El lado oscuro de las GAFAM: monopolización de los datos y pérdida de privacidad

CARLOS SAURA GARCÍA*
Universitat Jaume I (España)
al340993@uji.es

Resumen

El rápido avance de la digitalización y la *hiperconectividad* de las sociedades modernas en los últimos años ha dado lugar a la *datificación* de la vida de las personas y a la revolución del *big data*. Estos dos fenómenos presentan un gran potencial que puede originar múltiples beneficios en multitud de aspectos de la vida de los ciudadanos, pero también hay que tener en cuenta las implicaciones y los peligros de estos. Este artículo se centra en los peligros provocados por las grandes corporaciones tecnológicas del planeta, las denominadas GAFAM (Google, Amazon, Facebook, Apple y Microsoft). Los procesos de extracción, almacenamiento y análisis de datos de la vida de las personas llevados a cabo por las GAFAM presentan diversas implicaciones peligrosas para los propios ciudadanos entre las que destacan la monopolización de sus datos y la perdida de privacidad. En este artículo se profundizará en estas dos amenazas. Utilizando una metodología hermenéutico-crítica basada en el análisis de fuentes bibliográficas, el objetivo de este artículo es examinar los principales riesgos de las GAFAM para la ciudadanía y proponer soluciones a estos.

Palabras claves: datificación, big data, datos, monopolización, privacidad, GAFAM.

The dark side of GAFAM: Monopolization of data and loss of privacy

Abstract

The rapid advance of digitization and hyperconnectivity of modern societies has led to the datafication of people's lives and the big data revolution. These two phenomena have great potential that can cause multiple benefits in many aspects of citizens' lives, but their implications and dangers must also be considered. This article focuses on the dangers caused by the large technology corporations on the planet, the so-called GAFAM (Google, Amazon, Facebook, Apple, and Microsoft). The processes of extraction, storage, and analysis of data from people's lives carried out by the GAFAM have various dangerous implications for the citizens themselves, among which stand out the monopolization of their data and the loss of privacy. This article will delve into these two threats. Using a hermeneutical-critical methodology based on the analysis of bibliographic sources, the objective of this article is to examine the main risks of GAFAM for citizens and propose solutions to them.

Key words: *datafication, big data, data, monopolization, privacy, GAFAM.*

---

* Graduado en Economía (Universitat Jaume I), máster en Ética y Democracia (Universitat Jaume I y Universidad de Valencia) e investigador predoctoral en el programa de Doctorado en Ética y Democracia (Universitat Jaume I y Universitat de Valencia).

Recibido: 31/enero/2022 - Aceptado: 9/mayo/2022

INTRODUCCIÓN

El fenómeno del *big data* se desarrolla en la gran cantidad de datos extraídos y almacenados diariamente tanto por empresas como por organizaciones públicas y hace referencia al aprovechamiento de grandes conjuntos de datos para obtener nuevas percepciones y formas de creación de valor destinadas a conseguir información sobre determinados temas, realizar predicciones futuras y descubrir relaciones ocultas (Boyd y Crawford, 2012; Mayer-Schönberger y Cukier, 2013; Kitchin, 2014; Puyol, 2014). El nacimiento de este fenómeno está ligado al desarrollo de la esfera sociotecnológica que se ha producido en los primeros años del siglo XXI; los avances producidos en los campos de la tecnología, la computación y el tratamiento de datos han sido fundamentales para el surgimiento y crecimiento de la industria del *big data* (Suárez Gonzalo y Guerrero Solé, 2016; Delgado, 2018).

Una de las principales consecuencias de la revolución del *big data* ha sido la *datificación* de la vida privada de millones de personas (Van Dijck, 2014; Cardon, 2018; Mejias y Couldry, 2019). Los primeros académicos en acuñar el concepto de *datificación* fueron Mayer-Schönberger y Cukier (2013), estos definieron el fenómeno de la *datificación* como la recopilación de información sobre cualquier aspecto de la sociedad y su transformación a un formato de datos para poder cuantificarlos, efectuar un seguimiento a tiempo real y realizar un análisis predictivo. Desde su irrupción la *datificación* ha ido evolucionando hasta convertirse en un fenómeno que va mucho más allá de las funciones descritas por Mayer-Schönberger y Cukier (2013). En los últimos años han apareciendo diversas perspectivas que han criticado y transformando la definición inicial del concepto de *datificación* (Van Dijck, 2014; Mejias y Couldry, 2019; Zuboff, 2019; Southerton, 2020).

El gran crecimiento de los datos y el desarrollo de nuevas técnicas de tratamiento de datos han originado multitud de beneficios para los ciudadanos, pero también algunos peligros (Mayer-Schönberger y Cukier, 2013). La *datificación* y el desarrollo de las técnicas de análisis de datos masivos presentan un gran potencial que puede originar múltiples beneficios en numerosos aspectos de la vida de los ciudadanos, como, por ejemplo, la posibilidad de curar enfermedades, incrementar la seguridad de las sociedades o mejorar el funcionamiento de las ciudades (Mayer-Schönberger y Cukier, 2013; Mejias y Couldry, 2019). No obstante, hay que tener en cuenta las implicaciones y los peligros originados por la expansión de estos fenómenos.

La nueva lógica mundial basada en la extracción y el tratamiento de grandes conjuntos de datos de carácter privado supone un gran riesgo

para las libertades de las personas (Baruh y Popescu, 2017; Zuboff, 2019; Han, 2021). Dos de las cuestiones que más preocupación suscitan en la actualidad en las sociedades modernas son la monopolización de los datos y la perdida de privacidad. Por una parte, la monopolización de los datos hace referencia a la capacidad de extraer y analizar datos, de mantener su control exclusivo y de utilizar estos datos en el momento y finalidad deseada por un grupo reducido de grandes corporaciones tecnológicas (Ramge y Mayer-Schönberger, 2021). La monopolización de los datos por parte de un pequeño grupo de empresas tiene implicaciones negativas para las sociedades modernas como la limitación de la innovación (Ramge y Mayer-Schönberger, 2021), la concentración de poder económico y político (Zuboff, 2019; Webb, 2021) o la limitación de la información, la creación de burbujas de opinión y la manipulación de los propios sistemas democrático (Pariser, 2011; Lanier, 2018; Wylie, 2019). Por otra parte, la perdida de privacidad se vincula con la ruptura de los principios básicos de la privacidad dentro del contexto tecnológico de las sociedades modernas y el aumento de la vigilancia hacia todas las facetas de la vida de los ciudadanos (Suárez Gonzalo, 2019; Zuboff, 2019). La pérdida de privacidad supone entre otras cosas el acceso no autorizado a conjuntos de datos privados (Adams, 2017), la violación de la intimidad y la limitación del desarrollo personal de la ciudadanía (Véliz, 2020) o la creación de un tejido de control y vigilancia social basado en la extracción de datos (Lyon, 2018; Zuboff, 2019; Han, 2021).

El objetivo de este artículo será mostrar y buscar solución a los principales peligros relacionados con la extracción, almacenamiento y análisis de grandes conjuntos de datos de las vidas privadas de los ciudadanos por parte de las GAFAM. Para lograr este objetivo, en un primer momento se expondrá la monopolización de los datos de los ciudadanos por las GAFAM, a continuación, se profundizará en la perdida de privacidad de las personas y finalmente se propondrán soluciones a los principales problemas relacionados con la extracción, almacenamiento y análisis de datos privados de la ciudadanía.

1. LA MONOPOLIZACIÓN DE LOS DATOS

La rápida y amplia introducción de la industria del *big data* en las sociedades modernas han provocado que los datos crezcan exponencialmente y que se conviertan en la principal materia prima para el funcionamiento de esta (Mayer-Schönberger y Cukier, 2013; Delgado, 2018). Un claro ejemplo de esta expansión es la duplicación del volumen de datos que se intercambian a través de las conexiones de red a nivel planetario cada año y medio (López-Portillo, 2018). Esta situación está siendo

aprovechada por las grandes empresas y los gobiernos para obtener y estudiar grandes cantidades de datos (Suárez Gonzalo, 2017). El tratamiento de datos se ha convertido en uno de los grandes negocios del siglo XXI. Sara Suárez Gonzalo expone que:

> […] la explotación de datos masivos tiene dos objetivos principales: el beneficio económico y el control social. El primero suele atribuirse a las empresas y a los mercados (en un sentido amplio, los poderes privados), mientras que el segundo, al ejercicio de las instituciones y gobiernos (poderes públicos) (Suárez Gonzalo, 2019: 15).

Esta cita muestra que tanto las empresas privadas como las instituciones públicas tienen interés en la utilización de la tecnología del *big data* para aumentar su poder, por medio de aumentar sus beneficios en el caso del sector privado y para controlar a los ciudadanos en el caso del sector público. Aunque los objetivos del sector privado y del sector público son diferentes, la relación que existe entre ambos no tiene un carácter excluyente, ambos colectivos mantienen una colaboración continua que facilita la consecución de sus propósitos (Morozov, 2011; Webb, 2021; Westerlund et al., 2021). La rápida digitalización de las sociedades modernas y el desarrollo de la industria del *big data* han potenciado el nacimiento y crecimiento de grandes empresas tecnológicas que ofrecen diversos servicios digitales con la finalidad de extraer y tratar grandes conjuntos de datos. Las principales empresas tecnológicas son las denominadas gigantes digitales, conocidas también como GAFAM (Google, Amazon, Facebook, Apple y Microsoft), estas empresas monopolizan la gran mayoría de datos que difunde diariamente millones de ciudadanos alrededor del mundo (Cardon, 2018; Miguel de Bustos y Moreno Cano, 2018). Las GAFAM se han convertido en las corporaciones más importantes del siglo XXI gracias a la omnipresencia de internet y a las nuevas posibilidades del *big data* (McChesney, 2013; Zuboff, 2019; Webb, 2021). A pesar de que internet fue concebido como una herramienta que daría a la población un poder nunca antes visto en la historia, con el paso del tiempo se ha convertido en una plataforma monopolizada por las grandes empresas tecnológicas (Srnicek, 2018; Ramge y Mayer-Schönberger, 2021). A continuación, se expondrán los principales fenómenos que potencian la monopolización de los datos de la ciudadanía por las grandes corporaciones digitales. Según Suárez Gonzalo (2019) estas tendencias son:

- Los efectos de red
- La publicidad basada en datos

- La importancia de los estándares técnicos
- La codependencia público/privada

Los efectos de red están vinculados con los procedimientos que provocan que cuanto mayor sea el número de personas que utilizan una aplicación o plataforma, más útil y más atractiva será está para los usuarios y más se incrementará los costes de exclusión para los ciudadanos que no la utilizan (McChesney, 2013; Srnicek, 2018). Los efectos de red causan que las aplicaciones y plataformas secundarias tiendan a fracasar y provocan un círculo vicioso en el cual las GAFAM se enriquecen cada vez más y el resto de empresas rivales desaparecen (Suárez Gonzalo, 2019). Este fenómeno termina dando lugar a un incremento de las personas y de las interacciones en las principales aplicaciones y plataformas administradas por las GAFAM, produciendo un incremento de los datos que extraen estas corporaciones y una mayor personalización de contenidos que potencia continuamente los efectos de red (Antunes y Maia, 2018).

El segundo factor que potencia el monopolio de las grandes corporaciones digitales es el negocio de la publicidad basada en datos. La digitalización y la *hiperconectividad* de las sociedades modernas han causado un debilitamiento de los medios de comunicación tradicionales y han originado un nuevo formato comunicativo multidireccional asentado en las aplicaciones y plataformas online (Van Dijck, 2016; De Aguilera y Casero-Ripolles, 2018). Este nuevo formato comunicativo hace posible analizar grades conjuntos de datos y crear campañas publicitarias personalizadas gracias a las innovadores técnicas de procesamiento de estos datos (Nair et al., 2017). Las GAFAM monopolizan el negocio de la publicidad a nivel mundial e ingresan grandes cantidades de dinero gracias a su capacidad para producir una publicidad eficaz y personalizada basada en el análisis y procesamiento del continuo flujo de datos sobre preferencias, comportamientos e intereses que introducen diariamente los ciudadanos en las principales aplicaciones y plataformas de la red (Suárez Gonzalo, 2019).

El tercer factor está vinculado con la importancia de la homogeneización de los estándares técnicos para llevar a cabo un procesamiento de grandes conjuntos de datos que permita extraer de estos una gran cantidad de valor. La creación de estándares en los procedimiento de extracción de datos ha provocado el nacimiento de un potente sector financiero basado en la mercantilización de los datos y liderado por las grandes empresas digitales (Suárez Gonzalo, 2019). El rápido desarrollo de un sector económico basado en la extracción y mercantilización de los datos

de las personas ha creado una nueva forma de capitalismo paralela al capitalismo económico, este es el denominado capitalismo de la vigilancia (Zuboff, 2015).

El capitalismo de la vigilancia es una nueva vertiente del capitalismo que ha introducido los datos como materias primas de la economía, produciendo una mercantilización de los datos privados de la ciudadanía con el objetivo de extraer valor de estos y conseguir beneficios económicos (Zuboff, 2019). El nuevo capitalismo de la vigilancia ha dado lugar a una sociedad constantemente controlada en la que las grandes empresas e instituciones conocen detalladamente las características de los ciudadanos gracias a la extracción y tratamiento de sus datos y los utilizan para conseguir beneficios tanto económicos como políticos (Lyon, 2018; Zuboff, 2019). Las principales propiedades del capitalismo de la vigilancia son la *financierización*, la privatización y la mercantilización de diversos bienes entre los que destacan especialmente los datos de las personas (Suárez Gonzalo, 2019). La *financierización* hace referencia a la extracción de valor a partir de procesos de especulación y depredación, y la privatización y mercantilización se vinculan con la conversión de derechos de propiedad de la ciudadanía en derechos de propiedad privada exclusivos de las corporaciones. La digitalización, la *hiperconectividad* y la *datafication* de todas las facetas de la vida han provocado que las GAFAM se coloquen entre las principales empresas de la economía actual (Srnicek, 2018; Webb, 2021). La base del crecimiento de estas corporaciones es un capitalismo depredador que produce un nivel de riqueza muy bajo y se dedica a incautar el trabajo y los datos privados de las personas para conseguir ganancias (Arruzza et al., 2019). En esta nueva forma de capitalismo, las empresas privadas extraen y utilizan todo tipo de datos de las vidas de los ciudadanos para convertirlos en datos patentados y después crear paquetes de datos predictivos que se ponen a la venta en los mercado para anticipan las decisiones futuras de los posibles consumidores y para mejorar productos y servicios (Zuboff, 2019).

La cuarta y última tendencia es la codependencia entre el sector público y el sector privado. Este es otro de los efectos que incrementa el poder de las GAFAM. La importancia de los datos masivos en las sociedades modernas ha dado lugar a un conjunto de intereses comunes entre las grandes corporaciones tecnológicas y las instituciones públicas (Suárez Gonzalo, 2019; Westerlund et al., 2021). Es importante destacar que la monopolización de internet y el enorme poder que atesoran las grandes empresas digitales sería imposible sin la colaboración entre estas corporaciones y las organizaciones públicas (Morozov, 2011; Ramge y Mayer-Schönberger, 2021). La cooperación entre el sector privado y el sector público permite la obtención de beneficios a ambos grupos. Por

un lado, los gobiernos, las agencias de seguridad nacionales y los ejércitos utilizan los grandes conjuntos de datos que extraen y almacenan las GAFAM, que les sería imposible conseguir de otra forma, para llevar a cabo operaciones de vigilancia y control de la población (Lyon, 2018; Westerlund et al., 2021; Coeckelbergh, 2022). Por otra parte, la obtención de los beneficios de las grandes corporaciones tecnológicas está muy relacionada con las legislaciones estatales y las políticas de privacidad y protección de datos, por lo que la colaboración con las instituciones públicas es fundamental para su viabilidad. Además, hay que destacar también que la estrecha cooperación entre los estados y los gigantes digitales fortalece el poder de ambos sectores por medio de la promulgación de leyes poco exigentes en el campo de la privacidad y la mercantilización de datos por parte de los gobiernos en favor de las GAFAM y la distorsión de la opinión pública por parte de las GAFAM en favor de los gobiernos de los estados (Morozov, 2011; Zuboff, 2019).

En esta primera sección se han mostrado las principales tendencias que hacen posible la monopolización de grandes cantidades de datos de las sociedades modernas por parte de las GAFAM. La monopolización de los datos ha dotado de un gran poder a estas grandes corporaciones y ha dado lugar a un clima de vigilancia y control social. Esta situación se ha visto potenciada por la ruptura de los principios básicos de la privacidad en el actual contexto tecnológico y la consecuente pérdida de privacidad de la ciudadanía. Por ello, la siguiente sección se centrará en las consecuencias de la perdida de privacidad de las personas y en la pasividad de los ciudadanos de las sociedades modernas respecto a la perdida su privacidad.

2. PÉRDIDA DE PRIVACIDAD DE LOS CIUDADANOS

La transformación de todas las facetas de la vida de las personas en datos y su recopilación permite a los gigantes digitales y los poderes públicos tener una gran cantidad de información respecto a la vida de los ciudadanos, como, por ejemplo, cuáles son sus preferencias, dónde gastan su dinero, quiénes son sus amigos, qué hacen en su tiempo libre o cómo se desplazan (Nardi y Ekbia, 2017; Véliz, 2020). Esta gran cantidad de información se consigue gracias al análisis de los datos personales de los ciudadanos. Los datos de carácter personal hacen referencia a una lista amplia y abierta que permite la identificación de las personas y que abarcan desde el nombre y apellidos, el número de la seguridad social y las direcciones, hasta datos sobre la voz, los "me gusta" de Facebook o la forma de caminar (Gil, 2016). La continua creación de datos personales provocada por la interacción de las personas con las nuevas tecnologías,

y la extracción, recopilación y análisis de estos por las grandes corporaciones tecnológicas están causando un gran impacto en la privacidad de la ciudadanía (Adams, 2017; Zuboff, 2019; Coeckelbergh, 2022).

Para poner de manifiesto las principales consecuencias vinculadas con la extracción, recopilación y análisis de los datos personales de los ciudadanos, cabe exponer los resultados de dos investigaciones. La primera de ellas es la investigación titulada *How Pizza Night Can Cost More in Data Than Dollars*, publicada en *The Wall Street Journal,* en la que se analiza la gran cantidad de datos facilitados por dos personas durante una noche en la cual se juntan para comer una pizza y ver una película (Stamm et al., 2018). La segunda investigación es el trabajo realizado por Douglas Schmidt para la Universidad de Vanderbilt, en el que se analiza la recopilación de datos de los servicios de Google (Schmidt, 2018).

En lo que respecta a la primera investigación, esta analiza la extracción de datos desde que dos persones conocidas se ponen en contacto hasta que acaban de ver una película. Este periodo de tiempo abarca las siguientes acciones: intercambiar mensajes de texto vía Apple iMessage, pedir una pizza en la aplicación de Domino's pizza vía Amazon Echo, desplazarse con el coche utilizando Google Maps, hacer un *selfie* y subirlo a Facebook y ver una película en la Apple TV. La investigación determinó que en este conjunto de acciones realizadas de forma cotidiana por muchas personas en la actualidad se habrían facilitado 53 campos de información diferentes. De estos 53 campos, 15 conjuntos de datos se facilitaron de forma consciente por las personas implicadas, y 38 habían sido extraídos por parte de las empresas que intervinieron en el proceso (de los cuales 23 fueron extraídos por Facebook a partir del *selfie*) sin que los usuarios fueran conscientes de ello y en la práctica sin su consentimiento explicito debido a que los consumidores raramente leen las políticas de privacidad. El mismo estudio expone que las políticas de privacidad de las empresas que intervinieron (Amazon, Apple, Facebook, Google y Domino's) ocupan 76.069 palabras y su lectura requeriría como mínimo 5 horas de lectura continuada, además de estar actualizándose de forma continua.

Y en lo que respecta a la segunda investigación, esta profundiza en los conjuntos de datos recopilados por Google y en las diversas formas de recopilación de estos. Este estudio expone que Google realiza una recopilación de datos de los usuarios tanto de forma activa como pasiva. La extracción de datos de forma activa se realiza cuando los usuarios se comunican de forma directa con Google o sus múltiples aplicaciones (YouTube, Gmail, Chrome, Android, etc.) y la recopilación de datos de forma pasiva se lleva a cabo a partir de los datos intercambiados en segundo plano entre el dispositivo y las aplicaciones de Google, en la ma-

yoría de los casos sin el conocimiento ni ninguna notificación al usuario. Este trabajo destaca que dos tercios de la información recopilada diariamente por Google se extrae de una forma pasiva, entre esta información destacan especialmente datos vinculados con la actividad, la ubicación o el historial de navegación. Respecto a este tema, se subraya también que los dispositivos con sistema operativo Android transfieren a Google aproximadamente 900 muestras de datos de forma diaria que transmiten alrededor de 4,4 megabytes de información diaria que suman un total de 130 megabytes al mes, de los cuales más de un treintaicinco por ciento están vinculados con la ubicación de los usuarios.

Estos dos ejemplos permiten ver el enorme poder de las empresas tecnológicas, la gran extracción de datos en la que se ven envueltas las vidas de las personas y el impacto que tienen sobre su privacidad. Otros trabajos como son Cabañas et al. (2018) y Llaneza (2019) subrayan también aspectos de la extracción de datos llevada a cabo por las GAFAM. Estas investigaciones destacan, entre otras cosas, que Facebook tiene acceso a datos sensibles del veinticinco por ciento de las personas que viven en Europa y que el sesenta por ciento de los navegadores web del mundo están vinculados a Google y sus aplicaciones. A pesar de que actualmente controlan una cantidad de datos mastodóntica procedente de millones de personas, las grandes compañías tecnológicas continúan desarrollando formas de extracción de datos privados de forma pasiva como son, por ejemplo, la detección de sonidos y el registro de imágenes de las cámaras de ordenadores y móviles con la finalidad de analizar el entorno psíquico de los ciudadanos y realizar un reconocimiento facial respectivamente (Lyon, 2018).

La normalización en la vida diaria de las personas de la utilización de una amplia gama de aplicaciones que extraen un flujo constante de datos de la vida de los ciudadanos ha disminuido la percepción de los peligros que conlleva la pérdida de privacidad causada por la utilización de estas aplicaciones y plataformas de las grandes empresas tecnológicas (Véliz, 2020). La mayoría de los usuarios no son conscientes de que las empresas proveedoras de estas tecnologías acceden de forma continuada a sus datos sensibles y que esta información es utilizada para controlar y manipular sus comportamientos y para limitar sus libertades (O'Neil, 2016; Zuboff, 2019). Esta situación se ve agravada en las generaciones nacidas en los últimos veinte años, ya que estas personas han sufrido una extracción de sus datos privados y disponen de una huella digital desde el día de su nacimiento (Llaneza, 2019). Respecto a este tema, Paloma Llaneza manifiesta que:

> 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 tienen medios para mantener la información segura, ni si tienen 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'. (Llaneza, 2019: 63)

Este fragmento pone de manifiesto que, aunque actualmente las personas dicen tener una alta preocupación por su privacidad, la realidad muestra una situación de despreocupación total respecto a la gran cantidad de información privada de millones de personas que almacenen las grandes empresas tecnológicas, la utilización que hacen de estos datos las empresas o el impacto que pueden provocar la utilización de estos datos en la vida de la ciudadanía en un futuro. La actitud actual de las personas respecto a la privacidad hace necesario profundizar en la llamada paradoja de la privacidad.

La paradoja de la privacidad hace referencia a la dicotomía entre la actitud de las personas y su verdadero comportamiento en relación con la privacidad cuando hacen uso de navegadores, aplicaciones o dispositivos relacionados con internet y las nuevas tecnologías (Llaneza, 2019). Los usuarios dicen estar muy preocupados por su privacidad, pero en cambio no hacen ninguna cosa para protegerla (Lastra-Anadón y Rubio, 2020). Estos son conscientes de los riesgos de la extracción de grandes cantidades de información privada por parte de las grandes corporaciones tecnológicas para su privacidad, sin embargo continúan usándolas constantemente (Barth y De Jong, 2017). Hay que destacar esta actitud en la utilización de las redes sociales. En este caso los usuarios utilizan diversas estrategias de protección de la privacidad para limitar los datos que pueden ser vistos por sus amigos o seguidores, como, por ejemplo, restringir el acceso a su perfil, prohibir el envío de mensajes privados o condicionar el etiquetado en las fotografías. Este tipo de medidas muestran una preocupación de la privacidad respecto a las otras personas que utilizan las plataformas digitales, pero expone una total despreocupación por la extracción y utilización de los datos por parte de la empresa gestora de la plataforma (Young y Quan-Haase, 2013). En el libro *Datano-*

mics (2019), Llaneza resume a la perfección el valor de la privacidad para los ciudadanos de las sociedades modernas:

> 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. (Llaneza, 2019: 67-68)

Esta cita destaca que, aunque en la ciudadanía de las sociedades modernas sí que existe una conciencia de los riesgos que supone la extracción y utilización de los datos por parte de las empresas para su propia privacidad, en la actualidad los usuarios obvian estos riesgos para poder continuar utilizando los omnipresentes programas, plataformas y dispositivos tecnológicos y no verse excluidos de la vida social.

Ahora bien, la actitud actual de los usuarios frente a la privacidad es uno de los principales problemas que están afectando a la pérdida de la privacidad, pero no es el único: la normativa y el marco jurídico de la privacidad también tiene múltiples deficiencias que potencian los problemas vinculados con la extracción, recopilación y análisis de datos de las personas (Gil, 2016). La próxima sección se centrará en analizar los principales problemas relacionados con la extracción, recopilación y análisis de datos privados de las personas por parte de las grandes corporaciones y los poderes públicos, y en la propuesta de soluciones a estos problemas.

3. Problemas relacionados con la extracción, recopilación y análisis de datos privados y posibles soluciones

Los principales problemas relacionados con la extracción, recopilación y análisis de datos privados de las personas se centran en la laxitud de la actual normativa y marco jurídico de la privacidad para hacer frente al actual contexto de la explotación de datos (Suárez Gonzalo, 2019). Las carencias y la falta de mecanismos de control dificultan conocer cómo, quién y de qué forma se están utilizando los datos privados de los ciudadanos (Gil, 2016; Suárez Gonzalo, 2017). El impacto del fenómeno de *big data* en la actual normativa y el marco jurídico en relación a los datos supone una amenaza para la privacidad de los ciudadanos debido a diversos motivos. Entre estos motivos hoy en día destacan especialmente los

relacionados con la minimización de datos, el consentimiento informado a la hora de tratar con datos personales y la anonimización de los datos personales (Gil, 2016).

El primer motivo tiene que ver con el incumplimiento del principio de minimización de datos. Este principio implica que solamente se debe recopilar la cantidad mínima necesaria para la finalidad con la que se recogen. Además de no cumplirse, la minimización de datos choca frontalmente contra las principales tendencias de la industria del *big data* y se contrapone con el estudio de cantidades masivas de datos que realiza el *big data*. Hay que destacar que los avances introducidos por la tecnología del *big data* se han conseguido gracias a la extracción de la mayor cantidad de datos activos y pasivos posibles, aunque no tengan ninguna utilidad para su función primaria, para ser analizados y tratados diversas veces para múltiples actividades y propósitos secundarios.

El segundo motivo se vincula con el consentimiento informado y el tratamiento de datos. El marco jurídico actual confía demasiado en el consentimiento informado de los individuos para la recopilación y tratamiento de sus datos personales. La mayoría de los ciudadanos no leen las políticas de privacidad antes de prestar su consentimiento y los que las leen no las entienden. Otorgar el consentimiento se ha convertido en un ejercicio vacío (Barocas y Nissenbaum, 2014). El actual contexto tecnológico dominado por el *big data* ha puesto en entredicho los procesos utilizados para recabar el consentimiento de los usuarios para el tratamiento de sus datos personales. El fenómeno del *big data* ha originado dos grandes problemas vinculados con los consentimientos informados: la constante reutilización de datos y el entendimiento de las políticas de privacidad. Por un lado, la constante utilización de los datos de diversas formas y para diversos propósitos secundarios dificulta la obtención del consentimiento para reutilizarlos (Gil, 2016). Por otro lado, la desmesurada extensión de las políticas de privacidad ha provocado que la gran mayoría de los usuarios no lean estas políticas de privacidad y que los pocos que lo hacen no las entiendan (Barocas y Nissenbaum, 2014).

El tercer y último motivo se relaciona con la anonimización de los datos personales. La anonimización ha demostrado tener límites (Llaneza, 2019). La anonimización de los datos es un pilar fundamental para la protección de los datos privados, pero en los últimos años han aparecido técnicas capaces de desanonimizar bases de datos. Las nuevas posibilidades del *big data* han hecho posible la reidentificación de datos que en un principio habían sido anonimizados poniendo en peligro la privacidad de las personas (Mayer-Schönberger y Cukier, 2013). El análisis de grandes conjuntos de datos hace posible identificar personas a partir de datos anónimos o datos que no contienen información perso-

nal. El desarrollo de las nuevas capacidades de la tecnología del *big data* sumado a la gran cantidad de datos producida diariamente y a la imposibilidad de una anonimización de datos con absolutas garantías provoca que la reidentificación de personas sea cada vez más sencilla (Gil, 2016; Llaneza, 2019). Uno de los ejemplos más destacados de desanonimización de datos fue el que se llevó a cabo por investigadores de la Universidad de Texas en 2006 a partir de datos anonimizados de Netflix y de la base de datos publica sobre información vinculada con películas denominada Internet Movie Database (IMB) (Narayanan y Shmatikov, 2006). A partir de estos datos los investigadores lograron la identificación de personas y el descubrimiento de información privada como las preferencias políticas y la orientación sexual.

Encontrar una solución a los problemas de minimización de datos, consentimiento y anonimización descritos se ha convertido en un verdadero problema para el adecuado funcionamiento de las sociedades modernas y la potenciación de las libertades de la ciudadanía. A continuación, se analizará la situación actual de cada uno de estos problemas y se expondrán diversas medidas propuestas para abordar cada uno de estos problemas y hacer frente a la monopolización de los datos y a la perdida de privacidad de las personas.

La nueva reglamentación introducida por la Unión Europea respecto a la protección de datos en 2018, la denominada General Data Protection Regulation (GDPR) ha supuesto una mejora en la reglamentación del campo de la minimización de datos. Esta reglamentación ha supuesto la limitación de la extracción y utilización de datos solamente a los datos personales que se vayan a tratar, en el momento que se vayan a tratar y para una finalidad declarada. Esta normativa ha supuesto un mejora de los problemas vinculados con la minimización de datos, entre otras cosas, porque la normativa que regulaba este campo databa de los años noventa y había quedado totalmente anacrónica (Benjamins & Salazar García, 2020). En los últimos años han aparecido diversas posturas críticas con la normativa de minimización de datos del GDPR que han realizado propuestas para hacer frente a este problema desde otras perspectivas.

Por una parte, el presidente de la Free Software Foundation, Richard Stallman, en un artículo en *The Guardian* en el año 2018 afirma que la GDPR ha sido positiva para la protección de datos, pero no es la solución a los problemas de privacidad de las sociedades modernas. Stallman (2018) defensa que las reglas de la GDPR son demasiado laxas y que la única solución para hacer cumplir la minimización de datos y fortalecer la privacidad de las personas es la aplicación de una ley que impida la recopilación de datos personales y que tenga como principio básico la

creación de los datos estrictamente necesarios para el funcionamiento de los sistemas y la eliminación de cualquier dato que ponga en peligro la privacidad de los ciudadanos. La postura de Stallman propone restringir totalmente la extracción de datos privados, esta restricción podría limitar el poder de las GAFAM y también lograr revertir la perdida de privacidad de las personas.

Por otra parte, los académicos Thomas Ramge y Viktor Mayer-Schönberger (2021) defienden que las medidas de minimización de datos de la GDPR son demasiado restrictivas y que no tiene ningún sentido aplicar normativas de minimización de datos en unas sociedades modernas basadas en la extracción y utilización de datos. Proponen aplicar el acceso libre a grandes conjuntos de datos anonimizados de las grandes empresas digitales a todas las personas, empresas y organizaciones para de esta forma limitar el poder de las GAFAM y potenciar la innovación y la prosperidad en las sociedades modernas. La puesta en marcha de las medidas propuestas por Ramge y Mayer-Schönberger serían otra forma de reducir la monopolización de los datos y los efectos negativos del capitalismo de la vigilancia sin descuidar la privacidad de las personas y la protección de datos.

El consentimiento individual informado es la pieza fundamental del tratamiento de datos en la GDPR y el principal mecanismo de protección de los datos personales. En la actual situación de extracción y utilización masiva de datos este consentimiento se ha convertido en una práctica totalmente ineficaz debido a la complejidad de la lectura de las políticas de privacidad y a la imposibilidad de saber la totalidad de los usos que pueden tener los datos a lo largo de su vida (Llaneza, 2019; Suárez Gonzalo, 2019). Para hacer frente a los problemas de consentimiento en los últimos años se han planteado diversas propuestas. Los autores Soheil Human y Florian Cech han propuesto reforzar los mecanismos de obtención del consentimiento actuales a partir de la aplicación de una obtención del consentimiento centrada en el empoderamiento de los usuarios teniendo en cuenta las dimensiones cognitivas, colectivas y contextuales de las personas (Human y Cech, 2021). Otra propuesta relacionada con el consentimiento y los datos personales es la desarrollada por Ira S. Rubinstein. Rubinstein defensa la aplicación de un nuevo modelo de negocio basado en el empoderamiento de los individuos en el cual son las propias personas las que gestionan sus datos y comparten con las empresas la información que desean, en el momento que lo desean y de la forma que lo desean (Rubinstein, 2012). La aplicación de cualquiera de las propuestas de Human y Cech (2021) o de (Rubinstein, 2012) permitiría mejorar los problemas vinculados con el consentimiento

informado, empoderar a los individuos sobre sus datos y limitar la monopolización de datos y el poder de las GAFAM.

El desarrollo de la tecnología del *big data* ha tenido un gran impacto en la anonimización de datos, ha cambiado radicalmente la concepción de este campo y ha colocado a la anonimización como un proceso central en el tratamiento de datos (Gil, 2016). Rubinstein y Hartzog (2016) defienden que, vistas las dificultades en los procesos de anonimización total de los datos que existen en la actualidad, las políticas de anonimización se tienen que endurecer e ir enfocadas hacia un proceso de minimización de los riesgos basada en la responsabilidad de las empresas en la protección de los datos de sus clientes, la importancia de los datos según su contexto y la aplicación de sanciones a las compañías que incumplan sus responsabilidades y las prácticas de reidentificación de datos. La puesta en marcha de medidas que refuercen la anonimización de datos y la seguridad de los grandes conjuntos de datos es fundamental para limitar las consecuencias negativas del fenómeno del *big data* para las personas, para su privacidad y para su libertad (Llaneza, 2019).

La constante puesta en marcha de medidas para hacer frente a la extracción, recopilación y análisis de datos privados de las personas por parte de las GAFAM es fundamental para potenciar los beneficios del fenómeno del *big data* y limita sus consecuencias negativas. Los ciudadanos tienen que ser conscientes de las implicaciones negativas del fenómeno del *big data* para sus vidas y exigir transparencia, explicabilidad y responsabilidad en la extracción, recopilación y análisis de los datos y la implementación de medidas vinculadas con la minimización de los datos, la obtención del consentimiento de las personas y la anonimización de datos que limiten el actual modelo de negocio de datos y que defiendan a la ciudadanía y su privacidad del dominio de las grandes empresas tecnológicas y de los gobiernos. La creación de una normativa a nivel internacional que regule la utilización de datos y la privacidad de las personas y que empodere a los individuos respecto a sus propios datos es de suma necesidad en el actual contexto social y económico basado en el capitalismo de la vigilancia. Otro aspecto fundamental es la actualización periodica de esta normativa para adaptarse lo máximo posible al cambiante contexto tecnológico y digital de las sociedades modernas y de esta forma limitar el poder de las GAFAM y reforzar la privacidad de las personas.

CONCLUSIÓN

La monopolización de los datos y la perdida de privacidad se han convertido en grandes problemas del funcionamiento de las sociedades

modernas como consecuencia de la gran cantidad de datos que se crean y se mueven diariamente y la importancia de estos datos para vigilar, controlar y manipular a la ciudadanía (Suárez Gonzalo, 2019; Zuboff, 2019; Véliz, 2020). La creación constante de datos, la continua extracción de datos de la vida privada de las personas, la perdida de privacidad de la ciudadanía y el gran poder de las corporaciones digitales han potenciado un clima de vigilancia y control social que permite crear mecanismos destinados a la manipulación de la población y de su forma de pensar. Los avances en el análisis de datos masivos, en las técnicas de personalización de contenidos y en las operaciones de distorsión informativa están afectando seriamente a la capacidad de raciocinio y opinión de la ciudadanía, creando una sociedad fácilmente manipulable y causando un importante daño a la libertad de la ciudadanía (Suárez Gonzalo, 2018).

El clima de constante vigilancia en el que vive actualmente la ciudadanía está afectando negativamente al funcionamiento de las sociedades modernas (Zuboff, 2015; Lyon, 2018). La industria del *big data* está amenazando la libertad de los ciudadanos (Han, 2021). Todas sus acciones, emociones y pensamientos son constantemente analizados y procesados con finalidades comerciales o políticas (Gil, 2016; Polo Roca, 2020; Han, 2021). Las denominadas GAFAM han conseguido un enorme poder económico y político gracias al tratamiento de la gran cantidad de datos personales que los usuarios introducen en sus plataformas (Miguel de Bustos y Izquierdo-Castillo, 2019).

La puesta en marcha de acciones en este escenario es fundamental para revertir la actual situación de vigilancia y control sobre la ciudadanía e impedir la manipulación de la forma de pensar y de actuar de las personas por medio del tratamiento de sus datos privados.

REFERENCIAS

Adams, M. (2017). Big Data and Individual Privacy in the Age of the Internet of Things. *Technology Innovation Management Review*, *7*(4), 12-24.

Antunes, D. C. y Maia, A. F. (2018). Big Data, ubiquitous exploitation, and targeted advertising: New facets of the cultural industry. *Psicologiía USP*, *29*(2), 189-199.

Arruzza, C., Bhattacharya, T., & Fraser, N. (2019). *Manifiesto de un feminismo para el 99%*. Barcelona: Herder Editorial.

Barocas, S., y Nissenbaum, H. (2014). Big Data's End Run around Anonymity and Consent. En J. Lane, V. Stodden, S. Bender, & H. Nissenbaum (Eds.), *Privacy, Big Data, and the Public Good: Frameworks for Engagement* (pp. 44-75). New York: Cambridge University Press.

Barth, S., y De Jong, M. D. T. (2017). The privacy paradox – Investigating discrepancies between expressed privacy concerns and actual online

behavior – A systematic literature review. *Telematics and Informatics*, *34*(7), 1038-1058.

Baruh, L., y Popescu, M. (2017). Big data analytics and the limits of privacy self-management. *New media & society*, *19*(4), 579-596.

Benjamins, R. y Salazar García, I. (2020). *El mito del algoritmo : cuentos y cuentas de la inteligencia artificial*. Madrid: Anaya Multimedia.

Boyd, D. y Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society*, *15*(5), 662-679.

Cabañas, J. G., Cuevas, Á. y Cuevas, R. (2018). *Facebook Use of Sensitive Data for Advertising in Europe*. Recuperado de https://arxiv.org/abs/1802.05030

Cardon, D. (2018). *Con qué sueñan los algoritmos. Nuestras vidas en el tiempo de los big data*. Madrid: Dado Ediciones.

Coeckelbergh, M. (2022). *The political philosophy of AI: an introduction*. Cambridge: Polity Press.

De Aguilera, M. y Casero-Ripolles, A. (2018). ¿Tecnologías para la transformación? Los medios sociales ante el cambio político y social. *Icono 14*, *16*(1), 1-21.

Delgado, A. (2018). *La sociedad hiperdigital*. Barcelona: Libros de Cabecera.

Gil, E. (2016). *Big data, privacidad y protección de datos*. Madrid: Agencia Estatal Boletin Oficial del Estado.

Han, B. C. (2021). *Psicopolítica: Neoliberalismo y nuevas técnicas de poder*. Barcelona: Herder.

Human, S. y Cech, F. (2021). A Human-Centric Perspective on Digital Consenting: The Case of GAFAM. En A. Zimmermann, R. Howlett, and L. Jain (Eds.), *Human Centred Intelligent Systems. Smart Innovation, Systems and Technologies* (Vol. 189, pp. 139-159). Singapore: Springer.

Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data and Society*, *1*(1), 1–12.

Lanier, J. (2018). *Diez razones para borrar tus redes sociales de inmediato*. Barcelona: Debate.

Lastra-Anadón, C. y Rubio, D. (2020). *European Tech Insights 2020*. Recuperado de https://docs.ie.edu/cgc/CGC-European-Tech-Insights-2020.pdf

Llaneza, P. (2019). *Datanomics: Todos los datos personales que das sin darte cuenta y todo lo que las empresas hacen con ellos*. Barcelona: Deusto.

López-Portillo, J. R. (2018). *La gran transición: retos y oportunidades del cambio tecnológico exponencial*. México: Fondo de Cultura Económica.

Lyon, D. (2018). *The culture of surveillance: Watching as a way of life*. Cambridge: Polity Press.

Mayer-Schönberger, V. y Cukier, K. (2013). *Big data: La revolución de los datos masivos*. Madrid: Turner Publicaciones.

McChesney, R. W. (2013). *Digital Disconnect: How Capitalism is Turning the Internet Against Democracy*. New York: The New Press.

Mejias, U. A. y Couldry, N. (2019). Datafication. *Internet Policy Review*, *8*(4), 1-10.

Miguel de Bustos, J. C. y Izquierdo-Castillo, J. (2019). ¿Quién controlará la Comunicación? El impacto de los GAFAM sobre las industrias mediáticas

en el entorno de la economía digital. *Revista Latina de Comunicacion Social*, *74*, 803-821.

Miguel de Bustos, J. C. y Moreno Cano, A. M. (2018). Los señores de los datos: Google-Alphabet, Amazon, Facebook, Apple y Microsoft. *Boletín del Centro de Documentación Hegoa*, (53), 1-12.

Morozov, E. (2011). *The net delusion: the dark side of internet freedom*. New York: PublicAffairs.

Nair, L. R., Shetty, S. D. y Shetty, S. D. (2017). Streaming big data analysis for real-time sentiment based targeted advertising. *International Journal of Electrical and Computer*, *7*(1), 402-407.

Narayanan, A., & Shmatikov, V. (2006). *How To Break Anonymity of the Netflix Prize Dataset*. Recuperado de http://arxiv.org/abs/cs/0610105

Nardi, B., y Ekbia, H. R. (2017). *Heteromation, and Other Stories of Computing and Capitalism*. Cambridge: MIT Press.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. London: Penguin.

Pariser, E. (2011). *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think*. New York: Penguin Books.

Polo Roca, A. (2020). Sociedad de la Información, Sociedad Digital, Sociedad de Control. *Inguruak*, *68*, 50-77.

Puyol, J. (2014). Una aproximación a big data. *Revista de derecho UNED*, *14*, 472-505.

Ramge, T. y Mayer-Schönberger, V. (2021). *Fuori i dati!: Rompere i monopoli sulle informazioni per rilanciare il progresso*. Milano: Egea.

Rubinstein, I. S. (2012). Big Data: The End of Privacy or a New Beginning? *SSRN Electronic Journal*, *3*(2), 74-87.

Rubinstein, I. S. y Hartzog, W. (2016). Anonymization and Risk. *Washington Law Review*, *91*(2), 703-760.

Schmidt, D. C. (2018). *Google Data Collection*. Vanderbilt. Recuperado de https://digitalcontentnext.org/wp-content/uploads/2018/08/DCN-Google-Data-Collection-Paper.pdf

Southerton, C. (2020). Datafication. En L. A. Schintler y C. L. McNeely (Eds.), *Encyclopedia of Big Data* (pp. 1-4). Cham: Springer International Publishing.

Srnicek, N. (2018). *Capitalismo de plataformas*. Buenos Aires: Caja Negra Editora.

Stallman, R. (3 de abril de 2018). A radical proposal to keep your personal data safe. *The Guardian*. Recuperado de https://www.theguardian.com/commentisfree/2018/apr/03/facebook-abusing-data-law-privacy-big-tech-surveillance

Stamm, S., Mickle, T. y Kuronen, J. (10 de abril de 2018). How Pizza Night Can Cost More in Data Than Dollars. *The Wall Street Journal*. Recuperado de https://www.wsj.com/graphics/how-pizza-night-can-cost-more-in-data-than-dollars/

Suárez Gonzalo, S. (2017). Big social data: límites del modelo notice and choice para la protección de la privacidad. *El Profesional de la Información*, *26*(2), 283-292.

Suárez Gonzalo, S. (2018). Tus likes ¿tu voto? Explotación masiva de datos personales y manipulación informativa en la campaña electoral de Donald Trump a la presidencia de EEUU 2016. *Quaderns del CAC*, *XXI*(44), 27-36.

Suárez Gonzalo, S. (2019). *Big data, poder y libertad Sobre el impacto social y político de la vigilancia masiva*. Tesis Doctorals. Barcelona: Universitat Pompeu Fabra. Recuperado de Universitat Pompeu Fabra website: http://www.tdx.cat/handle/10803/668235

Suárez Gonzalo, S. y Guerrero Solé, F. (2016). La conversación sobre big data en Twitter. Una primera aproximación al análisis del discurso dominante. *Comunicació: revista de recerca i d'anàlisi*, *33*(2), 113-131.

Van Dijck, J. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance and Society*, *12*(2), 197-208.

Van Dijck, J. (2016). *La cultura de la conectividad: una historia crítica de las redes sociales*. Buenos Aires: Siglo Veintiuno Editores.

Véliz, C. (2020). *Privacy is Power: Why and How You Should Take Back Control of Your Data*. London: Bantam Press.

Webb, A. (2021). *Los nueve gigantes: cómo las grandes tecnológicas amenazan el futuro de la humanidad*. Barcelona: Península.

Westerlund, M., Isabelle, D. A. y Leminen, S. (2021). Perspectives from Higher Education: Applied Sciences University Teachers on the Digitalization of the Bioeconomy : The Acceptance of Digital Surveillance in an Age of Big Data. *Technology Innovation Management Review*, *11*(3), 32-44.

Wylie, C. (2019). *Mindf*ck. Inside Cambridge Analytica's Plot to Break the World.* London: Profile Books.

Young, A. L. y Quan-Haase, A. (2013). Privacy Protection Strategies on Facebook. *Information, Communication & Society*, *16*(4), 479-500.

Zuboff, S. (2015). Big other: surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, *30*(1), 75-89.

Zuboff, S. (2019). *The age of surveillance capitalism: the fight for a human future at the new frontier of power*. New York: Public Affairs.