier persona».

Siri, ¿Dios existe?

Si escribir a mano es una práctica en extinción, los asistentes virtuales como medio de comunicación con nuestra red, nuestro IoT casero o nuestros equipos informáticos, van a acabar definitivamente con los teclados y nos van a obligar a todos a mejorar nuestra dicción. El lenguaje es el modo más rápido y más natural de comunicarnos con nuestro entorno, de dar órdenes y de recibirlas, de expresar pensamientos, ilusiones y opiniones. Los asistentes virtuales, como dispositivos IoT, son y están dotados de micrófonos y sensores de sonido o movimiento, y de una inteligencia mínima para hacer los procesos locales más específicos. Es en las plataformas de datos donde está la IA capaz de procesar el lenguaje, la entonación y de tener preparadas o improvisar las respuestas a nuestras preguntas y peticiones. Aunque buena parte de lo que hacen por ahora es producto de una programación para dar respuesta rápida a una amplia panoplia de situaciones, la

sensación del usuario es de estar hablando con alguien que le entiende, aunque no pasasen el test de Turing[141] con un experto. Esto genera, como le ocurre a Theodore Twombly, la falsa sensación de que el asistente nos entiende y de que sabe a qué nos referimos.

La integración total del lenguaje natural en los asistentes virtuales nos acerca un poco más a aquella promesa que nos hizo Phil Schiller. Aunque estos asistentes nunca llegarán a sustituir una conversación con otra persona, algunas empresas están creando la ilusión de que sí es posible a través de aplicaciones como Xiaoice (desarrollado por Microsoft para China). Hay que recordar que las máquinas no son capaces de entender lo que decimos y que las respuestas que nos dan, ingeniosas, divertidas o cortantes, son fruto de una base de datos preescrita por periodistas, guionistas y lingüistas. No hay ninguna espontaneidad. No son capaces de crear lenguaje. Por ahora.

La combinación de estas dos tecnologías, la IA y el reconocimiento de voz, con el internet de las cosas, ha dado lugar a nuevos artilugios, como Echo de Amazon, el Google Home o el Apple Pod. Estos dispositivos cuentan con una serie de micrófonos que están constantemente «escuchando» lo que ocurre a su alrededor y en el momento en el que captan el comando de activación se ponen a nuestro servicio.

La voz, por si no lo ha adivinado ya el lector, es un dato personal conforme a la legislación de la UE y un elemento de autenticación biométrica que se usa en entornos de seguridad en sustitución de un par de claves o un tocken de acceso. Por tanto, saber si las conversaciones que tenemos con nuestro asistente virtual se graban, se suben a la plataforma del prestador (Google, Amazon, Apple...) y se analizan no sólo para entrenar a la IA sino para agregarlo al perfil que ya tienen de nosotros, no es una cuestión menor.

Behshad Behzadi,[142] uno de los principales ingenieros de Google Assistant, mantiene que las conversaciones se desechan, pues «aunque el asistente esté todo el día escuchando, esa información no se envía a ninguna parte. El asistente de Google sólo se activa en el momento en el que dices “Ok Google”, y a partir de ahí sí recoge las peticiones y las envía». Según Behzadi, una parte de su archivo es privada, y sólo se utiliza para almacenar datos personales de su dueño, lo que no deja de suponer una integración de la voz con todos sus elementos (tonalidad, acento, idioma hablado) junto con el contexto (ruidos de fondo de una lavadora, otras personas hablando, un niño jugando) en los datos que Google ya tiene de nosotros por otros medios.

OK, Google, ¿qué es un Whopper?

Otra cuestión es el estado de escucha activa. Según todos los fabricantes, el dispositivo sólo se activa cuando se dicen las palabras mágicas para que se pongan en funcionamiento, pero para que eso ocurra ha de estar atento, conectado, escuchando, para discernir si se está teniendo otra conversación o si se le está dando una orden. La escucha activa, en definitiva, requiere que el micrófono esté conectado y atento.

Un amigo se despertó sobresaltado de madrugada al escuchar la voz de una mujer que se dirigía a él desde detrás de la mesilla. Cuando fue a coger el teléfono para llamar al 112 comprobó que era Siri quien le hablaba, contestando a sus peroratas inconscientes. Así descubrió de una sola tacada que hablaba en sueños y que tenía a Siri configurada para activarse con cualquier voz.

No es el único. Una pareja de Portland afirmó que su dispositivo Amazon Echo grababa sus conversaciones privadas. Como prueba presentaron una concreta conversación privada entre ellos que acabó siendo enviada, íntegramente, a uno de sus contactos quien les alertó inmediatamente de lo sucedido. Para ellos esto resultaba

una prueba suficiente de que los asistentes personales nos espían y graban nuestras conversaciones de manera habitual.[143] Las explicaciones de Amazon[144] al respecto son perfectamente razonables aunque un poco rocambolescas: una palabra en una conversación de fondo que habría sonado parecida a «Alexa» habría provocado la activación de Echo, que escuchó la siguiente conversación como una solicitud de «enviar mensaje». En ese momento, Alexa habría dicho en voz alta «¿a quién?», y por la razón que sea, la conversación de fondo se interpretó como un nombre en la lista de contactos de los vecinos de Portland. Alexa habría preguntado de nuevo en voz alta: «[nombre de contacto], ¿no?», y de algún modo que nadie se explica habría interpretado la conversación de fondo como «correcto». Y así fue como Amazon Echo, según Amazon, grabó una conversación privada, y tras varias verificaciones mal entendidas, envió la grabación a un tercero.

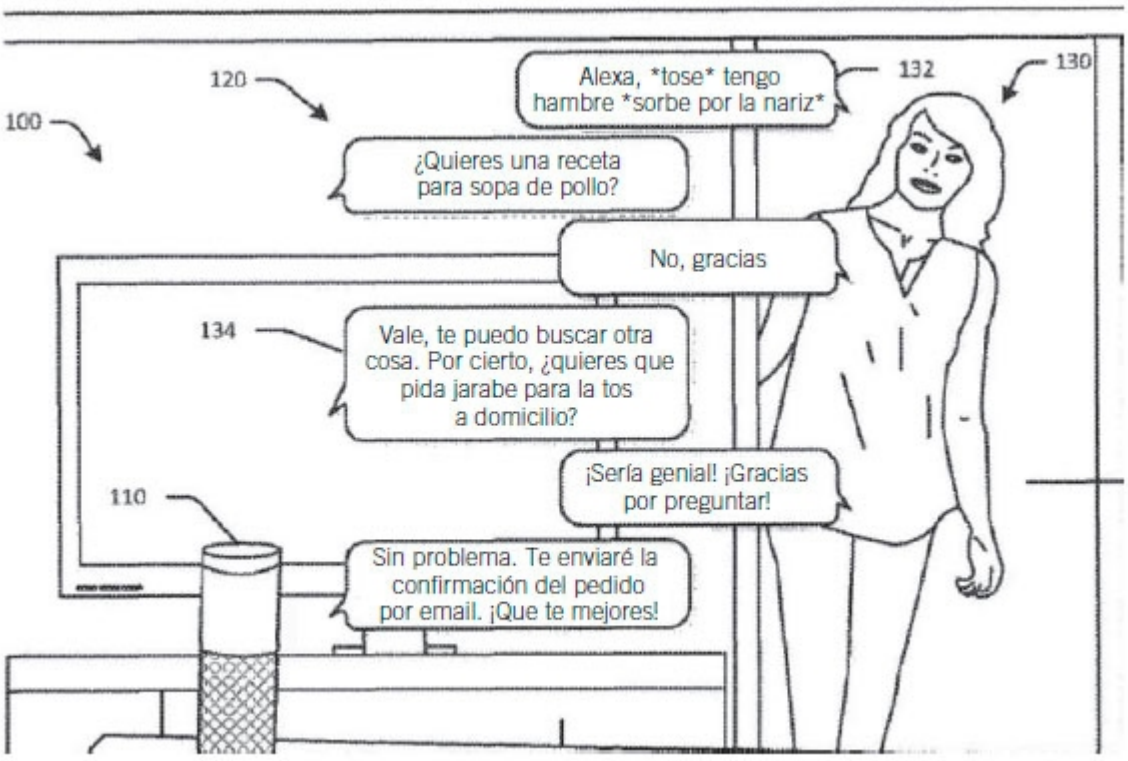
Creamos o no las explicaciones de Amazon, lo cierto es que la Alexa futura está pensada para estar cotilleando de manera permanente. Amazon ha registrado una patente para que Alexa escuche el entorno y detecte si el usuario está enfermo, aburrido o infeliz.[145] La patente permite a Alexa detectar automáticamente cuando alguien está enfermo y ofrecerse a comprarle medicamentos en la página de Amazon, obviamente. Según la patente, Alexa analizaría el habla e identificaría otros signos de enfermedad o emoción.

Alexa primero le sugiere un poco de sopa de pollo para curar su resfriado para, a continuación, ofrecerse a comprar un antitusivo en Amazon. Además, si Alexa escucha a los usuarios llorar, conforme a esta patente, los clasificaría como que están experimentando una «anormalidad emocional». Esto se usaría no sólo para comprar en Amazon sino para servirles publicidad ajustada a su enfermedad o incomodidad, ya que los anuncios se mostrarían en atención a su estado de ánimo o salud.

Nos espíen o no, lo cierto es que Alexa, como Siri y los dispositivos IoT que las representan en el mundo físico, cuentan con una seguridad muy precaria que puede ser sorteada de la manera más sencilla. Un grupo de estudiantes de la Universidad de California, Berkeley y la Universidad de Georgetown en 2016, demostraron que podían ocultar los comandos con el ruido blanco que se escuchaba en los altavoces y en los vídeos de YouTube para que los dispositivos inteligentes activaran el modo avión o abrieran un sitio web;[146] enviando comandos ocultos no detectables para el oído humano a Siri de Apple, a Alexa de Amazon, y a Google's Assistant pudieron activar secretamente los sistemas de inteligencia artificial de los asistentes virtuales, y los usaron para desbloquear puertas,[147] transferir dinero o comprar online,[148] insertando esos códigos inaudibles en la música que se reproducían en la radio.

FIGURA 2

Ejemplo que aparece en la patente de Alexa que representa a una mujer que tose y resopla mientras habla con su dispositivo Amazon Echo



Esto ilustra cómo la inteligencia artificial, incluso cuando está haciendo grandes avances, aún puede ser engañada y manipulada. Se puede engañar a un ordenador para que identifique a un avión como un gato simplemente cambiando unos pocos píxeles de una imagen digital, y los investigadores pueden hacer que un vehículo autónomo se desvíe o acelere simplemente mediante pegativas adheridas a las señales de tráfico confundiendo así el sistema de visión computarizada del mismo.

Con los ataques de audio, los investigadores están explotando la brecha entre el reconocimiento de voz humano y la máquina. Los sistemas de reconocimiento de voz generalmente traducen cada sonido a una letra, para compilarlos primero en palabras y luego en frases. Al realizar leves cambios en los archivos de audio, los investigadores pudieron cancelar el sonido que se suponía que el sistema de reconocimiento de voz debía escuchar y reemplazarlo con un sonido que las máquinas transcribirían de manera diferente. Y todo ello con sonidos casi indetectables para el oído humano.[149]

La proliferación de aparatos activados por voz aumenta las implicaciones a estos hackeos. Amazon dijo que no revela las medidas de seguridad específicas, pero que ha tomado algunas para garantizar que su altavoz inteligente Echo sea seguro. Google considera la seguridad como un enfoque continuo, y aseguró que su asistente tiene características para mitigar comandos de audio no detectables. Los asistentes de ambas compañías emplean tecnología de reconocimiento de voz para evitar que los dispositivos actúen con ciertos comandos a menos que reconozcan la voz del usuario. Apple mantiene que su altavoz inteligente, HomePod, está diseñado para evitar que los comandos realicen acciones como el desbloqueo de puertas, y señaló que los iPhones y iPads deben desbloquearse antes de que Siri actúe sobre los comandos que acceden a datos confidenciales o abran aplicaciones y sitios web, entre otras medidas.

Sin embargo, muchas personas dejan sus teléfonos inteligentes desbloqueados y, al menos por ahora, los sistemas de reconocimiento de voz son notoriamente fáciles de engañar. Ya existe un historial de dispositivos inteligentes que están siendo explotados para obtener ganancias comerciales a través de comandos hablados.

Burger King usó esta debilidad en una acción publicitaria: un anuncio online que preguntaba «OK, Google, ¿qué es un Whopper?». Los dispositivos Android respondieron leyendo la página de Wikipedia de Whopper. El anuncio se canceló después de que los espectadores comenzaron a editar la página de Wikipedia sólo para que el asistente leyese textos subidos de tono o humorísticos. Unos meses más tarde, la serie *South Park* le dedicó un episodio a los comandos de voz que provocaron que los asistentes de los espectadores repitieran las mismas obscenidades de los adolescentes en tiempo real.

Bares y fútbol

La sensación de ser escuchados, ya no por nuestros asistentes virtuales sino por nuestros teléfonos, se extiende como una leyenda urbana.

Un usuario toma un café con un amigo para organizar su viaje a Cuenca. Ambos tienen el teléfono inteligente encima de la mesa y ambos reciben, al cabo de unas horas, información sobre hoteles en la ciudad o billetes de tren baratos para Cuenca. ¿Los móviles nos escuchan?

Al descargar una aplicación, el usuario debe dar permiso a los requerimientos que le llegan y aceptar, por ejemplo, que la app acceda a su micrófono, cámara de fotos o geolocalización. En 2016, la periodista de Tecnología de la BBC, Zoe Kleinman[150] retó a dos expertos en ciberseguridad de la compañía Pen Test Partners, en Buckingham (Reino Unido), a desarrollar una app capaz de escuchar a los usuarios a través del micrófono del móvil. Su intención era comprobar si era viable desde el punto de vista técnico que una aplicación estuviese en modo de escucha activa, reconociese determinados parámetros y los subiese a una plataforma desde la que ofrecer servicios o servir publicidad. Después de dos días de trabajo, los ingenieros crearon un prototipo para Android, lo instalaron en un teléfono y le dieron acceso total al micrófono. Consiguieron que la app transmitiera en tiempo real y, aunque «con algunas dificultades», fueron capaces de identificar las palabras clave de las conversaciones.[151] Los investigadores concluyeron que no podían asegurarlo pero que resultaba perfectamente posible que las aplicaciones en las que activamos el micrófono sin restricciones, puedan escuchar, grabar y procesar nuestras conversaciones.

El caso de la aplicación de La Liga de Fútbol española para Android apuntalaría la idea de que las aplicaciones en las que no se gestionen adecuadamente los permisos de acceso al micrófono darían lugar a un dispositivo espía, en este caso, de La Liga, intentando averiguar qué bares no estarían pagando por la emisión de los partidos.

Según la información facilitada por La Liga, el micrófono sólo captaría un código binario de fragmentos de audio —con la única finalidad de poder conocer si está viendo partidos de fútbol de competiciones disputadas por equipos de La Liga—, pero nunca se accedería al contenido de la grabación que, según informaron, se convertiría en código binario en el propio terminal y, como tal, se transmitiría de inmediato. Esto significa que no se grabaría y enviarían las grabaciones completas a un servidor gestionado por La Liga. Además, ésta indicó que sólo activaría el micrófono y el geoposicionamiento del dispositivo móvil durante las franjas horarias de partidos en los que compitieran equipos de la competición.

Sin embargo, un análisis técnico[152] dejó en evidencia las afirmaciones de La Liga, ya que resultaba improbable que sólo subiera archivos binarios, y sobre todo que el micrófono grabara fragmentos de audio y los subiera tal cual a un servidor para su reconocimiento *a posteriori*.[153]

Persons of Interest

Resulta inquietante la sensación de estar siendo escuchado en todo momento, de estar siendo seguido por las cámaras de seguridad allí donde nos encontremos. Y más inquietante aún es que haya una inteligencia artificial que recibe toda esa información y la use con fines oscuros a favor de alguna agencia gubernamental secreta como en la serie *Persons of Interest.* Sin poder afirmar ni negar que exista una entidad similar, lo que sí he comprobado con mis propios ojos es que hay sistemas de reconocimiento facial, de objetos y de movimientos en tiempo real, que son capaces de captar a una persona bajándose de un avión en Londres y seguirla por toda la ciudad durante todos los días de su estancia hasta que vuela de regreso a casa. Las imágenes son muy similares a las que se ven en *Persons of Interest*, y el sistema permite individualizar no sólo a la persona sino a los elementos que lleva consigo, para así saber si

ha dejado el maletín al lado de una papelera en la Estación Victoria, si se ha olvidado el móvil o la cartera en un café o si se ha cambiado de gabardina de manera sospechosa. El sistema era, además, capaz de analizar los gestos, el comportamiento de todos los ciudadanos de una gran urbe, y modelizar aquellos que se consideraran sospechosos y hacerlo con un algoritmo de reconocimiento. Así, cuando alguien se comportase sospechosamente, el sistema lo identificaría de inmediato y haría saltar una alarma para que la autoridad competente revisara el vídeo y, en su caso, actuase.

Ya no nos hacen falta los Precogs de *Minority Report* para tener sistemas de precrimen basados en el tratamiento de todo tipo de datos, incluidos los biométricos de voz e imagen. Según Joseba Elola, «alguien te puede hacer una foto por la calle y conseguir saber quién eres para contactarte. Ocurre en Rusia.[154] Alguien puede cruzar un paso de cebra cuando no toca y ver que las autoridades le multan y pegan su foto cruzando indebidamente en las paradas de autobús tras identificarle con la imagen captada por una cámara de seguridad. Ocurre en China. Uno puede recibir la inoportuna visita de la Policía porque el algoritmo falló y lo identificó erróneamente. Ocurrió en Estados Unidos, en cinco ocasiones, con cinco ciudadanos, en 2015, según ha admitido la Policía de Nueva York. Todo ello podría haber ocurrido en otros momentos de la historia, pero nunca ha sido tan fácil como ahora. La tecnología del reconocimiento facial viene cargada de comodidades, sí, de promesas de una mayor seguridad, vale. Pero, en paralelo, la expansión de toda una industria de seguridad que pivota en torno a ella convierte la pesadilla orwelliana de una sociedad de ciudadanos controlados en algo más que una posibilidad de futuro. Así, el mercado del reconocimiento facial mueve ya más de 3.300 millones de dólares (2.800 millones de euros) en el mundo y podría llegar a 7.700 millones de dólares en 2022, según la consultora MarketsandMarkets. Bancos, compañías aéreas, de telefonía, fabricantes de ordenadores... todos se abren a esta nueva forma de

identificación biométrica que supone un salto adelante frente a la huella dactilar y el iris. Pero la cara no es lo mismo que la huella dactilar. Cuando acudimos a renovar el carnet, consentimos en ceder ese dato biométrico a las autoridades. Pero nuestra cara la puede captar quien quiera sin nuestro consentimiento. Con cualquier cámara que haya en la calle, en cualquier lugar».[155]

Esta tecnología se divide en dos grandes ramas: la de detección facial (*face detection*), que permite autenticar una identidad a través de los rasgos biométricos del rostro y que sería la que usa el iPhone como clave de desbloqueo; y la de *face search* o buscador de caras, que permite cruzar una imagen con las almacenadas en una base de datos en búsqueda de coincidencias. Para evitar falsos positivos o falsos negativos, se necesita que la IA esté entrenada para no confundirse, y de ahí que los grandes repositorios de datos de las redes sociales[156] o de Google usen estas fotos para entrenar sus programas de reconocimiento facial que, en caso de necesidad, pueden acabar en manos del Gobierno estadounidense con fines menos lúdicos que los de marcar a tus amigos en las fotos.

Más allá de una foto, los sensores frontales del iPhone X escanean 30.000 puntos para hacer un modelo 3D de nuestro rostro. Así es cómo el iPhone X usa nuestra cara como clave de desbloqueo. Ahora que un teléfono puede escanear nuestra cara con un nivel de precisión de 30.000 puntos, el sistema podría observarme y rastrear mis expresiones para juzgar si estoy deprimida, contenta, angustiada o enamorada; establecer mi sexo, raza, e incluso sexualidad. De hecho, ya se ha usado una IA para determinar con bastante precisión la orientación sexual[157] a partir de fotografías normales de los individuos estudiados. El estudio, como todos, tenía sus limitaciones, pero aun así la velocidad de cambio y las obvias implicaciones para la privacidad son ya un problema que necesitaba solución.

Facebook cuenta con una patente que permite servir contenido basado en las emociones detectadas.[158] Apple, en 2016, compró una startup llamada Emotient especializada en la detección de emociones.

Con cámaras no especialmente sofisticadas, compañías como Kairos desarrollan software para identificar el género, la etnia y la edad, así como el sentimiento de las personas. En los últimos doce meses, Kairos ha leído 250 millones de rostros para clientes que buscan mejorar su marketing y los productos que venden.

En China, por su parte, los ciudadanos que se escaneen el rostro con la aplicación móvil Xiaohua Qianbao pueden pedir un préstamo al banco virtual operado por la propia Xiaohua, o pagar con una sonrisa un cubo del Kentucky Fried Chicken de Hangzhou gracias al sistema desarrollado para la aplicación de pagos en línea Alipay.

Nuestro iPhone podría, incluso, tratar nuestra imagen con otros datos y rastrearnos en nuestra vida diaria sin mayor complicación.[159] Teniendo en cuenta que el negocio de Apple es el hardware y no los datos, cederle nuestro rostro no tendría por qué causarnos mayor preocupación. Pero, como sabemos, cualquier tecnología es producto del trabajo de más de un equipo de más de una compañía: Apple podría haber compartido mapas faciales con los fabricantes de aplicaciones que a lo mejor no están tan dispuestos a honrar sus compromisos de confidencialidad con la empresa de la manzana mordida como debieran.

Los teléfonos con Android, por su parte, han contado con características de desbloqueo facial durante años, pero la mayoría no ha ofrecido el mapeo facial 3D como el iPhone. Al igual que iOS, Android no hace una distinción entre las cámaras delantera y trasera. La Play Store de Google no prohíbe que las aplicaciones utilicen la cámara frontal para la comercialización o la creación de bases de datos, siempre que pidan permiso.

Nuestras caras, como un dato biométrico que no puede cambiarse (a no ser que se sea un narcotraficante colombiano dispuesto a operarse), tienen un enorme valor, pues son el elemento de identificación no intrusiva definitivo. La mitad de todos los adultos estadounidenses tienen sus imágenes almacenadas en, al menos, una base de datos que la Policía puede buscar, generalmente con pocas restricciones.

Facebook y Google utilizan IA para identificar caras en imágenes que cargamos en sus servicios fotográficos. En abril de 2015, Facebook recibió una demanda colectiva[160] por su sistema de etiquetado mediante reconocimiento facial por violación de la Illinois's Biometric Information Privacy Act[161] que ha sido admitida a trámite[162] en contra de las peticiones de la compañía. No es la única: el activista austríaco Max Schrems,[163] conocido por haber conseguido anular el sistema de transferencia de datos personales entre UE y Estados Unidos llamado coloquialmente «Safe Harbour»,[164] ha conseguido también que la compañía anulase parcialmente algunas de sus funcionalidades relacionadas con el reconocimiento facial.

El mejor ejemplo de *face search* lo ofrece la aplicación FindFace, que permite que cualquier persona saque una foto de cualquiera sin su permiso para que el algoritmo de la aplicación cruce la imagen con las existentes en la red social Vkontakte (con más de 400 millones de perfiles) e identifique con un 70 por ciento de probabilidades a la persona en cuestión.

En China, la tecnología avanza con paso firme de la mano de Face ++, startup que derrotó a finales de octubre a Facebook, Google y Microsoft en pruebas de reconocimiento de imagen en la Conferencia Internacional de Visión por Ordenador celebrada en Italia. Un Estado totalitario como el chino está usando las técnicas de recogida de datos empleadas hasta ahora por las empresas privadas, tanto chinas como estadounidenses, como medio para establecer un sistema de control férreo de sus ciudadanos, mucho más eficiente que el operado por el Partido Comunista durante el

Gran Salto Adelante del líder Mao Zedong. Veremos cómo el Estado del País del Centro ha sido capaz de aprender todas las técnicas de adicción, de desarrollar tecnología propietaria de recogida de datos, de análisis e IA, con la finalidad de establecer el sistema orwelliano perfecto.

Las autoridades policiales de todo el mundo están desarrollando y utilizando tecnologías para identificarnos a partir de nuestros datos biométricos, incluidos el rostro, las huellas dactilares, el ADN, la voz, el iris y la manera de andar o comportarnos. Empezamos admitiendo la recogida de marcadores de identidad únicos en los controles de pasaportes de los aeropuertos para entrar en determinados países, aceptamos luego que se escanease nuestra huella y se nos saque una foto, y no sé si en la siguiente vuelta de tuerca nos van a pedir el ADN. En un reciente viaje de vuelta de Latinoamérica experimenté el último invento de nuestras autoridades para dejarme entrar en mi propio país: para pasar los tornos tuve que mostrar mi pasaporte electrónico, dejar mi huella y mi foto desmayada después de trece horas de vuelo y caminar en estado de vigilia hacia los ascensores bajo la mirada irritada de los policías del control. El sistema era tan nuevo y la gente estaba tan despistada, que había policías abriendo los tornos y quejándose de que el invento les duplicaba el trabajo. Algo muy español, pensé. Pero no dudo que en cuanto el sistema se engrase, junto con los más de 30 datos que vuelan conmigo cada vez que subo a un avión, iré dejando mi biometría en diversas jurisdicciones, incluida la española.

Aadhaar

La seguridad frente a los ataques yihadistas es siempre el argumento. El mismo utilizado en el proyecto biométrico más importante del mundo, una solución multimodal (iris, huellas dactilares y rostro) que afecta a más de mil millones de indios.

Nandan Nilekani, el presidente de Infosys que dejó su trabajo para crear el sistema, conocido como Aadhaar, ha prometido al Gobierno indio ahorrarle aproximadamente 9 mil millones de dólares al eliminar las identidades duplicadas y falsas en las listas de beneficiarios del Gobierno.

Gracias a Aadhaar, más de 500 millones de indios han conectado sus identificaciones digitales directamente a una cuenta bancaria, lo que permite al Gobierno hacer pagos de más de 12 mil millones de dólares sin el riesgo de fraude, robo, o para evitar que los maltratadores se hagan con los pagos dirigidos a sus víctimas. Para muchos indios que viven en aldeas no pobladas o en barrios marginales, una identificación digital les otorga personalidad oficial, tanto como el DNI o un número de seguridad social en los países desarrollados.

Aadhaar, en su corta vida, se ha enfrentado a problemas legales y de seguridad. Por un lado, si bien acogerse al sistema es voluntario, se convierte en obligatorio para cualquier persona que necesite acceder a los servicios gubernamentales, abrir una cuenta bancaria o contratar un teléfono móvil. La Corte Suprema india declaró en 2017 ilegal su uso obligatorio: «El derecho a la privacidad [...] es una parte intrínseca del derecho a la vida y la libertad personal». Si bien, reconoció que el Gobierno indio podía limitar el derecho a la privacidad por motivos necesarios, siempre que la limitación fuese razonable y proporcional al fin buscado.

Más preocupante es que Aadhaar no es seguro. En enero de 2018, reporteros del diario indio *The Tribune* pagaron 500 rupias (poco menos de 8 dólares) para obtener un nombre de usuario y contraseña que les permitiera acceder al nombre, dirección, código postal (PIN), foto, número de teléfono y correo electrónico de todas las personas dadas de alta en el sistema. Por 300 rupias más, los reporteros pudieron imprimir (y comenzar a usar) copias de las tarjetas de identidad únicas de cualquier persona de la base de datos.[165]

Esta biometría aumenta la probabilidad del panóptico de Jeremy Bentham, la distopía de un Estado que todo lo ve. China no hace ningún esfuerzo por ocultar su uso de la biométrica y la inteligencia artificial (IA) para vigilar a su población. Menos conocido, sin embargo, es el uso avanzado de la biométrica en las democracias liberales.

En Estados Unidos, un estudio realizado en 2016 por el Centro de Privacidad y Tecnología en el Centro de Derecho de la Universidad de Georgetown determinó que las imágenes faciales de más de 117 millones de estadounidenses, casi la mitad de todos los adultos de Estados Unidos estaban a disposición del FBI. En el último trimestre de 2018, el Departamento de Aduanas y Protección Fronteriza de Estados Unidos —US Customs and Border Protection (CBP)— comenzó a utilizar una nueva tecnología de reconocimiento facial como parte de Traveler Verification Service (TVS), en concreto, cuando se sale del territorio estadounidense. El Programa de Salida Biométrica —«Biometric Exit program»— se teme que acabe en las bases de datos privadas de algunas empresas.[166]

En el Reino Unido, la Base de Datos de la Policía Nacional (NPD) almacena las imágenes faciales de 12,5 millones de personas, cientos de miles no culpables de ningún delito. Por su parte, el HM Customs and Revenue (HMRC) tiene más de cinco millones de grabaciones sin consentimiento a pesar de que el Tribunal Supremo británico ordenara al Ministerio del Interior en 2012 eliminar la biometría personal y de voz de los detenidos liberados sin cargos o absueltos, de acuerdo con la ley local que exige la eliminación del ADN y las huellas dactilares en esos casos. La Policía de Gales del Sur y la metropolitana se enfrentan a acciones legales emprendidas por Liberty y Big Brother Watch, respectivamente, por el uso indiscriminado del reconocimiento facial automático. En Estados Unidos, la ciudad de Orlando, Florida, ha tenido que abandonar la prueba del software de reconocimiento facial que les había proporcionado Amazon.

La recopilación y almacenamiento de datos biométricos de los ciudadanos cambia fundamentalmente la relación entre éstos y el Estado. Una vez fuimos «presuntos inocentes», para ser ahora, en palabras siniestras del exsecretario de Interior del Reino Unido, Amber Rudd, «personas no procesadas», individuos que aún no han sido declarados culpables de un delito. Los avances tecnológicos en los últimos años han puesto de relieve no sólo los beneficios de los grandes datos, sino también la necesidad de aceptar los peligros que representan para nuestra privacidad, libertades civiles y derechos humanos.[167]

A diferencia de nuestros nombres de usuario o contraseñas, los datos biométricos no se pueden reestablecer y los errores son aún más difíciles de corregir. Y cuando se combinan con otros datos sobre nosotros (financieros, profesionales y sociales), nuestros datos biométricos pueden incorporarse a algoritmos y utilizarse para negarnos préstamos, seguros de salud y empleos, adivinar nuestra sexualidad o preferencias políticas y predecir nuestra probabilidad de comprometernos o de cometer delitos, todo ello enteramente sin nuestro conocimiento.

Sólo tienes una cara, así que será mejor que la ocultes bien.

12

Tracking

Geoposicionamiento y monitoreo personal

La última vez que Anthony Aiello habló con su hijastra, llevó pizza y bizcochos caseros a su casa en San José, California, para una breve visita. El señor Aiello, de noventa años, dijo a los investigadores que luego lo acompañó a la puerta y le entregó dos rosas en agradecimiento. Pero un observador inadvertido en la casa reveló más tarde que su encuentro terminó en asesinato. Cinco días después, la hijastra del señor Aiello, Karen Navarra, de sesenta y siete años, fue descubierta muerta por una compañera de trabajo en su casa con heridas en la cabeza y el cuello. Karen usaba una pulsera de ejercicios Fitbit, que según los investigadores, mostró cómo su ritmo cardíaco había aumentado significativamente alrededor de las 15.20 del 8 de septiembre, cuando el señor Aiello estaba allí. Más tarde la pulsera registró que su ritmo cardíaco se desaceleraba rápidamente y se detenía a las 15.28, unos cinco minutos antes de que el señor Aiello saliera de la casa. El señor Aiello fue arrestado por asesinato e ingresado en la cárcel del condado de Santa Clara

Aunque originalmente tenían la intención de motivar a las personas a tomar control de su estado físico y de su salud, los dispositivos de acondicionamiento físico han encontrado su camino

en la caja de herramientas tecnológicas que los expertos de la ley utilizan para resolver delitos, junto con vídeos, dispositivos GPS y teléfonos móviles.

No sin mi Fitbit

Pegados al cuerpo de una persona, los dispositivos están sentados en la primera fila de la vida de sus anfitriones, documentando inadvertidamente para éstos los encuentros mundanos y peligrosos mientras registran los latidos del corazón, los patrones de sueño y el esfuerzo físico.[168]

Los datos de ubicación de Fitbit se usaron en un caso de agresión sexual en Pennsylvania en 2015 y en un caso de lesiones personales en Canadá en 2014. Un GPS de Garmin Vivosmart registró una pelea de dos mujeres con una atacante en Seattle en 2017. El mismo año, los investigadores utilizaron datos del Fitbit de una mujer de Connecticut para acusar de asesinato a su marido.

En 2018, investigadores en Iowa, con la ayuda de expertos del FBI, revisaron los datos de Fitbit de Mollie Tibbetts, una estudiante de veinte años que estuvo desaparecida durante aproximadamente un mes antes de que su cuerpo fuera descubierto en agosto.

El geoposicionamiento de los dispositivos de entrenamiento o de los móviles que nos acompañan a todas partes como un nuevo órgano, han sido igualmente usados en España para la resolución de diversos casos mediáticos de asesinato. De hecho, si uno hace caso a las informaciones, tan malo es llevar el móvil que te ubica en un sitio a una hora, como no llevarlo, para evitar que te ubique.

El geoposicionamiento y los datos recogidos por un dispositivo llamado vestible (del inglés *weareble)* son especialmente sensibles para quien los usa, pero no sólo.

El Pentágono de Estados Unidos ha dado una orden estricta prohibiendo a sus tropas militares y al personal de defensa en bases sensibles o en zonas de guerra de alto riesgo utilizar un rastreador

de ejercicios o aplicaciones de teléfonos móviles que puedan revelar su ubicación.[169] Nunca antes el movimiento de tropas se ha podido seguir por el enemigo con tanta precisión. No sólo se pueden ver los datos juntos en Global Heat Map, sino que puedes seguir los progresos de tu combatiente favorito a través de sus redes sociales. Muchas aplicaciones de entrenamiento publican los resultados de manera automática en las redes sociales de su propietario, para proporcionar un postureo sin esfuerzo.

El memorándum en cuestión publicado por The Associated Press, prohíbe los rastreadores de ejercicios u otros dispositivos electrónicos a los militares desplegados en zonas de conflicto, que a menudo están vinculados a aplicaciones de teléfonos móviles o relojes inteligentes y pueden proporcionar los detalles de GPS y de ejercicio de los usuarios a las redes sociales. Para el Pentágono, las aplicaciones en dispositivos personales representan un «riesgo significativo» para el personal militar, por lo que esas capacidades deben desactivarse en ciertas áreas operativas.

Bajo la nueva orden, los líderes militares podrán determinar si las tropas bajo su mando pueden usar la función GPS en sus dispositivos, según la amenaza de seguridad en esa área o en esa base. «Estas capacidades de geolocalización pueden exponer información personal, ubicaciones, rutinas y números de personal del DOD (US Department of Defense), y potencialmente crear problemas de seguridad no deseados así como un mayor riesgo para la fuerza y la misión conjunta», establece el memorandum.

El personal de defensa que no se encuentre en áreas sensibles podrá usar las aplicaciones de GPS si los comandantes concluyen que no representan un riesgo. Por ejemplo, las tropas que se ejercitan en las principales bases militares de todo el país, como Fort Hood en Texas o la Estación Naval de Norfolk en Virginia, probablemente podrán usar el software de ubicación en sus teléfonos o dispositivos de entrenamiento físico. Sin embargo, las tropas que se encuentran en misiones en lugares más sensibles,

como Siria, Irak, Afganistán o algunas regiones de África, tendrían prohibido utilizar los dispositivos o verse obligadas a desactivar cualquier función de ubicación.

Las preocupaciones sobre los rastreadores de ejercicio y otros dispositivos electrónicos llegaron a un punto crítico en enero tras las revelaciones de que un mapa interactivo online, el Global Heat Map, identificaba las ubicaciones de las tropas, las bases y otras áreas sensibles de todo el mundo. El Pentágono señaló que las señales electrónicas podrían revelar la ubicación de las tropas que se encuentran en lugares secretos o clasificados o en pequeñas bases de operaciones avanzadas en áreas hostiles.

El Global Heat Map, publicado por la compañía de rastreo de GPS Strava, usó información satelital para mapear las ubicaciones de los suscriptores del servicio de fitness de Strava. En ese momento, el mapa mostraba actividad desde 2015 hasta septiembre de 2017. El nivel de detalle era implacable. En zonas como Estados Unidos o Europa, donde se encuentran la mayoría de los 27 millones de personas que tienen una pulsera Fitbit o similar, hay amplias zonas iluminadas. Pero en otras partes, como Irak, Siria o regiones de conflicto en África, había pequeñas manchas. Un «zoom» en el mapa permitía ver esas manchas de actividad con toda nitidez. Muchos se encontraban en los aledaños de las bases militares estadounidenses, como Kandahar (Afganistán), o en zonas de guerra como Irak o Siria, señalando la presencia de personal militar que utiliza pulseras o dispositivos de entrenamiento físico. Se podían ver otros puntos de luz, rodeando misteriosas pistas para aviones, que ponían en evidencia la posible ubicación de bases secretas. En la costa de Somalia, un país en conflicto desde hace más de veinticinco años, un periodista experto en el terrorismo informó en Twitter de que las pequeñas iluminaciones en el mapa eran las bases de los equipos de operaciones especiales en Sahel.

En un segundo memorándum, sobre el uso de teléfonos móviles y otros dispositivos electrónicos, el Pentágono estableció nuevas restricciones para el uso de teléfonos móviles y otros

dispositivos inalámbricos móviles dentro del propio Pentágono.

Ese memorándum insistía en las buenas prácticas de seguridad y en su aplicación más estricta que requiere que los teléfonos se dejen en cajetines con efecto de jaula de Faraday[170] fuera de las áreas seguras o donde se discuten asuntos delicados. No llegó a prohibir los dispositivos pero dejó claro que los teléfonos móviles aún pueden usarse en áreas comunes y otras oficinas en el Pentágono siempre que no haya información clasificada. Estas medidas incluyen funciones de GPS en rastreadores de ejercicios, teléfonos, tabletas, relojes inteligentes y otras aplicaciones.

Lo curioso de este asunto, es que fue el propio Departamento de Defensa quien había incentivado el uso de estos accesorios para aumentar el ejercicio en las bases. En 2013, según *The Washington Post*[171] el Pentágono distribuyó 2.500 de estas pulseras como parte de un proyecto para combatir la obesidad.

El sencillo dato de «dónde estamos» en combinación con otros como ser militar y estar en una zona de conflicto, puede suponer un problema de seguridad nacional y permitir que se organicen atentados con esta información que fluye libre en la sociedad de la apariencia. Ésta es la importancia de generar datos y de regalárselos a empresas de las que no sabemos, tan siquiera, si tienen una oficina física ni dónde está ubicada.

El poder de los datos que ya avizoraban lo dataístas no ha pasado desapercibido para las aseguradoras, las empresas especializadas en hacer cálculos complejos sobre futuros siniestros, y calcular cuánto tienen que cobrar para que resarcirlos cuando ocurran les salga rentable. Podemos decir sin exagerar que son los futurólogos del mundo empresarial.

En el Reino Unido ya hay compañías de seguros de riesgo de la conducción que han obligado a sus clientes a incorporar sensores y GPS a sus coches para analizar su conducción, cuántas horas pasa el coche aparcado, si está cuidado en un entorno que favorezca su mantenimiento, la presión de las ruedas, si hay distintas maneras de conducción lo que indicaría más de un conductor no asegurado,

etc., todo eso para reducir la incertidumbre y establecer con mayor precisión la tasa de siniestros pero, lo que es más importante, el coste de la prima que puede ser evolutivo o cambiante.

Una gran aseguradora obligará a usar medidores de actividad física. John Hancock le ofrecerá un seguro de vida.

John Hancock, una de las aseguradoras de vida más antiguas y más grandes de Estados Unidos dejará de vender el seguro de vida tradicional y, en su lugar, venderá sólo pólizas interactivas que rastreen la condición física y los datos de salud a través de dispositivos portátiles y teléfonos inteligentes.[172] La decisión de la aseguradora, con 156 años y propiedad de la canadiense Manulife Financial, ha marcado un cambio importante para la compañía, que dio a conocer su primera póliza de seguro de vida interactiva en 2015. Ahora está aplicando el modelo en toda su cobertura de vida. El seguro de vida interactivo, iniciado por el socio de John Hancock, Vitality Group, ya se venía empleando en Sudáfrica y Gran Bretaña y se está generalizando en Estados Unidos.

Los titulares de pólizas obtienen descuentos Premium por alcanzar objetivos de ejercicio que son monitorizados en sus dispositivos Fitbit o Apple Watch y obtienen tarjetas de regalo para tiendas y otros beneficios al registrar sus entrenamientos. En teoría, todos ganan: los asegurados estarían incentivados para adoptar hábitos saludables y las compañías de seguros cobrarían unas primas más ajustadas y pagarían menos siniestros o los postergarían si los clientes viven más tiempo y mejor.

Esto es la teoría, en un mundo idílico en el que todos los asegurados siguen la dieta mediterránea como dogma de fe. Pero sabemos que el ser humano no es así, se salta las dietas, corre en exceso para su edad y condición física, y se lesiona, o no se levanta del sillón en todo el fin de semana. O se excede en sus salidas de cañas con sus amigos.

Fotografiar langostinos y cervezas de manera recurrente en Instagram, subir los resultados del running a Twitter, regalarle a nuestra pulsera nuestro peso y medidas, junto con los parámetros

de sueño y vigilia, nos va a hacer acreedores de un salvaje aumento de la póliza del seguro médico que no vamos a poder discutir porque son decisiones opacas y no transparentes para el usuario. Mientras, permitimos (o no, eso es cuestión de que lo investigue la Agencia de Protección de Datos) a las aseguradoras usar los datos para seleccionar a los clientes más rentables, mientras que las primas aumentan para aquellos que no lo sean, siendo penalizados sin capacidad de discusión.

Internet de la emoción

Este seguimiento nos llevará, según el escenario 5 del Center for Long-Term Cybersecurity de la University de Berkeley, a una escala nueva de control: al internet de la emoción.[173] En 2020, los dispositivos portátiles no se preocuparán por la cantidad de pasos que demos; se preocuparán por nuestro estado emocional en tiempo real. Con dispositivos que rastrearán los niveles hormonales, la frecuencia cardíaca, las expresiones faciales, el tono de voz, internet pasará a convertirse en un lector de las emociones más íntimas de la psicología humana. Estas tecnologías permiten rastrear y manipular los estados mentales, emocionales y físicos subyacentes de las personas. Los ciberdelincuentes y los gobiernos hostiles encontrarán nuevas formas de explotar los datos sobre las emociones. Los términos de la ciberseguridad necesitarán ser redefinidos ya que la gestión y protección de la «imagen pública emocional» y «la apariencia de salud mental» (dos nuevos conceptos) se convertirán en otros activos a proteger.

Los mayores beneficios no se esperan de los hogares inteligentes y conectados, sino que lo que supondrá desde un punto de vista comercial y político serán los datos que se obtengan a través de una tecnología que mide cómo se sienten las personas: cómo invocan y experimentan los estados mentales, donde y cómo reconocer el amor, el odio, los celos, la ambición, el dominio, la

competitividad y otros aspectos emocionales humanos básicos. Aprenderemos la palabra *biosensing*, la intersección entre los indicadores físicos y las medidas de ondas cerebrales, y no hablaremos más que de ella como oportunidad de negocio. En este mundo, la ciberseguridad y la seguridad emocional se entrelazarán inextricablemente. Los ciberdelincuentes, las empresas y los gobiernos no sólo aprovecharán el seguimiento de las emociones humanas, sino que también comenzarán a manipular sutilmente esas emociones para obtener ganancias lícitas e ilícitas.

Este escenario representa un mundo en el que la detección emocional se convierte en una característica central de las tecnologías de internet. Ya tenemos aquí algunos adelantos: el movimiento del «Yo cuantificado», una tendencia aficionada a utilizar la tecnología para medir aspectos inesperados de la vida diaria; las «métricas personales» que ya permiten rastrear empíricamente patrones de comportamiento.

Cuando la temperatura corporal, la actividad de las ondas cerebrales, el seguimiento ocular y la dilatación de la pupila, la transpiración, los niveles endocrinos y de glucosa, los niveles altos de endorfina y otras variables se puedan medir a través de dispositivos portátiles y entre grupos de personas que están interactuando en un entorno particular, será posible el análisis de los sentimientos.

Esto marcará el rápido lanzamiento de un nuevo campo de investigación, que combinará aspectos de la psicología clínica y la informática y se centrará en los «afectos» individuales o impresiones superficiales del estado mental de un individuo. Piénsese en los esfuerzos actuales para deducir de los gestos familiares emociones o sentimientos; con las métricas personales mejoradas las empresas estudiarán la productividad y los patrones de desempeño con total precisión.

Los datos sobre los estados emocionales serán la clave que desbloquee el valor latente de los datos personales y profesionales ya recopilados. En otras palabras, el análisis en la intersección de

resultados internos (personales) y externos (ambientales) revelará detalles extraordinarios sobre cómo las personas se responden unas a otras y los estímulos de su entorno. Los investigadores podrán medir y registrar el panorama de las emociones humanas, sus condiciones, factores desencadenantes y efectos. El interés en ideas agregadas, el «internet emocional», comenzará a suplantar el interés en el afecto individual a medida que el análisis se haga más sofisticado. Las impresiones superficiales de las emociones de las personas ya no serán interesantes, ya que las emociones subyacentes pueden medirse a escala y con una precisión muy alta.

No sé si os he convencido en este capítulo anterior de ocultar vuestro rostro y poner de moda a los embozados del XIX de nuevo. Tal vez tras leerlo no encontréis tan necesaria la pulsera de entrenamiento y la dejéis morir en casa como yo maté a mi Tamagochi. Deberíais hacerlo si no queréis que el siguiente paso sea que os lean el pensamiento.

13

Dataveillance

Vigilancia privada y crédito social

Lacie está pendiente de agradar a todo el mundo. Con su coleta pelirroja y su flequillo cortado a tazón, puntúa cada pequeña interacción con los demás (un café, unos buenos días), intentando que su calificación no baje de cuatro estrellas sobre cinco. En un mundo de color pastel, todos viven con el miedo de recibir una mala calificación, con la angustia de que una respuesta inadecuada o el deseo de revancha provoquen una caída en su rating por debajo de las tres estrellas. Podría estar describiendo cualquier red social llena de likes, favoritos, corazones, retuits o llamas. Pero se trata del primer episodio de la tercera temporada de la serie británica *Black Mirror* «Nosediv», Caída en Picado. Este sistema de puntuación, inicialmente inocuo, pasa a ser usado como la medida de confiabilidad en cada uno de los ciudadanos, cayendo en peor fortuna al ritmo del hundimiento de sus estrellas, negándose las empresas privadas a prestar servicios a personas con menos de tres estrellas, alquilarles casas, llevarlas en un taxi, servirles una comida, convirtiéndose así en apestados sociales. Es como si el número de seguidores en Instagram o los «me gusta» de Twitter establecieran nuestra capacidad para acceder a un puesto de trabajo, comprar una casa o tener permiso para procrear.

Armonía y crédito social

Alguien que valorara la armonía social sobre todas las cosas aprobaría un sistema basado en el miedo y en la hipocresía: gente comportándose bien, aparentando felicidad, y llevando unas vidas penosas. Así es en buena medida el mundo del postureo, pero los chinos, que desean la armonía de un pueblo sometido a la dictadura del proletariado, no han dejado pasar la oportunidad de vigilar y controlar a todo un pueblo por voluntad propia implantando un nuevo sistema (Social Credit System, SCS), que permite evaluar la capacidad crediticia de sus ciudadanos, como ya hacen las entidades financieras occidentales con todos nosotros.

El sistema, ferozmente ambicioso, autoritario, tecnológicamente sofisticado y disruptivo, ha sido diseñado, según un documento oficial, como «un componente importante del sistema de economía de mercado socialista», ya que «sus requisitos inherentes fijan la idea de una cultura de sinceridad, lo que permite promover la sinceridad y las virtudes tradicionales». Esa redacción vaga pero con una gran carga de reproche (¿quién no ha escuchado cuando uno se queja de la vigilancia electrónica diciendo «yo es que no tengo nada que ocultar»?) en realidad habla de control de los ciudadanos: con la figura del «crédito social» las autoridades chinas planean hacer algo más que conocer la capacidad de endeudarse de sus ciudadanos, quieren evaluar la confiabilidad de los mismos en todas las facetas de su vida.

En la mayoría de los países, la existencia de un sistema de crédito no es controvertido. La información financiera pasada se usa para predecir si las personas pagarán sus hipotecas o la factura de la tarjeta de crédito en el futuro. Pero China está llevando todo el concepto unos pasos más allá al construir un sistema de «crédito social» omnipotente destinado a calificar la confiabilidad de cada ciudadano.

En 1956, el ingeniero eléctrico Bill Fair y el matemático Earl Isaac crearon una pequeña empresa de tecnología en un piso de San Francisco. La llamaron Fair, Isaac and Co., y acabó siendo conocido como FICO. Su principal innovación fue usar los ordenadores y el análisis estadístico para traducir los detalles personales y el historial financiero de la gente en una puntuación simple, un rating de solvencia, que predeciría la probabilidad de que los préstamos fueran devueltos. Antes de FICO, las agencias de crédito dependían de los chismes de los propietarios, vecinos y tenderos locales, de la posición social, raza, sexo o amaneramiento junto con las ausencias dominicales en la iglesia respectiva. Este sistema restringía el crédito a las clases pudientes, a los dueños de la moral y a los amigos de los banqueros locales, e impedía no sólo que tuvieran acceso al mismo quien tenía capacidad de pago sino al propio banco ampliar su base de clientes y hacer más negocio. La puntuación algorítmica, según Fair e Isaac, era una alternativa científica más equitativa a un realidad injusta. FICO se hizo popular entre las agencias de crédito (TransUnion, Experian y Equifax), y en 1989 introdujo la calificación crediticia que conocemos hoy, permitiendo a millones de estadounidenses conseguir hipotecas y vivir del crédito de las tarjetas de crédito.

Sesame Credit, la base del crédito social chino, se basa en gran medida en FICO.

¿De dónde salen los datos para alimentar el algoritmo que permite puntuar a los ciudadanos chinos? De una combinación de varios servicios online privados.

En 2015, el Gobierno chino comprobó que todo el mundo pagaba con sus teléfonos. En McDonald’s, la tienda de conveniencia, incluso en los restaurantes familiares. El efectivo, en retirada, se había ido reemplazando por dos aplicaciones para teléfonos inteligentes: Alipay y WeChat Pay.

Para darse de alta en Alipay sólo es necesario facilitar el número de su teléfono móvil y escanear la tarjeta de identificación nacional, que cuenta con buena reputación, evitaba tener que

enfrentarse con la tediosa tarea de ir a un banco administrado con indiferencia perezosa y sin servicio de atención al cliente. Además, es fácil y usable, casi divertido. Alipay permite pagar prácticamente todo y funciona, además, como una red social. Puedes pagar el seguro de automóvil, pedir citas para los médicos, añadir amigos, pagar las facturas de electricidad, gas e internet en la sección de «Servicio de la ciudad» de la aplicación, o el estacionamiento con My Car de Alipay, facilitando el carnet de conducir, el número de matrícula y el número de bastidor del coche.

Alipay propiedad de Ant Financial, filial de Alibaba, es una superaplicación. Su principal competidor, WeChat, pertenece al gigante social y de juegos Tencent. Alipay y WeChat no son aplicaciones al uso sino que funcionan como ecosistemas enteros, que permiten acceder a una plétora de servicios que pueden ser pagados desde la aplicación inicial. Comparándolo con ejemplos conocidos es como si Amazon hubiera engullido eBay, Apple News, Groupon, American Express, Citibank y YouTube, y pudiera extraer datos de todos estos servicios.

Zhima Credit (o Sesame Credit) «es la personificación del crédito personal», se puede leer en la pestaña que se abre en la aplicación de Alipay. Y continúa: «Usamos el big data para realizarte una evaluación objetiva. Cuanto más alta sea tu puntación, mejor puntuación tendrás y mejor será tu crédito». Sólo hay que clicar el siguiente botón para «comenzar mi viaje de crédito».[174]

El Banco Popular de China, el regulador de la banca central del país, mantiene registros sobre millones de consumidores, pero a menudo contienen poca o ninguna información. Hasta hace poco, era difícil obtener una tarjeta de crédito con cualquier otro banco por lo que el efectivo era la regla. El aumento de los precios de la vivienda hizo que el uso de efectivo fuera insostenible. Sin embargo, los esfuerzos para establecer un sistema de crédito confiable fracasaron porque China a finales de 2011 carecía de una entidad de calificación crediticia como FICO. En cambio, contaba con 356 millones de usuarios de teléfonos inteligentes. Ese año, Ant

Financial lanzó una versión de Alipay con un escáner incorporado para leer códigos QR. WeChat Pay, que se lanzó en 2013, tiene un escáner incorporado similar. Escanear un código QR es algo sencillo y, tanto hace las funciones de una tarjeta de embarque como te lleva a una web, como te abre una aplicación o te conecta con la página de Facebook o el perfil de Twitter de una persona.

A partir de ese momento, empezó un furor en China por los códigos QR: se incorporaban a las tumbas para que, quien tuviera interés, los escaneara y obtuviese más información sobre el fallecido, o en las camisas de los camareros, que se escaneaban para darles propina. Incluso los mendigos imprimieron códigos QR y los pusieron en la calle. Gracias a los QR, los pagos móviles de Alipay alcanzaron casi 70 mil millones de dólares en sólo un año.

En 2013, los ejecutivos de Ant Financial se retiraron a las montañas de Hangzhou para discutir la creación de una serie de nuevos productos, entre ellos, Zhima Credit. Se dieron cuenta de que podían recopilar datos de Alipay para calcular un rating de crédito basado en las actividades de un individuo. Ant Financial no fue la única entidad interesada en usar datos para medir el valor de las personas. Un año más tarde, vaya usted a saber si por casualidad, el Gobierno chino anunció su proyecto de «crédito social». En 2014, el Consejo de Estado, el gabinete del Gobierno de China, pidió públicamente el establecimiento de un sistema de seguimiento a nivel nacional para calificar la reputación de individuos, empresas, e incluso funcionarios del Gobierno. El objetivo era que todos los ciudadanos chinos tuvieran para el año 2020 un archivo con sus datos —accesible mediante huellas digitales y otras características biométricas—, recopilados de fuentes públicas y privadas. El Consejo de Estado lo denominó como un «sistema de crédito que cubre a toda la sociedad».[175] Para el Partido Comunista Chino, el crédito social es un intento de autoritarismo suave e invisible, porque, recordemos, la percepción de intrusión es distinta cuando jugueteas con una aplicación que cuando un señor con cuello mao llama a tu puerta, se sienta a tu

mesa, te observa mientras ves la tele, mira cómo duermes y te sigue al trabajo y allí donde te queden ganas de ir. Es mucho más sencillo espiar y controlar una población del tamaño de China con los aspectos coercitivos que se imponen los usuarios entre sí alentados desde el diseño del propio sistema.

En 2015, Ant Financial obtuvo, junto con otras ocho empresas tecnológicas, la aprobación del Banco Popular de China para desarrollar sus propias plataformas privadas de calificación crediticia. Zhima Credit, su sistema de rating crediticio, se incorporó a la aplicación de Alipay poco después.

Zhima Credit registra el comportamiento del usuario en la aplicación y otorga un rating entre 350 y 950. Los que tengan mejor puntuación, como si se tratase de una tarjeta de fidelización, reciben beneficios y recompensas. El algoritmo de Zhima Credit no sólo tiene en consideración si pagas tus deudas, sino también analiza qué compras, tus estudios y la puntuación de tus amigos. Con la misma filosofía que Fair e Isaac, pero con distinto subtexto, los ejecutivos de Ant Financial publicitaron que un enfoque basado en datos abriría el sistema financiero a las personas que habían sido bloqueadas, como los estudiantes y los habitantes de zonas rurales de China. Para los más de 200 millones de usuarios de Alipay, que han optado por Zhima Credit de manera voluntaria, las ventajas son claras y los perjuicios difusos y distantes. Como en todo lo que tiene que ver con los datos.

Participar en Zhima Credit es voluntario, y no está claro si su uso afecta a la calificación del sistema que el Gobierno está trabajando, aparentemente, en paralelo. Sin embargo, Ant Financial declaró en un comunicado de prensa en 2015 que la compañía planeaba «ayudar a construir un sistema de integridad social», lo que ya ha hecho al facilitar al Gobierno chino la lista negra de más de 6 millones de personas que han dejado de pagar sus multas. El cruce y la cesión de datos se da por descontado. ¿Quién le diría que no a un Gobierno de poder omnímodo?

El Consejo de Estado chino ha señalado que usando el sistema nacional de crédito social se penalizará a los que propaguen rumores online, entre otros comportamientos que se definan (imaginamos que todos ellos relacionados con libertades públicas y derechos fundamentales), y ha dejado claro que aquellos que sean considerados «muy poco dignos de confianza» deben prepararse para recibir servicios de calidad inferior. Ant Financial está en la misma línea. Lucy Peng, directora ejecutiva de la compañía, ha dejado claro que Zhima Credit «se asegurará de que las personas malas de la sociedad no tengan dónde ir, mientras que las personas buenas pueden moverse libremente y sin obstrucciones».[176]

El algoritmo detrás del rating de crédito Zhima es un secreto corporativo. Ant Financial enumera oficialmente cinco categorías amplias de información que se incorporan al rating, pero proporciona poco detalle de cómo se cocinan estos ingredientes juntos. Como cualquier sistema de calificación crediticia convencional, Zhima Credit controla el historial de gastos y si los préstamos se pagan a tiempo. Pero usa otras categorías que tienen más que ver con los sistemas que las empresas occidentales están usando ya para puntuar a sus clientes y usuarios:

– La categoría llamada «conexiones» tiene en consideración el crédito de los contactos en la red social de Alipay. El sistema te premia si todos tus amigos son personas de rating alto perjudicándote los que lo tengan malo.
– Otra tiene en cuenta el tipo de automóvil, y en ella se trabaja en qué escuela se estudió.
– La categoría «comportamiento» analiza los matices de la vida del consumidor, concentrándose en aquellas acciones que se correlacionan con un buen crédito. El director de tecnología de la compañía, Zhima Credit, Li Yingyun, dijo a la revista china *Caixin* que un comportamiento de gasto

como comprar pañales, por ejemplo, podría mejorar la puntuación, mientras que jugar con videojuegos durante horas podría reducirla.

Para comprender el atractivo que la ingeniería social tiene para los líderes chinos, hay que retroceder a los años posteriores a la Revolución Comunista de 1949. El Gobierno asignó a todos a unidades de trabajo locales, que se convirtieron en el lugar de vigilancia y control; unos controladores que espiaban a sus vecinos mientras hacían todo lo posible por evitar entrar en la lista negra. Mantener el sistema requería, como en el caso de la Stassi de la Alemania del Este, un esfuerzo y supervisión masivos del Estado. A medida que las reformas económicas en la década de 1980 llevaron a millones de personas a abandonar sus aldeas y emigrar a las ciudades, el sistema de unidades de trabajo se vino abajo. La migración también tuvo un efecto secundario: ciudades llenas de extraños. No pasó mucho tiempo hasta que el Gobierno consideró «gamificar» el comportamiento afecto al partido. A fines de la década de 1990, un grupo de trabajo en un instituto de la Academia de Ciencias de China desarrolló los conceptos básicos detrás del sistema de crédito social. Pero la tecnología no estaba lo suficientemente madura como para seguir el ritmo del diseño político.

En 2010, ensayaron el sistema de crédito social en Suining. Los funcionarios evaluaban a los ciudadanos conforme a una serie de criterios, incluido el nivel de educación, el comportamiento online y cómo de cumplidores eran con las leyes locales. A cada uno de los 1,1 millones de ciudadanos mayores de catorce años de Suining se les otorgó 1.000 puntos de partida y se establecieron puntos por tareas o por conductas indeseadas: el cuidado de los mayores de la familia, 50 puntos; ayudar a los pobres, 10 puntos; ayudar a los pobres con cobertura mediática, 15 puntos; conducir borracho, 50 puntos menos; sobornar a un funcionario, otros 50 puntos menos. La suma de los puntos obtenidos se repartía en cuatro categorías,

siendo A la mejor y D la peor. Los ciudadanos A tendrían prioridad de acceso en la educación, mientras que a los ciudadanos D se les negarían licencias, permisos y acceso a algunos servicios sociales.

El sistema de Suining era rudimentario, y provocó un breve debate nacional sobre qué criterios deberían incluirse en una puntuación de crédito social, pero proporcionó un campo de pruebas para lo que podría funcionar a nivel nacional.

Desde el piloto de Suining, docenas de ciudades han desarrollado sus propios sistemas con la ayuda de la tecnología que ya entonces acompañaba en el proceso. Todos estos pilotos se integrarán en el sistema de crédito social del Gobierno nacional. Para homogeneizar las bases de datos y traducir los sistemas de rating, el Gobierno chino ha contratado a Baidu, que le ayudará a desarrollar la base de datos de crédito social para la fecha límite de 2020.

Los chinos, que han sido calificados como poco confiables, ya están sufriendo en sus propias carnes el sistema de crédito social unificado. Si uno está en la lista negra del Tribunal Supremo Popular (la lista de personas deshonestas que es, curiosamente, la misma de Zhima Crédito) tiene prohibido el uso de la mayoría de los medios de viaje; sólo puede reservar las clases más bajas en los trenes más lentos; tiene limitado el acceso a ciertos bienes de consumo, completamente vedado el acceso a los hoteles de lujo, y no podrá obtener préstamos bancarios de cuantía elevada. Uno puede entrar en la lista negra por no pagar una multa fijada por un tribunal (que bien puede ser por haber escrito algo que no ha agradado al régimen) o por haberse equivocado en la cuenta de ingreso.[177]

La puntuación algorítmica está en todas partes

Cuando nos cuentan cómo funciona el sistema de crédito chino, miramos con la superioridad del que se cree que está en una democracia representativa. Pensamos que es el tipo de comportamiento que se esperaría del Gobierno autoritario chino pero que es algo, sin embargo, que no nos va a alcanzar. Pero pocos son los algoritmos, datos y lógica del sistema que usa Alipay o Zhima Credit que sean exclusivos de China. El sistema, que a cualquier demócrata preocuparía muy seriamente, no se diferencia mucho de los algoritmos, sistemas, plataformas o herramientas que usamos sin criterio y que nos puntúan y nos clasifican en todos los momentos de nuestra vida. Todos nosotros estamos siendo categorizados y juzgados por algoritmos similares, tanto por las empresas con las que tratamos como por nuestros gobiernos. Mientras que el objetivo de estos sistemas podría no ser el control social, a menudo es su subproducto. Se ha hecho muy difícil volar por debajo del radar. Y todo apunta que cada vez será más difícil sin concitar toda serie de sospechas.

El juicio humano está siendo reemplazado por algoritmos automáticos, y eso trae consigo enormes beneficios y riesgos. La tecnología está permitiendo una nueva forma de control social, a veces deliberadamente y otras veces como efecto secundario. Y a medida que el internet de las cosas inicia una era de más sensores, más datos, y más algoritmos, debemos asegurarnos de que obtenemos los beneficios y evitamos los daños.

La mayoría de nosotros, sin saberlo, es puntuado conforme diversos parámetros, muchos de ellos extraídos de métricas de comportamiento y demográficas similares a las utilizadas por Zhima Credit, tanto por compañías como por las administraciones quienes, a pesar de las exigentes obligaciones que la legislación europea de protección de datos establece, no da a los usuarios la oportunidad de optar por no participar.

Publico mis pensamientos y sentimientos en Twitter, hago mis compras en Amazon, me compro el traje de novia por eBay, viajo con Cabify, y en algunas estancias he usado Airbnb. En todos estos

servicios alguien me puntúa, el sistema infiere de mis comportamientos un rating que me permite seguir usando el servicio en determinadas condiciones. Y tengo suerte por ser europea y saber que lo que hagan las empresas será bajo el riesgo de recibir una sanción de entre el 2 y 4 por ciento de su cifra de negocios anual mundial. Aun así, no puedo estar segura de que empresas verticales que prestan muchos servicios o empresas que colaboren entre sí, se «presten» las puntuaciones y comparen ratings. Lo que es cierto es que ni en Estados Unidos ni en Europa se cuenta aún con una aplicación o servicio como el chino que permita compilar estas puntuaciones, pero si nos inducen a usar de manera cruzada servicios y a consentir legalmente que se intercambien nuestros datos, incluidas nuestras puntuaciones, no está lejos el día en que tengamos una nota única como consumidor/ciudadano.

Richard Stallman, presidente de la Free Software Foundation al igual que Bruce Sneier, considera que «la vigilancia que se nos impone hoy es peor que en la Unión Soviética. Necesitamos leyes para evitar que se recopilen estos datos en primer lugar».[178]

Entiende Stallman que la vigilancia es más amplia que la de la URSS en tanto se extiende a todos los sistemas de vigilancia y a todos los servicios que recibimos en conexión y en movilidad; y más profunda, porque requiere pasar de regular el uso de datos a regular la acumulación de datos. Debido a que la vigilancia es tan generalizada, restablecer la privacidad es necesariamente un gran cambio y requiere medidas poderosas en su opinión. Hay tantas maneras de usar datos para perjudicar a cualquiera de nosotros que la única base de datos segura es la que nunca se recopiló. Por tanto, en lugar del enfoque de la UE de regular principalmente cómo se pueden usar los datos personales (en su Reglamento General de Protección de Datos), Stallman propone una ley para evitar que los sistemas recopilen datos personales.

La forma robusta de hacerlo —y que no puede dejarse de lado por el capricho de un Gobierno—, en su opinión, sería exigir que se construyan sistemas para no recopilar datos sobre una persona. El

principio básico sería que un sistema debe estar diseñado para no recopilar ciertos datos, si su función básica puede llevarse a cabo sin esos datos.

Propone volver a la programación de la escasez, en la que los datos que generaban los sistemas eran los estrictamente necesarios para su funcionamiento. E incluso ésos, si ponen en riesgo la privacidad de los ciudadanos, habría que borrarlos cuando ya no fueran necesarios.

El mejor dato es, pues, el que no se recoge. Porque si se recoge, puede ser pirateado.

14

Data breach

La seguridad de los datos

Nos hemos empeñado tanto en decir que los datos son el nuevo petróleo, el oro del siglo XXI, el nuevo dinero, que ya estaban tardando en venir a robárnoslos.

No hacemos más que recibir avisos de empresas, siempre estadounidenses —las únicas obligadas, hasta ahora, a notificárnoslo— animándonos a cambiar nuestras claves porque han sufrido una pérdida de datos o algún tipo inespecífico de incidente. Es la manera amable de comunicarnos que la peor gente de la Tierra ha puesto sus manos sobre nuestra vida, que ha sufrido un ataque de proporciones bíblicas y que los atacantes se han llevado todo o parte de la base de datos de clientes. Nos animan a poner sistemas de doble autenticación y a no usar las mismas claves para todos los servicios, cuando la realidad es que en los ataques no se roban perfiles individuales, sino que se descargan cientos de miles de ellos. Y que eso no habría pasado si no se hubiesen empeñado en almacenar tantos datos, o si hubieran protegido tanto empeño con las medidas técnicas y organizativas de seguridad adecuadas.

Yahoo, LinkedIn, Equifax, Target, Facebook, Google, Twitter... todos han sufrido un ataque y todos han perdido datos que acaban siendo vendidos en forma de paquetes de tarjetas de crédito, o de acceso a los perfiles pirateados y, por tanto, a los permisos dados

por los usuarios a sus teléfonos, micrófonos, cámaras. Cuando consentimos que las empresas que nos prestan servicios sin mediar pago en dinero, no sólo troyanicen nuestras vidas sino nuestros dispositivos, estamos dejando la puerta abierta a grupos organizados de delincuentes informáticos que entran en nuestra casa y en nuestras vidas.

En el internet profundo, el acceso a las cámaras de los móviles o los portátiles se vende por una media de entre 5 a 15 dólares dependiendo de si lo que se ve al otro lado es el dormitorio de una adolescente y de cuantas veces se haya revendido el acceso.

Los piratas no se conforman ya con hacer cargos indebidos a una tarjeta de crédito o con hacer chantaje usando lo que encuentran. Construyen una vida con los datos que roban y los ponen a comprar, pedir tarjetas, dar de altas líneas de móvil, abrir cuentas corrientes en bancos que no requieren una gestión presencial. Hacen ingeniería social con ellos. Secuestran nuestro doppelgänger y lo ponen a trabajar.

Todos los días desayunamos con noticias como éstas: «La aplicación MyFitnessPal fue víctima de un ataque informático que ha filtrado los datos de sus 150 millones de usuarios» o «La app de citas Grindr compartió los datos del VIH de sus usuarios».

Veamos los ataques más famosos y sus consecuencias. Adelantamos que no es que las empresas estadounidenses tengan peor seguridad que sus homólogas europeas, sino que están obligadas legalmente a hacer públicos los incidentes de seguridad que involucren datos de clientes y de comunicárselo a ellos personalmente. Esta obligación no existía en Europa hasta la entrada en vigor del Reglamento General de Protección de Datos el 25 de mayo de 2018 y de la Directiva NIS.

Target

Target, el gigante minorista estadounidense, fue pirateado entre el Blackfriday del 27 de noviembre de 2013 y el 15 de diciembre de ese mismo año. El incidente, probablemente, se debió a un ataque al sistema de punto de venta (POS), mediante la inyección de código malicioso, que hacía que los datos que se introducían en las cajas se transmitieran, en tiempo real, a los piratas que las habían infectado.[179]

Los piratas informáticos robaron datos de hasta 40 millones de tarjetas de crédito y débito de compradores que habrían visitado sus tiendas durante la temporada de vacaciones navideñas de 2013. La investigación, dirigida por los fiscales generales de Connecticut e Illinois, había descubierto que los ciberatacantes habían accedido a un gateway (enlace entre el servidor del Target y proveedores terceros) mediante credenciales robadas a un proveedor externo.

En mayo de 2017, casi cuatro años después, Target llegó a un acuerdo de 18,5 millones de dólares para resolver las reclamaciones y demandas que había recibido de 47 estados y finalizar la investigación multiestatal iniciada a raíz del ataque.

Junto con esta indemnización menor, en comparación con la gravedad del ataque y de la cuenta de resultados de Target, ésta habría gastado 202 millones de dólares en todas las cuestiones relativas al ataque. Este coste se reparte en indemnizaciones a clientes, gastos de comité de crisis, comunicaciones, abogados y pleitos.

Equifax

El 7 de septiembre de 2017, Equifax reveló que sus bases de datos habían estado comprometidas durante meses. Es importante resaltar que Equifax es la empresa que guarda la información y rating crediticio de todos los estadounidenses merecedores del mismo. En total, la información de más de 148 millones de personas

se vio comprometida en un ataque cuya duración es incapaz de precisar públicamente. La compañía esperó seis semanas para revelarlo públicamente.

La información comprometida iba desde los datos de la tarjeta de crédito, el número del carnet de conducir y de la seguridad social (que en Estados Unidos hacen las veces de documento de identidad) hasta la fecha de nacimiento, número de teléfono y dirección de correo electrónico de las personas de su base de datos.[180]

La Oficina de Contabilidad General de Estados Unidos (GAO, por sus siglas en inglés) publicó en septiembre de 2018 un informe exhaustivo en el que examinó la violación masiva de datos sufrido por Equifax. En él se resumen los errores de Equifax, en gran parte relacionados con la falta del mínimo uso de las mejores prácticas de seguridad conocidas, así como la falta de controles internos y revisiones rutinarias de seguridad.

El informe de GAO confirmaba que el ataque se realizó sobre un servidor web conectado a internet con software desactualizado y que duró, al menos, 76 días en los que Equifax no se enteró de que estaba siendo pirateada. Los atacantes realizaron 9.000 consultas que pasaron desapercibidas debido a un fallo del mantenimiento de los controles de red (IDS) que llevaban diez meses sin funcionar. Por si esto fuera poco, una vez que ganaron el acceso, los atacantes se encontraron una base de datos que contenía credenciales sin cifrar y que se usaban para acceder a otras bases de datos internas.

Consumer Union, editores de «Don’t Let Equifax Put Americans At Risk Again»,[181] se queja de que un ataque de la profundidad y extensión como el de Equifax no haya dado lugar a cambios sustanciales en las empresas que tratan datos personales masivos: «Los estadounidenses permanecen en gran medida en la oscuridad sobre las prácticas de la industria de informes crediticios y, en general, en gran medida son incapaces de controlar el uso que se

hace de su información personal. La propia Equifax ha sufrido mínimas consecuencias y continúa haciendo negocios más o menos igual que antes».

La compañía anunció, coincidiendo con el informe de GAO, que había presupuestado 200 millones de dólares adicionales para seguridad y tecnología, pero no dio más información sobre lo que había gastado hasta el momento o en qué se lo iba a gastar. Habida cuenta del estado lamentable de su seguridad detectado por GAO, puede que esos 200 millones sean los primeros que se gasten en muchos años.

Ocho supervisores bancarios estatales están detrás de Equifax para asegurarse de que mejore su seguridad y la audite.

California aprobó una ley a principios de 2018 que obliga a revelar información sobre la recopilación de datos personales e impone multas significativas por pérdida de los mismos de hasta 750 dólares por pérdida. Queremos creer que se trata de por cada usuario afectado. Lamentablemente, esta ley no entrará en vigor hasta el 1 de enero de 2020. Alabama y Dakota del Norte, por su parte, aprobaron leyes que obligan a la notificación y que sancionan en caso de que la misma no se produzca o se produzca tardíamente. Los plazos son generosos: en Alabama, 60 días, y en Dakota del Norte, 45.

A nivel federal, hay un proyecto presentado en mayo de 2018 que permite a un particular «congelar» y «descongelar» su perfil de crédito en las tres agencias más grandes de Estados Unidos: TransUnion, Experian y Equifax. La congelación impediría el acceso a un archivo de crédito, lo que disuadiría a los ladrones de identidad de abrir cuentas nuevas a nombre de alguien que ha congelado su perfil. La misma norma permitiría extender la alerta sobre un perfil comprometido de 90 días a un año. La agencia que reciba la petición ha de trasladarla a las otras dos, y todas deberán adoptar controles reforzados y tomar medidas adicionales para verificar la identidad.

TeenSafe

La aplicación móvil para iOS y Android, TeenSafe, pensada para que los padres pudieran controlar de manera segura a sus hijos, sufrió un ataque a mediados de 2018, resultando no ser tan segura ni para unos ni para otros.

La aplicación permitía a los padres ver los mensajes de texto y la ubicación de sus hijos, monitorizar a quién estaban llamando y cuándo, acceder a su historial de navegación web y comprobar qué aplicaciones habían instalado en su móvil.

El ataque habría comprometido al menos a un servidor en donde se almacenaban las direcciones de correo electrónico de ID de Apple de los adolescentes y las contraseñas en texto plano, esto es, sin ningún tipo de cifrado o protección.[182] Debido a que la aplicación requería que la autenticación de dos factores estuviese deshabilitada, con la clave se podía acceder a la cuenta del menor, a sus datos y contenido personal.

En el servidor se habrían almacenado también el nombre del dispositivo del niño, que a menudo suele ser su verdadero nombre, el identificador único de su dispositivo y los mensajes de error asociados con una acción de cuenta fallida. La compañía mantiene que en el servidor no había datos de contenido, tales como fotos o mensajes, ni la ubicación de los padres o los hijos.

El ataque habría afectado al millón de padres y a sus respectivos hijos que usaban el servicio.

Facebook

Por si Facebook no había tenido suficiente con el escándalo de Cambridge Analytica, las fake news y los bots rusos, cerró el año 2018 con un ataque informático que comprometió entre 50 y 90 millones de cuentas. El ataque, continuado, se llevaba produciendo

desde hacía 14 meses. En la lista de damnificados se contarían los perfiles de su CEO y fundador, Mark Zuckerberg, y el de Sheryl Sandberg, directora de operaciones de la red social.[183]

¿Cómo entraron los atacantes? El formulario para felicitar el cumpleaños en un perfil permitía (por error) subir un vídeo. Al colgar el clip, en ese formulario se podía ver el token (clave de acceso) para un usuario distinto que permitía acceder a ese perfil usando la herramienta «ver cómo» (que te deja observar cómo otro usuario ve tu perfil). A partir de ese momento, se tenía acceso completo a ese usuario de Facebook.

Los tokens de acceso son cadenas únicas de números que se pueden usar para identificar personas, aplicaciones o páginas en Facebook. Una vez se inicia sesión en una cuenta de Facebook, se crea un token de acceso que confirma la identidad en el dispositivo. Este mismo sistema se usa para también para aplicaciones y servicios como Tinder que permiten a los usuarios darse de alta o autenticarse usando su usuario y clave de acceso a Facebook.

Hay que decir que ese token es lo que se usa para tener, por ejemplo, continuamente abierto el perfil de Facebook desde una aplicación móvil, sin tener que loguearse cada vez que se abre. Un apunte al margen: la insistencia de muchos servicios de que usemos su aplicación en lugar de la web se debe, precisamente a que, gracias a este mecanismo, nos tienen conectados todo el tiempo y en todo lugar, con acceso a los permisos activados, que suelen ser los de acceso a la cámara, al micrófono, a los archivos y las fotos.

El número de perfiles afectados se ha calculado en atención al número de sesiones en móviles que Facebook cerró el día que se supo del ataque, pero podría ser mayor. Dado que el sistema de token se utiliza para otros servicios de la red social, como Instagram y WhatsApp, los perfiles de estas redes también podrían verse afectados.

El ataque a Facebook en su *annus horribilis* ha creado varios efectos: no sólo ha dejado expuesta la vida de millones de personas que literalmente viven en Facebook sino que, además, ha abierto la

puerta a su vida a otros servicios. Con el pirateo de las cuentas de millones de usuarios de Facebook se han dejado expuestos otros miles de accesos a servicios: aquellos en los que los usuarios —por mera pereza o porque quieren federar su identidad (algo, me temo nada probable)— se logueen con sus usuarios de Facebook. Los tokens de acceso permiten acceder a sitios web de terceros que Facebook usa para iniciar sesión. La red social introdujo su característica de «inicio de sesión único» o Facebook Connect en 2010, y es ampliamente utilizada por aplicaciones como Tinder, Spotify o Airbnb.

Este mecanismo, además de ser inseguro, no sale gratis al usuario. Conecta los datos de la plataforma a la que se accede con Facebook, que facilita un puñado de datos del prestador del cliente y, a cambio, tiene acceso a la actividad de su usuario en todos los servicios a los que se conecte con su identificador de Facebook. Y todo ello, como se dice en el argot, de manera transparente para el usuario, es decir, sin que sepa que este intercambio existe y sin que lo consienta, en la mayor parte de los casos.

Es esto, y no los 90 millones de cuentas comprometidas, lo que hace que el ataque a Facebook sea particularmente grave.

Éste es el motivo del consejo, tan repetido y tan desoído, de tener claves distintas para cada servicio o al menos, entre los perfiles «recreativos» y las cosas importantes (bancos, medios de pago, servicios de salud, etc.).

Para rematar el año 2018, Facebook podría enfrentarse a causa de este incidente a una multa récord por virtud del Reglamento europeo, que le sería de aplicación en cuanto a los perfiles de las personas afectadas que se encontrasen en el territorio de la Unión en algún momento del ataque. Algunos calculan que ascendería a aproximadamente 1.400 millones de euros que podrían quedar reducidos si Facebook colabora con las autoridades europeas y toma medidas drásticas e inmediatas.[184]

Google

Google, que fabrica sus propios servidores para asegurarse la seguridad de los mismos, tampoco ha quedado al margen del mal de nuestro tiempo, si bien el ataque ha afectado a un servicio por el que no parece tener mucho afecto: la red social de Google+.

El ataque habría ocurrido entre 2015 y marzo de 2018, cuando los investigadores internos lo descubrieron, y habría comprometido 500.000 perfiles que, en palabras de algún malicioso, serían todos los usuarios que habría tenido jamás la fallida red social lanzada en 2011 para competir con Facebook.

Sin embargo, la dirección de Google decidió no comunicar hasta octubre de 2018 la existencia ni la extensión del ataque. La falta de transparencia de Google y la tardanza en publicar el incidente se debería a su coincidencia con la entrada en vigor del RGPD en Europa. Supongo que la preocupación de que se quisiere dar un escarmiento con ellos y el consiguiente daño en su reputación les llevaría a posponer una publicación que, curiosamente, coincidió en el tiempo con el pirateo de las cuentas de Facebook.[185]

Poco más se sabe del ataque, habida cuenta de la opacidad con la que opera la compañía.

Entre las medidas para mitigarlo, Google anunció el cierre permanente de todas las funciones de Google+, lo que supone el cierre *de facto* de un producto que ha dado para tantos chistes.

Cuánto cuesta una fuga de datos

«La seguridad total no existe» es un mantra repetido hasta la saciedad por compañías que tratan datos para justificar, en muchos casos, dejadez, falta de inversión y presión de las áreas de negocio para relajar los controles a favor de la productividad. Sin embargo, si

la seguridad en la protección de los datos es una preocupación, las inversiones se notan y desaniman a los piratas, que tienden a atacar sitios con menos controles y más rendimiento.

Un gran poder comporta una gran responsabilidad. Hemos visto cómo las empresas tecnológicas atesoran datos como el avaro Scrooge dickensiano adora su dinero, pero no ponen la misma diligencia en proteger los datos de los que viven, ni parece preocuparles en exceso el daño que una filtración de datos puede causar a una persona cualquiera.

Y es, simple y llanamente, porque hasta fecha muy reciente no ha habido incentivo.

Los ataques son algo que las empresas no comparten a no ser que sean tan graves que acaben saliendo de los límites de su control. En un análisis de riesgo y de coste-beneficio, se calcula la posibilidad de que el ataque sea conocido, de que el regulador se entere, del coste de la sanción en caso de que decida sancionarles, y el coste de indemnizar a los afectados. Todo eso dividido por el tiempo transcurrido.

En general, la seguridad es cara y es un tostón: ralentiza los procesos y muchos la viven como una limitación a la agilidad del negocio. Además, si observamos los ejemplos anteriores, el coste después de publicarse el incidente puede ser menor que lo que hubiera costado implantar y mantener las medidas de seguridad necesarias para evitarlo.

Un total de 242,7 millones de dólares es lo que Equifax se ha gastado desde el incidente, un ataque que expuso información personal y financiera confidencial de casi 148 millones de consumidores. Todo por no cifrar y dejar la información de los consumidores en claro.[186] Esta cantidad, algo menos de 2 dólares por perfil atacado, y comparada con la promesa de invertir 200 millones en seguridad, parece de risa. Y aun así es una cantidad importante si la comparamos con la del ataque a Target: en sólo siete meses, Equifax ha gastado casi lo que Target gastó (252 millones de dólares) en dos años después de su incidente de 2013.

Durante muchos años, los analistas y los profesionales de la seguridad han tratado de calcular cuánto puede costarle a una empresa un ataque con exposición de los datos de sus clientes y usuarios. Hay que tener en consideración el coste de tener que actualizar la infraestructura y la seguridad de TI, pagar los honorarios legales, las multas del regulador y otros tangibles e intangibles que pesan en la cuenta de resultados. Está el impacto en el precio de las acciones, la erosión de la confianza de los clientes —sobre todo si se trata de servicios financieros— o la caída de los ejecutivos involucrados. El presidente y CEO de Target renunció meses después del incidente, y el CEO de Equifax hizo lo mismo a las pocas semanas de saberse el ataque.

Se han realizado muchos estudios para calcular el coste de los incidentes de fuga de datos pero nos quedaremos con los resultados del último informe anual de Ponemon. Patrocinado por IBM Security y llevado a cabo por Ponemon Institute, el «Estudio 2018» estableció que el coste promedio de una brecha de datos a nivel mundial estaría en unos 3,86 millones, lo que supone un aumento de un 6,4 por ciento comparado con el informe de 2017.[187] En Estados Unidos, el coste es casi el doble, esto es, unos 7,35 millones, lo que está muy lejos del coste de los grandes incidentes como Target o Equifax.

Pero ¿estos informes de investigación miden realmente lo que le costará una brecha de datos a una empresa? Equiparar los daños a un coste «por registro» hace que las violaciones de datos sean sólo un ejercicio actuarial de riesgo aceptable. Lo que no anima a gastar en seguridad.

Hay cuatro actividades o centros de coste que identifica el Instituto Ponemon relacionadas con el proceso de gestión de un incidente y sus gastos asociados:

1. Detección y escalado. Aquí se incluyen las actividades que permiten a una empresa detectar y reportar la violación de seguridad al personal apropiado dentro de un determinado período de tiempo, como, por ejemplo, las actividades forenses y de

investigación; los servicios de evaluación y auditoría; la gestión de equipos de crisis, y las comunicaciones a la dirección ejecutiva y el consejo de administración.

2. Costes de notificación a las personas afectadas por la brecha de datos, como por ejemplo, remisión de correos electrónicos, cartas, llamadas telefónicas o, en general, cualquier notificación sobre la información personal que se ha perdido o robado; y la comunicación al regulador.

3. Respuesta posterior a la brecha de datos y, en concreto, la ayuda a los afectados por la fuga de datos para resolver los problemas que les haya causado como, por ejemplo, servicios de seguimiento para saber si se ha usado la información en perjuicio del afectado, servicios de protección de identidad o de apertura de nuevas cuentas o emisión de tarjetas de crédito, los gastos legales en los que hayan incurrido así como posibles descuentos en los productos o servicios de la empresa.

4. Pérdida de ingresos por pérdida de confianza de los clientes, pérdidas por la paralización de los sistemas informáticos, así como los costes de pérdida de clientes y los gastos necesarios para recuperarlo o conseguir nuevos clientes. También se incluye el cálculo de la pérdida de negocio por la afectación de la reputación de la compañía.

Sufrir una pérdida de datos de clientes conlleva no sólo los costes indicados sino, además, un riesgo legal importante. Las multas, que para las grandes tecnológicas pueden ser molestas pero no trágicas, son suficientemente potentes como para cerrar muchas startup, sobre todo aquellas que basan su modelo de negocio en la recogida intensiva de datos y en las rondas de financiación.

El problema evidente de las cuantías sancionatorias es que las empresas vean más ventajas en ocultar las brechas de datos que en ponerlas honestamente en conocimiento de sus clientes. Si la propia Google ha tardado varios meses en admitir la brecha de Google+, es de esperar que, como ya ocurre, no haya incentivo para las

tecnológicas con menos músculo para informar a los afectados, que se verán doblemente dañados por la fuga de sus datos y por su uso en manos de delincuentes.

Esa idea de «recoge todos los datos que ya veremos qué hacemos con ellos, en algún momento, en el futuro, para algún negocio que aún ni siquiera se nos ha ocurrido» tiene que ir acompañada, en compensación, de un coste proporcional en seguridad. Y los costes de seguridad son muy elevados. De ahí que en el mundo de «cuantos más datos mejor», sea conveniente reflexionar desde el negocio si todos los datos que se recaban son necesarios y útiles, y si no sería más barato y generaría menos riesgo recoger menos datos y con más criterio.

Por tanto, como usuarios deberíamos esperar, junto con un comportamiento ético y transparente de las empresas en cuanto a nuestros datos, unas medidas de seguridad adecuadas para que no acaben en las manos equivocadas.

Dicho esto, no nos dejemos llevar por el mensaje de las tecnológicas intensivas de datos, de que la responsabilidad es del usuario que debería leer las condiciones legales de privacidad y aguantar con ellas si las ha aceptado.

Como veremos en el capítulo siguiente, «he leído y aceptado las condiciones legales y de privacidad» es la gran mentira de internet.

15

La gran mentira

Los problemas del consentimiento

En 1936, dos profesores del MIT, Robert F. Elder y Louis Woodruff, inventaron el audímetro. Según la definición dada por la Wikipedia,[188] este artilugio, también denominado «*people meter*» («medidor de personas», en inglés) es un aparato que se conecta a algunos televisores y mide la audiencia de manera permanente y automática, usando sus datos con fines estadísticos de los que se extrapolan los datos de audiencia de los distintos canales de televisión. En un primer momento, la función de este aparato era realizar mediciones del número de aparatos de radio encendidos y registrar la emisora que sintonizaban. A tal fin, se conectaban al dial de la radio y grababan en un rollo de papel los datos obtenidos.

Los años pasaron, Nielsen se hizo con el control de las audiencias de radio y televisión e inventó algunos artefactos de medición como un audímetro con forma de cojín que se activaba cada vez que alguien se sentaba encima, que, como era de esperar, daba muchos falsos positivos. En la actualidad parece que hay 4.755 audímetros en toda España y que no es panelista (a quien se le adjudica uno) quien quiere sino quien puede. Sólo basta una visita a la web de la empresa española de medición, Kantar Media, para

comprobar que no se puede pedir un audímetro, sino que se adjudican en base a estudios estadísticos de la población para que la muestra sea homogénea.

Durante años, he buscado desesperadamente a los propietarios de la figura mitológica del audímetro sin ningún éxito: he preguntado a todo con el que he tenido una reunión, tomado un café, compartido asiento en el metro; he usado las conferencias que he dado para preguntar a la audiencia si tenían un audímetro o si conocían a alguien que lo tuviese. Nada. Nadie parecía tener un audímetro ni conocer a nadie que lo tuviera.

Con las condiciones de privacidad, cláusulas de consentimiento y legalidades similares, pasa lo mismo: por mucho que pregunto nadie se las ha leído antes de haberlas aceptado. Ni mis amigos, ni los amigos de mis amigos, ni yo misma, hemos leído las condiciones de privacidad de los servicios y productos que contratamos online.

Toda una vida me llevaría leerte

Una respuesta simple a nuestros problemas de privacidad sería que todos los que hacemos uso de las mismas nos informáramos al máximo sobre la cantidad y calidad de los datos que se recaban, almacenan, ceden, analizan y agregan sobre nosotros cada vez que usamos un servicio web, instalamos un programa o descargamos una aplicación. Como diligentes usuarios, tendríamos que leer todas las políticas de privacidad de esos servicios, programas o aplicaciones.

No es ningún secreto que, a menudo, se marca esa casilla sin leer el aviso de privacidad. De hecho, la renuencia a leer los avisos de privacidad se encuentra bien documentada. En este sentido resulta revelador el trabajo realizado por la Universidad de Berkeley[189] en el que plantearon un ejercicio para determinar si la presencia de un link a una política legal de privacidad en una web o en la descarga de una app tenía alguna influencia significativa en la

voluntad de los usuarios para facilitar información personal. El estudio partía de la base de que los usuarios no son conscientes de la elección que realizan y de las consecuencias e implicaciones que para su vida privada puede tener. El resultado fue demoledor: comprobaron que ninguno de los participantes hizo clic en el link que llevaba a las condiciones de privacidad.

Esto es especialmente dramático si tenemos en cuenta que nuestro sistema de protección de datos se sustenta, esencialmente, alrededor del consentimiento informado, que supone que el usuario acepta libremente, tras haber sido adecuadamente informado, las consecuencias de su aceptación. Informar al interesado previamente a recabar sus datos es una construcción irreprochable desde un punto de vista jurídico. El problema es que las empresas y sus abogados (*mea culpa*) no están especialmente interesados en que se entienda lo que se lee, sino que la compañía que recaba los datos lo haga cumpliendo con la legalidad estricta, y se cumple con ella poniendo a disposición del usuario las condiciones legales como paso previo a contratar el servicio. Una vez que hace clic, a nadie le preocupa si el texto era incomprensible para los pobres mortales, ni a éstos lo que se vaya a hacer con sus datos mientras no se le cobre por el servicio.

La elección sobre si facilitar o no datos personales (el consentimiento informado) parte de la premisa de que los usuarios somos seres racionales capaces de leer en profundidad documentos legales para evaluar las prácticas de privacidad y protección de datos de una web o una aplicación móvil. Múltiples estudios, además del mencionado, ya han confirmado que la mayoría de nosotros sabemos que estas condiciones legales se leen muy raramente;[190] que leer las políticas de privacidad de cada web que se visita nos llevaría una cantidad de tiempo poco razonable,[191] y que están, a menudo, escritas en un lenguaje tan complejo que se alejan de la compresión lectora de la mayor parte de los usuarios.

Sin embargo, sería erróneo inferir que las personas son indiferentes a cómo se utiliza su información o que el requisito de proporcionar información de privacidad es irrelevante o inaplicable; la actitud de la gente hacia el uso de sus datos es más compleja de lo que un enfoque simplista podría sugerir. Más bien, el problema radica en el formato de muchos avisos de privacidad. Es comprensible que las personas no estén dispuestas a leer los avisos de privacidad que son largos y están escritos en términos legales incomprensibles o destinados a proteger a la organización en lugar de informar al titular de los datos.

La gente no lee los avisos legales. Es un hecho. Son demasiado largos, tediosos y complejos, y una barrera que los aleja del deseado acceso al servicio, a la aplicación o al dispositivo al que muchos quieren acceder de manera inmediata y sin mayor dilación. Muchos, o por lo que indican los estudios, casi todos, consideran que no vale la pena leer contratos adhesivos que no pueden negociar y en los que no se determina un sistema claro y sencillo para establecer las condiciones de privacidad de manera independiente. Así marcar la casilla de «he leído y acepto los términos y condiciones» parece haberse convertido en una de las mayores mentiras de internet.[192]

Las políticas de privacidad y protección de datos son documentos legales escritos por abogados para abogados y no para usuarios finales. Una encuesta de 2009 determinó que la mayoría de los participantes creían que las políticas de privacidad y protección de datos protegían sus derechos, cuando, en realidad, están orientadas a proteger al responsable del tratamiento, que asegura que la recogida de datos y su cesión son conformes a la normativa de protección de datos. De hecho, si uno lee la extensión de muchas cláusulas de recogida del consentimiento o del ámbito en el que se permite que las cesiones se produzcan, aun cumpliendo la ley escrupulosamente, suponen *de facto* una invasión de la intimidad de sus usuarios.[193]

En 2008, dos investigadoras, Lorrie Faith Cranor y Aleecia McDonald, decidieron calcular cuánto tiempo les llevaría leer cada política de privacidad de los servicios que usaban.[194] Hicieron un sencillo cálculo: si la longitud media de una política de privacidad de los 75 sitios web principales en 2008 era de unas 2.514 palabras, a una tasa de lectura estándar según la literatura académica de aproximadamente 250 palabras por minuto, cada política de privacidad individual debería de llevar a un individuo acostumbrado a leer con frecuencia textos largos una media de diez minutos. A continuación, identificaron los sitios web —cada uno con su propia política de privacidad— visitados por el estadounidense promedio y, aunque la estimación no fue sencilla (usaron varias fuentes, incluidos sus propios servicios y e encuestas) obtuvieron un rango de entre 1.354 y 1.518 de servicios web, lo que suponía una media de 1.462 webs con sus consiguientes políticas de privacidad. Conforme a este cálculo, llegaron a la conclusión de que si cada usuario de internet tuviera que leer cada política de privacidad en cada sitio web que visitase, pasaría 25 días al año sólo leyendo las políticas de privacidad. Si su trabajo consistiera en leer políticas de privacidad con una jornada laboral de ocho horas diarias, le llevaría 76 días hábiles completar la tarea, lo que llevado a números nacionales supondrían 53.800.000 horas de tiempo requerido en leer estas políticas.

En Gran Bretaña, el Comité de Ciencia y Tecnología de la Cámara de los Comunes llegó a la misma conclusión: evidenció que la lectura de todos los términos y condiciones que se encuentran en internet nos llevaría, al menos, un mes de cada año de vida.[195] Sólo para usar Amazon Echo, el dispositivo IoT que incorpora el asistente virtual de Amazon, Alexa, hay que leer 17[196] contratos de unas 94 páginas que hay que revisar con frecuencia, porque los cambios que se hagan afectan al aceptante como si los hubiese leído o aceptado.

Para poner en contexto esta enorme cantidad de tiempo, Cantor y McDonald siguieron algunos procedimientos estándar que la literatura económica sugiere que podrían usarse para calcular el

coste de oportunidad de las tareas, resultando que el coste de leer las políticas de privacidad en Estados Unidos en el año 2008 era superior al PIB de Florida de ese mismo período. La de Florida es la cuarta economía estatal más grande de Estados Unidos.

Teniendo en cuenta que el trabajo se realizó en 2008, es más que probable que la cantidad de servicios con sus condiciones legales sea mayor a fecha de hoy, no sólo por el aumento de la población de internet en Estados Unidos en los últimos 10 años, sino por la incorporación de la movilidad, la hiperconectividad y la multiplicación de servicios. Para Cantor y McDonald, el peso colectivo es tan grande que nadie puede mantener una relación responsable con sus propios datos.

En Estados Unidos el informe de la Casa Blanca sobre grandes volúmenes de datos[197] observó el fenómeno de la «fatiga de la privacidad»: a pesar de que se facilitaba información sobre el tratamiento, pocos usuarios la leían o la entendían.

Hasta tal punto de complejidad y falta de tiempo hemos llegado, que se ha desarrollado una IA que lee las políticas de privacidad: Polisis,[198] un marco automatizado de análisis de políticas de privacidad alimentado por IA,[199] en cuyo centro hay un modelo lingüístico sobre privacidad, construido con políticas de privacidad de 130.000 sitios web y una novedosa jerarquía de clasificadores de redes neuronales que atiende los aspectos de alto nivel y los puntos detallados de las prácticas de privacidad.[200] Polisis va acompañado de un bot gratuito, PriBot, que responde a las preguntas sobre las políticas de privacidad.

La ineficacia del consentimiento como método de recogida legal de datos es especialmente significativo en un contexto de grandes datos o de nuevas tecnologías como las móviles, los dispositivos IoT o los sistemas ciberfísicos, en los que el consentimiento explícito según requiere el Reglamento General de Protección de Datos Europeo va a tener que ser la manera de legitimar el tratamiento. Esto puede ser problemático, ya que los avisos de privacidad «clásicos» no son factibles en ninguno de estos entornos. Además,

el tipo de análisis de caja negra que se da en el caso de los grandes datos es demasiado difícil de explicar en términos que los usuarios puedan comprender, sobre todo si tenemos en cuenta que el análisis de grandes datos a menudo implica la reutilización de los mismos. En este supuesto, ni el propio responsable del tratamiento puede prever, desde el principio, todos los usos que se pueden hacer de los datos obtenidos.

Parecería, por tanto, que un sistema basado en la recogida de consentimiento pueda no estar protegiendo los derechos fundamentales a la intimidad y a la protección de datos de los individuos, o que la manera en que se presenta la información previa al consentimiento no se hace para el mejor interés del usuario.

Como constructo legal, es irreprochable basar en la voluntad del individuo todo el sistema de protección. Es una pena que quien pensó que el consentimiento era la solución no viese las ventajas que da a quien cumple haciendo una extracción salvaje de datos, ni los problemas para recogerlo con todas las garantías en entornos IoT o en los que el resultado del análisis es desconocido (como en el caso del machine learning no supervisado).

Una vez que el lector ha pasado en diagonal por este capítulo puede marcar sin rubor:

☐ He leído y entendido este libro

16

Nada que ocultar

Coda

Una fiscal de la Audiencia Nacional, en una comida privada entre amigos, llama «nenaza» a un juez, «maricón» a otro, y manifiesta que no le gustan nada los juzgados formados por mujeres. Prefiere trabajar con tíos. Está en confianza. Aparentemente. Uno de los asistentes a la comida, miembro de un cuerpo policial, graba la conversación porque tiene esa mala costumbre.

Al cabo de los años, esa fiscal es nombrada ministra de justicia en el primer Gobierno de la democracia española con mayoría femenina y el policía, en la cárcel por sus chanchullos (en los que mucho tiene que ver su manía de grabarlo todo), monta un medio de comunicación online y publica los audios de ese lejano encuentro de 2009.

Ninguno de nosotros podría salir airoso si tuviéramos un comisario pegado permanentemente, grabando lo que decimos, los mensajes que mandamos, sacándonos fotos de donde nos encontramos, debiéramos estar ahí o no, observando nuestro comportamiento al milímetro. El móvil de casi ninguno de nosotros aguantaría el escrutinio ajeno. Sólo hay que abrir el WhatsApp, consultar el chat con la persona o amigos de mayor confianza, y repasar los chistes o los comentarios sobre terceros hechos en la confianza de que no serían leídos por esos terceros. Porque,

cuando lo hacemos, esperamos que se mantengan dentro de la esfera en la que lo hicimos, tenemos una expectativa de privacidad y confidencialidad sobre lo dicho.

Alguien podría mantener que 2009 es mucho tiempo, y que es improbable que nadie tenga memoria de lo dicho o hecho diez años atrás. Esa percepción es errónea. Porque todos llevamos un «comisario político» en el bolsillo que registra cada segundo de nuestra vida con total precisión y que lo guarda a perpetuidad. Las dimisiones a costa de un mal tuit son sólo la punta del iceberg de lo que una filtración de nuestros datos podría suponer para cualquiera de nosotros. Sería el final de asociaciones, de negocios, de matrimonios y familias, si nuestros familiares, socios o amigos supieran lo que pensamos de ellos.

Por eso no deja de sorprenderme que, cuando prevengo a alguien de los riesgos que para la privacidad tienen las redes sociales, o del mal uso o errónea configuración de sus dispositivos móviles llenos de aplicaciones que troyanizan su vida privada, la respuesta que recibo en casi todos los casos es «no tengo nada que ocultar». Esta afirmación, tan protestante, no sólo es incierta, sino que ha sido el eslogan de muchas dictaduras y de alguna limpieza étnica o religiosa, como parte que es de ese brocardo tan inquisitorial de «si nada temes, nada tienes que ocultar» que, en el imaginario colectivo, asociamos por tradición y cultura, con los pecados de la carne y los comportamientos heréticos. En muchas de esas respuestas, de orgullo protestante, se esconde la convicción de no llevar una vida turbia, llena de sexo extraño y drogas ilegales, y de que el interlocutor que tanto se empeña en proteger su intimidad es porque su existencia no soporta la luz, los taquígrafos ni la memoria. El querer mantener tu vida fuera de la mirada ajena es fuente de sospecha. Cuando, en un ejercicio inútil, advierto de la afectación que para su vida y su futuro tiene el uso indiscriminado y sin sentido de servicios, programas y aplicaciones

gratuitas, mis interlocutores parecen pensar que mi prevención se debe a mis muchos vicios, vida desordenada o sexo escandaloso. Que oculto porque temo.

El lector que cree no tener ningún vicio cristianamente reprochable, que piense en la fiscal-ministra y considere si esa conversación de WhatsApp en la que pone a escurrir a un compañero de trabajo, o a su jefe, le gustaría que llegara al correo electrónico de éste; que cuando está luchando por un ascenso, se filtrase que busca información sobre benzodiacepinas para soportar la presión, o que mira el horóscopo del día para saber si lo va a hacer bien o mal en la mesa de operaciones.

Porque la intimidad alcanza desde mi estado de ánimo hasta mis miedos más profundos. Nuestros miedos, anhelos y arrepentimientos nos representan más que cualquier opción sexual o adicción, y es, precisamente, lo que menos deseamos que los demás conozcan. El análisis de datos que la tecnología actual permite pone al descubierto nuestro verdadero yo como objeto de manipulación y mercantilización, pero no traslada sus resultados a nuestro círculo cercano, se queda en una nube lejana que parece no tener impacto negativo en nuestra vida diaria. Más al contrario. Mientras que el efecto real y negativo es imperceptible, el usuario sólo percibe de manera muy potente las ventajas de usar un servicio inteligente, sencillo y gratuito. Ser Casandra siempre ha supuesto pontificar en el desierto.

De esta ausencia de percepción de peligro, surge la segunda afirmación más repetida: «¿Quién va a querer mis datos? No soy tan importante». En una combinación de vagancia adolescente y sensación de beneficio en el uso del servicio, aplicación o herramienta, los usuarios son impermeables a la evidencia de que los servicios que reciben tienen un coste enorme y que no se pagan sólo con servirles publicidad (que por otro lado se encargan de bloquear con diligencia).

Y es así cómo cala el mantra de que la privacidad ha muerto cuando, curiosamente, sólo beneficia a quien hace su negocio de enterrarla.

Pero ¿qué es eso de la privacidad?

En diciembre de 1890, la *Harvard Law Review*[201] publicó un artículo titulado «El derecho a la vida privada» en el que, nada más empezar, planteó cómo los seres humanos evolucionamos en el apetito de nuevos derechos según vamos cubriendo las necesidades más básicas, modelándolos sobre necesidades más sofisticadas. En una época en que se moría tras tener un furúnculo infectado, después de dormir en la misma cama que tus padres o por compartir jergón con tus hermanos y las bestias, no era la mayor de las preocupaciones. Pongo este ejemplo porque es uno de los más usados por los *Privacy Killers*, que nos venden un mundo poético y pastoril en donde un ser humano, más simple y «natural», no necesitaba intimidad, ni penicilina, ni derecho al voto o libertad individual. Sin embargo, el bienestar se da en todos los ámbitos de la vida y no se puede pretender tener derecho a voto al mismo tiempo que convences de que volver al «agua va» en términos de derecho a la intimidad es un gran avance.

Cuando la miniaturización llegó a las radios portátiles, los adolescentes de los años 50 ganaron su independencia de la sala y la radio común pudiendo escuchar lo que quisieran sin las restricciones paternas en sus propias habitaciones. La búsqueda de la independencia, del propio espacio y los gustos que nos acompañarán en la edad adulta, es una evolución social que va unida a la necesidad de que se reconozca ese derecho a estar en soledad y a no tener que compartir con otros, que pueden burlarse o impedir el acceso a la música o la información que nos dé la gana.

No deja de sorprender, por tanto, la modernidad de la reflexión de Warren y Brandeis, autores del trabajo, que escriben esta pieza a finales del siglo XIX. Ya entonces, les preocupaba que las innovaciones técnicas y comerciales, afectaran al derecho de los seres humanos a ser dejados en paz. Porque de eso va la

privacidad: de tener el derecho a ser dejado en paz.[202] Decían los autores del trabajo que «las invenciones recientes y los métodos comerciales llaman la atención sobre el próximo paso que debe darse para la protección de la persona y para asegurar al individuo lo que el juez Cooley llama el derecho "a ser dejado en paz"».

Mucho de lo que susurramos en las redes sociales acaba en el tejado de la toma de decisiones sobre nosotros. Un estudio de ProPublica descubrió que quince empleadores, incluido Uber, habrían anunciado puestos de trabajo en Facebook exclusivamente para un sólo sexo conforme los clásicos estereotipos de género:[203] hombres como conductores de Uber[204] o como Policía del estado de Pensilvania. Una empresa de camiones con sede en Míchigan publicó anuncios que apuntaban no sólo a hombres, sino a hombres interesados en el fútbol estadounidense universitario. Y un centro de salud comunitario en Idaho buscó enfermeras y asistentes médicos certificados, y limitó su audiencia a las mujeres. El Tribunal Supremo de Estados Unidos estableció en 1973 que es ilegal que un empleador publique anuncios de trabajo en los periódicos con parámetros tales como: «Se busca ayuda: sólo hombres».[205]

La orientación por sexo es apenas una de las formas en que Facebook y otras compañías de tecnología permiten que los anunciantes puedan centrar sus búsquedas en ciertos usuarios y excluyan a otros. Esos datos que damos porque no tenemos nada que ocultar y porque a quién le vamos a importar tanto como para que se tomen la molestia, convierten a Facebook en la plataforma favorita de los anunciantes, quienes, gracias a su enorme capacidad de conocer hasta el último átomo de nuestro cuerpo y mente, pueden gastar sus presupuestos de marketing o de buscador de personal sólo en aquellos a los que quieren llegar, eligiendo de manera casi eugenésica a sus candidatos.[206]

Esa capacidad de granularizar sus esfuerzos proporciona a los anunciantes el poder de elegir y discriminar. La propia ProPublica anunció en 2016 que Facebook permitía a los anunciantes excluir a los usuarios por razón de la raza,[207] y en 2017 detalló cómo los

anuncios de empleo en Facebook permitían excluir a los candidatos de mayor edad.[208] Desde que ProPublica sacara a la luz estas prácticas, Facebook eliminó la posibilidad de que los anunciantes excluyesen ciertas categorías de personas por raza, religión, problemas sociales y origen nacional. Sin embargo, Facebook no ha hecho cambios similares por edad o sexo.[209]

Un informe interno de los ejecutivos de Facebook, revelado por el diario australiano *The Australian*,[210] afirma que la compañía es capaz de monitorizar las publicaciones y fotos de sus usuarios con la finalidad de determinar su comportamiento y estado de ánimo. Esto no resulta nuevo ni impactante. Ya sabemos que el análisis de comportamiento es un básico en los análisis que llevan a efecto los grandes prestadores de servicios sin precio (que gratuitos no son). La alarma saltó cuando se supo que estos esfuerzos de análisis se dirigían a determinar si los más jóvenes se sentían «estresados», «derrotados», «inútiles», «inseguros», «sin valor», «abrumados», «ansiosos», «nerviosos», «estúpidos», «tontos» o si «necesitaban un impulso de confianza». Todos estos estados de ánimo se mostraban a los anunciantes para que pudieran personalizar la publicidad y servírsela a estos atribulados jóvenes ajustada a su lamentable espíritu. El informe fue elaborado para uno de los principales bancos de Australia, y en éste se informaba de que la red social contaba con una base de datos de jóvenes usuarios australianos que ofrecer conforme a estos parámetros. 1,9 millones de estudiantes de secundaria, 1,5 millones de estudiantes técnicos y 3 millones de trabajadores jóvenes fueron sujetos al escrutinio de sus sentimientos.

Este informe, según Facebook, se basó en «investigaciones realizadas por Facebook y posteriormente compartidas con un anunciante» y estaban «destinadas a ayudar a los profesionales de marketing a comprender cómo se expresan las personas». Facebook se negó a descartar si se habían realizado investigaciones similares sobre la vulnerabilidad emocional de los adolescentes para publicidades en mercados fuera de Australia.

Según *The Australian*, los datos que se ponían a disposición de los anunciantes incluirían el estado de la relación de los jóvenes, su ubicación, el número de amigos en la plataforma y la frecuencia con la que accedían al sitio desde dispositivos móviles o desde la web. El periódico informó que Facebook también tenía datos sobre los usuarios que hablaban sobre temas tales como «verse bien y con confianza corporal» o «hacer ejercicio y perder peso». Facebook es capaz de analizar de manera profunda cómo los usuarios representan sus emociones y las comunican visualmente, con lo que no hace un análisis ramplón de frases de desánimo, sino que analiza fotos de hojas otoñales o canciones melancólicas para determinar si tenemos un mal día o una depresión acusada. El informe afirmaba que Facebook era capaz de determinar la evolución de las emociones a lo largo de una semana. Así, resultaba más probable que las emociones anticipadas se expresasen a principios de semana, mientras que las emociones reflexivas aumentasen durante el fin de semana. Según estos parámetros, de lunes a jueves construimos confianza y el fin de semana contamos lo estupendos que somos.

La filtración de este controvertido informe llegó tres años después de que la compañía se enfrentara a una reacción negativa significativa después de publicar los resultados de un vasto experimento en el que manipuló la información publicada en las páginas de inicio de 689.000 usuarios, y descubrió que podía cambiar los sentimientos de las personas a través de un proceso de «contagio emocional». La compañía llevó a cabo la investigación en secreto, algo de dudosa ética, dado que su filtrado expuso a propósito a «contenido emocional negativo» a sus usuarios.

Tal vez por esto, por el escándalo ruso o por la filtración de Cambridge Analytica, los estadounidenses estarían cambiando su actitud hacia Facebook. Algo más de la mitad de sus usuarios por encima de los dieciocho años (54 por ciento) manifiestan haber cambiado sus ajustes de privacidad en los últimos 12 meses, según un reciente informe de *Pew Research Center*. Un 42 por ciento

dicen haberse cogido unas vacaciones de Facebook y no haberlo consultado en un período de varias semanas, mientras que un 26 por ciento asegura haber borrado la app de Facebook de sus teléfonos inteligentes. Según el estudio, un 74 por ciento de los usuarios de Facebook habría tomado alguna de estas tres decisiones en el último año. Es un comienzo.[211]

Así que la próxima vez que decida apoyar una campaña en Facebook, quejarse del tiempo o expresar su estado de ánimo en Twitter, subir una selfie a Instagram o compartir canciones melancólicas en Spotify, piénselo. Aunque no tenga nada que ocultar.

Bibliografía

Acquisti, Alessandro, Ralph Gross, y Fred Stutzman. 2014. «Face Recognition and Privacy in the Age of Augmented Reality.» *Journal of Privacy and Confidentiality*. <https://doi.org/10.29012/jpc.v6i2.638>.

Acquisti, J. Grossklags. «Privacy and rationality in individual decision making.» IEEE Secur. Priv., 3 (1) (2005), pp. 26-33.

Acquisti, A. «Privacy in electronic commerce and the economics of immediate gratification.» *EC '04 de la 5th ACM Conference on Electronic Commerce*. 2004.

«Advances in AI are used to spot signs of sexuality.» *The Economist*, septiembre de 2017. <https://www.economist.com/science-and-technology/2017/09/09/advances-in-ai-are-used-to-spot-signs-of-sexuality>.

Aleecia M. Mcdonald y Lorrie Faith Cranor. «The Cost of Reading Privacy Policies. I/S: A Journal of Law and Policy for the Information Society», vol. 4, n.º 3 (2008), pp. 543-568. *Ohio State University. Moritz College of Law*. <https://kb.osu.edu/bitstream/handle/1811/72839/ISJLP_V4N3_543.pdf>.

A. L. Young, A. *Quan-Haase. Privacy protection strategies on Facebook*. Inf. Commun. Soc., 16 (4) (2013), pp. 479-500.

«An executive's guide to AI.» *McKinsey Analytics*. <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/an-executives-guide-to-ai>.

Angwin, Julia; y Parris, Terry. «Facebook lets advertisers exclude users by race.» *ProPublica*, 28 de octubre de 2016. <https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race>.

Angwin, Julia (*ProPublica*); Scheiber, Noam (*The New York Times*); y Tobin, Ariana (*ProPublica*). «Dozens of companies are using Facebook to exclude older workers from job ads.» *ProPublica*, diciembre de 2017. <https://www.propublica.org/article/facebook-ads-age-discrimination-targeting>.

A. N. Joinson, U.-D. Reips, T. Buchanan, C. B. «Paine Schofield. Privacy, trust, and self-disclosure online. Hum.-Comput.» *Interact*, 25 (2010), pp. 1-24.

«Are your phones listening to you?» *Pent Test Partners*. <https://www.pentestpartners.com/security-blog/are-your-phones-listening-to-you>.

Aytes, Ayhan. *Return of the Crowds: Mechanical Turk and Neoliberal States of Exception. Digital Labor: The Internet as Playground and Factory*. Ed. Routledge. 2012.

Barnes, S. B. «A privacy paradox: Social networking in the United States.» *First Monday*, 11(9). <http://firstmonday.org/article/view/1394/1312>. 2006.

Barth, Susanne; y De Jong, Menno D. T. «The privacy paradox — Investigating discrepancies between expressed privacy concerns and actual online behavior — A systematic literature review.» *Elseviers. Telematics and Informatics*, vol. 34, n.º 7, noviembre de 2017, pp. 1038-1058 <https://www.sciencedirect.com/science/article/abs/pii/S0736585317302022>.

Bell, Karissa. «Google+ isn't dead and these are the people astill using it the most.» *Mashable*, enero de 2017. <https://mashable.com/2017/01/18/who-is-using-google-plus-anyway/#_lP0PXr.eiqZ>.

Bergen, Mark; y Surane, Jennifer. «Google and Mastercard Cut a Secret Ad Deal to Track Retail Sales.» *Bloomberg*, agosto de 2018. <https://www.bloomberg.com/news/articles/2018-08-30/google-and-mastercard-cut-a-secret-ad-deal-to-track-retail-sales>.

Biddle, Sam. «Long-Secret Stingray Manuals Detail How Police Can Spy on Phones.» *The Intercept*, septiembre de 2016. <https://theintercept.com/2016/09/12/long-secret-stingray-manuals-detail-how-police-can-spy-on-phones/>.

Big data, artificial intelligence, machine learning and data protection. 20170904 Version: 2.2. ICO (Information Commissioner's Office).

Biersdorfer, J. D. «Getting social with Waze.» *The New York Times*, septiembre de 2017. <https://www.nytimes.com/2017/09/20/technology/personaltech/getting-social-with-waze.html>.

Borges, Jorge Luis. *Funes el memorioso*. Ed. Ficciones. 1944.

Brandom, Russell. «Lawsuit challenging Facebook's facial recognition system moves forward.» *The Verge*, 5 de mayo de 2016. <https://www.theverge.com/2016/5/5/11605068/facebook-photo-tagging-lawsuit-biometric-privacy>.

Brandom, Russell. «The Facebook hack could be Europe's first big online privacy battle. The 50 million compromised accounts are the first major test of the GDPR», octubre de 2018. <https://www.theverge.com/2018/10/1/17922946/facebook-breach-gdpr-lawsuit-privacy-commissioner-europe>.

Bruell, Alexandra. «Ad Fraud Declines Offer Hope as Marketers Fight Sophisticated Bots.» *The Wall Street Journal*, mayo de 2017. <https://www.wsj.com/articles/ad-fraud-declines-offer-hope-as-marketers-fight-sophisticated-bots-1495645098>.

Buck, C., Horbel, C., Germelmann, C. C., Eymann, T. *The unconscious app consumer: Discovering and comparing the information-seeking patterns among mobile application consumers*. Twenty Second European Conference on Information Systems, Tel Aviv, Israel, pp. 1-14, 2004.

Busby, Mattha. «Social media copies gambling methods "to create psychological cravings".» *The Guardian*, mayo de 2018 <https://www.theguardian.com/technology/2018/may/08/social-media-copies-gambling-methods-to-create-psychological-cravings>.

Burgess, Matt. «Here's what you need to do after the huge Facebook hack.» *Wired*, septiembre de 2018. <https://www.wired.co.uk/article/facebook-hack-data-breach-news-what-to-do>.

C. Flender, G. Müller. «Type indeterminacy in privacy decisions: the privacy paradox revisited.» J. Busemeyer, F. Dubois, A. Lambert-Mogiliansky, M. Melucci (eds.), Quantum interaction. Lecture notes in Computer Science, SpringerVerlag, Berlín, Heidelberg (2012), pp. 148-159.

Chad Couser in Enterprise Security. «The Cost of a Data Breach.» *Gemalto*, abril de 2018 <https://blog.gemalto.com/security/2018/04/30/data-breach-costs/>.

Cherry, Kendra. «What Is a Schedule of Reinforcement? What impact do schedules of reinforcement have on learning?» *VeryWellMind*, diciembre de 2017. <https://www.verywellmind.com/what-is-a-schedule-of-reinforcement-2794864>.

Citton, Yves. *The Ecology of Attention*. Ed. Cambridge Polity. 2017.

Coen, Rena; King, Jennifer; y Wong, Richmond, *The Privacy Policy Paradox*. UC Berkeley School of Information.

Cook, James. «Amazon patents new Alexa feature that knows when you're ill and offers you medicine.» *Telegraph*, octubre de 2018. <https://www.telegraph.co.uk/technology/2018/10/09/amazon-patents-new-alexa-feature-knows-offers-medicine/>.

«Cybersecurity Futures 2020.» *Center for Long-Term Cybersecurity University of California, Berkeley*. 2016. <https://cltc.berkeley.edu/scenarios/>.

Dans, Enrique. «Asistentes de voz, espionaje y la navaja de Occam.» <https://www.enriquedans.com/2018/05/asistentes-de-voz-espionaje-y-la-navaja-de-occam.html>.

«Data workers of the world, unite. What if people were paid for their data?» *The Economist*. <https://www.economist.com/the-world-if/2018/07/07/what-if-people-were-paid-for-their-data>.

Del Rey, Jason. «Here's Amazon's explanation for the Alexa eavesdropping scandal.» *Recode*, mayo de 2018. <https://www.recode.net/2018/5/24/17391480/amazon-alexa-woman-secret-recording-echo-explanation>.

Elola, Joseba. «Nos hemos quedado con tu cara.» *El País*, enero de 2018.

Executive Office of the President. «Big data: seizing opportunities, preserving values.» Casa Blanca, mayo de 2014. <https://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_5.1.14_final_print.pdf>.

Eyal, Nir. *Hooked: How to Build Habit-Forming Products*. Ed. Ryan Hoover. 2014.

Farivar, Cyrus. «Chicago man sues Facebook over facial recognition use in "Tag Suggestions".» *Arstechina*, agosto de 2015. <https://arstechnica.com/tech-policy/2015/04/chicago-man-sues-facebook-over-facial-recognition-use-in-tag-suggestions/>.

Fakhoury, Hanni. «When a Secretive Stingray Cell Phone Tracking "Warrant" Isn't a Warrant.» *EFF*, marzo de 2013. <https://www.eff.org/homepage-feature/stingray>.

Fleishman, Glenn. «Equifax Data Breach, One Year Later: Obvious Errors and No Real Changes, New Report Says.» *Fortune*, septiembre de 2018. <http://fortune.com/2018/09/07/equifax-data-breach-one-year-anniversary/>.

Fox, Maggie. «Drug giant Glaxo teams up with DNA testing company 23andMe. DNA test results will be used to develop drugs under a new agreement.» *NBC*, julio de 2018. <https://www.nbcnews.com/health/health-news/drug-giant-glaxo-teams-dna-testing-company-23andme-n894531>.

Fowler, Geoffrey A. «Apple is sharing your face with apps. That's a new privacy worry.» *The Washington Post*, noviembre de 2017. <https://www.washingtonpost.com/news/the-switch/wp/2017/11/30/apple-is-sharing-your-face-with-apps-thats-a-new-privacy-worry/>.

«FTC Complaint, request for investigation, injunction, and other relief submitted.» *The Electronic Privacy Information Center*. <https://epic.org/privacy/ftc/google/EPIC-FTC-Google-Purchase-Tracking-Complaint.pdf>.

Ghorayshi, Azeen; y Ray, Sri. «Grindr Is Letting Other Companies See User HIV Status And Location Data.» *Buzzfeed*, abril de 2018. <http://www.buzzfeed.com/azeenghorayshi/grindr-hiv-status-privacy>.

González Cabañas, José; Cuevas, Ángel y Rubén. «Facebook Use of Sensitive Data for Advertising in Europe.» Universidad Carlos III de Madrid.

«Hablar con la máquina. La tecnología de reconocimiento de voz transforma nuestro modo de relacionarnos con los dispositivos.» *El País*, enero de 2018. <https://elpais.com/tecnologia/2018/01/05/actualidad/1515173663_103498.html>.

Hasbrouck, Edward. «CBP expands partnership with airlines on facial recognition.» *Papers, please! The identity project*. <https://papersplease.org/wp/2018/08/28/cbp-expands-partnership-with-airlines-on-facial-recognition/>.

Harkous1, Hamza; Fawaz2, Kassem; Lebret1, Rémi; Schaub3, Florian; Shin3, Kang G.; y Aberer1, Karl. «Polisis: Automated Analysis and Presentation of Privacy Policies Using Deep Learning». (1 Ecole Polytechnique Fédérale de Lausanne (EPFL); 2 University of Wisconsin-Madison; 3 University of Michigan.) <https://pribot.org/files/Polisis_Technical_Report.pdf>.

Harris, Tristan. *Ensayos*. <http://www.tristanharris.com/essays/>.

Harris, Tristan. «How Technology Hijacks People's Mindsfrom a Magician and Google's Design Ethicist.» *Tristan Harris*, mayo de 2016. <http://www.tristanharris.com/2016/05/how-technology-hijacks-peoples-minds%E2%80%8A-%E2%80%8Afrom-a-magician-and-googles-design-ethicist/>.

Harari, Yuval Noah. *Homo Deus. Breve historia del mañana*. Ed. Debate. 2016.

Hauser, Christine. «Police Use Fitbit Data to Charge 90-YearOld Man in Stepdaughter's Killing.» *The New York Times*, octubre de 2018. <https://www.nytimes.com/2018/10/03/us/fitbit-murder-arrest.html>.

House of Commons Science and Technology Committee. «Responsible use of data. HC245.» *Publicatons Parliament*, noviembre de 2014. <http://www.publications.parliament.uk/pa/cm201415/cmselect/cmsctech/245/245.pdf>.

Hughes-Roberts, T., 2012. *A cross-disciplined approach to exploring the privacy paradox: explaining disclosure behaviour using the theory of planned behavior*. UK Academy for Information Systems Conference Proceedings, Paper 7.

Hvistendahl, Mara. «Inside China's Vast New Experiment in Social Ranking.» *Wired*, diciembre de 2017. <https://www.wired.com/story/age-of-social-credit/>

I. Oomen, I. Leenes. *Privacy risk perceptions and privacy protection strategies*. E. de Leeuw, S. Fischer-Hübner, J. Tseng, J. Borking (eds.), *Policies and Research in Identity Management*, Springer Verlag, Boston (2008), pp. 121-138.

Jackson, Sean. «Big data in big numbers - it's time to forget the 'three Vs' and look at real-world figures.» *Computing*, febrero de 2016. <http://www.computing.co.uk/ctg/opinion/2447523/big-data-in-big-numbers-its-time-to-forget-the-three-vs-and-look-at-real-world-figures>.

Jané, Carmen. «Los señores de los datos.» *El Periódico*, abril de 2015. <https://www.elperiodico.com/es/sociedad/20150426/los-senores-de-los-datos-4134655>.

Jensen, C., Potts, C., y Jensen, C. (2005). «Privacy practices of Internet users: self-reports versus observed behavior.» *International Journal of Human-Computer*. Studies, 63(1), p. 212.

J. Nagy, P. Pecho. *Social networks security*. Third International Conference on Emerging Security Information, Systems and Technologies, IEEE, Atenas/Glyfada, Grecia (2009), pp. 740746.

Jones, Rhett. «Roomba's Next Big Step Is Selling Maps of Your Home to the Highest Bidder.» *Gizmodo*, julio de 2017. <https://gizmodo.com/roombas-next-big-step-is-selling-maps-of-your-home-to-t-1797187829>.

Kalra, Aditya; y Shah, Aditi. «India's antitrust watchdog fines Google for abusing dominant position.» Reuters, febrero de 2018. <https://www.reuters.com/article/us-india-google-antitrust/indias-antitrust-watchdog-fines-google-for-abusing-dominant-position-idUSKBN1FS2AD>.

Keach Sean. «Google Allo just got a major upgrade but do we really need it?» *Trusted Reviews*, agosto de 2017. <http://www.trustedreviews.com/news/google-allo-3261336>.

Kleinman, Zoe. «Is your smartphone listening to you?» BBC News, marzo de 2016. <https://www.bbc.com/news/technology-35639549>.

Krebs, Brian. «Target Investigating Data Breach.» *Krebs on security*, diciembre de 2013. <https://krebsonsecurity.com/2013/12/sources-target-investigating-data-breach/>.

Krishna, Swapna. «Google Duo allows you to call people who don't have the app.» *Engadget*, enero de 2018. <https://www.engadget.com/2018/01/12/google-duo-call-android-users-without-app>.

Lardinois, Frederic. «Google says there are now 2 billion active Chrome installs.» *TechCrunch*, noviembre de 2016. Disponible en: <https://techcrunch.com/2016/11/10/google-says-there-are-now-2-billion-active-chrome-installs/>.

Larson, Jeff; Mattu, Surya; Kirchner, Lauren; y Julia Angwin. «How We Analyzed the COMPAS Recidivism Algorithm.» *ProPublica*, mayo de 2016. <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>.

Levin, Sam. «Facebook told advertisers it can identify teens feeling "insecure" and "worthless".» *The Guardian*, mayo de 2017 <https://www.theguardian.com/technology/2017/may/01/facebook-advertising-data-insecure-teens>.

Lewis-Kraus, Gideon. «The great A.I. awakening.» *The New York Times*, diciembre de 2016. <https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html>.

Lynn, Barry; y Stoller, Matt. «How to stop Google and Facebook from becoming even more powerful.» *The Guardian*, noviembre de 2017. <https://www.theguardian.com/commentisfree/2017/nov/02/facebook-google-monopoly-companies>.

Madden, Mary; Lenhart, Amanda; Cortesi, Sandra; Gasser, Urs; Duggan, Maeve; Smith, Aaron; y Beaton, Meredith. «Teens, Social Media, and Privacy», 21 de mayo de 2013. Pew Research Center. <http://www.pewinternet.org/2013/05/21/teens-social-media-and-privacy/> y <http://www.pewinternet.org/files/2013/05/PIP_TeensSocialMediaandPrivacy_PDF.pdf>.

MacMillan, Douglas y Robert. «Google Exposed User Data, Feared Repercussions of Disclosing to Public.» *The Wall Street Journal*, octubre de 2018. <https://www.wsj.com/articles/google-exposed-user-data-feared-repercussions-of-disclosing-to-public-1539017194>.

Maheshwari, Sapna. «YouTube Is Improperly Collecting Children's Data, Consumer Groups Say.» *The New York Times*, abril de 2018. <https://www.nytimes.com/2018/04/09/business/media/youtube-kids-ftc-complaint.htm>.

Martínez Ron, Antonio. «El oráculo de los test de saliva.» *Vozpópuli*. <https://www.vozpopuli.com/altavoz/next/oraculo-test-saliva_0_1104189582.html>.

McCabe, David; y Fischer, Sara. «Facebook tells advertisers more scrutiny is coming.» *Axios*, 7 de octubre de 2017. <https://www.axios.com/facebook-tells-advertisers-more-scrutiny-is-coming-1513306024-90689d7e-8424-40d0-ba0a-30c09db495cc.html>.

Mcdonald, A. M., y Cranor, L. F. «The Cost of Reading Privacy Policies.» I/S: *A Journal of Law and Policy for the Information Society 2008 Privacy Year in Review Issue*, 4, 543. 2008.

Menárguez, Ana Torres. «Los renegados de Silicon Valley que ahora quieren frenar a las tecnológicas.» *El País*, agosto de 2018.

Meyer, David. «Facebook faces class action privacy suit in Austria, with most of the world invited to join in.» *Gigaom*, 1 de agosto de 2014. <https://gigaom.com/2014/08/01/facebook-faces-class-action-privacy-suit-in-austria-with-most-of-the-world-invited-to-join-in/>.

McNamee, Roger. «How to Fix Facebook—Before It Fixes Us.» *Washington Monthly*, enero-marzo de 2018. <https://washingtonmonthly.com/magazine/january-february-march-2018/how-to-fix-facebook-before-it-fixes-us/>.

Milne, G. R. and Culnan, M. J. (2004), «Strategies for reducing online privacy risks: Why consumers read (or don't read) online privacy notices.» *Journal of Interactive Marketing*, 18(3), pp. 15-29.

M. J. Manier, M. O'Brien Louch. *Online social networks and the privacy paradox: a research framework*. Issues Inf. Syst., XI (1) (2010), pp. 513-517.

Nietzsche, Friedrich Wilhelm. *De la utilidad y de los inconvenientes de los estudios históricos para la vida*. Ed. Tecnos. 2018.

Oetzel, M. C., Gonja, T., 2011. *The online privacy paradox: A social representations perspective*. Paper presented at the Proceedings of the 2011 Annual Conference Extended Abstracts on Human factors in Computing Systems, Vancouver, BC, Canadá.

O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Ed. Random House. 2016.

O'Reilly, Lara. «Google Issuing Refunds to Advertisers Over Fake Traffic, Plans New Safeguard.» *The Wall Street Journal*, agosto de 2017. <https://www.wsj.com/articles/google-issuing-refunds-to-advertisers-over-fake-traffic-plans-new-safeguard-1503675395>.

O'Sullivan. Fergus. «What-exactly is zero knowledge in the cloud and how does it work?» *Cloudwards*, junio de 2017. <https://www.cloudwards.net/what-exactly-is-zero-knowledge-in-the-cloud-and-how-does-it-work>.

P. A. Norberg, D. R. Horne, D. A. Horne. «The privacy paradox: personal information disclosure intentions versus behaviors J. Consum.» *Affairs*, 41 (1) (2007), pp. 100-126.

Pariser, Eli. *El filtro burbuja: Cómo la web decide lo que leemos y lo que pensamos*. Ed. Taurus. 2017.

Perrin, Andrew. «Americans are changing their relationship with Facebook.» *Pew Research*, septiembre de 2018. <http://www.pewresearch.org/fact-tank/2018/09/05/americans-are-changing-their-relationship-with-facebook/>.

Ponemon Institute. «Cost of a Data Breach Study 2018.» <https://securityintelligence.com/series/ponemon-institute-cost-of-a-data-breach-2018/>.

Pontin, Jason. «Artificial Intelligence, With Help From the Humans.» *The New York Times*, marzo de 2007. <https://www.nytimes.com/2007/03/25/business/yourmoney/25Stream.html>.

Posner, Eric A., y Weyl, E. Glen. «Radical Markets: Uprooting Capitalism and Democracy for a Just Society Hardcover.» *Princeton University Press*, 15 de mayo de 2018.

Remnick, David. «Obama Reckons with a Trump Presidency.» *The Newyorker*. <https://www.newyorker.com/magazine/2016/11/28/obama-reckons-with-a-trump-presidency>.

Reyes, Irwin; Wijesekera, Primal; Reardon, Joel; Bar On, Amit Elazari; Razaghpanah. Abbas; Vallina-Rodriguez, Narseo; y Egelman, Serge. «Won't Somebody Think of the Children?» *Examining COPPA Compliance at Scale*. <https://petsymposium.org/2018/files/papers/issue3/popets-2018-0021.pdf>.

Romm, Tony. «Google's location privacy practices are under investigation in Arizona.» *The Washington Post*, septiembre de 2018. <https://www.washingtonpost.com/technology/2018/09/11/googles-location-privacy-practices-are-under-investigation-arizona>.

Russell Brandom. «Google backs off on previously announced Allo privacy feature.» *The Verge*, septiembre de 2016. Disponible en: <https://www.theverge.com/2016/9/21/12994362/allo-privacy-message-logs-google>.

Schmidt, Douglas C. «Google Data Collection.» Vanderbilt University, agosto de 2018. <https://digitalcontentnext.org/wp-content/uploads/2018/08/DCN-Google-Data-Collection-Paper.pdf>.

Schüll, Natasha Dow. «Addiction by Design. Machine Gambling in Las Vegas.» Princeton University Press. 2014.

Silverman, Craig; y Alexander, Lawrence. «How Teens In The Balkans Are Duping Trump Supporters With Fake News.» *BuzzFeed*. <https://www.buzzfeednews.com/article/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo#.fu2okXaeKo>.

Slaughter, Anne-Marie. «Our Bodies or Ourselves.» *Project Syndicate*, julio de 2018. <https://www.project-syndicate.org/commentary/dangers-of-biometric-data-by-anne-marie-slaughter-and-stephanie-hare-2018-07>.

Sly, Liz. «U.S. soldiers are revealing sensitive and dangerous information by jogging.» *The Washington Post*, enero de 2018. <https://www.washingtonpost.com/world/a-map-showing-the-users-of-fitness-devices-lets-the-world-see-where-us-soldiers-are-and-what-they-are-doing/2018/01/28/86915662-0441-11e8-aa61-f3391373867e_story.html>.

Smith, Craig S. *Alexa and Siri Can Hear This Hidden Command. You Can't*, 10 de mayo de 2018.

S. Pötzsch. «Privacy awareness: a means to solve the privacy paradox?» M. Vashek, S. Fischer-Hübner, D. Cvrc ˇ ek, P. Švenda (eds.), *The Future of Identity in the Information Society*, Springer-Verlag, Berlín, Heidelberg (2009), pp. 226-236.

Stallman, Richard. «A radical proposal to keep your personal data safe.» *The Guardian*. <https://amp.theguardian.com/commentisfree/2018/apr/03/facebook-abusing-data-law-privacy-big-tech-surveillance>.

Stamm, Stephanie; Mickle, Tripp; y Kuronen, Jessica. «How Pizza Night Can Cost More in Data Than Dollars.» *The Wall Street Journal*, abril de 2018. <https://www.wsj.com/graphics/how-pizza-night-can-cost-more-in-data-than-dollars/>.

Subramanian, Samanth. «Inside the Macedonian Fake-News Complex.» *Wired*, febrero de 2017. <https://www.wired.com/2017/02/veles-macedonia-fake-news/>.

Sundar, S. S., Kang, H., Wu, M., Go, E., Zhang, B., 2013. «Unlocking the privacy paradox: Do cognitive heuristics hold the key?» Proceedings of CHI'13 Extended Abstracts on Human Factors in Computing Systems, Francia, pp. 811-816.

Swartz, Jon. «Soros: Beware IT Monopolies Facebook, Google», 25 de enero de 2018. *Barrons*. <https://www.barrons.com/articles/soros-beware-it-monopolies-facebook-google-1516923914>.

Tilley, Aaron. «Amazon Alexa Can Now Unlock Your Doors. Forbes», febrero de 2017. <https://www.forbes.com/sites/aarontilley/2017/02/16/amazon-alexa-can-now-unlock-your-front-door>.

Tsai, J., Cranor, L., Acquisti, A., Fong, C., 2006. «What's it for you? A survey of online privacy concerns and risk.» NET Institute Working Paper, n.º 06-29, pp. 1-20.

Tobin, Ariana; y Merrill, Jeremy B. «Facebook is letting job advertisers target only men.» *ProPublica*, septiembre de 2018. <https://www.propublica.org/article/facebook-is-letting-job-advertisers-target-only-men>.

Turow, J., King, J., Hoofnagle, C. J., Bleakley, A., y Hennessy, M. «Americans reject tailored advertising and three activities that enable it.» *Papers*. 2009. <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1478214>.

Warren, Samuel D.; y Brandeis, Louis D.. *The Right to Privacy*. Harvard Law Review, vol. 4, n.º 5 (15 de diciembre de 1890), pp. 193-220. <http://www.jstor.org/stable/1321160>.

Whitaker, Roger M.; Colombo, Gualtiero B.; y Rand, David G. «Indirect Reciprocity and the Evolution of Prejudicial Groups.» *Scientific Reports*, 2018; 8 (1) <DOI: 10.1038/ s41598-018-31363-z>.

Whittaker, Zack. «Teen phone monitoring app leaked thousands of user passwords.» *Zdnet*, mayo de 2018. <https://www.zdnet.com/article/teen-phone-monitoring-app-leaks-thousands-of-users-data/>.

Wittes, Ben; y Liu, Jodie. «The Privacy Paradox: The Privacy Benefits of Privacy Threats.» <https://www.brookings.edu/wp-content/uploads/2016/06/Wittes-and-Liu_Privacy-paradox_v10.pdf>.

Yoo, C. W., Ahn, H. J., Rao, H. R., 2012. *An exploration of the impact of information privacy invasion*. Proceeding of Thirty Third International Conference on Information Systems, Orlando, Florida, pp. 1-18.

Zhang, Sarah. «The Coming Wave of Murders Solved by Genealogy.» *The Atlantic*, mayo de 2018.

Zuboff, Shoshana. «Big Other: Surveillance Capitalism and the Prospects of an Information Civilization.» *Journal of Information Technology* 30, n.º 1, 1 de marzo de 2015, pp. 75-89. <https://doi.org/10.1057/jit.2015.5>.

Notas

1. «Cybersecurity Futures 2020.» Center for Long-Term Cybersecurity University of California, Berkeley. 2016. <https://cltc.berkeley.edu/scenarios/>.

2. Harari, Yuval Noah. *Homo Deus. Breve historia del mañana*. Debate, 2016, p. 400.

3. La evolución darwiniana y el procesamiento complejo de Turing.

4. Harari, Yuval Noah. «La tecnología permitirá “hackear” a seres humanos.» *El País Semanal*, 20 de agosto de 2018. <https://elpais.com/elpais/2018/08/20/eps/1534781175_639404.html>.

5. Pérez-Chirinos, César. *Algoritmos y programación*.

6. «Data workers of the world, unite. What if people were paid for their data?» *The Economist*. <https://www.economist.com/the-world-if/2018/07/07/data-workers-of-the-world-unite>.

7. Posner, Eric A.; y Weyl, E. Glen. «Radical Markets: Uprooting Capitalism and Democracy for a Just Society Hardcover.» Princeton University Press, 15 de mayo de 2018.

8. Lee, Dave. «Why Big Tech pays poor Kenyans to teach self-driving cars.» BBC News, 3 de noviembre de 2018. <https://www.bbc.com/news/technology-46055595>.

9. Aytes, Ayhan. «Return of the Crowds: Mechanical Turk and Neoliberal States of Exception.» *Digital Labor: The Internet as Playground and Factory*. Trebor Scholz (ed.). Londres: Routledge, 2012.

10. Pontin, Jason. «Artificial Intelligence, With Help From the Humans.» *The New York Times*, 25 de marzo de 2007. <https://www.nytimes.com/2007/03/25/business/yourmoney/25Stream.html>.

11. Aytes, *op. cit.*

12. Los doppelgänger aparecen en diversas obras literarias de ciencia ficción, fantasía y horror típicas de finales del siglo XIX como *El extraño caso del Dr. Jekyll y Mr. Hyde* (1886) de Robert Louis Stevenson, *Él o El terror* de Guy de Maupassant (1850-1893) o los *Cuentos Góticos* de Edgar Allan Poe.

13. La figura del doble apareció antes en nuestra mitología en la figura del griego Narciso, que se enamora de su reflejo. Esto está grabado en nuestro subconsciente colectivo, como ya señaló el propio Freud. En el psicoanálisis freudiano, la visualización de un doppelgänger por parte del paciente le permite enfrentarse a aspectos de su personalidad que está luchando por controlar.

14. González Cabañas, José, Cuevas, Ángel y Rubén. «Facebook Use of Sensitive Data for Advertising in Europe.» Universidad Carlos III de Madrid.

15. La coreana Vonvon es un referente en los test absurdos para Facebook, como aquéllos en los que se pregunta «¿qué animal eres: gorila, mono u orangután?» o «¿puedes pasar este test increíblemente difícil?». Vonvon elabora estos tests en 14 idiomas, y son completados por 100 millones de usuarios únicos de 50 países diferentes. La empresa admitió a la BBC que su única finalidad era recabar datos. Otras compañías que se dedican a la extracción de datos a través de los test de Facebook son Chiquitests, Nametests o Kuezz.

16. Jané, Carmen. «Los señores de los datos.» *El Periódico*, 27 de abril de 2015. <https://www.elperiodico.com/es/sociedad/20150426/los-senores-de-los-datos-4134655>.

17. Data Trust, Aristotle o PK List Marketing, entre otras, tienen como único objetivo recabar datos a partir de perfiles psicográficos, que pueden conseguirse, según Amnistía Internacional, a partir de 7 céntimos de euro el perfil. Una de ellas, MedBase200, especializada en marketing médico, estuvo acusada de vender datos sobre pacientes de alcoholismo, problemas de erección y víctimas de violación. Facebook, que mantiene que no vende la información de sus perfiles, fue acusada por *The Guardian* de afirmar que podía encontrar «adolescentes inseguros», como veremos más adelante.

18. Stamm, Stephanie; Mickle, Tripp; y Kuronen, Jessica. «How Pizza Night Can Cost More in Data Than Dollars.» *The Wall Street Journal*. <https://www.wsj.com/graphics/how-pizza-night-can-cost-more-in-data-than-dollars/>. 10 de abril de 2018.

19. A continuación, hay una lista de algunos de los tipos de datos que Apple podría haber recopilado conforme a su política de privacidad:

- Nombre, dirección postal, número de teléfono, dirección de correo electrónico, preferencias de contacto, información de la tarjeta de crédito, fecha de nacimiento, uso del servicio/dispositivo, ubicación, ocupación, tienda de aplicaciones, actividades, servicios, consultas de búsqueda, teléfono, operador, idioma, país, código postal, código de área, sistema operativo, tipo de navegador, proveedor de servicios de internet, url de referencia, identificador único del dispositivo, zona horaria, dirección IP, correo electrónico abierto de Apple, nombre de amigo o familiar, dirección de correo electrónico de amigo o familiar, correo electrónico de amigo o familiar, números de teléfono de amigo o familia...
- País de residencia, método de pago, ID de Apple, dispositivo actividad, ubicación, uso...

Las políticas de Apple establecen que «la compañía no cree que las empresas deban construir perfiles detallados de clientes. A menudo (Apple) disocia la información de los usuarios y la usa para mejorar los dispositivos que vende, pero no vende información personal a los anunciantes».

20. A continuación, hay una lista de algunos de los tipos de datos que Amazón podría haber recopilado conforme a su política de privacidad.

- Información que se facilita en la web: nombre, número de teléfono, dirección postal , información de la tarjeta de crédito, nombres de las personas a quienes se han enviado las compras, direcciones de las personas a las que se han enviado compras, número de teléfono de las personas a las que se han enviado las compras, direcciones de correo electrónico de los amigos, contenido de reseñas, contenido de correos electrónicos a la empresa.
- Perfil: descripción personal, fotografía, número de identificación personal (seguridad social o carnet de conducir en Estados Unidos, DNI en España), dirección de correo electrónico, contraseña, historial de compras, productos vistos, búsqueda de dirección IP, zona horaria, tipo de navegador, versión del navegador, complementos del navegador, sistema operativo.
- Clickstream: número de teléfono utilizado para llamar, compañía, abrió o no un correo electrónico, tiempo de respuesta de la página de la empresa, errores de descarga, duración de las visitas, interacción de la página (desplazamientos, clics...), métodos utilizados para navegar desde la página, uso de la aplicación, identificadores del dispositivo de ubicación, nombre.
- Específico de Alexa: número de teléfono, detalles de contacto, tareas pendientes, listas de compras, lista de reproducción de música, información de pago predeterminado, envío predeterminado, características de la voz, contactos del teléfono móvil, si se importaron, contenido de las solicitudes, historial de interacción, tipos de compras, código postal, cadenas de música personalizadas, información del producto auxiliar, dispositivo de inicio inteligente, tipo hogar inteligente, nombre del dispositivo, dispositivo de inicio inteligente, características dispositivo de inicio inteligente, conectividad de red smart home, ubicación del dispositivo, mensajes de voz, contacto comunicado con la mayoría...

La política de Amazon establece: «Nuestro aviso de privacidad describe qué información recabamos y cómo la utilizamos. Nunca vendemos la información personal de nuestros clientes. Ciframos los datos en tránsito y en reposo, y ofrecemos al cliente la posibilidad de activar la autenticación de múltiples factores».

21. A continuación, hay una lista de algunos de los tipos de datos que Domino’s podría haber recopilado conforme a su política de privacidad.

– Nombre de usuario registrado, dirección postal, número de teléfono, dirección de correo electrónico, información de facturación, áreas de interés, uso del producto, información de la tarjeta de crédito, transacciones con contraseña, ubicación, naturaleza de la compra, cantidad de compra, precio de compra, transcripción de las conversación, uso del servicio, comunicaciones secundarias, ruido de fondo, identificador del dispositivo, tipo de dispositivo, sistema operativo, tipo de navegador, configuración de hardware, estadísticas de rendimiento, nombre del servidor, dirección IP, proveedor de servicios de internet, información geográfica general, fecha y hora, páginas accesibles al sitio, acceso dentro del sitio o URL de aplicación, URL de salida, historial de transacciones, fuentes instaladas, objetos javascript, contenido de redes sociales, hashtag de Domino’s...

La política de Domino’s establece que «cualquier información del cliente que recopilamos en un pedido online se usa para completar el pedido o para mejorar la experiencia del usuario».

22. A continuación, una lista de algunos de los tipos de datos que Google podría haber recopilado conforme a sus políticas de privacidad.

– Registro: nombre de la cuenta, contraseña, iniciar sesión, dirección de correo electrónico, número de teléfono.
– Perfil: foto, género, fecha de nacimiento, país, uso del servicio, idioma, preferencias, análisis de interacciones con Google, servicios, información de la tarjeta de crédito, contactos, comentarios, escribir comentarios, publicar ubicación, historial de ubicación, mapa, búsquedas, dirección de velocidad de desplazamiento, Travel Voice Search, cargas de fotos y vídeos, fecha y hora de las fotos, vídeos tomados, ubicación, información de fotos, edad (confirmada mediante transacción con tarjeta de crédito), historial de navegación, fecha y hora de búsqueda, historial de búsquedas, frecuencia de visitas, anuncios vistos, categorías comprometidas, interesados en Gmail, Mensajes o Gchat, reconocimiento facial, contenido de Google Drive (es decir, documentos), historial de reproducción de llamadas, número de teléfono de llamada, número de reenvío, historial de llamadas, conversaciones grabadas, hora y fecha de las llamadas, mensajes de correo de voz, saludos de correo de voz, duración de las llamadas, tipos de llamadas, información de enrutamiento, identificadores de dispositivo, configuración de hardware, informe de bloqueo, tipo de navegador, marcadores, extensiones instaladas, pestañas del navegador, URL de referencia, eventos del calendario, ubicaciones de inicio de sesión, fecha y hora de solicitud, contactos compartidos con la mayoría de direcciones IP vinculadas a URL, registro visitado de descargas de sitio web, potencia del wifi o señal del móvil.

La política de privacidad de Google establece que «para poder elegir las opciones de privacidad adecuadas, es esencial que los usuarios puedan comprender y controlar sus datos de Google. A lo largo de los años, hemos desarrollado herramientas como My Account expresamente para este propósito, y alentamos a todos a que la revisen con regularidad».

23. Los receptores IMSI son piezas intrusivas de tecnología de vigilancia que imitan las estaciones base de los teléfonos móviles con el fin de engañar a los teléfonos dentro de su radio de acción para que se conecten a ellos. IMSI responde a Identidad de Suscriptor Móvil Internacional (International Mobile Subscriber Identity, en inglés), un número único asociado a la tarjeta SIM. Una vez que se engaña al teléfono para que se conecte a un receptor IMSI, éste muestra este identificador único a partir del que se puede inferir la identidad, la ubicación (incluida la intensidad de la señal que permite afinarla), intervenir las llamadas, mensajes de texto y el tráfico de internet.

24. Estas cajas se conocen como Stingray, y existen en relativa oscuridad desde la década pasada. Los fabrica una empresa llamada Harris Corporation y son altamente controvertidas, ya que la policía y otros cuerpos de seguridad las utilizan sin pasar por la preceptiva orden judicial. Se puede ver el aspecto y características de uno en el Catálogo de vigilancia masiva de The Intercept. Stingray I/II Ground Based Geo-Location (Vehicular). Fabricante Harris Corporation. Precio: 134.952 dólares. <https://theintercept.com/surveillance-catalogue/stingray-iii/>.

25. Biddle, Sam. «Long-Secret Stingray Manuals Detail How Police Can Spy on Phones.» *The Intercept*, 12 de septiembre de 2016. <https://theintercept.com/2016/09/12/long-secret-stingray-manuals-detail-how-police-can-spy-on-phones/>.

26. Uno de los primeros informes del uso de los receptores IMSI por parte de la Policía del Reino Unido es de 2011, cuando *The Guardian* publicó un artículo que indicaba que el Servicio de Policía Metropolitana (Met) había adquirido uno de ellos (<https://www.theguardian.com/uk/2011/oct/30/metropolitan-police-mobile-phone-surveillance>). En los años siguientes, otros medios sugirieron que el uso de receptores IMSI se había extendido a todo Londres.

En octubre de 2016, *The Bristol Cable* publicó un artículo que revelaba que otras seis fuerzas policiales británicas habrían adquirido receptores IMSI (Avon y Somerset, South Yorkshire, Staffordshire, Warwickshire, West Mercia y West Midlands) (<https://thebristolcable.org/2016/10/imsi/>).

Los registros de compra indicaron que la Policía metropolitana había gastado más de un millón de libras en receptores IMSI en 2015. Revelaron, además, que en un período similar, la Policía de Avon y Somerset se habría gastado casi 170.000 libras y la Policía de South Yorkshire 144.000 libras en la adquisición de receptores IMSI. La Policía británica «ni confirma ni niega» que estén usando receptores IMSI.

No serían sólo los británicos los que estarían usando esta tecnología para controlar a la población; las fuerzas del orden público de Estados Unidos en Baltimore, Milwaukee, Nueva York, Tacoma, Anaheim y Tucson también habrían utilizado receptores IMSI en vehículos o aeronaves para identificar y localizar sospechosos.

Algunas ciudades como Washington D.C. estarían llenas de estos dispositivos (<http://www.microsiervos.com/archivo/seguridad/dispositivos-rastreadores-telefonos-moviles.html>). Decenas de rastreadores de teléfonos móviles que se habrían, a su vez, rastreado en las calles de Washington D.C. La NBC ha recoge la actividad de un investigador que habría detectado hasta 40 dispositivos de localización y rastreo de teléfonos móviles mediante lo que se conoce como «simuladores de antenas de telefonía». Estos dispositivos se hacen pasar por antenas de las empresas de telecomunicaciones, y cuando los teléfonos móviles de las personas de interés a las que se está investigando pasan cerca «acaban revelando su posición aproximada mediante triangulación».

27. Fakhoury, Hanni. «When a Secretive Stingray Cell Phone Tracking “Warrant” Isn’t a Warrant.» *Electronic Frontier Foundation*, 28 de marzo de 2013. <https://www.eff.org/homepage-feature/stingray>.

28. Madden, Mary; Lenhart, Amanda; Cortesi, Sandra; Gasser, Urs; Duggan, Maeve; Smith, Aaron; y Beaton, Meredith. «Teens, Social Media, and Privacy.» *Pew Research Center*, 21 de mayo de 2013. <http://www.pewinternet.org/2013/05/21/teens-social-media-and-privacy/>. <http://www.pewresearch.org/wp-content/uploads/sites/9/2013/05/PIP_TeensSocialMediaandPrivacy_PDF.pdf>.

29. Éstas son algunas de las principales conclusiones de este estudio basado en una encuesta a 802 adolescentes que examina su gestión de la privacidad en las redes sociales.

1. La Generación Z comparte más información acerca de sí mismos en redes sociales de lo que lo hicieron los millennials:
   – Un 91 por ciento publica selfies, frente al 79 por ciento de los millennials.
   – Un 71 por ciento publica el nombre de su universidad, frente al 49 por ciento de los millennials.
   – Un 71 por ciento publica la ciudad o pueblo en el que viven, frente al 61 por ciento de los millennials.
   – Un 53 por ciento publica su dirección de correo electrónico, frente al 29 por ciento de los millennials.
   – Un 20 por ciento publica su número de teléfono móvil, frente al 2 por ciento de los millennials.
   – Un 92 por ciento publica su nombre real en su perfil.
   – Un 84 por ciento publica sus intereses reales: películas, música o libros que les gustan.
   – Un 82 por ciento facilita su fecha de nacimiento.
   – Un 62 por ciento publica su estado civil.
   – Un 24 por ciento publica vídeos personales.
2. Los miembros más mayores de la Generación Z es más probable que compartan determinados contenidos que los más jóvenes, si bien no se aprecia diferencia entre chicos y chicas, que tienden a publicar lo mismo.
3. Un 16 por ciento de los usuarios de redes sociales ha configurado su perfil para incluir su ubicación por defecto en sus publicaciones.
4. La Generación Z usa más Twitter que los millennials.
5. Las cuentas públicas en Twitter son la norma.
6. El usuario medio de Facebook tiene 300 amigos, mientras que el usuario medio de Twitter tiene 79 followers.
7. Los perfiles de la Generación Z en Facebook son un reflejo de su vida offline: siete de cada diez tienen a sus padres como amigos en la red social.
8. Los mayores de la Generación Z tienen mayor variedad de amigos en Facebook, mientras que los más jóvenes tienden a ser más cautelosos con la gente que no conocen en la vida real.
9. Un 60 por ciento de los usuarios de Facebook configuran sus perfiles como privados y están convencidos de que saben cómo gestionar la herramienta de configuración de privacidad.
10. Las chicas suelen proteger sus perfiles más que los chicos.

11. La Generación Z es consciente de la importancia de la reputación online y adoptan medidas para protegerla.
12. La Generación Z revisa y borra contenidos de sus redes como parte de la gestión de su identidad online.
13. Un 74 por ciento ha borrado alguna vez a un amigo de su perfil y un 58 por ciento ha bloqueado a alguien en una red social.
14. Un 26 por ciento ha publicado información falsa (nombre, edad o localización) para proteger su privacidad.
15. No manifiestan una especial preocupación por la cesión a terceros de sus datos personales. Sólo un 9 por ciento estaría muy preocupado.
16. La Generación Z no sería consciente del acceso de terceros a los datos que comparten en redes sociales.
17. Sin embargo, sus padres estarían muy preocupados sobre cuánta información obtienen las empresas de publicidad acerca del comportamiento de sus hijos en la red.
18. Aquellos de la Generación Z más preocupados por el acceso a sus datos por terceros es más probable que gestionen su reputación online.
19. Más de la mitad de los entrevistados han decidido no publicar un contenido, preocupados por su reputación.

30. YouTube manifiesta que su servicio no está dirigido a niños que cuentan con una aplicación específica, YT Kids, dotadas de filtros. A pesar de ello, YouTube no sólo recababa datos de los menores como de cualquier otro usuario, sino que sabía cuando un menor accedía, ya que le mostraba anuncios dirigidos específicamente a ellos.

La app YT Kids, por su parte, recibió quejas paternas debido a vídeos creados por algoritmos ideados por productores que habían encontrado la manera de conseguir más clics de los menores. Estos vídeos incluían personajes infantiles en situaciones inapropiadas que, como es natural, tenían un éxito arrollador entre los menores que los consumían sin supervisión.

Como respuesta a estas críticas, YouTube borró millones de vídeos y anunció que los vídeos de la aplicación para los menores, YT Kids, serían moderados por humanos y no por algoritmos.

Por su parte, la demanda presentada contra Disney por Amanda Rushing, tendría su base en la recogida y venta a terceros de datos de menores con fines publicitarios. La demanda se dirige contra 42 aplicaciones para niños que recababan sus datos con la finalidad de generar perfiles y servirles publicidad dirigida. Estas aplicaciones permitían detectar la actividad del niño a través de diferentes apps, plataformas y dispositivos, generando una cronología completa. La demanda incluía a compañías tecnológicas como Upsight, Unity y Kochava.

Maheshwari, Sapna. «YouTube Is Improperly Collecting Children's Data, Consumer Groups Say.» *The New York Times*, 9 de abril de 2018. <https://www.nytimes.com/2018/04/09/business/media/youtube-kids-ftc-complaint.htm>.

31. Reyes, Irwin; Wijesekera, Primal; Reardon, Joel; Bar On, Amit Elazari; Razaghpanah, Abbas; Vallina-Rodriguez, Narseo; y Egelman, Serge. «Won't Somebody Think of the Children?» *Examining COPPA Compliance at Scale*. <https://petsymposium.org/2018/files/papers/issue3/popets-2018-0021.pdf>.

Los investigadores de International Computer Science Institute (ICSI) de la Universidad de Berkeley dejan en evidencia que las aplicaciones gratuitas para menores, igual que las orientadas al público adulto, recogen datos, los tratan y, con toda seguridad, los venden a terceros para análisis agregados.

De 5.855 aplicaciones de Android dirigidas a niños, un 57 por ciento incumple la ley federal que protege a los niños estadounidenses, conocida como Children's Online Privacy Protection Act (COPPA). Esto supone que un 4,8 por ciento de las apps comparte la localización o información de contacto sin consentimiento de los padres, un 40 por ciento envía datos personales en blanco o sin seguridad en la transmisión —lo que permitiría que cualquiera accedies a esos datos— y un 18,8 por ciento envía identificadores de usuarios a terceros para mostrar anuncios dirigidos. Un 39 por ciento incumple sus obligaciones contractuales para proteger la privacidad infantil.

32. La web de citas Ashley Madison sufrió un ataque informático que dejó a la vista la información más personal de los 34 millones de usuarios registrados en su web, y en evidencia a la propia compañía, que utilizaba bots femeninos como señuelos para que los usuarios masculinos (en mayoría aplastante) consumieran y gastasen en la página. Los datos publicados fueron usados por una empresa española Tecnilógica —posteriormente adquirida por Accenture en 2017— para realizar un mapa que localizó por ciudades a los usuarios que buscaban, a través de esta página, sexo extramatrimonial. El resultado permitió comprobar que Ashley Madison contaba con clientes en más de 50.000 municipios repartidos por 48 países y que la mayoría de ellos eran hombres, el 86,2 por ciento del total.

33. Ghorayshi, Azeen; y Ray, Sri. «Grindr Is Letting Other Companies See User HIV Status And Location Data.» *Buzzfeed*, 2 de abril de 2018. <http://www.buzzfeed.com/azeenghorayshi/grindr-hiv-status-privacy>.

34. Acquisti, A. «Privacy in electronic commerce and the economics of immediate gratification.» *EC '04 de la 5th ACM Conference on Electronic Commerce*. 2004.

Barnes, S.B. A privacy paradox: *Social networking in the United States*. First Monday, 11(9). <http://firstmonday.org/article/view/1394/1312>. 2006.

35. Barth, Susanne; y De Jong, Menno D. T. *The privacy paradox – Investigating discrepancies between expressed privacy concerns and actual online behavior – A systematic literature review*. Elseviers. Telematics and Informatics, vol. 34, n.º 7, noviembre de 2017, pp. 1038-1058 <https://www.sciencedirect.com/science/article/abs/pii/S0736585317302022>.

36. A. N. Joinson, U.-D. Reips, T. Buchanan y C. B. Paine Schofield. «Privacy, trust, and self-disclosure online.» *Hum.-Comput. Interact,* 25 (2010), pp. 1-24.

S. Pötzsch. «Privacy awareness: a means to solve the privacy paradox?» M. Vashek, S. Fischer-Hübner, D. Cvrc˘ek, P. Švenda (Eds.). *The Future of Identity in the Information Society*. Springer-Verlag. Berlín, Heidelberg (2009), pp. 226-236.

Tsai, J.; Cranor, L.; Acquisti, A.; y Fong, C., 2006. «What's it for you? A survey of online privacy concerns and risk.» *NET Institute Working Paper*, n.º 06-29, pp. 1-20.

37. P. A. Norberg, D. R. Horne y D. A. Horne. *The privacy paradox: personal information disclosure intentions versus behaviors*, J. Consum. Affairs, 41 (1) (2007), pp. 100-126.

38. C. Flender, G. Müller. *Type indeterminacy in privacy decisions: the privacy paradox revisited*. J. Busemeyer, F. Dubois, A. Lambert-Mogiliansky, M. Melucci (eds.), Quantum interaction. *Lecture notes in Computer Science*, Springer-Verlag, Berlín, Heidelberg (2012), pp. 148-159.

39. A. Acquisti y J. Grossklags. *Privacy and rationality in individual decision making*. IEEE Secur. Priv., 3 (1) (2005), pp. 26-33.

Sundar, S. S.; Kang, H.; Wu, M.; Go, E.; y Zhang, B., 2013. *Unlocking the privacy paradox: Do cognitive heuristics hold the key*?. Proceedings of CHI'13 Extended Abstracts on Human Factors in Computing Systems, Francia, pp. 811-816.

40. Wittes, Ben; y Liu, Jodie. «The Privacy Paradox: The Privacy Benefits of Privacy Threats.» *Brookings*. (<https://www.brookings.edu/wp-content/uploads/2016/06/Wittes-and-Liu_Privacy-paradox_v10.pdf>.) Este artículo analiza cómo las nuevas tecnologías que amenazan la privacidad serían al mismo tiempo de ayuda para conseguir su protección:

> «Al pensar en el big data y en las tecnologías digitales como una calle de sentido único que erosiona la privacidad simplificamos mucho la verdadera naturaleza de la relación privacidad/ tecnología [...] pocas tecnologías implican una pérdida de privacidad pura. Internet, después de todo, no es sólo una serie de tecnologías de vigilancia. Las capacidades de vigilancia, más bien, son en gran parte consecuencias colaterales de su elemento principal: la apertura de nuevos canales de comunicación. Y esos nuevos canales implican, en primera instancia, un beneficio para la privacidad, o más bien, una serie de beneficios para la privacidad, ya que permiten una mayor elección sobre cómo los seres humanos se comunican entre sí [...]. Permite la posibilidad de comunicación verbal remota, la capacidad de mantener una conversación confidencial a distancia.
>
> »Buena parte del desarrollo tecnológico sigue este patrón básico. Google, Microsoft o Yahoo! permiten buscar información de forma privada, con la recopilación de datos por parte de las empresas y la posible recuperación por parte de otros actores como consecuencia. Amazon permite comprar todo tipo de productos sin que nadie más lo sepa, pero Amazon guarda el historial de compras que usa para generar patrones. Nuestros teléfonos inteligentes nos permiten tener la capacidad de cómputo de un ordenador en el bolsillo y llevarla con nosotros, lo que les permite geoposicionarnos [...]. Facebook nos permite identificar de manera discreta grupos de personas con quienes compartir contenidos, pero almacena nuestras acciones para procesarlas y recuperarlas según las necesite.
>
> »En nuestra contabilidad mental de ganancias y pérdidas, tendemos a contar sólo una de las columnas, embolsándonos lo que ganamos como si no tuviera ningún valor en términos de privacidad mientras lamentamos lo que hemos abandonado [...]. Cuando reconocemos los beneficios, tendemos a redefinirlos como beneficios distintos a la privacidad, esto es, como meros beneficios de conveniencia o de eficiencia: una descripción desdeñosa que implica que hemos ganado algo intrascendente o que hemos ahorrado tiempo.
>
> »Pero esa construcción nos deja con una perspectiva distorsionada y en exceso sombría sobre el impacto de la tecnología en nuestras vidas. Sí, la tecnología implica beneficios en conveniencia y eficiencia, pero no son los únicos.

»No argumentamos que la tecnología mejore la privacidad en su conjunto, o que la tecnología no erosione la privacidad. Por el contrario, nuestro punto general es que la interacción entre la tecnología y la privacidad es menos clara de lo que comúnmente reconoce el debate, que no hacemos un cálculo correcto y que el calculo de la privacidad real no está nada claro».

Las tecnologías que pueden ser utilizadas para evitar la vigilancia y que permiten nuevos tipos de vigilancia son un arma de doble filo desde una perspectiva de privacidad. Pueden agregar privacidad o quitársela, dependiendo de cómo se usen y de qué reglas y prácticas legales establecen los supervisores.

41. I. Oomen y I. Leenes. *Privacy risk perceptions and privacy protection strategies*. E. de Leeuw, S. Fischer-Hübner, J. Tseng y J. Borking (eds.), *Policies and Research in Identity Management*, Springer Verlag, Boston (2008), pp. 121-138.

42. A. L. Young y A. Quan-Haase. *Privacy protection strategies on Facebook*. Inf. Commun. Soc., 16 (4) (2013), pp. 479-500.

43. Hughes-Roberts, T., 2012. «A cross-disciplined approach to exploring the privacy paradox: explaining disclosure behaviour using the theory of planned behavior.» *UK Academy for Information Systems Conference Proceedings*, Paper 7.

M. J. Manier y M. O'Brien Louch. «Online social networks and the privacy paradox: a research framework.» Issues Inf. Syst., XI (1) (2010), pp. 513-517.

J. Nagy y P. Pecho. «Social networks security. Third International Conference on Emerging Security Information, Systems and Technologies, IEEE.» Atenas/Glyfada, Grecia (2009), pp. 740-746.

Yoo, C.W., Ahn, H. J. y Rao, H.R., 2012. «An exploration of the impact of information privacy invasion. Proceeding of Thirty Third International Conference on Information Systems.» Orlando, Florida, pp. 1-18.

44. Buck, C., Horbel, C., Germelmann, C. C. y Eymann, T., 2014. «The unconscious app consumer: Discovering and comparing the information-seeking patterns among mobile application consumers. Twenty Second European Conference on Information Systems.» Tel Aviv, Israel, pp. 1-14.

45. Oetzel, M. C. y Gonja, T., 2011. «The online privacy paradox: A social representations perspective. Paper presented at the Proceedings of the 2011 Annual Conference Extended Abstracts on Human factors in Computing Systems.» Vancouver, BC, Canadá.

46. Borges, Jorge Luis. *Funes el memorioso*. Ficciones. 1944.

47. Nietzsche, Friedrich Wilhelm. *De la utilidad y de los inconvenientes de los estudios históricos para la vida*. Editorial Tecnos. 2018.

48. El escritor lo resume así: «Pensar es olvidar diferencias, es generalizar, abstraer. En el abarrotado mundo de Funes no había sino detalles, casi inmediatos. Funes era un precursor de los superhombres. Sus recuerdos no eran simples; cada imagen visual estaba ligada a sensaciones musculares, térmicas, etcétera. Podía reconstruir todos los sueños, todos los entresueños. Dos o tres veces había reconstruido un día entero; no había dudado nunca, pero cada reconstrucción había requerido un día entero. Me dijo: “Más recuerdos tengo yo solo que los que habrán tenido todos los hombres desde que el mundo es mundo”. Y también: “Mis sueños son como la vigilia de ustedes”. Y también, hacia el alba: “Mi memoria, señor, es como vaciadero de basuras”».

49. «Google y Facebook refuerzan el control del mercado publicitario digital de Estados Unidos.» *EMarketer*, 21 de septiembre de 2017. Disponible en: <https://www.emarketer.com/Article/Google-Facebook-Tighten-Grip-on-US-Digital-Ad-Market/1016494>.

50. Schmidt, Douglas C. «Google Data Collection.» Vanderbilt University, agosto 2018. <https://digitalcontentnext.org/wp-content/uploads/2018/08/DCN-Google-Data-Collection-Paper.pdf>.

51. Schmidt, Douglas C., *op. cit.*

52. «Market share or leading internet browsers in the United States and worldwide as of February 2018.» *Statista*, febrero de 2018. Disponible en: <https://www.statista.com/statistics/276738/worldwide-and-us-market-share-of-leading-internet-browsers/>.

53. «Global OS market share in sales to end users from 1st quarter 2009 to 2nd quarter 2017.» *Statista*, agosto de 2017. Disponible en: <https://www.statista.com/statistics/266136/global-market-share-held-by-smartphone-operating-systems/>.

54. «Google 10K filings with the SEC, 2017.» <https://abc.xyz/investor/pdf/20171231_alphabet_10K.pdf>.

55. «My activity». Disponible en: <https://myactivity.google.com/myactivity>.

56. «Download your data.» Google. Disponible en: <https://takeout.google.com/settings/takeout?pli=1>.

57. Dave Burke. «Android: celebrating a big milestone together with you.» Google, 17 de mayo de 2017. disponible en: <https://www.blog.google/products/android/2bn-milestone/>.

58. Yubal, F. M.: «He mirado todos los datos que Google tiene sobre mí, y confirmo que es el Gran Hermano definitivo». Xataka, 5 de abril de 2018. Actualizado el 5 de septiembre de 2018. <https://www.xataka.com/privacidad/he-mirado-todos-los-datos-que-google-tiene-sobre-mi-y-confirmo-que-es-el-gran-hermano-definitivo>.

59. Éstos son los datos guardados:
- Ubicación en todo momento: Google Maps cuenta con una función cronológica en la que registra todos los movimientos. Requiere que esté habilitado, lo que no es inhabitual ya que la aplicación lo tiene activado por defecto. Muestra de manera visual, ordenada y cronológica dónde se ha estado, cada parada realizada aunque sea delante de un cajero para sacar efectivo o el tipo de transporte utilizado.
- Datos de toda la actividad física: Google guarda las actividades registradas en Google Fit. con todos los ejercicios, horas y coordenadas.
- Todos los datos de navegación: los datos aparecen organizados cronológicamente y van acompañados de enlaces. De un vistazo se comprueba lo que se ha buscado y se puede recuperar el resultado de las búsquedas, incluida la navegación profunda más allá de la home. También almacena los hilos de Feedly por su título con el link para acceder a ellos.
- Toda la actividad del móvil: el uso de móviles con sistema operativo Android permite a Google acceder al uso que se realiza de las aplicaciones instaladas en el móvil, cuántas veces se utiliza cada aplicación o los perfiles en los que se ha entrado dentro de cada aplicación.
- Toda la actividad realizada dentro de los servicios propios de Google:
  - YouTube, cuando se ha entrado con el usuario de Google, almacena un historial con las búsquedas de vídeos, todos los comentarios dejados o los vídeos concretos visitados mientras.
  - Google Calendar se almacena eternamente, lo que permite acceder a la agenda desde que la aplicación existe.
  - Google Drive conserva los archivos generados, todas las notas de Google Keep, y todas y cada una de las fotos subidas a Google Fotos ordenadas cronológicamente, incluidas las enviadas desde Hangout.
- Todos los contactos telefónicos del móvil Android. En la carpeta Voz se guardan todos los contactos en formato vcf —que se puede importar a cualquier gestor de correo—. En la carpeta Contactos se almacenan las fotografías asignadas a cada uno.
- Todas las apps utilizadas en Android y cuándo se utilizan. Google guarda una cronología completa de cada aplicación abierta en sus dispositivos Android. Así, y en palabras del periodista, «Google sabe por ejemplo que el día 2 de abril a las 15.11 abrí Telegram, y que sólo 11 minutos antes había abierto Slack. También sabe las horas a la que he encendido el móvil, cuándo he estado mirando fotos de la galería, cuándo se realizan acciones de mantenimiento del dispositivo o cuándo instalo una aplicación». Cuenta también con un historial de aplicaciones con permiso

de acceso en el que aparecen todas las apps desde las que se ha dado acceso a las credenciales de Google, especificando el tipo de datos a que accede cada una.

60. La política de privacidad de Google, bajo el título «Datos que procesamos cuando usa Google», dice que: «Cuando busca un restaurante o ve un vídeo en YouTube, por ejemplo, procesamos información sobre esa actividad, incluida información como el vídeo, identificaciones de dispositivos, direcciones IP, datos de cookies y ubicación».

Sin embargo, no se refiere explícitamente a dispositivos Android, sólo a servicios de Google. También establece: «Cuando utilizas los servicios de Google, podemos recopilar y procesar información sobre tu ubicación real».

61. En un dispositivo Android, incluso si el usuario apaga el wifi, la ubicación del dispositivo aún se rastrea. Para evitar dicho seguimiento, el escaneo de wifi debe estar explícitamente deshabilitado en una acción de usuario separada. La ubicuidad de los concentradores wifi ha hecho que el seguimiento de la ubicación sea bastante frecuente. Según el estudio de Schmidt, durante una breve caminata de quince minutos por un barrio residencial, un dispositivo Android envió nueve solicitudes de ubicación a Google. La solicitud contenía colectivamente ~ 100 BSSID únicos de puntos de acceso wifi públicos y privados.

62. Quartz capturó transmisiones de información del «historial de ubicaciones» en tres teléfonos de diferentes fabricantes, ejecutando varias versiones recientes de Android. Para lograrlo, crearon una red wifi portátil conectada a internet que podía escuchar y transmitir todas las transmisiones de los dispositivos conectados a ella. Ninguno de los dispositivos tenía tarjetas SIM insertadas.

Se pasearon por áreas urbanas, centros comerciales, tiendas, restaurantes y bares. El equipo de Quartz registró todas las solicitudes de red realizadas por los tres móviles objeto del estudio: Google Pixel, Samsung Galaxy S8 y Moto Z Droid. De acuerdo con su análisis de las transmisiones de los teléfonos, observaron que, al menos, los dispositivos comunicaban periódicamente a los servidores de Google para alimentar el «historial de ubicaciones» la siguiente información:

– Una lista del tipo de movimientos que el teléfono cree que el usuario podría estar haciendo, por ejemplo, caminar un 51 por ciento, en bicicleta un 4 por ciento, o en un vehículo a motor 3 por ciento.
– La presión barométrica.
– Si está conectado o no al wifi.
– La dirección MAC, que es un identificador único, del punto de acceso wifi al que está conectado.
– La dirección MAC, la potencia de la señal y la frecuencia de cada punto de acceso wifi cercano.
– La dirección MAC, el identificador, el tipo y las dos medidas de intensidad de señal de cada baliza Bluetooth cercana.
– El nivel de carga de la batería del teléfono y si el teléfono se está cargando o no.
– El voltaje de la batería.
– Las coordenadas GPS del teléfono.
– La elevación GPS.

Esta prueba demostró que Google posee datos más extensos sobre los usuarios de Android de lo que los usuarios creen o sospechan. Esa información se utiliza para vender publicidad dirigida pero también para rastrear en las tiendas a aquellos consumidores que vieron anuncios publicitarios en sus teléfonos u ordenadores. Eso supone que los Estados, sus fuerzas policiales o los tribunales pueden solicitar datos detallados sobre el paradero de una persona.

63. El 26 de febrero de 2018 se publicó un documento inicial abriéndose un período de contribuciones hasta el 3 de abril de 2018. El informe preliminar debe presentarse al Tesorero antes del 3 de diciembre de 2018, y el informe final debe presentarse antes del 3 de junio de 2019.

64. Google emitió un comunicado en el que afirmó que el historial de ubicaciones es completamente optativo, y sólo se usa para mejorar los productos y servicios de Google, como las predicciones de tráfico o la localización de un teléfono perdido. Agregó que los datos enviados y recibidos desde los dispositivos Android pueden enviarse a través de una red wifi o móvil, y que los tipos y la cantidad de datos recopilados dependen de su configuración y de los productos o servicios que estén utilizando los usuarios.

Si se usa un dispositivo Android con aplicaciones de Google, ese dispositivo se comunica periódicamente con los servidores de Google para proporcionar información sobre el dispositivo y la conexión a Google. Esta información incluye elementos como el tipo de dispositivo, el nombre del operador, los informes de fallos y las aplicaciones instaladas.

Estas transferencias de datos no sólo generan nuevas preocupaciones de privacidad sino que suponen un coste para los usuarios que pagan las comunicaciones que Google usa para recibir los datos sin conocimiento del titular de la línea.

65. Romm, Tony. «Google's location privacy practices are under investigation in Arizona.» *The Washington Post*, 11 de septiembre de 2018. <https://www.washingtonpost.com/technology/2018/09/11/googles-location-privacy-practices-are-under-investigation-arizona>.

66. Lardinois, Frederic. «Google says there are now 2 billion active Chrome installs.» *TechCrunch*, 10 de noviembre de 2016. Disponible en: <https://techcrunch.com/2016/11/10/google-says-there-are-now-2-billion-active-chrome-installs/>.

67. «Using Gmail? You will be force logged into Chrome.» *Hacker News*. <https://news.ycombinator.com/item?id=17942252>.

68. Según el estudio de Schmidt, Douglas C. (Google Data Collection), durante un período de 24 horas, el dispositivo Android comunicó aproximadamente 900 muestras de datos a una variedad de puntos finales del servidor de Google. De éstos, aproximadamente un 35 por ciento estaba relacionado con la ubicación.

Los dominios de anuncios de Google recibieron aproximadamente un 3 por ciento del tráfico, lo que se debe principalmente al hecho de que el navegador móvil no se usó activamente durante el período de recopilación.

Aproximadamente, el 62 por ciento restante de las comunicaciones con los dominios del servidor de Google se dividieron entre las solicitudes a la tienda Play App de Google, las cargas de Android de datos relacionados con el dispositivo (como informes de fallos y la autorización del dispositivo) y otros datos predominantemente de la categoría llamadas y actualizaciones de servicios de Google.

El dispositivo iPhone se comunicó con los dominios de Google en más de un orden de magnitud de frecuencia más baja que el dispositivo Android (aproximadamente 50 x). Google no recopiló ninguna ubicación de usuario durante el período de prueba de 24 horas a través de iPhone. Los datos de ubicación representan una fracción muy pequeña (aproximadamente un 1 por ciento) de los datos netos enviados a los servidores de Apple desde el iPhone. Apple recibió las comunicaciones relacionadas con la ubicación una vez al día en promedio. En términos generales, los teléfonos Android comunicaron 4,4 MB de datos por día (aproximadamente 130 MB por mes) a los servidores de Google, lo que es seis veces más de lo que los servidores de Google comunicaban a través del dispositivo iPhone.

Como recordatorio, este experimento se realizó utilizando un teléfono estacionario sin interacción del usuario. A medida que un usuario se vuelve móvil y comienza a interactuar con su teléfono, la frecuencia de las comunicaciones con los servidores de Google aumenta considerablemente.

69. Google recopiló la ubicación a una tasa más alta 1.4 x en comparación con las pruebas en un teléfono estacionario sin interacción del usuario. Los servidores de Google comunicaron 11,6 MB de datos por día (o 0,35 GB/mes) con el dispositivo Android. Este experimento sugiere que incluso si un usuario no interactúa con ninguna aplicación clave de Google, éste puede recopilar información considerable a través de sus productos publicitarios y editores.

70. DoubleClick y Google Analytics (GA) son los productos líderes de Google en el seguimiento del comportamiento de los usuarios y análisis de tráfico de páginas web en PC de escritorio y dispositivos móviles. GA es utilizado por aproximadamente un 75 por ciento de los 100.000 sitios web más visitados.

Las cookies de DoubleClick están asociadas a más de 1,6 millones de sitios web. GA utiliza piezas cortas de código de seguimiento (llamadas «etiquetas de página») integradas en el código HTML de un sitio web. Después de que una página web se cargue, el código GA llama a un archivo «analytics.js» en los servidores de Google. Este programa transfiere una instantánea «predeterminada» de los datos del usuario en ese momento, que incluye la dirección de la página web visitada, el título de la página, la información del navegador, la ubicación (derivada de la dirección IP) y la configuración del idioma del usuario. Los scripts de GA usan cookies para rastrear el comportamiento del usuario.

El script GA, la primera vez que se ejecuta, genera y almacena una cookie específica del navegador en el ordenador del usuario. Esta cookie tiene un identificador de cliente único o ID de cliente.

Google usa el identificador único para vincular las cookies almacenadas previamente que capturan la actividad de un usuario en un dominio en particular, siempre y cuando la cookie no caduque o el usuario no borre las ya almacenadas en su navegador. Si bien una cookie de GA es específica del dominio particular del sitio web que visita el usuario (llamada «cookie de primera parte»), una cookie de DoubleClick suele estar asociada a un dominio de terceros (como doubleclick.net). Google utiliza dichas cookies para rastrear la interacción del usuario en varios sitios web de terceros. Cuando un usuario interactúa con un anuncio publicitario en un sitio web, las herramientas de seguimiento de conversiones de DoubleClick (por ejemplo, Floodlight) coloca cookies en el ordenador del usuario y genera una ID de cliente única. A partir de entonces, si el usuario visita el sitio web anunciado, el servidor de DoubleClick accede a la información de cookies almacenada, registrando así la visita como una conversión válida.

71. AdSense, AdWords y AdMob.

AdSense y AdWords son herramientas de Google que publican anuncios en sitios web y en los resultados de búsqueda de Google, respectivamente. Más de 15 millones de sitios web tienen instalado AdSense para mostrar anuncios patrocinados. Asimismo, más de 2 millones de sitios web y aplicaciones que forman la Red de Display de Google (GDN) y que llegan a más del 90 por ciento de los usuarios de internet muestran anuncios de AdWords.

AdSense recopila información sobre si un anuncio se ha mostrado en la página web del editor. También recopila cómo el usuario interactuó con el anuncio, si ha hecho clic en un anuncio o cómo ha movido el cursor sobre el anuncio.

AdWords permite a los anunciantes publicar anuncios en la búsqueda de Google, mostrarlos en páginas de editores y anuncios superpuestos en vídeos de YouTube. Para rastrear las tasas de conversión y clics del usuario, los anuncios de AdWords colocan una cookie en los navegadores de los usuarios para identificar al mismo usuario si luego visitan el sitio web del anunciante o completan una compra.

Si bien AdSense y AdWords recopilan datos en dispositivos móviles, su capacidad para obtener información del usuario en estos dispositivos es limitada debido a que las aplicaciones móviles no comparten datos de cookies entre ellos, una técnica de aislamiento conocida como *sandboxing* —que dificulta a los anunciantes rastrear el comportamiento del usuario a través de aplicaciones móviles.

Para abordar este problema, Google y otras empresas utilizan «bibliotecas de anuncios» móviles (como AdMob), que sus desarrolladores integran en las aplicaciones para publicar anuncios en aplicaciones móviles. Estas bibliotecas se compilan y ejecutan con las aplicaciones, y envían a Google datos que son específicos de la aplicación a la que pertenecen, incluidas las ubicaciones de GPS, la marca del dispositivo y el modelo del dispositivo cuando las aplicaciones tienen los permisos adecuados.

Como se observa a través del análisis del tráfico de datos, y se confirma a través de las propias páginas web de desarrolladores de Google, dichas bibliotecas también pueden enviar datos personales del usuario, como edad y sexo a Google siempre que los desarrolladores transfieran explícitamente estos valores a la biblioteca.

72. Sin embargo, se necesita un análisis adicional para determinar si dicho intercambio de información ocurre con una cierta periodicidad o si se desencadena por actividades específicas en los teléfonos. La identificación de las cookies de DoubleClick se conecta con la información personal del usuario en la cuenta de Google.

Los datos anónimos de los equipos de sobremesa/portátiles se recopilan a través de identificadores basados en cookies (por ejemplo, Cookie ID), que normalmente se generan mediante los productos publicitarios y editores de Google (por ejemplo, DoubleClick) y se guardan en el almacenamiento local del usuario.

El análisis de Schmidt en su documento Google Data Collection propone un experimento para evaluar si Google puede conectar tales identificadores (y, por tanto, la información asociada a ellos) con la información personal de un usuario. A tal fin:

1. Se abrió de nuevo el navegador sin cookies, en modo privado o de incógnito.
2. Se visitó un sitio web de terceros que usaba la red publicitaria DoubleClick de Google.
3. Se visitó a continuación un sitio web de un servicio de Google ampliamente utilizado, Gmail en este caso.
4. Entrar en Gmail.

Después de completar los pasos 1 y 2, como parte del proceso de carga de la página, el servidor de DoubleClick recibió una solicitud cuando el usuario visitó por primera vez el sitio web de terceros. Esta solicitud formaba parte de una serie de peticiones que comprendían el proceso de inicialización de DoubleClick por parte del sitio web del editor, que dio como resultado que el navegador Chrome configurara una cookie para el dominio de DoubleClick. Esta cookie permanece en el ordenador del usuario hasta que expire o hasta que el usuario borre las cookies manualmente a través de la configuración del navegador.

A continuación, en el paso 3, cuando el usuario visitó Gmail, se le solicitó que iniciase sesión con sus credenciales de Google. Google administra la identidad utilizando una arquitectura de «inicio de sesión único (SSO)», mediante la cual se proporcionan credenciales a un servicio de cuenta (accounts.google. com) a cambio de un «token de autenticación», que luego se puede presentar a otros servicios de Google para identificar a los usuarios.

En el paso 4, cuando un usuario accede a su cuenta de Gmail, inicia sesión efectivamente en su cuenta de Google, que luego le proporciona a Gmail un token de autorización para verificar la identidad del usuario.

En el último paso de este proceso de inicio de sesión, se envía una solicitud al dominio de DoubleClick. Esta solicitud contiene tanto el token de autenticación proporcionado por Google como el conjunto de cookies de seguimiento cuando el

usuario visitó el sitio web de terceros en el paso 2. Esto permite que Google conecte las credenciales de Google del usuario con un ID de cookie de DoubleClick.

Por tanto, si los usuarios no borran regularmente las cookies del navegador, su información de navegación en páginas web de terceros que usan servicios de DoubleClick podría asociarse con su información personal en la cuenta de Google.

73. Es el motor de buscador web más popular en el mundo, con más de 11 mil millones de consultas por mes en Estados Unidos. Además de publicar resultados de páginas web clasificadas en respuesta a consultas generales de los usuarios, Google opera con otros buscadores basados en herramientas, como Google Finance, Vuelos, Noticias, Académico, Patentes, Libros, Imágenes, Vídeos y Hoteles.

74. YouTube proporciona a los usuarios una plataforma para subir y ver contenido de vídeo. Tiene más de 180 millones de usuarios sólo en Estados Unidos. Y tiene la distinción de ser el segundo sitio web más visitado en Estados Unidos, sólo detrás del buscador de Google. Entre las empresas de medios de transmisión en línea, YouTube tiene casi el 80 por ciento del mercado en términos de visitas mensuales de usuarios.

75. Ante algunas preguntas de senadores estadounidenses, Google reconoció que permite que terceras partes obtengan y compartan datos de cuentas de su servicio de correo electrónico Gmail. Google permite que desarrolladores de software analicen los contenidos de los correos electrónicos para conocer mejor las prácticas de los usuarios mediante la detección de palabras clave.

«Google reconoce que sigue permitiendo a terceros acceder a datos de Gmail.» *La Vanguardia*, 20 de septiembre de 2018. <https://www.lavanguardia.com/tecnologia/20180920/451935626701/google-gmail-privacidad-datos-correos.html>.

76. Biersdorfer, J. D. «Getting social with Waze.» *The New York Times*, 20 de septiembre de 2017. Disponible en: <https://www.nytimes.com/2017/09/20/technology/personaltech/getting-social-with-waze.html>.

77. Ayuda, privacidad, Google Cloud. Disponible en: <https://support.google.com/googlecloud/answer/6056650?hl=es>.

78. Después de que un usuario se registre en una cuenta de Google, se le ofrece espacio de almacenamiento gratuito (actualmente 15 GB) en Google Drive para compartir productos como Gmail, Fotos y Documentos.

Los términos de servicio de Google indican que Google establece una licencia para usar los datos almacenados de varias maneras, incluida la reproducción, modificación, comunicación y publicación. («Términos de servicio de Google Drive», Google. Disponible en: <https://www.google.com/drive/terms-of-service>.)

Aunque todos los datos de usuario almacenados en Google Drive están cifrados, no se trata de un «cifrado de cero conocimiento», ya que Google administra la clave de cifrado de datos.

O'Sullivan. Fergus. «What-exactly is zero knowledge in the cloud and how does it work?» *Cloudwards*, 16 de junio de 2017. Disponible en: <https://www.cloudwards.net/what-exactly-is-zero-knowledge-in-the-cloud-and-how-does-it-work>.

79. Centro de aprendizaje de GSuite . Disponible en: <https://gsuite.google.com/learning-center/products/hangouts/get-started/#!/>.

80. Además del Hangouts centrado en la empresa, Google ofrece una aplicación de mensajería instantánea llamada Allo, disponible para Android e iOS o a través de navegadores web.

Keach Sean. «Google Allo just got a major upgrade — but do we really need it?» *Trusted Reviews*, 16 de agosto de 2017. Disponible en: <http://www.trustedreviews.com/news/google-allo-3261336>.

Google registra y almacena todos los mensajes que se comunican a través de Allo de manera predeterminada (a menos que el usuario lo use en incógnito).

Russell Brandom. «Google backs off on previously announced Allo privacy feature.» *The Verge*, 21 de septiembre de 2016. Disponible en: <https://www.theverge.com/2016/9/21/12994362/allo-privacy-message-logs-google>.

Google ofrece una aplicación llamada Google Duo dedicada al videochat móvil. Está disponible tanto para Android como para iOS, pero no en ordenadores de escritorio o portátiles. Los usuarios pueden llamar o videochatear e incluso conectarse con usuarios del sistema operativo Android que no hayan descargado la aplicación.

Krishna, Swapna. «Google Duo allows you to call people who don’t have the app.» *Engadget*, 12 de enero de 2018. Disponible en: <https://www.engadget.com/2018/01/12/google-duo-call-android-users-without-app>.

Una vez instalado, Google Duo tiene acceso a los datos de perfil del usuario, contactos, cámara, micrófono, información de conexión wifi, ID de dispositivo e información de llamada (Google informa al usuario de este acceso cuando el usuario descarga la aplicación). Google también almacena las marcas de tiempo de cuando se usan las aplicaciones y proporciona esta información a los usuarios a través de Google Takeout.

81. Bell, Karissa. «Google+ isn’t dead and these are the people astill using it the most.» *Mashable*, 18 de enero de 2017. Disponible en: <https://mashable.com/2017/01/18/who-is-using-google-plus-anyway/#_lP0PXr.eiqZ>.

82. Google Cloud, pricing. <https://cloud.google.com/translate/pricing>.

83. Los esfuerzos de Google para desplegar redes extensas de cables de fibra óptica se vieron obstaculizados por los costes físicos y las negociaciones con los municipios locales. Por tanto, Google Fiber ahora se está expandiendo a nuevas ciudades a través de Webpass (un proveedor de servicios de internet adquirido por Google), que ofrece velocidades similares a través de las tecnologías inalámbricas y Ethernet existentes.

84. AMP convierte el código HTML y JavaScript convencional en una versión más simplificada desarrollada por Google y guarda en el caché las páginas web validadas por AMP en la red de servidores de Google para un acceso más rápido.

AMP ofrece enlaces de páginas a través de los resultados de búsqueda de Google, así como de plataformas de terceros, como LinkedIn y Twitter.

Como informa la propia página de AMP: «El ecosistema de AMP incluye 25 millones de dominios, más de 100 proveedores de tecnología y plataformas líderes, que abarcan las áreas de publicación, publicidad, comercio electrónico, empresas locales y pequeñas, ¡y más!».

85. AMP está centrado en el usuario, es decir, ofrece una experiencia de navegación mucho más rápida y mejorada a los usuarios sin el desorden de ventanas emergentes y barras laterales.

Aunque AMP es un cambio importante en la forma de almacenar el contenido y entregarlo a los usuarios, la política de privacidad de Google asociada con AMP es bastante general. En particular, Google puede recabar información sobre el uso de la página web (por ejemplo, registros del servidor y dirección IP) a los servidores de caché AMP. Además, las páginas se convierten en AMP mediante el uso de AMP API.

Así pues, Google puede acceder a aplicaciones o sitios web («clientes API») y utilizar cualquier información enviada a través de la API de acuerdo con sus políticas generales. Al igual que las páginas web comunes, las páginas web de AMP realizan un seguimiento de los datos de uso a través de Google Analytics y DoubleClick. En particular, recopilan información sobre datos de las páginas (por ejemplo, dominio, ruta y título de página), datos de usuario (por ejemplo, ID de cliente, zona horaria), datos de navegación (por ejemplo, identificador único de vista de página y referencia), información del navegador e interacción y datos de eventos. Aunque los modos de recopilación de datos de Google no han cambiado con AMP, la cantidad de datos recopilados ha aumentado, ya que los visitantes gastan un 35 por ciento más de tiempo en contenido web que carga con Google AMP en comparación con páginas móviles estándar.

86. Aunque Chromebook tiene una fracción muy pequeña del mercado de PC, está creciendo rápidamente, especialmente en dispositivos usados en escuelas donde cuenta con un 59,8 por ciento del mercado según los datos del segundo trimestre de 2017.

87. Los Chromebooks también permiten cookies de Google y de terceros para rastrear la actividad de los usuarios, de forma similar a cualquier otro dispositivo portátil o de PC. Muchas escuelas usan Chromebooks para acceder a los productos de Google a través de su servicio GSuite for Education.

Google afirma que los datos recopilados durante dicho uso no se utilizan para publicidad dirigida. Sin embargo, los anuncios de los estudiantes se muestran si utilizan servicios adicionales (como YouTube o Blogger) en los Chromebooks proporcionados a través de sus instituciones educativas.

88. «FTC Complaint, request for investigation, injunction, and other relief submitted.» *The Electronic Privacy Information Center*. Disponible en: <https://epic.org/privacy/ftc/google/EPIC-FTC-Google-Purchase-Tracking-Complaint.pdf>.

89. Bergen, Mark; y Surane, Jennifer. «Google and Mastercard Cut a Secret Ad Deal to Track Retail Sales.» *Bloomberg*, 30 de agosto de 2018. <https://www.bloomberg.com/news/articles/2018-08-30/google-and-mastercard-cut-a-secret-ad-deal-to-track-retail-sales>.

90. Escenario 2: «Omega. Cybersecurity Futures 2020». Center for LongTerm Cybersecurity University of California, Berkeley. 2016. <https://cltc.berkeley.edu/scenarios/>.

91. Gartner IT Big data glosario. <http://www.gartner.com/it-glossary/big-data>.

92. Jackson, Sean. «Big data in big numbers - it's time to forget the "three Vs" and look at real-world figures.» *Computing*, 18 de febrero de 2016. <https://www.computing.co.uk/ctg/opinion/2447523/big-data-in-big-numbers-its-time-to-forget-the-three-vs-and-look-at-real-world-figures>.

93. «Big data, artificial intelligence, machine learning and data protection.» 20170904 Version: 2.2. ICO (Information Commissioner’s Office).

94. Pérez-Chirinos, César. *Algoritmos y programación*.

95. «An executive's guide to AI.» *McKinsey Analytics*. <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/an-executives-guide-to-ai>.

96. Citada en la Guía de McKinseyAnalytics.

97. 2017: Google presenta TPU actualizado que acelera los procesos de aprendizaje automático. Google presenta su unidad de procesamiento (TPU) en 2016, que ejecutaba sus propios modelos de aprendizaje automático de 15 a 30 veces más rápido que las GPU y las CPU.

En 2017, Google anunció una versión mejorada del TPU más rápida (180 millones de teraFLOPS, más cuando se combinan múltiples TPU), que podría usarse para entrenar modelos además de ejecutarlos y se ofrecería al público que paga a través de la nube. La disponibilidad de TPU podría generar aún más (y más potentes y eficientes) aplicaciones empresariales basadas en el aprendizaje automático.

En el mismo año, los usuarios de dispositivos electrónicos generaron 2,5 quintillones de bytes de datos por día. Según esta estimación, alrededor del 90 por ciento de los datos mundiales se produjeron en los últimos dos años. Cada minuto, los usuarios de YouTube vieron más de cuatro millones de vídeos y los usuarios móviles enviaron más de 15 millones de mensajes de texto.

También en 2017, AlphaZero venció a AlphaGo Zero tras aprender a jugar tres juegos diferentes en menos de 24 horas (Go, ajedrez y shogi...). Si bien la creación de software de inteligencia artificial con inteligencia general completa permaneció desactivada, DeepMind de Google dio un paso más hacia la inteligencia general con AlphaZero.

A diferencia de AlphaGo Zero, que recibió instrucciones de expertos humanos, AlphaZero aprendió estrictamente jugando solo: tras ocho horas, fue capaz de derrotar a su antecesor AlphaGo Zero así como a algunos de los mejores jugadores de ajedrez del mundo.

98. Citton, Yves. *The Ecology of Attention*. Cambridge, Reino Unido: Polity, 2017.

99. Zuboff, Shoshana. «Big Other: Surveillance Capitalism and the Prospects of an Information Civilization.» *Journal of Information Technology 30*, n.º 1, 1 de marzo de 2015, pp. 75-89. <https://doi.org/10.1057/jit.2015.5.>.

100. Esa cuantificación a menudo se basa en fundamentos muy limitados: conjuntos de datos como AVA que muestra principalmente a mujeres en la categoría de «jugar con niños», y hombres en la categoría de «dar patadas a una persona».

101. O’Neil, *Cathy.Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Random House LCC US. 2016.

102. Larson, Jeff; Mattu, Surya; Kirchner, Lauren; y Julia Angwin. «How We Analyzed the COMPAS Recidivism Algorithm.» *Propublica*, 23 de mayo de 2016. <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>.

103. Whitaker, Roger M.; Colombo, Gualtiero B.; y Rand, David G. «Indirect Reciprocity and the Evolution of Prejudicial Groups.» *Scientific Reports*, 2018; 8 (1). <DOI: 10.1038/s41598-018-31363-z>.

104. «National non-discrimination and equality Tribunal of Finland. Decision: Assessment of creditworthiness, authority, direct multiple discrimination, gender, language, age, place of residence, financial reasons, conditional.» *Plenary session (voting)*, 21 de marzo de 2018. <https://www.yvtltk.fi/en/index/opinionsanddecisions/decisions.html>.

105. Sus declaraciones originales y completas fueron: «*It literally changes your relationship with society, with each other. It probably interferes with productivity in weird ways. God only knows what it's doing to our children's brain ... How do we consume as much of your time and conscious attention as possible?... —the "like" button would give users— a little dopamine hit —to encourage them to upload more content— [...]. It's a social-validation feedback loop [...] exactly the kind of thing that a hacker like myself would come up with, because you're exploiting a vulnerability in human psychology*».

«Ex-Facebook president Sean Parker: site made to exploit human "vulnerability".» *The Guardian*, 29 de noviembre de 2017. <https://www.theguardian.com/technology/2017/nov/09/facebook-sean-parker-vulnerability-brain-psychology>.

106. Busby, Mattha. «Social media copies gambling methods “to create psychological cravings”», 8 de mayo de 2018. <https://www.theguardian.com/technology/2018/may/08/social-media-copies-gambling-methods-to-create-psychological-cravings>.

107. Schüll, Natasha Dow. *Addiction by Design. Machine Gambling in Las Vegas*. Princeton University Press. 2014.

108. Eyal, Nir. «Hooked: How to Build Habit-Forming Products», 4 de noviembre de 2014. Ryan Hoover.

109. Cherry, Kendra. «What Is a Schedule of Reinforcement? What impact do schedules of reinforcement have on learning?», 13 de diciembre de 2017. <https://www.verywellmind.com/what-is-a-schedule-of-reinforcement-2794864>.

110. Mark Griffiths, profesor de adicción conductual y director de la Unidad de Investigación de Juegos de la Universidad de Nottingham (Trent University's International Gaming Research Unit).

111. Harris, Tristan. *Ensayos*. <http://www.tristanharris.com/essays/>.

112. Harris, Tristan. «How Technology Hijacks People's Minds—from a Magician and Google's Design Ethicist», 19 de mayo de 2016. <http://www.tristanharris.com/2016/05/how-technology-hijacks-peoples-minds-from-amagician-and-googles-design-ethicist/>.

113. «Food and Brand Lab. Cornell University.» *Bottomless bowls*. <http://foodpsychology.cornell.edu/discoveries/bottomless-bowls>.

114. *Ensayos*, *op. cit*.

115. Menárguez, Ana Torres. «Los renegados de Silicon Valley que ahora quieren frenar a las tecnológicas.» *El País*, 9 de agosto de 2018.

116. Subramanian, Samanth. «Inside the Macedonian Fake-News Complex», 15 de febrero de 2017. <https://www.wired.com/2017/02/veles-macedonia-fake-news/>.

117. Silverman, Craig; y Alexander, Lawrence. «How Teens In The Balkans Are Duping Trump Supporters With Fake News.» *BuzzFeed*. <https://www.buzzfeednews.com/article/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo#.fu2okXaeKo>.

118. Remnick, David. «Obama Reckons with a Trump Presidency.» *The Newyorker*. <https://www.newyorker.com/magazine/2016/11/28/obama-reckons-with-a-trump-presidency>.

119. O’Reilly, Lara. «Google Issuing Refunds to Advertisers Over Fake Traffic, Plans New Safeguard.» *The Wall Street Journal*, 25 de agosto de 2017. <https://www.wsj.com/articles/google-issuing-refunds-to-advertisers-over-fake-traffic-plans-new-safeguard-1503675395>.

120. Bruell, Alexandra. «Ad Fraud Declines Offer Hope as Marketers Fight Sophisticated Bots.» *The Wall Street Journal*, 24 de mayo de 2017. <https://www.wsj.com/articles/ad-fraud-declines-offer-hope-as-marketers-fight-sophisticated-bots-1495645098>.

121. McNamee, Roger. «How to Fix Facebook—Before It Fixes Us.» *Washington Monthly*, enero-marzo de 2018. <https://washingtonmonthly.com/magazine/january-february-march-2018/how-to-fix-facebook-before-it-fixes-us/>.

122. Eli Pariser acuñó el término en una charla TED en 2011 y lo desarrollo en su libro *El filtro burbuja: Cómo la web decide lo que leemos y lo que pensamos* (Taurus, 2017), traducción de su obra en 2011 *The Filter Bubble: What the Internet Is Hiding from You* (Penguin Press, Nueva York) sobre las formas en que las plataformas digitales que gobiernan nuestro mundo controlan y dan forma a lo que vemos y deciden a qué prestamos atención.

123. Cruz, Eva. «Así era el futuro hace cincuenta años.» *El País*, 6 de febrero de 2018.

124. Resolución del Parlamento Europeo, de 25 de octubre de 2018, sobre la utilización de los datos de los usuarios de Facebook por parte de Cambridge Analytica y el impacto en la protección de los datos (2018/2855[RSP]) <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P8-TA-2018-0433+0+DOC+XML+V0//ES&language=ES>.

125. Swartz, Jon. «Soros: Beware IT Monopolies Facebook, Google.» *Barrons*, 25 de enero de 2018. <https://www.barrons.com/articles/soros-beware-it-monopolies-facebook-google-1516923914>.

126. Kalra, Aditya; y Shah, Aditi. «India's antitrust watchdog fines Google for abusing dominant position.» Reuters, 8 de febrero de 2018. <https://www.reuters.com/article/us-india-google-antitrust/indias-antitrust-watchdog-fines-google-for-abusing-dominant-position-idUSKBN1FS2AD>.

127. La investigación, iniciada en abril de 2015, se concretó en un pliego de cargos en el que se acusaba a Google de:

- Obligar a preinstalar su buscador y su navegador (Chrome) a los fabricantes de móvil que quisieran utilizar Android.
- Exigir a los fabricantes que se comprometiesen a no modificar el sistema operativo si querían utilizar las aplicaciones desarrolladas por Google.
- Conceder incentivos financieros a los fabricantes de teléfonos y de tabletas, así como a los operadores de telefonía, para que se comprometan a utilizar exclusivamente sus productos.

128. Lynn, Barry; y Stoller, Matt. «How to stop Google and Facebook from becoming even more powerful.» *The Guardian*, 2 de noviembre de 2017. <https://www.theguardian.com/commentisfree/2017/nov/02/facebook-google-monopoly-companies>.

129. Proyecto de ley S. 1989: «To enhance transparency and accountability for online political advertisements by requiring those who purchase and publish such ads to disclose information about the advertisements to the public, and for other purposes». <https://www.congress.gov/bill/115th-congress/senate-bill/1989/text>.

130. En <https://www.ancestry.mx/cs/privacyphilosophy> encontraremos los usos y finalidades que Ancestry da a los datos proporcionados:

> «Ancestry utiliza la información respecto a tu persona principalmente con el fin de proporcionar, personalizar, actualizar, expandir y mejorar los servicios que te ofrecemos. Por ejemplo, utilizamos la información relacionada contigo para crear tu cuenta, brindarte nuestros sitios web y servicios, ayudarte a elaborar árboles genealógicos y realizar pruebas de tu ADN. También podríamos emplear tu información en proyectos de investigación genealógica o genómica, con el objeto de mejorar o desarrollar nuevos productos y servicios, y para fines internos de la empresa. También podríamos utilizar tu información personal para verificar tu identidad, comunicarnos contigo y ofrecerte anuncios publicitarios; mejorar la información de seguridad de Ancestry; ayudarte a crear y brindarte perspectivas sobre tus árboles genealógicos en función de los datos que se encuentran en nuestras bases de datos; ofrecer encuestas y cuestionarios con el fin de recopilar más información de los usuarios para usarla en los servicios y en la publicidad de nuevos productos y ofertas procedentes de nuestra empresa o de socios comerciales que realicen estudios científicos, estadísticos e históricos; detectar y proteger de errores, fraudes u otras actividades criminales y malintencionadas; así como facilitar iniciativas de desarrollo e investigación de productos».

131. Ancestry publica en su web algunos datos sobre las peticiones de acceso a los datos por parte de las autoridades estadounidenses o extranjeras. En su «Informe de transparencia» de 2017 recoge los siguientes datos:

– Recibió 34 solicitudes válidas de autoridades de la ley en relación con usuarios en 2017, y respondió facilitando la información solicitada en 31 de esos 34 casos.
– Todas las solicitudes tenían relación con investigaciones respecto al uso indebido de tarjetas de crédito y con el robo de identidad.
– Rechazaron solicitudes debido a que el solicitante no siguió el proceso legal apropiado.
– No recibieron ninguna solicitud de información relacionada con la salud ni la genética de ningún miembro de Ancestry, y no revelaron, según su información, ningún dato de ese tipo a ninguna agencia encargada del cumplimiento de la ley.
– En cuanto a las solicitudes de seguridad nacional, no recibió ninguna petición de información confidencial de conformidad con las leyes de seguridad de Estados Unidos o de ningún otro país.

132. Martínez Ron, Antonio. «El oráculo de los test de saliva.» *Next*. <https://www.vozpopuli.com/altavoz/next/oraculo-test-saliva_0_1104189582.html>.

133. Fox, Maggie. «Drug giant Glaxo teams up with DNA testing company 23andMe. DNA test results will be used to develop drugs under a new agreement.» *NBC*, 25 de julio de 2018. <https://www.nbcnews.com/health/health-news/drug-giant-glaxo-teams-dna-testing-company-23andme-n894531>.

134. Zhang, Sarah. «The Coming Wave of Murders Solved by Genealogy.» *The Atlantic*, 19 de mayo de 2018.

135. DNA Doe Project. <http://dnadoeproject.org/>.

136. La compañía de Jeff Bezos presentó en septiembre de 2018 varios productos relacionados con su asistente de voz para llevar a Alexa a todos los aspectos de la vida de sus usuarios:

– Echo Auto para incorporar a Alexa al coche con puerto de auriculares y Bluetooth para conectarse con el coche. Amazon ha llegado a acuerdos con Audi para incorporar Alexa en el nuevo E-Tron, sumando éste a los acuerdos que ya tiene con otros fabricantes como SEAT, Volkswagen, BMW o Toyota.

– Un microondas (60 dólares) que llevaría incorporado a Alexa.

– Un enchufe inteligente (25 dólares) sin micrófono pero con conexión a otros dispositivos inteligentes del hogar y con Alexa para su encendido y apagado.

– Altavoces Echo Sub, un subwoofer para añadir a los Echo actuales y dos amplificadores: Echo Amp y Echo Link Amp, con entradas coaxial y óptica.

– También nuevos Echo Dot (3.ª versión) y Echo Plus (2.ª), con mejores altavoces, y nuevo Echo Show de 10” para acompañar al actual de 7”.

Alexa incorpora un nuevo «lenguaje de programación» llamado APL, que permite mejorar las habilidades de sus Echo, tanto con pantalla como sin ella.

137. Jones, Rhett. «Roomba's Next Big Step Is Selling Maps of Your Home to the Highest Bidder.» *Gizmodo*, 24 de julio de 2017. <https://gizmodo.com/roombas-next-big-step-is-selling-maps-of-your-home-to-t-1797187829>.

138. Tenemos definiciones para todos los gustos, desde las que ponen el acento en la comunicación hasta las que se centran en sus elementos o sus funcionalidades.

Del primer tipo es la definición de IoT de la UIT (UIT y el IoT European Research Cluster [IERC]) que la considera como: «Una infraestructura de red global dinámica con capacidades de autoconfiguración basadas en protocolos de comunicación estándar e interoperables donde las cosas físicas y virtuales tienen identidades, atributos físicos y personalidades virtuales, y usan interfaces inteligentes de manera integrada en el red de información».

La consultora Gartner opta por una definición a partir de sus funcionalidades, y define el IoT como «la red de objetos físicos con tecnología embebida que les permite comunicar, sentir o interactuar con su estado interno o con su entorno exterior».

Por su parte, el estudio «Análisis de los vectores de ataque del internet de las cosas (IoT)», considera que un dispositivo IoT es «todo objeto físico o dispositivo inteligente, con capacidades de computación, es decir, que dispone de electrónica y/o de un ordenador embebido (habitualmente de reducido tamaño), y con capacidades de interconexión con redes de datos, ya sean internas (por ejemplo, en el hogar, empresa o entorno industrial) o a internet. El IoT incluye la agrupación y trabajo coordinado de éstos en diversos entornos, el doméstico, los edificios de oficinas, o en movilidad (coches, trenes, ciudades, etc.)».

Así contempla el IoT el CCN-CERT (CCN-CERT BP-05/16 Buenas Prácticas en Internet de las Cosas [junio de 2017]) que la considera como un concepto evolutivo y no estático definiéndolo tanto por su parte física y de conectividad, cosas que sienten, actúan y se comunican (artefactos, vehículos, edificios, electrodomésticos, atuendos, o implantes), como por sus funcionalidades e interdependencias entre sus componentes, las comunicaciones y las plataformas.

139. «FDA approves pill with sensor that digitally tracks if patients have ingested their medication.» *New tool for patients taking Abilify*, noviembre de 2017. <https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm584933.htm>.

140. «Cybersecurity Futures 2020.» *Center for Long-Term Cybersecurity University of California, Berkeley*. 2016. <https://cltc.berkeley.edu/scenarios/>.

141. Según la Wikipedia, «el test de Turing es una prueba de la habilidad de una máquina para exhibir un comportamiento inteligente similar al de un ser humano o indistinguible de éste. Alan Turing propuso que un humano evaluara conversaciones en lenguaje natural entre un humano y una máquina diseñada para generar respuestas similares a las de un humano. El evaluador sabría que uno de los participantes de la conversación era una máquina, y los intervinientes serían separados unos de otros. La conversación estaría limitada a un medio únicamente textual como un teclado de ordenador y un monitor por lo que sería irrelevante la capacidad de la máquina de transformar texto en habla. En el caso de que el evaluador no pueda distinguir entre el humano y la máquina acertadamente (Turing originalmente sugirió que la máquina debía convencer a un evaluador, después de cinco minutos de conversación, el 70 por ciento del tiempo), la máquina habría pasado la prueba. Esta prueba no evalúa el conocimiento de la máquina en cuanto a su capacidad de responder preguntas correctamente, pues sólo se toma en cuenta la capacidad de ésta de generar respuestas similares a las que daría un humano.

»Turing propuso esta prueba en su ensayo “Computing Machinery and Intelligence” en 1950 mientras trabajaba en la Universidad de Mánchester (Turing, 1950; p. 460). Lo inicia así: “Propongo que se considere la siguiente pregunta, ‘¿pueden pensar las máquinas?’”. Como es difícil definir la palabra “pensar”, Turing decide “reemplazar la pregunta con otra que está estrechamente relacionada y en palabras no ambiguas”, la nueva pregunta de Turing es: “¿Existirán ordenadores digitales imaginables que tengan un buen desempeño en el juego de imitación?”. Turing creía que esta pregunta sí era posible de responder y en lo que resta de su ensayo se dedica a argumentar en contra de las objeciones principales a la idea de que “las máquinas pueden pensar”.

»Desde que fue creada por Turing en 1950, la prueba ha demostrado ser altamente influyente y, a la vez ampliamente criticada, además de transformarse en un concepto importante en la filosofía de la inteligencia artificial».

<https://es.wikipedia.org/wiki/Test_de_Turing>.

142. «Hablar con la máquina. La tecnología de reconocimiento de voz transforma nuestro modo de relacionarnos con los dispositivos.» *El País*, 7 de enero de 2018. <https://elpais.com/tecnologia/2018/01/05/actualidad/1515173663_103498.html>.

143. Dans, Enrique. «Asistentes de voz, espionaje y la navaja de Occam.» <https://www.enriquedans.com/2018/05/asistentes-de-voz-espionaje-y-la-navaja-de-occam.html>.

144. Del Rey, Jason. «Here's Amazon's explanation for the Alexa eavesdropping scandal.» *Recode*, 24 de mayo de 2018. <https://www.recode.net/2018/5/24/17391480/amazon-alexa-woman-secret-recording-echo-explanation>.

145. Cook, James. «Amazon patents new Alexa feature that knows when you're ill and offers you medicine.» *Telegraph*, 9 de octubre de 2018. <https://www.telegraph.co.uk/technology/2018/10/09/amazon-patents-new-alexa-feature-knows-offers-medicine/>.

146. Smith, Craig S. «Alexa and Siri Can Hear This Hidden Command. You Can’t», 10 de mayo de 2018.

147. Tilley, Aaron. «Amazon Alexa Can Now Unlock Your Doors.» *Forbes*, 16 de febrero de 2017. <https://www.forbes.com/sites/aarontilley/2017/02/16/amazon-alexa-can-now-unlock-your-front-door>.

148. «Venmo using a simple voice command with Siri.» YouTube. <https://www.youtube.com/watch?v=Jwb4kdDk2wg>.

149. No existe una ley estadounidense contra la transmisión de mensajes subliminales a los humanos, y mucho menos a las máquinas. La Comisión Federal de Comunicaciones desalienta la práctica como «contraria al interés público», y el Código de Televisión de la Asociación Nacional de Radiodifusión prohíbe «transmitir mensajes por debajo del umbral de la consciencia normal». Tampoco se establece nada sobre los estímulos subliminales provenientes de los dispositivos inteligentes.

150. Kleinman, Zoe. «Is your smartphone listening to you?» BBC News, 2 de marzo de 2016. <https://www.bbc.com/news/technology-35639549>.

151. «Are your phones listening to you?» *Pent Test Partners*. <https://www.pentestpartners.com/security-blog/are-your-phones-listening-to-you>.

152. Revisión del código fuente de la aplicación en su versión 6.4.7 publicada el dia 8 de junio de 2018. La Liga alega que la funcionalidad de recopilar información del micrófono y ubicación se ha habilitado el 8 de junio. La version 6.4.0 publicada el 21 de febrero de 2018 con SHA1 efd50120f73c0d674492126ce9e9198da57c8287 tiene la capacidad de recopilar información tanto del micrófono como de la ubicación exactamente de la misma forma que la ultima versión disponible. La primera versión de la aplicación de La Liga que incluye el SDK de la empresa Fluzo para grabar el micrófono es la 6.3.0 publicada el 21 de noviembre de 2017, si bien la primera versión que envía la ubicación y el audio a sus servidores sin informar al usuario es la 6.4.5 publicada el 7 de junio de 2018.

153. Resulta dudoso que fuese capaz de generar un hash progresivo con su aplicación después de grabar el audio y en el propio terminal, lo mismo que hace Shazam —que es capaz de construir una base de datos sobre canciones que son un conjunto finito y concreto—. Parece más problemático que el sonido ambiente correspondiente a un partido de fútbol en un bar, que es algo dinámico y en directo, cuente con una base de datos previa. Lo más probable es que la operación no se realice en el móvil sino que el sonido se envíe a otro servicio para identificarlos.

154. Romero, Pablo. «Cuando el reconocimiento facial destapa tu vida como actriz porno.» *El Español*.

En este artículo se detalla como un fotógrafo ruso, Egor Tsvetkov, llevó a cabo un experimento en el que tomaba fotografías de desconocidos y buscaba, mediante un programa de reconocimiento facial, datos sobre ellos en la ubicua red social rusa Vkontakte, utilizada habitualmente por más de 350 millones de personas en Rusia y en muchas de las antiguas repúblicas soviéticas. En un caso no relacionado, los usuarios del foro ruso Dvach procesaron biométricamente las caras de actrices pornográficas y de fotografías de mujeres desnudas encontradas en distintas páginas, buscaron más información sobre ellas en la misma red, y enviaron dosieres a sus familias siguiendo tácticas de acoso con fines supuestamente moralistas.

El programa utilizado por Tsvetkov, FindFace, ya había sido objeto de un artículo en el blog de la empresa de seguridad Kaspersky, titulado «You can’t replace your face, says face recognition», en el que planteaban precisamente ese problema, la facilidad para simplemente tomar una foto de un desconocido, buscarla mediante sus rasgos biométricos, y encontrar coincidencias en la red que permitiesen obtener más datos para, por ejemplo, llevar a cabo tácticas de doxing o acoso de diversos tipos.

155. Elola, Joseba. «Nos hemos quedado con tu cara.» *El País*, 5 de enero de 2018.

156. Acquisti, Alessandro; Ralph Gross y Fred Stutzman. «Face Recognition and Privacy in the Age of Augmented Reality.» *Journal of Privacy and Confidentiality*. 2014. <https://doi.org/10.29012/jpc.v6i2.638>.

Con una webcam y un buen software de reconocimiento facial, pudieron identificar a uno de cada tres alumnos que circulaban por el campus de la Universidad Carnegie Mellon. No tuvieron más que cruzar la imagen obtenida con las que les ofrecía el motor del buscador de Google o los perfiles de Facebook. En algunos casos, el algoritmo permitía incluso acceder al número de seguridad social del interfecto.

157. «Advances in AI are used to spot signs of sexuality.» *The Economist*, 9 de septiembre de 2017. <https://www.economist.com/science-and-technology/2017/09/09/advances-in-ai-are-used-to-spot-signs-of-sexuality>.

158. El texto de la patente (<https://patents.google.com/patent/US20150242679>) es el siguiente:

«Técnicas para la detección de emociones y entrega de contenido. Se describen técnicas para la detección de emociones y la entrega de contenido. En una implantación, por ejemplo, un componente de detección de emociones puede identificar al menos un tipo de emoción asociado con al menos una característica de emoción detectada. Un componente de almacenamiento puede almacenar el tipo de emoción identificado. Un componente de interfaz de programación de aplicaciones (API) puede recibir la solicitud de una o más aplicaciones para el tipo de emoción y, en respuesta a la solicitud, devuelve el tipo de emoción identificada. Una o más aplicaciones pueden identificar contenido para mostrar según el tipo de emoción identificada. La identificación de contenido para mostrar por una o más aplicaciones basadas en el tipo de emoción identificada puede incluir la búsqueda entre una pluralidad de elementos de contenido, pudiendo cada elemento de contenido estar asociado con uno o más tipos de emociones. Se describen y reivindican otras implantaciones».

159. Fowler, Geoffrey A. «Apple is sharing your face with apps. That's a new privacy worry.» *The Washington Post*, 30 de noviembre de 2017. <https://www.washingtonpost.com/news/the-switch/wp/2017/11/30/apple-is-sharing-your-face-with-apps-thats-a-new-privacy-worry/>.

160. Farivar, Cyrus. «Chicago man sues Facebook over facial recognition use in “Tag Suggestions”.» *Arstechnica*, 4 de agosto de 2015. <https://arstechnica.com/tech-policy/2015/04/chicago-man-sues-facebook-over-facial-recognition-use-in-tag-suggestions/>.

161. Biometric Information Privacy Act. <http://www.ilga.gov/legislation/ilcs/ilcs3.asp?ActID=3004&ChapterID=57>.

162. Brandom, Russell. «Lawsuit challenging Facebook's facial recognition system moves forward.» *The Verge*, 5 de mayo de 2016. <https://www.theverge.com/2016/5/5/11605068/facebook-photo-tagging-lawsuit-biometric-privacy>.

163. Meyer, David. «Facebook faces class action privacy suit in Austria, with most of the world invited to join in.» *Gigaom*, 1 de agosto de 2014. <https://gigaom.com/2014/08/01/facebook-faces-class-action-privacy-suit-in-austria-with-most-of-the-world-invited-to-join-in/>.

164. Caso 362/14, Maximillian Schrems e Irish Data Protection Commissioner. Tribunal de Justicia de la Unión Europea. <https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:62014CJ0362&from=EN>.

165. «UIDAI's Aadhaar Software Hacked, ID Database Compromised, Experts Confirm. Skilled hackers disabled security features of Aadhaar enrolment software, circulated hack on Whatsapp.» *Huffington Post*, 11 de septiembre de 2018. <https://www.huffingtonpost.in/2018/09/11/uidai-s-aadhaar-software-hacked-id-database-compromised-experts-confirm_a_23522472>.

166. Hasbrouck, Edward. «CBP expands partnership with airlines on facial recognition. Papers, please! The identity project.» *Papersplease*. <https://papersplease.org/wp/2018/08/28/cbp-expands-partnership-with-airlines-on-facial-recognition/>.

167. Slaughter. «Anne-Marie. Our Bodies or Ourselves.» *Project Syndicate*, 23 de julio de 2018. <https://www.project-syndicate.org/commentary/dangers-of-biometric-data-by-anne-marie-slaughter-and-stephanie-hare-2018-07>.

168. Hauser, Christine. «Police Use Fitbit Data to Charge 90-Year-Old Man in Stepdaughter's Killing.» *The New York Times*, 3 de octubre de 2018. <https://www.nytimes.com/2018/10/03/us/fitbit-murder-arrest.html>.

169. «Pentagon Restricts Use of Fitness Trackers, Other Devices.» *Associated Press*, 7 de agosto de 2018.

170. Wikipedia: «Se conoce como jaula de Faraday al efecto por el cual el campo electromagnético en el interior de un conductor en equilibrio es nulo, anulando el efecto de los campos externos. Esto se debe a que cuando el conductor está sujeto a un campo electromagnético externo se polariza, de esta manera queda cargado positivamente en la dirección en que va el campo electromagnético, y cargado negativamente en el sentido contrario. Puesto que el conductor se ha polarizado, éste genera un campo eléctrico igual en magnitud pero opuesto en sentido al campo electromagnético, luego la suma de ambos campos dentro del conductor será igual a cero.

»Se pone de manifiesto en numerosas situaciones cotidianas, por ejemplo, el mal funcionamiento de los teléfonos móviles en el interior de ascensores o edificios con estructura de rejilla de acero. Una manera de comprobarlo es con una radio sintonizada en una emisora de onda media. Al rodearla con un periódico, el sonido se escucha correctamente. Sin embargo, si se sustituye el periódico por un papel de aluminio, la radio deja de emitir sonidos: el aluminio es un conductor eléctrico y provoca el efecto jaula de Faraday.

»Este fenómeno, descubierto por Michael Faraday, tiene una aplicación importante en aviones o en la protección de equipos electrónicos delicados, tales como discos duros o repetidores de radio y televisión situados en cumbres de montañas y expuestos a las perturbaciones electromagnéticas causadas por las tormentas».

<https://es.wikipedia.org/wiki/Jaula_de_Faraday>.

Las salas o cajetines de Faraday impiden la entrada de señales de móviles por lo que se usan cuando se quiere proteger de escuchas o interferencias ajenas un dispositivo o una reunión.

171. Sly, Liz. «U.S. soldiers are revealing sensitive and dangerous information by jogging.» *The Washington Post*, 29 de enero de 2018. <https://www.washingtonpost.com/world/a-map-showing-the-users-of-fitness-devices-lets-the-world-see-where-us-soldiers-are-and-what-they-are-doing/2018/01/28/86915662-0441-11e8-aa61-f3391373867e_story.html>.

172. «John Hancock will include fitness tracking in all life insurance policies.» Reuters, 19 de septiembre de 2018. <https://venturebeat.com/2018/09/05/new-government-rules-could-accelerate-telemedicine/>.

173. «Escenario 5. Sensorio (Internet de la Emoción) Cybersecurity Futures 2020.» Center for Long-Term Cybersecurity University of California, Berkeley. 2016. <https://cltc.berkeley.edu/scenarios/>.

174. Hvistendahl, Mara. «Inside China's Vast New Experiment in Social Ranking.» *Wired*, 14 de diciembre de 2017. <https://www.wired.com/story/age-of-social-credit/>.

175. Hvistendahl, Mara, *op. cit.*

176. Hvistendahl, Mara, *op. cit*.

177. Hvistendahl, Mara, *op. cit*.

178. Stallman, Richard. «A radical proposal to keep your personal data safe.» *The Guardian*. <https://amp.theguardian.com/commentisfree/2018/apr/03/facebook-abusing-data-law-privacy-big-tech-surveillance>.

179. Krebs, Brian. «Target Investigating Data Breach.» *Krebs on security*, 18 de diciembre de 2013. <https://krebsonsecurity.com/2013/12/sources-target-investigating-data-breach/>.

180. Fleishman, Glenn. «Equifax Data Breach, One Year Later: Obvious Errors and No Real Changes, New Report Says.» *Fortune*, 8 de septiembre de 2018. <http://fortune.com/2018/09/07/equifax-data-breach-one-year-anniversary/>.

181. «Don’t Let Equifax Put Americans At Risk Again. One year after the historic Equifax data breach, consumer credit data is still vulnerable to hackers.» *Consumers Union*, 7 de septiembre de 2018. <https://www.consumerreports.org/data-theft/dont-let-equifax-crisis-go-to-waste-equifax-data-breach/>.

182. Whittaker, Zack. «Teen phone monitoring app leaked thousands of user passwords.» *Zdnet*, 20 de mayo de 2018. <https://www.zdnet.com/google-amp/article/teen-phone-monitoring-app-leaks-thousands-of-users-data/>.

183. Burgess, Matt. «Here's what you need to do after the huge Facebook hack.» *Wired*, 29 de septiembre de 2018. <https://www.wired.co.uk/article/facebook-hack-data-breach-news-what-to-do>.

184. Brandom, Russell. «The Facebook hack could be Europe's first big online privacy battle. The 50 million compromised accounts are the first major test of the GDPR.» *The Verge*, 1 de octubre de 2018. <https://www.theverge.com/2018/10/1/17922946/facebook-breach-gdpr-lawsuit-privacy-commissioner-europe>.

185. MacMillan, Douglas y Robert. «Google Exposed User Data, Feared Repercussions of Disclosing to Public.» *The Wall Street Journal*, 8 de octubre de 2018. <https://www.wsj.com/articles/google-exposed-user-data-feared-repercussions-of-disclosing-to-public-1539017194>.

186. «Chad Couser in Enterprise Security. The Cost of a Data Breach», 30 de abril de 2018 <https://blog.gemalto.com/security/2018/04/30/data-breach-costs/>.

187. «Ponemon Institute Cost of a Data Breach Study 2018.» <https://securityintelligence.com/series/ponemon-institute-cost-of-a-data-breach-2018/>.

188. Definición de audímetro recogida en Wikipedia. <https://es.wikipedia.org/wiki/Aud%C3%ADmetro>.

189. Coen, Rena; King, Jennifer; y Wong, Richmond. «The Privacy Policy Paradox.» UC Berkeley School of Information.

190. Jensen, C.; Potts, C.; y Jensen, C. «Privacy practices of Internet users: self-reports versus observed behavior.» *International Journal of Human-Computer. Studies*, 63 (1), p. 212. 2005.

Milne, G. R.; y Culnan, M. J. «Strategies for reducing online privacy risks: Why consumers read (or don't read) online privacy notices.» *Journal of Interactive Marketing*, 18 (3), pp. 15-29. 2004.

191. Mcdonald, A. M.; y Cranor, L. F. «The Cost of Reading Privacy Policies. I/S: A Journal of Law and Policy for the Information Society 2008.» *Privacy Year in Review Issue*, 4, p. 543. 2008.

192. *The Biggest Lie*. <http://biggestlie.com/>.

193. Turow, J.; King, J.; Hoofnagle, C. J.; Bleakley, A.; y Hennessy, M. (2009). «Americans reject tailored advertising and three activities that enable it.» <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1478214>.

194. Aleecia M. Mcdonald y Lorrie Faith Cranor. «The Cost of Reading Privacy Policies. I/S: A Journal of Law and Policy for the Information Society», vol. 4, n.º 3 (2008), pp. 543-568. *Ohio State University. Moritz College of Law*. <https://kb.osu.edu/bitstream/handle/1811/72839/ISJLP_V4N3_543.pdf>.

195. House of Commons Science and Technology Committee. «Responsible use of data. HC245», noviembre de 2014. <http://www.publications.parliament.uk/pa/cm201415/cmselect/cmsctech/245/245.pdf>.

196. La lista de documentos es la siguiente: Alexa Terms of Use, Amazon. co.uk Conditions of Use y Sale, Amazon.co.uk Privacy Notice, Cookies y Internet Advertising, Amazon Prime Terms and Conditions, Amazon Music Terms of Use, Amazon Device Terms of Use, iMDb Conditions of Use, Amazon Video Terms of Use, Subscriptions/Promotions Terms, Amazon Assistant Conditions of Use, Alexa Voice Service Agreement, Kindle Store Terms of Use, Alexa Voice Service, Audible Service Conditions of Use, Amazon Photos ToU, Program Requirements y App Distribution Agreement.

197. Executive Office of the President. «Big data: seizing opportunities, preserving values.» Casa Blanca, 1 de mayo de 2014. <https://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_5.1.14_final_print.pdf.>.

198. San José, Julio. «Polisis: una IA que se lee las políticas de privacidad por ti.» *Derecho de la Red*, 20 de febrero de 2018.

199. *Pribot*. <https://pribot.org/>.

200. Harkous[1], Hamza; Fawaz[2], Kassem; Lebret[1], Rémi; Schaub[3], Florian; Shin[3], Kang G.; y Aberer[1], Karl. Polisis: «Automated Analysis and Presentation of Privacy Policies Using Deep Learning». (1. Ecole Polytechnique Fédérale de Lausanne (EPFL); 2. University of Wisconsin-Madison; 3. University of Michigan.) <https://pribot.org/files/Polisis_Technical_Report.pdf>.

201. Warren, Samuel D.; y Brandeis, Louis D. «The Right to Privacy.» *Harvard Law Review*, vol. 4, n.º 5, 15 de diciembre de 1890, pp. 193-220. <http://www.jstor.org/stable/1321160>.

202. Warren, Samuel D.; y Brandeis, Louis D., *op. cit*.: «Que el individuo debe tener plena protección tanto en su persona y como en su propiedad es un principio tan antiguo como la ley común; pero se ha considerado necesario, de vez en cuando, definir de nuevo la naturaleza exacta y el alcance de dicha protección.

»Los cambios políticos, sociales y económicos implican el reconocimiento de nuevos derechos, y el derecho común, en su eterna juventud, crece para satisfacer las demandas de la sociedad.

»Por tanto, en los primeros tiempos, la ley sólo buscó remedio para la interferencia física con la vida y la propiedad, por las transgresiones *vi et arrmis.* Así, mientras el "derecho a la vida" sirvió para proteger al sujeto frente a las diversas formas de agresión física, el derecho a la propiedad aseguró al individuo sus tierras y su ganado. Más tarde, llegó el reconocimiento de la naturaleza espiritual del hombre, de sus sentimientos y su intelecto. Poco a poco se amplió el alcance de estos derechos legales; y ahora el derecho a la vida ha llegado a significar el derecho a disfrutar de la vida, el derecho a ser dejado en paz, el derecho a la libertad garantiza el ejercicio de amplios privilegios civiles, y el término "propiedad" ha crecido para abarcar toda forma de posesión, tangible e intangible.

»Las invenciones recientes y los métodos comerciales llaman la atención sobre el próximo paso que debe darse para la protección de la persona, y para asegurar al individuo lo que el juez Cooley llama el derecho "a ser dejado en paz". Fotografías instantáneas y empresas de periódicos han invadido los recintos sagrados de la vida privada y doméstica, y numerosos dispositivos mecánicos amenazan con cumplir la predicción de que "lo que se susurra en el armario se proclamará desde el tejado"».

203. Tobin, Ariana; y Merrill, Jeremy B. «Facebook is letting job advertisers target only men.» *ProPublica*, 18 de septiembre de 2018. <https://www.propublica.org/article/facebook-is-letting-job-advertisers-target-only-men>.

204. De 91 anuncios de Uber, sólo uno se dirigía a mujeres y tres no se dirigían a un sexo específico.

205. ACLU, the Communications Workers of America y la firma Outten & Golden presentaron una queja ante la Comisión de Igualdad de Oportunidades en la oficina de empleo estadounidense sobre las prácticas de Facebook. La queja, que es el primer paso antes de presentar una demanda, menciona a diez empleadores que habrían anunciado empleos sólo para hombres. La queja sostiene que el propio Facebook habría violado la ley al publicar los anuncios.

206. En algunos casos, las empresas parecen estar apuntando anuncios de trabajo por sexo para diversificar su fuerza laboral o abordar las disparidades. ProPublica encontró anuncios de T-Mobile y Boeing que promocionaban carreras de ingeniería para mujeres.

Pero la queja encontró casos de anuncios que no parecían dirigidos a corregir desequilibrios históricos. Había diez anuncios para trabajar en industrias tradicionalmente dominadas por hombres que estaban dirigidos sólo a hombres, incluida una compañía de software, una empresa de mudanzas y un departamento de Policía.

La mayor parte de las ocasiones, a los anunciantes les costó reconocer que estaban excluyendo a usuarios por razón de su género. La ciudad de Omaha, por ejemplo, anunció un trabajo para un investigador de derechos civiles. Entre los deberes enumerados: «Investigar los cargos de discriminación» y «procesar las quejas que alegan discriminación en el empleo». El anuncio se marcó sólo para ser mostrado a las mujeres.

207. Angwin, Julia; y Parris, Terry. «Facebook lets advertisers exclude users by race.» *ProPublica*, 28 de octubre de 2016. <https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race>.

208. Angwin, Julia (*ProPublica*); Scheiber, Noam (*The New York Times*); y Tobin, Ariana (*ProPublica*). «Dozens of companies are using Facebook to exclude older workers from job ads.» *ProPublica*, 20 de diciembre de 2017. <https://www.propublica.org/article/facebook-ads-age-discrimination-targeting>.

209. McCabe, David; y Fischer, Sara. «Facebook tells advertisers more scrutiny is coming.» *Axios*, 7 de octubre de 2017. <https://www.axios.com/facebook-tells-advertisers-more-scrutiny-is-coming-1513306024-90689d7e-8424-40d0-ba0a-30c09db495cc.html>.

210. Levin, Sam. «Facebook told advertisers it can identify teens feeling “insecure” and “worthless”.» *The Guardian*, 1 de mayo de 2017. <https://www.theguardian.com/technology/2017/may/01/facebook-advertising-data-insecure-teens>.

211. Perrin, Andrew. «Americans are changing their relationship with Facebook.» *Pew Research Center*, 5 de septiembre de 2018. <http://www.pewresearch.org/fact-tank/2018/09/05/americans-are-changing-their-relationship-with-facebook/>.

Datanomics
Paloma Llaneza

Deusto es un sello editorial de Centro de Libros PAPF, SLU.
Av. Diagonal, 662-664
08034 Barcelona
www.planetadelibros.com

Primera edición en libro electrónico (epub): febrero de 2019

ISBN: 978-84-234-3026-0 (epub)

Conversión a libro electrónico: Newcomlab, S. L. L.
www.newcomlab.com