LA LIBERTAD DE EXPRESIÓN EN JAQUE, EL PANÓPTICO DEL SIGLO XXI. BIG DATA COMO AMENAZA PARA LA DEMOCRACIA
A propósito del caso Cambridge Analytica

FREEDOM OF SPEECH IN PERIL, THE TWENTY-FIRST CENTURY PANOPTICON. BIG DATA AS A THREAT TO DEMOCRACY
About the Cambridge Analytica case

Andrés Fernando Mejía*

RESUMEN: En el artículo se plantea que la libertad de expresión es vulnerada de manera distinta a los escenarios tradicionales en razón del uso del Big Data y otras tecnologías como el análisis predictivo de datos y la personalización de contenido. Esta vulneración resulta diferente de la habitual puesto que no existe imposibilidad o limitación a expresar lo que se quiere de manera espontánea, sino que, la libertad de expresión se ve coartada ex ante a través de la generación de circunstancias en las que lo que se expresa viene dado previamente a través de información adecuada de manera personalizada según el perfil de quien opina. Para lograr presentar esta tesis, se efectúa un marco de referencia en el que este fenómeno se contextualiza. Al final se realiza una propuesta para disminuir o eliminar esta vulneración a la libertad de expresión.

***ABSTRACT**: The article argues that freedom of speech is violated differently from traditional scenarios due to the use of Big Data and other technologies such as predictive data analysis and content personalization. This violation is different from usual ones since there is no impossibility or limitation to express what is wanted spontaneously, but rather, freedom of expression is restricted ex ante through the generation of circumstances in which what is expressed is previously given through appropriate personalized information according to the profile of the person who talks. In order to present this thesis, a frame of reference is made in which this phenomenon is contextualized. In the end a proposal is made to reduce or eliminate this violation of freedom of expression.*

PALABRAS CLAVE: libertad de expresión, big data, personalización de contenido, sicográficos, análisis predictivo, ciencia comportamental, Cambridge Analytica.

***KEYWORDS**: free of speech, big data, content personalization, psychographics, predictive analysis, behavioral science, Cambridge Analytica.*

Fecha de recepción: 17/10/2019
Fecha de aceptación: 13/04/2020

doi: https://doi.org/10.20318/universitas.2020.5512

* Abogado Universidad Libre Seccional Pereira, Licenciado en Derecho por homologación Universidad de Málaga (España), especialista en Derecho Procesal Contemporáneo Universidad de Medellín, Especialista en Administración Universidad EAFIT, Magíster en Derecho Procesal Universidad de Medellín, Magíster en Administración Universidad EAFIT. Candidato a Doctor en Derecho Procesal Contemporáneo Universidad de Medellín. Docente Programa de Derecho Fundación Universitaria del Área Andina seccional Pereira, vinculado al Grupo de estudios e investigaciones socio-jurídicas GEIS. Docente Universidad Libre Seccional Pereira. Gerente Construcciones C.I.R. S.A.S., Pereira. E-mail: amejia3@areandina.edu.co

1.- INTRODUCCIÓN

Si bien la tecnología ha traído consigo una gran cantidad de beneficios para la sociedad, no es menos cierto que ella implica importantes retos al igual que peligros. En ese sentido Kranzberg (1986) anunciaba en sus leyes que la tecnología no es intrínsecamente buena o mala ni tampoco neutral, y, que las consecuencias de esta van más allá de los propósitos inmediatos de los dispositivos y las prácticas en sí mismas consideradas[1] teniendo consecuencias impensadas en espectros diferentes como el ambiental y social.

El derecho constantemente se enfrenta a nuevos escenarios que desafían las regulaciones que hasta el momento se han previsto y que, por tanto, permiten que este sea dinámico en la medida que evoluciona al compás de los cambios sociales. El derecho en su ámbito funcional como instrumento previsto para evitar o resolver algunos conflictos entre individuos (Nino, 2003, p.3) se estructura a partir de los factores que condicionan el dictado de las normas y las reacciones que a estas se tiene por parte de la comunidad y las transformaciones que este "derecho en acción" en términos usados por Alf Ross, genera a nivel social y económico (Nino, 2003, p. 7). Estos factores han permitido que se presenten escenarios impensables para el derecho años atrás como la posibilidad de que un oso se proteja a través de Hábeas Corpus[2] y la titularidad de derechos para un río[3] o para una región[4].

La tecnología, las innovaciones y los emprendimientos que satisfacen necesidades humanas en muchas ocasiones se ven huérfanas de un marco jurídico que permita abordarlas de manera adecuada, de modo tal que las actividades que se desarrollen a partir de estas cuenten con un panorama claro en cuanto a la asunción de responsabilidades por parte de sus ejecutores y los límites respecto de una diversa gama de derechos, entre estos, los fundamentales. En efecto, hoy no existe claridad en materia regulativa acerca de tecnologías y emprendimientos disruptivos como Uber, Airbnb y el almacenamiento en la nube. Entre otros, el Big data se erige, tal y como se señaló, en tecnología que no solo requiere de regulación,

[1] Technology is neither good nor bad; nor is it neutral...technology's interaction with the social ecology is such that technical developments frequently have environmental, social, and human consequences that go far beyond the immediate purposes of the technical devices and practices themselves.

[2] A través de Hábeas Corpus se solicitó la protección del oso Chucho, medida que fue concedida en primera instancia por un magistrado de la sala civil de la Corte Suprema de Justicia, pero que posteriormente fue revocada por la sala laboral de la misma Corporación. Finalmente, la Corte Constitucional decidió que el oso no puede ser sujeto de protección de la libertad a través de esta acción constitucional.

[3] A través de la Sentencia T-622 de 2016 proferida por la Corte Constitucional de Colombia se reconoció al río Atrato como sujeto de derechos.

[4] Sentencia proferida por la Corte Suprema de Justicia, sala de casación civil de 5 de abril de 2018.

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

sino que la demanda de manera urgente gracias al potencial invasivo que implica para la privacidad de las personas.

Respecto del Big Data ha existido una multiplicidad de perspectivas acerca de los peligros implícitos a su implementación. El análisis se ha desarrollado desde diferentes aristas: la vulneración de la intimidad, los peligros en cuanto a la discriminación, el manejo de datos personales sensibles en materia de salud, el conflicto que se podría presentar entre las predicciones realizadas a través de técnicas analíticas relacionadas con el big data y el derecho al olvido, la tensión entre la libertad de información y la protección de datos personales y las implicaciones éticas del uso de datos por parte de los investigadores por citar algunos ejemplos. Sin embargo, a la fecha no se han evidenciado hallazgos acerca de los peligros que entraña en materia de libertad de expresión; efectivamente, aunque existen investigaciones que cuestionan la función del big data de cara a la democracia desde la limitación a la libertad general relacionada con la manipulación del electorado y la restricción al libre desarrollo de la personalidad, la específica esfera de la libertad de expresión no ha contado con mayor desarrollo. El presente artículo es un simple ejercicio reflexivo en cuanto a las dificultades que el Big Data comporta frente a este derecho a partir de las evidencias del caso Cambridge Analytica.

La tesis que aquí se sostiene consiste en que la libertad de expresión no se ve restringida *ex post* en virtud del Big Data como de manera habitual sucede. Esto es, la imposibilidad o limitación para manifestarse de manera espontánea, sino que, la vulneración se materializa *ex ante* a través de la generación de escenarios en los que lo que se expresa viene dado previamente a través de información adecuada de manera personalizada según el perfil de quien exterioriza su opinión. Lo que se manifiesta tiene como base la información de la que se dispone: información presentada según lo que al sujeto le interesa o le gusta, creando de esta manera, no solo un sesgo, sino también un aislamiento del debate público al que las ideas diferentes son inherentes. Se es preso de los pensamientos, gustos e intereses que se validan y refuerzan de manera exógena con información similar, creando la ilusión de que el mundo es exactamente como lo ve el sujeto.

El término "panóptico del siglo XXI" hace referencia a la forma en la cual a través del uso del big data acompañado de otras tecnologías, se realiza un continuo rastreo, observación, análisis, y, finalmente, presentación de información a los ciudadanos que accedemos al internet.

Para poder alcanzar el objetivo que se traza, resulta menester elaborar un corto marco de referencia que permita contextualizar el escenario en el cual se plantea la vulneración de derecho a la libertad de expresión.

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

2.- EL BIG DATA

El 'Big Data' al decir de Rubinstein (2012:74) se refiere a "formas novedosas en que las organizaciones, incluidos el gobierno y las empresas, combinan diversos conjuntos de datos digitales y luego usan estadísticas y otras técnicas de minería de datos para extraer de ellos tanta información oculta como correlaciones sorprendentes"[5], no obstante, el mismo autor advierte que si bien esta herramienta promete importantes beneficios económicos y sociales, también plantea serias preocupaciones en cuanto a la privacidad.

A pesar de que existen algunos elementos comunes al concepto de Big Data identificables en diferentes autores, este no deja de ser un término generalizado e impreciso; en ese sentido, podría afirmarse que el big data es un término pobre en muchos sentidos (Boyd y Crawford, 2011). Bajo este concepto se ha considerado históricamente una cantidad de datos que requieren necesariamente para su análisis de computadoras (o supercomputadoras en algunos casos). No obstante, para Lev Manovich (2011) el tamaño no resulta ser su característica definitoria, el elemento esencial lo constituye su "relacionalidad" con otros datos. En efecto, el big data está fundamentalmente conectado en red, "su valor proviene de los patrones que pueden derivarse haciendo conexiones entre datos, sobre un individuo, sobre individuos en relación con otros, sobre grupos de personas, o simplemente sobre la estructura de la información en sí misma"[6] (Boyd y Crawford, 2011: 2).

Existen, por supuesto, diversas interpretaciones acerca de esta nueva realidad, del lado positivo se ubican quienes consideran que el big data presenta muchas oportunidades para mejorar la sociedad moderna: permitiría optimizar la investigación científica en cuanto la vuelve más productiva y aceleraría el descubrimiento y la innovación; las personas podrían mejorar su salud y la atención médica sería más eficiente y efectiva; de igual modo, los Estados a través del correcto manejo de las bases de datos podrían mejorar la prestación de servicios gubernamentales y monitorear las amenazas a la seguridad nacional (Bollinger, 2010: 40).

Desde la orilla pesimista se ubican aquellos que consideran que las herramientas analíticas propias del big data erigen graves problemas de privacidad en los algoritmos predictivos ya que a menudo son impredecibles y sus efectos pueden no ser comprendidos por sus programadores: "Como los informáticos han demostrado, en muchos contextos, es imposible garantizar una privacidad diferencial cuando se utiliza un algoritmo de aprendizaje que extrae datos de una distribución continua" (Chaudhur y Hsu, 2011: 179). En otros términos, no se podría conocer de antemano cuándo un algoritmo predecirá PII (Personally identifiable information -Información de

[5] Traducción propia.
[6] Traducción propia.

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

identificación personal-), y, por tanto, no se podría predecir dónde y cuándo edificar protecciones de privacidad en torno a esos datos: Cuando una adolescente embarazada compra vitaminas, ¿ella podría predecir que cualquier compra o visita particular en un almacén desencadenaría que los algoritmos de un minorista la señalaran como una clienta embarazada? ¿en qué punto habría sido apropiado dar aviso y solicitar su consentimiento? (Chaudhur y Hsu, 2011: 179).

Teniendo claro que el big data es un instrumento positivo para la sociedad, pero que también entraña múltiples peligros acerca de su uso inadecuado e invasivo, resulta pertinente, a efectos de señalar su importancia a nivel electoral, y, por ende, para la democracia, recordar lo que Alistair Croll, fundadora de bitcurrent señalaba: "Después de John F. Kennedy no podrías ganar una elección sin televisión. Después de Obama, no podrías ganar una elección sin las redes sociales. Predigo que en 2012 no podrás ganar una elección sin big data" (González, 2017: 9). Esa predicción finalmente se materializó en las elecciones de los Estados Unidos de 2016 y en otro evento de elección popular previo como fue el Brexit. El big data jugó un papel preponderante en ambas, de hecho, antes de la elección presidencial de Estados Unidos, el director de Cambridge Analytica, Alexander Nix, afirmaba respecto del concepto aplicado de big data en el trabajo que adelantó con la campaña de Ted Cruz: "es realmente la agregación de tantos puntos de datos individuales posibles que puedas tener en tus manos y que son sintetizados en una base de datos utilizada para informar y crear información sobre tu público objetivo."[7] (Nix, youtube, 2016).

Ciertamente, entre los peligros que el big data trae inmersos consigo se encuentra el de la manipulación de votantes y, como se sostiene en este artículo, la trasgresión del derecho a la libertad de expresión, ya no entendida en términos tradicionales donde la prohibición o limitación es el mecanismo de vulneración, sino coartada a partir de la cuidadosa disposición de noticias, artículos y, en general, todo tipo de información que se ajusta a las preferencias de persona. De este modo, según Solove: "lo que expresan las personas pueden estar más controladas que nunca (y también irónicamente, ello se puede estructurar para hacerlas creer que tienen el control)".

Para terminar este aparte, es importante acotar que Alexander Nix terminando su intervención sostuvo: Ahora claramente la campaña de Cruz ha terminado, pero lo que les puedo decir es que de los 2 candidatos que quedan en esta elección, uno de ellos está

[7] Lo referido se puede escuchar en el minuto 5:41 del video en estos términos: "is really the aggregation of as many individual data points that you can possibly get your hands on which are the synthesized in one database of record used to inform and create insight on your target audience"

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

usando estás tecnologías y será muy interesante ver como impactan las siguientes 7 semanas.[8] (Nix, youtube, 2016).

3.- ANÁLISIS DE DATOS PREDICTIVO Y SICOGRÁFICOS

Entre las herramientas habituales de la investigación en mercadeo se encuentran los perfiles demográficos. En virtud de estos se determina la edad, ingreso, educación y otros indicadores de posición vital que tienen demasiada influencia en los comportamientos de consumo de los usuarios (Wells, 1975); no obstante, la necesidad de conocer de manera más profunda a los consumidores dio paso al surgimiento de la sicografía, técnica que se puede definir como: un tipo de investigación cuantitativa destinada a ubicar a los consumidores en dimensiones psicológicas, a diferencia de las demográficas. Porque va más allá de lo estándar y lo aceptado, ofrece la posibilidad de nuevos conocimientos y conclusiones inusuales. Debido a que es cuantitativo más que discursivo, abre el camino a muestras grandes y representativas de encuestados, y al análisis estadístico multivariado de los hallazgos.[9] (Wells, 1975: 197).

Los sicográficos entonces se refieren a datos como las costumbres, pasatiempos y hábitos de gasto. Los datos demográficos explican "quién" es tu comprador, mientras que los sicográficos explican "por qué" compran (Meredith, 2017).

El big data ha permitido que el uso de los sicográficos se incremente al punto de crear "mensajes sicográficos"; esto es, a la medida de cada persona de conformidad con su perfil sicológico. Así pues, los mensajes se *matizan* con el objeto de que "resuenen más eficazmente con esos grupos de audiencia clave"[10] (Nix, youtube, 2016). Con esta herramienta y la ayuda del análisis de datos no se requiere confeccionar soluciones creativas que puedan o no funcionar, sino que con cientos o miles de puntos de datos de los *objetivos* (los destinatarios del mensaje) "se conoce exactamente a qué tipo de mensajes se debe apelar para atraer a las audiencias mucho antes de que comience el proceso creativo" (Nix, youtube, 2016).

Para recolectar estos datos, en el caso de Cambrigde Analytica fue fundamental Facebook, pero en realidad el análisis de sicográficos no requiere de redes sociales. Las computadoras pueden clasificar a las personas psicológicamente utilizando miles de puntos de datos disponibles en el mercado (Burleigh, 2017).

---

[8] Lo referido se puede escuchar en el minuto 10:40 del video en estos términos: now clearly the Cruz campaign is over now but what I can tell you is that the two candidates left in this election, one of them is using these technologies and it`s going to be very interesting to see how they impact the next seven weeks.

[9] Traducción propia del texto original "Psychographics: A Critical Review".

[10] Lo referido se puede escuchar en el minuto 4:09 del video en estos términos: you can nuance your messaging to resonate more effectively with those key audience groups.

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

American Express usando sus bases de datos (big data y análisis predictivo a partir de los sicográficos) logró identificar comportamientos de sus clientes que le permitieron perfilarlos y proveer respuestas adecuadas frente a una contingencia: La empresa descubrió que las personas que acumulan grandes saldos en su tarjeta y luego registran una nueva dirección de envío en Florida tienen una mayor probabilidad de declararse en bancarrota. Esto debido a que este Estado tiene una de las leyes de bancarrota más laxas, lo que la convierte en un destino favorito para los deudores con problemas financieros. La identificación de tales correlaciones en los datos -un aumento vertiginoso de la tarjeta de crédito y una reubicación en Florida- puede desencadenar una investigación sobre la solvencia real del titular de la tarjeta. (Bollier, 2010: 21).

Y aunque estos mecanismos comportan múltiples beneficios, también generan preocupación respecto de su uso y alcance. Ejemplo de ello es la investigación que recientemente (febrero de 2018) concluyó que las caras contienen mucha más información acerca de la orientación sexual de lo que puede percibir o interpretar el cerebro humano. A través de redes neuronales profundas se extrajeron características de imágenes faciales a las que aplicaron regresiones logísticas para clasificar la orientación sexual "La precisión del algoritmo aumentó a 91% y 83%, respectivamente, con cinco imágenes faciales por persona" (Wang y Kosinski, 2018: 250). Si bien es cierto que con estos hallazgos se amplía la comprensión acerca de los orígenes de la orientación sexual y los límites de la percepción humana, como lo señalan los investigadores, "dado que las empresas y los gobiernos utilizan cada vez más los algoritmos de visión para detectar los rasgos íntimos de las personas, nuestros hallazgos exponen una amenaza a la privacidad y seguridad de los hombres y mujeres homosexuales"[11] (Wang y Kosinski, 2018: 255).

4.- IMPLICACIONES, PELIGROS Y RETOS QUE PRESENTA EL BIG DATA PARA LOS DERECHOS

La tecnología de datos se ha tornado tan invasiva, penetrante y difícil de comprender que cuestionamientos válidos se han erigido y a la fecha no tienen aún una respuesta clara ¿Cómo se protegerá la sociedad contra aquellos que abusen de las grandes bases de datos? ¿Qué nuevos sistemas regulatorios, innovaciones de derecho privado o prácticas sociales serán capaces de controlar conductas antisociales y cómo deberíamos incluso definir qué es social y legalmente aceptable cuando las prácticas habilitadas por Big Data son tan novedosas y a menudo misteriosas? (Bollier, 2010: 40). A continuación se señalan solo algunos de los peligros que entraña esta tecnología y los retos que respecto de la protección de derechos se

[11] Traducción propia del texto original "Deep neural networks are more accurate than humans at detecting sexual orientation from facial images".

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

presentan. Se ha decidido no realizar mención al peligro del uso del big data frente a la privacidad[12] ya que esta ha sido la arista que mayor desarrollo tiene en el ámbito investigativo.

4.1.- Acerca de la ética y la manipulación emocional

En 2014 facebook realizó un experimento en compañía de investigadores de la Universidad de Cornell en el que a 683.003 usuarios les realizaron un ajuste respecto de las noticias que les aparecían en la plataforma. El ajuste consistió en presentar un contenido emocional más positivo o más negativo. Se pretendía determinar si ello tendría un efecto emocional en las personas; en efecto, lo hizo. "Las personas expuestas a contenido más positivo tenían publicaciones que eran más positivas, y las que estaban expuestas a contenido más negativo tenían publicaciones que eran más negativas. Esto fue medido por los tipos de palabras que usaron"[13] (Solove, 2014).

Entorno de la naturaleza de los daños que puede producir el big data Cotino (2017: 137) afirma que el daño individual que produce el big data y la inteligencia artificial es imperceptible para el derecho fundamental desde la perspectiva del individuo titular del derecho. Sin embargo, la vulneración se da masivamente a los derechos fundamentales de sectores o conjuntos de la sociedad. En ese sentido, "Dogmáticamente considero que puede ser necesario trabajar con una dimensión colectiva de los derechos que no es la habitual". Esta realidad demanda una teoría jurídica distinta que permita abordar de manera efectiva estas afectaciones a los derechos.

4.2.- Respecto de las implicaciones de la personalización de la información

Turow y Sunstein han sostenido que la personalización de contenidos resultante de la aplicación de los sicográficos y el análisis predictivo de los datos limita el mercado de las ideas que a su vez son fundamentales para cualquier sociedad, esta personalización refuerza las posiciones particulares, generando indiferencia, falta de apertura y compromiso con lo diferente. Esta personalización de contenido implicaría en últimas la desaparición del foro público. (Cotino, 2017: 140).

4.3.- Sobre de las consecuencias de las equivocaciones de los algoritmos predictivos

[12] Especial atención ha recibido el asunto de la privacidad en materia de salud.
[13] Traducción propia del texto original "Facebook's Psych Experiment: Consent, Privacy, and Manipulation".

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

La policía del estado de Maryland aprovechó su acceso a los centros de mando unificado[14] para vigilar grupos de derechos humanos, activistas por la paz y opositores a la pena de muerte durante un período de diecinueve meses. 53 activistas políticos finalmente fueron clasificados como "terroristas", incluidas dos monjas católicas y un candidato demócrata para un cargo local. El centro de mando unificado compartió estas clasificaciones erróneas de terroristas con las bases de datos federales de combate de drogas y la NSA, todo ello sin otorgar a las personas oportunidad de conocer, y mucho menos corregir el registro (Gray y Citron, 2013: 81).

4.4.- Discriminación

El big data puede eventualmente desembocar en segregación de determinados grupos que han sido perfilados a través de análisis predictivo. Muestra de ello es que, si a través de esta herramienta se indica un pobre historial crediticio, el usuario ni siquiera verá una oferta de crédito de las principales instituciones de préstamos, y no se dará cuenta de que los préstamos están disponibles para ayudarlo con sus prioridades personales o profesionales actuales (Fertik, M, 2013).

4.5.- Análisis predictivo en materia delictual

El Departamento de Policía de Santa Cruz inició en 2011 con un ejercicio a partir de un algoritmo aplicado a los delitos con el objetivo de disminuir la tasa de ocurrencia de los mismos. El algoritmo era el resultado de una investigación del comportamiento antropológico y criminológico. Utilizaba matemática compleja para estimar el crimen y predecir las zonas de alta ocurrencia de estos en el futuro. Su uso fue tan exitoso que ciudades como Los Ángeles, y estados como Carolina del Sur y Arizona implementaron el programa. En este evento, más que un peligro (por el contrario, este es un buen ejemplo del uso adecuado y benéfico del big data y el análisis predictivo), se identifica un reto, toda vez que estas predicciones consistentes en áreas geográficas particulares con mayores probabilidades de ocurrencia de crímenes seguramente producirán más arrestos al ordenar a la policía que las patrulle de manera prioritaria. Esto, a su vez, generará más datos históricos delictivos para esas áreas y de esta manera, aumentar la probabilidad de patrullas. Para aquellos que viven allí, estas zonas bien pueden llegar a ser tanto PII (Información de identificación personal) como otra información demográfica. (Friend, 2013).

[14] Fusion Centers.

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

5.- PERSONALIZACIÓN DE CONTENIDO

Las mismas herramientas que permiten a los profesionales de marketing identificar y crear grupos de "gemelos estadísticos" o personas de ideas afines, y luego orientar los anuncios para venderles zapatos, viajes y lavadoras también permiten a los estrategas políticos crear "cámaras de eco" llenas de lemas e historias que la gente quiere escuchar, también conocidas como noticias falsas[15]. (Burleigh, 2017).

Para comprender de manera más profunda las implicaciones frente a la libertad de expresión que surgen con uso coordinado del big data, el análisis predictivo de datos y los sicográficos, se deben tener en cuenta 2 elementos adicionales: de una parte, las limitaciones cognitivas que son explotadas a través de la manipulación del mercado; por otro, la personalización de contenido que usa elementos de esta manipulación de mercado en la esfera digital. Respecto del primero se puede señalar:

En 1999, Jon Hanson y Douglas Kysar acuñaron el término "manipulación del mercado" para describir cómo las empresas explotan las limitaciones cognitivas de los consumidores. Por ejemplo, todo cuesta $ 9.99 porque los consumidores ven el precio más cerca de $ 9 que de $ 10. Aunque ampliamente citado por académicos, el concepto de manipulación del mercado ha tenido solo un impacto modesto en la ley de protección del consumidor. (Calo, 2014: 995).

Acerca del segundo elemento, Eli Pariser exhibe con claridad un ejemplo que da cuenta del peligro implícito de la personalización de contenido. Efectivamente, en una presentación acerca de su libro "The Filter Bubble: What The Internet Is Hiding From You" que se halla alojada en la página de London School of Economics evidencia la diferencia de resultados que obtienen Scott y Daniel al buscar la palabra Egipto en el motor de búsqueda de Google: Scott obtuvo páginas, imágenes y videos de "La crisis en Egipto", "las protestas de 2011" y "Lara Logan"; por su parte, Daniel obtuvo "Viajes y vacaciones", "la página de Egypt daily news" y "CIA world factbook".

De esto precisamente se trata la personalización de contenido, se muestra aquello que interesa al sujeto a partir del análisis de datos predictivo y sicográfico realizado y se aprovechan las limitaciones cognitivas a través de la manipulación de mercado digital. En algún momento lo dijo el propio Mark Zuckerberg: "Una ardilla muriendo frente a tu casa puede ser más relevante para tus intereses en este momento que las personas que mueren en África"[16]. Por su parte, Tapan Bhat, en su momento Vicepresidente de Yahoo

[15] Traducción propia del texto original "How big data mines personal info to craft fake news and manipulate voters".

[16] Pariser, E. (2011, 22 de mayo). When the Internet Thinks It Knows You. *The New York Times*. Recuperado de: https://www.nytimes.com

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

decía que "El futuro de la web se trataba de la personalización". En la misma línea, Eric Schmidt, ex director ejecutivo de Google aseveró: "será difícil para las personas mirar o consumir algo que, en cierto sentido, no se haya adaptado para ellos".

El propio Alexander Nix, en materia electoral aseguró que "la publicidad general está muerta, la idea de que todo el mundo recibe el mismo mensaje (...) mis hijos nunca entenderán el concepto de comunicación masiva"[17] y termina con un lapidario: "La comunicación ahora está dirigida, está individualizada para cada persona".

En gracia de discusión podría argumentarse que la personalización del contenido (información) no tiene nada de malo, puesto que se nos muestra aquello que consideran que quizá nos interese o guste más, pero, en primera instancia, es censurable que no se solicite consentimiento para hacerlo, y, en segunda, esta metodología implica que estás teniendo una versión parcializada de la realidad, una edición de la misma ajustada según tu personalidad. Es cierto que Google News proporciona un ejemplo al respecto, con la atractiva sugerencia: "Nadie puede leer todas las noticias que se publican todos los días, así que ¿por qué no configurar tu página para mostrarte las historias que mejor representan tus intereses?" (Sunstein, 2007: 4). Pero esto genera, en definitiva, consecuencias perjudiciales para la democracia deliberativa habermasiana, creándose "enclaves deliberativos" en los que sólo se accede a puntos de vista e identidades que refuerzan las posiciones individuales, contribuyendo a extremar y polarizar la esfera pública (Corredoira y Cotino, 2013: 44-45).

6.- LA CIENCIA COMPORTAMENTAL

En 2013 Robert Shiller[18] ganó el premio Nobel de economía en razón de sus aportes desde la teoría de la ineficiencia de los mercados a partir del estudio de las finanzas comportamentales. Su conocido libro "Exuberancia irracional" que debe su título a la frase que pronunciara Alan Greenspan, ex presidente de la Reserva Federal de los Estados Unidos en 1996 da cuenta de los sesgos que se incluyen en las decisiones financieras. Elementos determinantes desde la sicología, la sociología y la cultura inciden en las decisiones que, finalmente, son lejanas de la pretendida racionalidad de los mercados que cimienta la teoría de la eficiencia del mismo. Su concepción de los mercados le permitió vaticinar las burbujas .com y subprime.

[17] Lo referido se puede escuchar en el minuto 8:37 del video en estos términos: "blanket advertising is dead, the idea that everybody receives the same message… my children will never understand the concepto of mass communication".

[18] Es pertinente señalar que en el mismo año ganó también el premio Nobel de economía Eugene Fama quien sostiene la teoría de la eficiencia de los mercados que es contraria a la que pregona Shiller.

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

En 2002 el sicólogo Daniel Kahneman recibe también el premio Nobel de economía, ¿La razón?: su propuesta de análisis sicológico de la economía y puntualmente del tipo de decisión que se adopta en escenarios de incertidumbre en los que no tenemos en cuenta las probabilidades. Su planteamiento acerca del sistema 1[19] y el sistema 2[20] en la toma de decisiones resulta fundamental, así como ineludible es la lectura de su libro "pensar rápido, pensar despacio" en el que los refiere.

Otro reconocido autor que unió la sicología con las finanzas y la economía además de las políticas públicas es Richard Thaler, quien fue galardonado con el premio Nobel de economía en 2017 por sus aportes en materia de falta de autocontrol de los individuos, teoría que se ha materializado en el concepto "nudge" que ha sido exitosamente aplicado a políticas públicas.

Como queda evidenciado con los 3 casos relacionados, las ciencias comportamentales o de la conducta se han consolidado como área sobre y desde la que se construyen teorías interdisciplinarias que afectan las decisiones de los seres humanos en amplios ámbitos que abarcan desde lo que compramos y por qué lo compramos, las decisiones de inversión y ahorro (pensional por ejemplo), la salud, el medio ambiente hasta el matrimonio, entre otras. Estas ciencias también han sido usadas en materia electoral y, como se afirma en este texto, también han servido para sesgar, limitar y vulnerar la libertad de expresión *ex ante*.

Ciertamente, tal y como advierte Burleigh (2017), la tecnología big data ha superado los marcos legales y regulatorios y poco se pregunta sobre la ética de los mensajes políticos basados en evadir la cognición o el pensamiento racional[21]. Esta nueva realidad, desconocida para algunos y que resulta indiferente para otros, tiene consecuencias intrigantes y peligrosas para la democracia.

Alexander Nix en su presentación da un ejemplo de cómo se deben usar las ciencias comportamentales con el objeto de influir en la conducta: si se pretende que las personas no ingresen a una playa por ser privada se pueden plantear 2 tipos de avisos: a) Uno en el que advierta que hasta ese punto llega la playa pública y que, por tanto, nadie puede continuar a partir de allí. b) Otro que indique: "cuidado, tiburones". La más persuasiva a efectos de lograr el objetivo es evidentemente la segunda.

Dentro de las múltiples herramientas que Cambridge Analytica usó en las elecciones presidenciales estadounidenses de 2016 se encontraba la ciencia comportamental. Existe tal nivel de detalle en la identificación de las personas que a través de una aplicación, el mismo Nix podía filtrar por ciudades, filiación política, género, etc. "Finalmente, solo queda un nombre, incluyendo edad, dirección,

[19] Sistema automático.
[20] Sistema reflexivo
[21] Traducción propia.

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

intereses, personalidad e inclinación política. ¿Cómo dirige ahora Cambridge Analytica a esta persona con un mensaje político apropiado?" (Grassegger & Krogerus, 2017).

7.- EL MODELO OCEAN Y MICHAL KOSINSKI

Michal Kosinski estudió su doctorado en el centro de Psicometría de la Universidad de Cambridge. En compañía de David Stiwell y después de que éste lanzara una aplicación denominada MyPersonality en Facebook lograron mediante cuestionarios sicométricos que eran diligenciados por los usuarios de la plataforma completar su "perfil de su personalidad" basados en las denominadas 5 grandes (big five). Estos usuarios podían compartir este perfilamiento en la misma plataforma (Grassegger y Krogerus, 2017).

Kosinski y su equipo refinaron estos modelos predictivos y en 2012 demostró que sobre la base de un promedio de 68 "me gusta" de Facebook de un usuario, era posible predecir su color de piel (con un 95% de precisión), su orientación sexual (88% de precisión) y su afiliación al partido demócrata o republicano (85 por ciento). Pero no se detuvo allí. Inteligencia, afiliación religiosa, así como el consumo de alcohol, cigarrillos y drogas, todo podría determinarse. A partir de los datos, incluso fue posible deducir si los padres de alguien estaban divorciados. En poco tiempo, pudo evaluar a una persona mejor que un colega de trabajo promedio, simplemente sobre la base de diez "me gusta" de Facebook. Setenta "me gusta" eran suficientes para superar lo que sabían los amigos de una persona, 150 lo que sus padres sabían, y 300 "me gusta" lo que su pareja sabía. Más "Me gusta" podrían incluso superar lo que una persona creía saber sobre sí mismo. El día que Kosinski publicó estos hallazgos, recibió dos llamadas telefónicas: una amenaza de demanda y una oferta de trabajo. Ambas de Facebook. (Grassegger y Krogerus, 2017).

Pero más allá de los modelos algorítmicos detrás de estos hallazgos, se encuentra la cuestión de cómo, a partir de ciertos datos se podía dar cuenta del perfil sicológico de alguien de manera tan precisa. La respuesta a ello es el modelo OCEAN o de los 5 grandes rasgos de la personalidad; en efecto, OCEAN es el acrónimo con el que se denominan los siguientes 5 rasgos y que responden a la pregunta que se halla al frente de cada uno de ellos:

O: Openess (Apertura a nuevas experiencias, apertura al cambio): ¿Disfrutas de nuevas experiencias?

C: Conscientiousness (Responsabilidad): ¿Prefieres el orden y la planeación en tu vida?

E: Extraversion (Extroversión): ¿Te gusta pasar tiempo con otros? ¿Eres sociable?

A: Agreeable (Cordialidad o amabilidad): ¿Pones las necesidades de los demás antes que las tuyas?

N: Neuroticism (inestabilidad emocional o neuroticismo): ¿Te preocupas mucho?

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

A cada rasgo le corresponden diferentes escalas, que, leídas en conjunto, dan como resultado una predicción acerca de la personalidad de quien absuelve los cuestionarios. Tradicionalmente, esta era la metodología para evaluar la personalidad; sin embargo, a partir del modelaje de Kosinski, las preguntas no son necesarias. Bastan unos simples likes en Facebook para determinarla; de hecho, ya ni siquiera se requiere de esta red social (u otras) para poder aplicar el modelo. Kosinski concluye que nuestro teléfono inteligente es un "vasto cuestionario psicológico que constantemente estamos diligenciando, tanto consciente como inconscientemente" (Grassegger y Krogerus, 2017).

Sin embargo, más allá de esta innovación, queda un asunto complejo de abordar para el derecho, ya que esta herramienta también funciona en el sentido inverso: no solo se pueden crear perfiles psicológicos a partir de tus datos, sino que también se pueden usar al revés para buscar perfiles específicos: todos los padres ansiosos, todos aquellos enojados introvertidos, por ejemplo, ¿o tal vez incluso todos los demócratas indecisos? Esencialmente, lo que Kosinski ha inventado era una especie de motor de búsqueda de personas. Empezó a reconocer el potencial, pero también el peligro inherente, de su trabajo. (Grassegger y Krogerus, 2017).

Este desarrollo supone una clara amenaza para la libertad de individuo y de manera puntual para la libertad de expresión, que tal y como se ha referido, se ve vulnerada de manera previa; esto es, ex ante. Efectivamente, como asevera Nix: "La comunicación está personalizada para cada individuo: En caso de elecciones, aquellos temas que más te interesan pero que han sido matizados para reflejar la forma en que ves el mundo"[22] (Nix, youtube, 2016).

8.- EL CASO CAMBRIDGE ANALYTICA

En 2010 Facebook permitió que Apps recolectaran datos de los usuarios a partir del lanzamiento de "Open Graph" (a través de esta Kosinski y Stiwell aplicaron su metodología). En el centro de la controversia se encuentra Aleksandr Kogan, sicólogo y neuro-científico de la Universidad de Cambridge quien a principios de 2014 se acercó a Kosinski señalando que se encontraba investigando a nombre de una compañía (cuyo nombre no podía revelar), que estaban interesados en su método y que querían acceder a la base de datos de MyPersonality (tampoco reveló el propósito de ello). En principio Kosinski consideró la oferta, ya que ella aparejaba una gran suma de dinero para el instituto. Posteriormente Kogan reveló el nombre de la empresa para la cual realizaba la gestión, se trataba de SCL (Strategic Communication Laboratories). Kosinski buscó la

[22] Lo referido se puede escuchar en el minuto 8:55 del video en estos términos: "In case of elections, issues that you care about most but that have been nuanced in order to reflect the way you see the world"

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

compañía en Google y se dio cuenta que su objeto social era el de "administración de elecciones". SCL proporciona marketing basado en modelos psicológicos. Uno de sus enfoques principales: influir en las elecciones. (Grassegger y Krogerus, 2017).

A partir de esta búsqueda, Kosinski evidenció que empresas relacionadas con SCL habían participado en elecciones desde Ucrania a Nigeria y habían ayudado al monarca de Nepal contra los rebeldes. En 2013, SCL creó una nueva compañía para participar en las elecciones en los Estados Unidos: Cambridge Analytica. (Grassegger y Krogerus, 2017).

Según un informe de diciembre de 2015 en The Guardian y documentos internos de la compañía entregados a Das Magazin, se prueba que SCL se enteró del método de Kosinski a través de Kogan, quien había registrado una empresa (GSR Global Science Research) que hacía negocios con SCL. (Nature, 2018).

La información que Kogan entregó a Cambridge Analytica la recolectó a través de la aplicación "Thisisyourdigitallife", logrando acceder a datos de 87 millones de personas a partir de 300 mil personas que contestaron el cuestionario. Aplicando la metodología de Kosinski logró determinar la personalidad de estos (a través del método OCEAN) y, de esta manera, Cambridge Analytica logró influenciar a favor de Donald Trump, las elecciones presidenciales de 2016. (Nature, 2018).

Cambridge Analytica, según su director, Alexander Nix, es una compañía que usa datos para cambiar el comportamiento del público, con finalidades tanto comerciales, como políticas. Este objetivo se logra con el uso de los datos personales; presentando publicidad, noticias e información a la medida de cada votante.

Cambridge Analytica dividió la población de Estados Unidos en 32 tipos de personalidad y se enfocó en 17 Estados, descubriendo, por ejemplo, que preferir carros hechos en el país era un gran indicador de potenciales votantes por Trump. La decisión de enfocarse en las últimas semanas en Michigan y Wisconsin se adoptó con base en el análisis de datos según la investigación de Grassegger y Krogerus.

La cantidad de información personal que se encuentra en línea que luego es analizada mediante algoritmos determina qué tipo de personalización de mensajes políticos se requieren para grupos cada vez más pequeños de personas con ideas afines. Aunado a esto, las vastas y crecientes bases de datos recopiladas para el comercio y la policía también están a la venta para los políticos y sus estrategas, que ahora pueden saber más sobre usted que su cónyuge o sus padres. Los análisis sicográficos ni siquiera requieren Facebook. Las computadoras pueden clasificar a las personas psicológicamente utilizando miles de puntos de datos disponibles en el mercado. (Grassegger y Krogerus, 2017).

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

9.- PROTECCIÓN DE DATOS

La protección de datos tradicionalmente se ha enfocado en el consentimiento del titular de estos. Sin embargo, este elemento por si solo carece de efectividad: de una parte, el consentimiento se encuentra establecido por defecto y es habitual que las personas no lean los términos de concesión del mismo. De otra, la imposibilidad de renunciar al uso de las IT y la ausencia de una cultura de la privacidad, implica que materialmente la garantía es irreal o inefectiva (Rubinstein, 2013). Esto genera en últimas, que el consentimiento se torne en una carta blanca al descontrol del flujo de los datos personales (Cotino, 2017: 145).

Un elemento que no se ha tenido completamente en cuenta es la diferencia existente entre el simple uso de datos por parte de una empresa a su uso en el contexto de los macrodatos. En efecto, la inteligencia artificial y las decisiones automatizadas implican una gran dificultad respecto de las finalidades del uso de estos. De entrada, en muchas ocasiones no se conocen cuales son estas finalidades; siendo claro que tampoco los ciudadanos pueden consentir de manera informada respecto de lo que determinados algoritmos realizarán a partir de ellos.

Es, por tanto, una falacia afirmar que se pueda obtener el consentimiento para tratar una infinita cantidad de datos que se generan en el ámbito de modelos de negocios en los cuales se realiza un análisis comportamental y de predicción de los consumidores (Martínez, 2014: 89). "El usuario no suele ser consciente de que está dando todos esos permisos. Sociólogos norteamericanos han calculado que necesitaríamos 100 días para leer y entender todos los contratos de consentimiento que aceptamos por usar apps, redes sociales." (Valero, 2014: 47)

En esa línea, se ha argumentado que la regulación[23] existente se ha quedado corta frente a los nuevos modelos de negocios y usos de los datos personales (Rubinstein, 2012). Una de las herramientas propuestas para lograr evitar este incontrolable manejo de los datos personales y la vulneración de la intimidad de los usuarios (de su información de identificación personal[24]) ha sido la de la anonimización de los mismos. Esta técnica pretende eliminar los peligros inherentes al uso de datos privados. Por ejemplo, borrando los nombres y el documento de identificación de las personas, así como los números de cuentas bancarias y el código de identificación universitario (Ohm, 2009: 1703). No obstante, existen técnicas que permiten desanonimizar estos datos, y, finalmente, identificar a las personas titulares de los mismos.

[23] Aún la General Data Protection Regulation (GDPR) de 2016 proferida por la Unión Europea.

[24] PII Personally Identifiable Information.

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

10.- LIBERTAD DE EXPRESIÓN EN LA ERA DIGITAL

La libertad de expresión se enfrenta a desafíos que la dinámica social le presenta. Uno de ellos es el contexto digital en el que los parámetros tradicionales para salvaguardarla se ven en ocasiones obsoletos; en ese sentido, se ha propuesto, entre otras:

10.1.- Derechos de las máquinas

A partir del asistente personal inteligente de Apple Siri, Toni Massaro y Helen Norton se cuestionan, no solo si las máquinas tienen derechos, si no si estos pueden protegerse. De manera puntual les inquieta saber si lo que expresa una máquina (en este caso desarrollada bajo inteligencia artificial) puede ser cobijado por la Primera Enmienda de la Constitución de los Estados Unidos. Así, las tecnologías "pueden evadir las categorías legales convencionales de manera que empujarán a los tribunales a redefinir las categorías más antiguas, con efectos que nos resultan difíciles de imaginar anticipadamente"[25] (Massaro y Norton, 2016: 1171).

Plantean, que en algún momento, se podría imaginar que los altavoces de los computadores puedan estar lo suficientemente desconectados y que sean lo suficientemente inteligentes como para afirmar que el discurso que producen es el suyo, no el nuestro, sin un creador o director humano a la vista (Massaro y Norton, 2016: 1172). Si bien, puede que este escenario jamás se materialice, los autores brindan un argumento mucho más persuasivo a efectos de clamar por protección legal del discurso de los dispositivos inteligentes. Señalan que el riesgo reside en que un gobierno tipo "Orwelliano" pretenda la supresión de determinado discurso proferido por Inteligencia Artificial que no se alinea con el establecimiento, lo que implicaría, que los tribunales deberían interpretar la Primera Enmienda para proteger este tipo de discurso (Massaro y Norton, 2016: 1174). Sería, por tanto, una suerte de factor de conexidad con la libertad de expresión de los seres humanos, lo que le brindaría esta protección al discurso de los dispositivos inteligentes.

10.2.- Libertad de expresión artificial

Tim Wu se realiza 2 interesantes preguntas: ¿Las máquinas hablan?, si ello es así, ¿Son titulares del derecho constitucional a la libertad de expresión? (Wu, 2012). Google es una de las empresas que defienden la tesis de que ello, en efecto, es así. El profesor de derecho de la Universidad de California Eugene Volokh elaboró un artículo denominado "Google, Microsoft's Bing, Yahoo! Search, and other search engines are speakers" en el que sostuvo exactamente la

[25] Traducción propia del texto original "Siri-ously? Free Speech Rights and Artificial Intelligence".

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

misma afirmación. Sin embargo, Wu considera que: debemos vacilar antes de permitir que los principios más elevados de la Carta de Derechos se conviertan en poco más que simples herramientas de ventaja comercial. Dar a las computadoras los derechos destinados a los humanos es elevar nuestras máquinas por encima de nosotros mismos[26] (Wu, 2012).

Como queda en evidencia, la libertad de expresión en la era digital se halla lejos de esferas pacíficas en cuanto a su delimitación y alcance. Si a esto se suma la irrupción de los sicográficos, así como el análisis predictivo a partir del big data, la personalización de contenido y los intereses, no solo comerciales, sino también políticos. Se tiene que los derechos en general, y la libertad de expresión, en particular, deben ser re conceptualizados a efectos que lograr una protección efectiva de los mismos.

Como señala Pariser, respecto de los peligros que entraña para la democracia la personalización de contenido político a partir del análisis sicográfico de los votantes: de una parte, se dejan de escuchar los argumentos políticos y, por tanto, de debatir alrededor de estos, y, al tener tantos mensajes específicos para grupos de personas estadísticamente significativas diferentes, ni siquiera se comprende lo que se argumenta (Burleigh, 2017)[27].

De esta forma, si se interesa en la política, puede limitarse a determinados puntos de vista de aquellas personas con las que se encuentra de acuerdo. Y debido a que es tan fácil aprender sobre las opciones de "personas como usted", innumerables personas toman las mismas decisiones que otras personas como ellos (Sunstein, 2007: 2)[28].

Tal y como acertadamente propone Miralles (2013), deslindar la frontera entre "influir" en las decisiones y "generarlas", es tarea compleja, ¿en qué punto se dejan de brindar datos e información que permiten valorar y tomar una decisión libre y se pasa a inducir las decisiones a través de herramientas que coartan el elemento libertad?

La intervención de Cambridge Analytica en las elecciones presidenciales de los Estados Unidos de 2016 es un claro ejemplo de decisiones inducidas (que evidentemente coartan la libertad de expresión *ex ante*). La campaña Trump dirigió publicidad que solo podía ser vista por usuarios con perfiles específicos como afroamericanos, en la que Hillary Clinton a través de video se refería a ellos como depredadores; en tanto que, en Little Haiti en Miami se difundió información acerca del fracaso de la Fundación Clinton ayudando después del terremoto de Haití, lo que evitó que muchas de estas personas votaran por ella. Facebook demostró ser el arma

[26] Traducción propia del texto original "Free Speech for Computers?".
[27] Traducción propia del texto original "How big data mines personal info to craft fake news and manipulate voters".
[28] Traducción propia del texto original "Republica.com2".

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

definitiva y la mejor campaña electoral. (Grassegger y Krogerus, 2017).

11.- REGULACIÓN

Tal y como afirma Burleigh (2017), el Big Data ha superado el marco regulatorio disponible, no obstante ello, no existen discusiones a nivel legislativo con alta prioridad cuestionando lo que la minería de datos implica para la intimidad o acerca del manejo ético de datos en el marco de los mensajes políticos basados en evadir la cognición o el pensamiento racional. El 11 de abril de 2018, Mark Zuckerberg testificó ante el Congreso de los Estados Unidos y tanto él, como los legisladores parecían estar de acuerdo en regular el potencial de influencia de las redes sociales en materia electoral, sin embargo, a la fecha, no existe claridad acerca del tipo de regulación que se debe implementar.

Tan poca claridad se tiene acerca de la adecuada regulación en la materia, que en Estados Unidos, de momento, se ha elegido el camino de la autoregulación. En esa línea, empresas que tienen acceso a gran cantidad de datos, como facebook, han optado por brindar mayor seguridad a las cuentas de sus usuarios[29].

Existen diferencias regulatorias entre Europa y Estados Unidos puntualmente acerca de la opción por "default" que se aplica a los usuarios. En efecto, mientras que en Europa se requiere autorizar de manera expresa el uso y publicación de datos, en Estados Unidos, se puede realizar, a menos que el usuario opte por no permitirlo. En esta materia, los conceptos de "Nudge" (empujoncito) y de arquitectura de elección que ha planteado Richard Thaler, resultarían de valiosa aplicación en el marco del "paternalismo libertario". Elementos que, aunque de importancia capital para los días que corren, exceden la pretensión que en este corto texto se plantea.

Hoy, se pueden identificar diversos instrumentos que pretenden servir de marco regulatorio al uso de datos masivos en internet, entre estos:

11.1.- Reglamento General de Protección de Datos (General Data Protection Regulation (GDPR) -Reglamento 2016/679-)

Regula la privacidad y protección de datos en la Unión Europea, fue aprobada en 2016 y entró en vigor en 2018. A través de este reglamento se reemplazó la Data Protection Direction (DPD) (Directiva 45/96/CE de 1995).

[29] En abril de 2018 Facebook aclaró en un comunicado que habían tenido que expandir su enfoque de seguridad desde el comportamiento abusivo tradicional, como la piratería de cuentas, el malware, el spam y las estafas financieras, para incluir formas de abuso más sutiles e insidiosas, incluidos los intentos de manipular el discurso cívico y engañar a las personas.

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

Esta normatividad resulta novedosa por cuanto crea diferentes categorías de datos (genéticos por ejemplo), así como principios (enfoque de riesgo y responsabilidad proactiva) y nuevos derechos (portabilidad de datos y limitación de tratamiento -artículo 18-, otros como el derecho de supresión -artículo 17-, es conocido en Colombia como el derecho al olvido).

11.2.- Opinión 5/2014 de 10 de abril- Grupo de trabajo del artículo 29 sobre anonimización

Si bien, en desarrollo del artículo 68 del Reglamento General de Protección el Grupo de Trabajo del artículo 29 cesó en sus funciones para que asumiera las mismas la Junta Europea de Protección de Datos, el dictamen 5 de 2014 continúa siendo una herramienta relevante a través de la cual se pretendió regular la forma en la cual se anonimizan los datos de forma irreversible. Con esa finalidad, se identificaron, entre otros, los principales riesgos en el proceso de anonimización (singularización, vinculatoriedad e inferencia).

11.3.- European Parliament resolution of 14 March 2017 on fundamental rights implications of big data: privacy, data protection, non-discrimination, security and law-enforcement

En esta Resolución, el Parlamento Europeo señala la importancia de salvaguardar los derechos fundamentales en un marco poco regulado de las relaciones creadas a partir del uso del Big Data y el análisis de datos. Entre estos, como uno de los principales sujetos de tutela, el derecho a la libertad de expresión[30].

Al igual que en los 2 instrumentos regulatorios referidos previamente, las disposiciones se quedan cortas al desconocerse los alcances de estas nuevas tecnologías disruptivas. No obstante, es de resaltar la imperiosa necesidad que se explicita en cuanto al resguardo de garantías personales de los asociados.

[30] I. (...) but also entails significant risks, namely with regard to the protection of fundamental rights, such as the right to privacy, data protection and data security, but also freedom of expression and non-discrimination, as guaranteed by the EU Charter of Fundamental Rights and Union law; (...)
13. Calls on the Commission and Member States to ensure that data-driven technologies do not limit or discriminate access to a pluralistic media environment, but rather foster media freedom and pluralism; emphasises that cooperation between governments, educational institutions and media organisations will play a pivotal role in ensuring that digital media literacy is supported in order to empower citizens and protect their rights to information and freedom of expression;

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

11.4.- Honest Ads Act

Esta ley estadounidense (S.1989) de 19 de Octubre de 2017 tuvo su génesis en un informe[31] presentado por la Oficina del Director Nacional de Inteligencia en el que se afirmaba la influencia rusa en las elecciones de 2016. A través de esta norma, se pretende regular la publicidad y las campañas políticas llevadas a cabo a través de los medios masivos electrónicos tales como Google o Facebook.

Dentro de las medidas adoptadas, se encuentra la de obligar a las plataformas en línea a mantener y permitir el acceso al historial de compras de "publicidad política calificada" realizada por cualquier persona si supera los 500 dólares en el año calendario, lo anterior, a efectos de poder efectuar inspección pública de esta información. Por "publicidad política calificada", se entiende cualquier anuncio de cualquier tipo que es hecho por o en favor de algún candidato o que comunica un mensaje político de importancia nacional.

Aunque esta ley no pretende garantizar el derecho a la libertad de expresión, si resulta importante traerla a colación debido a la importancia manifiesta de las plataformas en línea en ejercicios de elección popular, al punto de requerirse de su regulación para garantizar el correcto debate en escenarios deliberativos informados sobre la base de hechos, datos e información fidedigna que asegure posiciones críticas por parte de la comunidad.

11.5.- Ley estatutaria 1581 de 2012

Esta ley colombiana que tiene como finalidad la protección de los datos personales de los ciudadanos, en su artículo 5 refiere el concepto de "datos sensibles". Dentro de estos, incluye aquellos que puedan revelar la orientación política o la pertenencia a asociaciones que promuevan derechos y garantías, así como datos biométricos, origen racial, etc. Sin embargo, esta normatividad, de carácter puramente formalista, es claramente insuficiente frente a los fenómenos de captación de información y su manejo a través de Big Data y análisis predictivo y sicográficos. Muestra de ello es su casi nula referencia al internet o a medios en línea.

Este compendio normativo no refiere en ninguno de sus apartes la protección del derecho a la libertad de expresión, como tampoco lo hace el Decreto reglamentario 1377 de 2013.

[31] En el informe se aseveró: (1) On January 6, 2017, the Office of the Director of National Intelligence published a report titled "Assessing Russian Activities and Intentions in Recent U.S. Elections", noting that "Russian President Vladimir Putin ordered an influence campaign in 2016 aimed at the U.S. Presidential election …". Moscow's influence campaign followed a Russian messaging strategy that blends covert intelligence operation—such as cyber activity—with overt efforts by Russian Government agencies, State-funded media, third-party intermediaries, and paid social media users or "trolls."

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

Hoy no se habla acerca de protección de datos en términos absolutos ¿dónde inicia la protección? ¿cómo se sabe a través de que instrumento, plataforma o empresa se logró el acceso a la misma?. Sin que esto quede claro, leyes como la 1581 serán ineficaces en su finalidad de proteger los datos. Naturalmente esta amenaza se encuentra vinculada intrínsecamente al derecho a la intimidad (y de manera concreta al derecho de protección de datos), la igualdad, el consentimiento en el manejo de datos y el habeas data, así como la regulación de actividades económicas privadas y el uso por parte del gobierno de la información obtenida.

Tal y como se ha evidenciado en las cortas referencias normativas vertidas en las líneas que anteceden, la reglamentación del Big Data, del análisis predictivo y de la personalización de contenido es casi inexistente y la que se ha elaborado es incipiente al desconocerse los límites materiales, implicaciones y peligros que significan para los derechos fundamentales y de manera puntual, acerca de la amenaza que representan frente al derecho a la libertad de expresión.

12.- PROPUESTA

Se ha pretendido con los elementos relacionados en estas líneas, fundamentar la tesis según la cual, la libertad de expresión se ve vulnerada en virtud del uso de tecnologías de recolección y análisis de datos. Sin embargo, esta trasgresión del derecho es sui generis y sutil, ya que no se enmarca en el tradicional escenario fáctico en el que se impide expresar lo que se piensa, sino que, se adecúa un escenario previo a la expresión para que, atendiendo a los gustos e intereses propios históricamente rastreados, se genere la sensación de que el mundo es tal y como se piensa por parte de cada sujeto, impidiendo contrastar ideas arraigadas con los pensamientos, posiciones y argumentos de otras personas o grupos de personas que tienen consideraciones distintas de quien se expresa.

Ello, aunado al hecho de las redes de contactos que se generan en torno al pensamiento generalizado de la pequeña comunidad, crea la sensación de una especie de "profecía autocumplida" en la que al verificarse que los demás piensan como el sujeto, se refuerza la idea inicial de que el mundo (posiciones políticas, consideraciones en cuanto a candidatos a cargos de elección popular, corporaciones públicas, instituciones jurídicas, cambio climático, ideología económica, etc.) es tal y como se concibe. Por tanto, lo que se expresa viene previamente dispuesto para que así se exteriorice. Se plantea que esta es una forma diversa y nueva de coartar la libertad de expresión, al sesgar y limitar la información de la cual se dispone y a la cual se accede.

A efectos de disminuir la afectación al derecho de libertad de expresión tal y como se ha evidenciado, sugiero las siguientes posibilidades:

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

1. Generar la posibilidad de que las personas se encuentren frente a posiciones, tesis e ideologías que no hayan elegido de manera anticipada o que les sean presentadas teniendo en cuenta sus gustos, búsquedas históricamente consideradas y afinidades exteriorizadas, tanto explícita, como implícitamente.

Esta idea halla fundamento en lo que expone Cass Sunstein en su obra Republic.com2 a efectos de garantizar una democracia en la que no se generen fragmentación y extremismos propios y derivados de la acción comunicativa exclusiva entre personas afines que solo hablen entre sí (Sunstein, 2007: 6).

Es claro que los gobiernos no pueden imponer una particular visión del mundo y enmarcada en esta prohibición, se encuentra, evidentemente, la de forzar a las personas a encontrarse con aquello que pretenden de antemano evitar. Para solucionar esta problemática, se permitiría la opción al ciudadano de manifestar de manera expresa su voluntad, y, por tanto, actuar en consecuencia, esto es, que no se le presente determinado contenido. Esta tesis se encuentra enmarcada en el concepto de "paternalismo libertario" (al cual suscribo) acuñado por Thaler y el propio Sunstein en la obra "Nudge". Se plantea un sistema en el cual la opción de encontrar información no elegida de antemano o que comulgue con los gustos e intereses propios sería la elegida "por defecto", a menos que, la persona elija solo ver aquello que le interese y vete determinado tipo de contenido que desee ignorar. Esta fórmula, permitiría, de una parte, tutelar la libertad de elección de los ciudadanos, a la vez que lograría que las personas se vean enfrentadas a posiciones que difieren de las personales. Ejercicio connatural e inherente a escenarios de debates democráticos.

Esta propuesta en últimas implica que los ciudadanos expresarían su posición, una vez ha sido contrastada con argumentos contrarios a los que se tienen arraigados. De ese modo, en primer término, se eliminaría la vulneración a la libertad de expresión ex ante aquí insinuada, y, en segunda instancia, se disminuirían las posibilidades de tener una sociedad extremista y fragmentada.

2. Permitir que las personas se vean expuestas a experiencias comunes puede incrementar los consensos en torno a los problemas que la sociedad enfrenta. Esto, aumentará la cohesión social (Sunstein, 2007. p.6), y permitirá que lo que se exprese cuente con un sustento diverso a las ideas arraigadas en cada persona (o grupo homogéneo de personas) así como un elemento subyacente común. Esta propuesta, de compleja materialización, claro está, encontraría la misma censura expuesta en la idea 1 de este acápite según la cual, ello significaría la imposición de un modelo de bien considerado por el Estado que coartaría la libertad individual. No obstante, la respuesta a tal objeción sería la misma esbozada anteriormente; es decir, partir del respeto por la elección individual, pero garantizando una opción por defecto para que las personas experimenten vivencias comunes.

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

Finalmente, una opción disruptiva respecto de la situación actual, pero que de momento parece prometedora, consiste en evitar que sean grandes empresas las que manejen los datos personales de los ciudadanos. Por el contrario, serían los mismos ciudadanos los que determinan qué tipo de datos comparten, con qué empresas y personas y cuáles de ellos desean que sean públicos. Así, la información estaría descentralizada y no dejaría huella, evitando el escenario de análisis predictivo y personalización de información a través del big data que es en últimas lo que genera la vulneración a la libertad de expresión ex ante. Ejemplo de esta iniciativa es la del proyecto Decode que es financiado por diferentes entidades de la Unión Europea y que tiene como ciudades piloto a Barcelona y Ámsterdam; así como la que adelanta Aral Balkan y su equipo en la ciudad belga de Ghent.

13.- REFERENCIAS

Bollier, David (2010). *The Promise and Peril of Big Data*. Washington The Aspen Institute. Disponible en:
https://assets.aspeninstitute.org/content/uploads/files/content/docs/pubs/The_Promise_and_Peril_of_Big_Data.pdf

Boyd, Danah and Crawford, Kate (September 21, 2011): Six Provocations for Big Data: A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society. *SSRN.* Disponible en:
http://dx.doi.org/10.2139/ssrn.1926431

Burleigh, N (2017, 6 de agosto): How big data mines personal info to craft fake news and manipulate voters. *Newsweek*. Disponible en: www.newsweek.com

Calo, Ryan (August 15, 2013): Digital Market Manipulation. *82 George Washington Law Review 995 (2014); University of Washington School of Law Research* Paper No. 2013-27. *SSRN*. Disponible en: https://ssrn.com/abstract=2309703

Chaudhuri, K., and Hsu, D.J. (2011): *Sample Complexity Bounds for Differentially Private Learning*. JMLR workshop and conference proceedings. pp 155-186. Disponible en: http://proceedings.mlr.press/v19/chaudhuri11a/chaudhuri11a.pdf

Corredoira L. y Cotino L (Dir.)(2013): *La selección y personalización de noticias por el usuario de nuevas tecnologías. En: Libertad de expresión e información en Internet*. Madrid, España: Centro de Estudios Políticos y Constitucionales pp 41-56.

Corte Constitucional de Colombia. Sentencia del 10 de noviembre de 2016. Magistrado Ponente. Jorge Iván Palacio. Sentencia T-622 de 2016

Corte Suprema de Justicia de Colombia. Sala de casación civil. Sentencia del 5 de abril de 2018. Magistrado Ponente: Luis Armando Tolosa. STC4360-2018.

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512

Corte Suprema de Justicia de Colombia. Sala de casación civil. Sentencia del 26 de Julio de 2017. Magistrado Ponente: Luis Armando Tolosa. STC4360-2017.

Corte Suprema de Justicia de Colombia. Sala de casación civil. Sentencia del 16 de Agosto de 2017. Magistrado Ponente: Fernando Castillo Cadena. STL12651-2017.

Cotino, L (2017). Big data e inteligencia artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales. Dilemata. *Revista Internacional de Éticas Aplicadas,* No 24. 131-150. Disponible en:

https://dialnet.unirioja.es/servlet/articulo?codigo=6066829.

Fertik, M (2013). The Rich See a Different Internet Than the Poor. Ninety-nine percent of us live on the wrong side of a one-way mirror. *Scientific American*. Available at: www.scientificamerican.com

Friend, Z (2013). Predictive Policing: Using Technology to Reduce Crime, FBI L. ENFORCEMENT BULL. Disponible en: https://leb.fbi.gov/articles/featured-articles/predictive-policing-using-technology-to-reduce-crime

González, R. J (2017). "Hacking the citizenry?: Personality profiling, 'big data' and the election of Donald Trump". *Anthropology Today*, vol. 33, N° 3: pp. 9-12. Disponible en: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1467-8322.12348

Grassegger, H. & Krogerus, M (2017, 28 de Enero):The Data That Turned the World Upside Down How Cambridge Analytica used your Facebook data to help the Donald Trump campaign in the 2016 election. *Motherboard*. Disponible en:

https://motherboard.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win

Gray, David C. and Citron, Danielle Keats (March 5, 2013): The Right to Quantitative Privacy. Minnesota Law Review, Vol. 98, 2013; U of Maryland Legal Studies Research Paper, 2013-23. *SSRN*. Disponible en: https://ssrn.com/abstract=2228919

Martínez, Ricard (2014): "Ética y privacidad de los datos", texto escrito de la Jornada: Big Data: de la investigación científica a la gestión empresarial, Fundación Ramón Areces, 3 de julio de 2014, http://sgfm.elcorteingles.es/SGFM/FRA/recursos/conferencias/ppt/1776180509_1472014102438.docx

Massaro, Toni M. and Norton, Helen L (October 4, 2016). Siri-ously? Free Speech Rights and Artificial Intelligence. 110 Northwestern University Law Review 1169 (2016), Arizona Legal Studies Discussion Paper No. 15-29. *SSRN.* Disponible en: https://ssrn.com/abstract=2643043

Kranzberg, M (1986). Technology and History: Kranzberg's Laws, The Johns Hopkins University Press and the Society for the History of Technology. vol. 27, N° 3: pp. 544-560. Disponible en: https://www.jstor.org/stable/3105385?seq=1#page_scan_tab_contents

Meredith, A (2017). How to Use Psychographics in Your Marketing: A Beginner's Guide, *Hubspot.* Avaliable at: https://blog.hubspot.com/insiders/marketing-psychographics

Nature (2018).Cambridge Analytica controversy must spur researchers to update data ethics. A scandal over an academic's use of Facebook data highlights the need for research scrutiny, *Nature.* Disponible en: https://www.nature.com

Nix, A (2016). YouTube. Video. Disponible en https://www.youtube.com/watch?v=n8Dd5aVXLCc

Nino, C. S (2003): *Introducción al análisis del derecho*. (Buenos Aires, Argentina, Editorial Astrea, 2ª edición ampliada y revisada.12ª reimpresión).

Ohm, Paul (August 13, 2009): Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization", UCLA Law Review, Vol. 57, p. 1701, 2010; U of Colorado Law Legal Studies Research Paper No. 9-12. *SSRN.* https://ssrn.com/abstract=1450006

Pariser, E. (2011, 22 de mayo). When the Internet Thinks It Knows You. *The New York Times*. https://www.nytimes.com

Pariser, E. (SF).The Filter Bubble: What The Internet Is Hiding From You. *LSE*. http://www.lse.ac.uk

Rubinstein, Ira (October 5, 2012). Big Data: The End of Privacy or a New Beginning?, International Data Privacy Law (2013 Forthcoming); NYU School of Law, Public Law Research Paper No. 12-56. http://dx.doi.org/10.2139/ssrn.2157659

Solove, Daniel J (November 4, 2012). Privacy Self-Management and the Consent Dilemma", 126 Harvard Law Review 1880 (2013); GWU Legal Studies Research Paper No. 2012-141; GWU Law School Public Law Research Paper No. 2012-141. *SSRN*. Disponible en: https://ssrn.com/abstract=2171018

Solove, D. (2014, 7 de Julio). Facebook's Psych Experiment: Consent, Privacy, and Manipulation. *Teachprivacy*. Disponible en: https://teachprivacy.com/facebooks-psych-experiment-consent-privacy-manipulation/

Sunstein, C. (2007). Republic.com 2.0. PRINCETON; OXFORD: Princeton University Press. http://www.jstor.org/stable/j.ctt7tbsw

Terry, Nicolas P (September 27, 2012). Protecting Patient Privacy in the Age of Big Data", Indiana University Robert H. McKinney School of Law Research Paper No. 2013-04; University of Missouri-Kansas City Law Review, Vol. 81, No. 2, 2012. *SSRN*. Disponible en: http://dx.doi.org/10.2139/ssrn.2153269

THALER, R. & Sunstein, C. (2009): Nudge: Improving Decisions about Health, Wealth, and Happiness. *Yale University Press, New Haven & London*.

VALERO, Mateo (2014): "El estado del arte del Big Data & Data Science. La revolución de los datos". Texto escrito de la Jornada: Big Data: de la investigación científica a la gestión empresarial, Fundación Ramón Areces, 3 de julio de 2014.

Wang, Y. y Kosinski, M. (2018): Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal Of Personality And Social Psychology*, *114*(2), 246-257. doi: 10.1037/pspa0000098

Wells, W. D. (1975). Psychographics: A Critical Review. Journal Of Marketing Research *(JMR)*. *JSTOR*. Vol *12, No* 2, pp.196-213.

Wu, T. (2012, 19 de junio).Free Speech for Computers? *The New York Times*. Disponible en: https://www.nytimes.com

Universitas, 2020, Nº 32/ pp.79-105
ISSN 1698-7950 / doi: https://doi.org/10.20318/universitas.2020.5512