Informe

Protección legal de datos personales inferidos

Fundación Vía Libre

Mayo 2023.

Autoría: Margarita Trovato
Dirección del proyecto: Beatriz Busaniche
Asesoramiento: Enrique Chaparro
Diseño editoral: Alexia Halvorsen

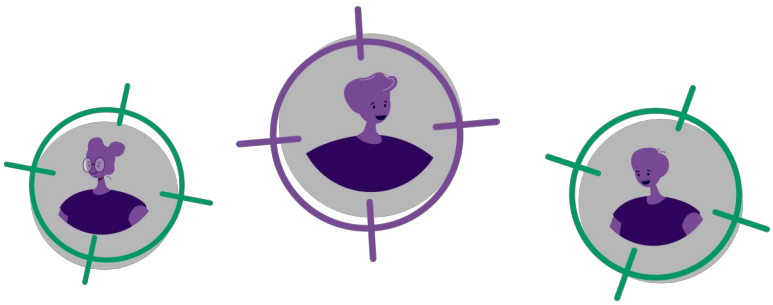
índice

Introducción

Estamos en pleno auge del Big Data: el avance de la tecnología permite no sólo la recolección masiva de información sino su análisis continuo –y generalmente automatizado– para intentar entender e incluso predecir el comportamiento humano.

Si bien en un principio este modelo tenía como objetivo únicamente direccionar mejor la publicidad (en tanto tiene la ambición de presuponer gustos o necesidades según comportamientos previos), nos encontramos ahora un paso más allá: la vigilancia constante de nuestros movimientos es el sustrato de un nuevo modelo de negocios, donde se busca "ofrecer servicios gratuitos a cambio de que el usuario conceda ser vigilado en sus interacciones digitales", en las que "(...) vuelquen informaciones relevantes para el emergente mercado de los datos" [1]. En otras palabras, es ahora la base de un modelo económico basado en el monitoreo y la vigilancia. Esto implica una lógica extractivista de datos acerca de los comportamientos al navegar en Internet, pero también sobre el uso de otros dispositivos digitales que nos rodean (teléfonos, relojes, televisores, altavoces, balanzas, heladeras, etc) para tener una visión más compleja y detallada de los comportamientos individuales y grupales, que permite, a la vez, establecer precios discriminatorios [2]. Así se habilita y alimenta este nuevo ámbito de negocios.

La existencia de este fenómeno implica distintos procesos: desde la recolección de datos (directa o indirecta) y el armado de perfiles en base a ellos, hasta la derivación de nueva información en función de aquella y su uso para seguir construyendo perfiles y para predecir futuros comportamientos [3]. Por ejemplo, establecer a qué partido votará una persona por saber qué libros compra periódicamente y en base a eso dirigir publicidad, presentar determinadas noticias primero, ofrecer ciertos servicios y/o excluirla de otros, etc.



[1] Barbudo, C. F., "Hacia una privacidad colectiva: repensar las bases teóricas de la distinción público/privado en la economía de la vigilancia", en Teknokultura. Revista de Cultura Digital y Movimientos Sociales, Eds. Complutense, diciembre 2019, pp. 72-73. Disponible en http://dx.doi.org/10.5209/tekn.66844

[2] Estas técnicas permiten y se utilizan para determinar el valor marginal que un bien tiene para una persona determinada y establecer en consecuencia, para esa persona y en esas circunstancias, el precio más alto que esté dispuesta a pagar. Al respecto ver, entre otros, Odlyzko, Andrew M. "Privacy, economics, and price discrimination on the Internet", ICEC2003: Fifth International Conference on Electronic Commerce, N. Sadeh, ed., ACM, 2003, pp. 355-366. Disponible en https://papers.ssrn.com/sol3/papers.cfm?abstract_id=429762.

[3] Blanke, J., "Protection for 'Inferences Drawn': A Comparison Between the General Data Protection Regulation and the California Consumer Privacy Act". Global Privacy Law Review 1(2):81-92, 2020, pp. 81-82, disponible en https://doi.org/10.54648/gplr2020080.

Nos referimos a las inferencias: datos que no son provistos ni cedidos por su titular sino creados por el responsable del tratamiento o por terceros, muchas veces sin siquiera conocimiento de aquel, pero que siguen refiriendo a un individuo identificado o identificable y que son usados como fundamento de decisiones que tienen un impacto directo en él. A lo largo de este informe nos proponemos entender mejor de qué hablamos cuando hablamos de datos inferidos, por qué se trata de datos personales, y presentar los problemas que este proceso y su resultado traen aparejados en términos de derechos y garantías. Veremos, por ejemplo, que se trata de información subjetiva y generalmente no verificable, que puede estar sesgada y acarrear decisiones discriminatorias, que los titulares no suelen estar al tanto de su existencia, entre otros aspectos preocupantes.

Si bien este informe no busca ser exhaustivo con respecto a todas las aristas problemáticas o potencialmente problemáticas de las inferencias, sí pretende transmitir la urgencia de contemplarlas y regularlas. Para ello ilustraremos cómo estamos rodeados de decisiones tomadas en función de ellas y los pocos o nulos mecanismos de protección y salvaguarda que nos asisten. Es que, por todo lo que permiten suponer sobre una persona y el enorme abanico de decisiones que fundamentan, “Se puede decir que las inferencias extraídas de los datos personales se han vuelto más peligrosas para la privacidad individual que la propia recopilación y almacenamiento de los datos.”[4] .

En una sociedad como la actual es indispensable reconocer la existencia y la amenaza que implican las inferencias sobre la privacidad de los ciudadanos, así como sobre todos los derechos que ésta posibilita. Para eso repasaremos algunas experiencias interesantes sobre regulación a nivel internacional y, finalmente, nos centraremos en la legislación argentina para proponer algunos estándares mínimos a la hora de sancionar una nueva norma. **Es que, como veremos a lo largo de este informe, hoy los datos no se recolectan, se infieren. Urge contemplar normativamente este proceso y proteger la información personal que de allí surge.**

Comienza el capítulo 1

¿De qué hablamos cuando hablamos de datos inferidos?

[4]Ibídem, p.1.

CAPÍTULO 1

¿De qué hablamos cuando hablamos de datos inferidos?

i. Algunas definiciones

Los datos inferidos pueden ser definidos como aquella "información relativa a un individuo identificado o identificable, creada a través de razonamiento o deducción y no sólo mediante la mera observación o recolección de sus datos personales" [5] . También como aquellos generados a través de "una operación intelectual que incluye comparación o deducción" [6] . Asimismo, fueron definidos como "la derivación de información, datos, presunciones o conclusiones a partir de hechos, evidencia u otra fuente de información o datos" [7] .

En otras palabras, lo que estas definiciones comparten es el énfasis en que:

- **no son provistos por su titular** ni pasiva ni activamente, sino que

- son **creados por el responsable del tratamiento de datos o por terceros**

- a partir de datos provistos por aquel, datos recolectados de conductas suyas sin que medie la cesión del dato (ej. en qué momento pausa un video en internet) y, en algunos casos,

- otra información de contexto [8] a través del **análisis, en general automatizado, de grandes volúmenes de datos** ("big data analytics").

[5] Wachter, S. and Mittelstadt, B. "A right to reasonable inferences: Re-thinking data protection law in the age of big data and AI", en Columbia Business Law Review (2019), disponible en https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3248829, p. 23; y Holder, A. E., "What We Don't Know They Know: What To Do About Inferences in European and California Data Protection Law", 2020, disponible en https://doi.org/10.15779/Z38MP4VP1V, p. 5. Traducción propia.

[6] European Court of Justice, Judgment of the Court (Grand Chamber) of 1 August 2022, case C-184/20 "OT v Vyriausioji tarnybinės etikos komisija". Disponible en https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A62020CJ0184.

[7] California Consumer Privacy Act (CCPA), art. (r). Disponible en https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?division=3.&part=4.&lawCode=CIV&title=1.81.5

[8] Así los define el Grupo de Trabajo del Art. 29.
Para más información, consultar Watcher, S. y Mittelstadt, B., Op. Cit., p. 23.

En este punto es importante aclarar que, incluso en el caso de aquellos datos provistos “con consentimiento” por el titular, éste no es pleno ni absoluto: en primer lugar, no es pleno porque no puede afirmarse que sea libre, en tanto vivimos en una sociedad en la que la interacción con ciertas tecnologías es inevitable (y a veces imperceptible), por lo que no es tan sencillo transpolar la visión “contractualista” del consentimiento. Por el otro, aún cuando sí lo sea, no es válido para cualquier tratamiento que se realice sobre esos datos: por ejemplo, no es lo mismo consentir que un portal de comercio electrónico sepa qué libros compro que aceptar que en base a eso asuma a quién voto. Al mismo tiempo, muchas veces el consentimiento brindado por una persona es aplicable a un conjunto de datos mayor que el suyo propio, brindando así acceso a datos de otros[9].

Sin embargo, veremos que este tipo de datos se construyen a partir de inferir preferencias de los usuarios, atributos sensibles (ej. género, orientación sexual, estado de salud) y opiniones (ej. posicionamientos políticos, visión sobre temas de actualidad) para predecir comportamientos (por ejemplo para orientar publicidad, segmentar campañas electorales o prever potenciales conductas como evasión de impuestos), es decir para tomar decisiones importantes, muchas veces incluso con efectos legales.

En términos generales, “(...) estas inferencias siguen la lógica de que si la persona P tiene los atributos p1, ..., pn existe una cierta probabilidad de que tenga (o no tenga) un cierto atributo pi no observado (por ejemplo, la propensión a adquirir ciertos bienes o servicios, o —con consecuencias más graves para la privacidad— que tenga una determinada condición de salud)”[10] , que servirá como base para determinadas acciones.

Es que las inferencias sirven para incitar a las personas a comportarse de determinada forma (marketing, publicidad), pero también para tomar decisiones sobre ellos, por ejemplo, al decidir sobre un préstamo, una beca o una contratación. El abanico de potenciales víctimas es cada vez más diverso, dado que los métodos modernos de análisis de datos permiten encontrar conexiones “pequeñas pero significativas” entre personas y construir perfiles grupales a partir de información personal, de terceros y/o anonimizada[11] .

El siguiente es un famoso caso que ilustra cómo se aplica el proceso de inferencias en la vida cotidiana:

[9] Nos referimos, entre otros supuestos, al consentimiento generacional (sobre muestras genéticas, por ejemplo) o a los “datos emergentes”, es decir aquella información que es posible deducir a partir de los datos que otros están suministrando. Sobre este punto ver Barbudo, C. F., Op. Cit., p. 72. Disponible en http://dx.doi.org/10.5209/tekn.66844
[10] Fundación Vía Libre, Comentario sobre el Anteproyecto de Reforma de la Ley de Protección de Datos Personales, 15/11/2022, pp. 4-5. Disponible en https://www.vialibre.org.ar/aporte-ley-datos/.
[11] Wachter, S. and Mittelstadt, B., Op. Cit., p. 13.

Un día de 2003, un hombre entró enfurecido en un comercio multirubro (“Target”) en el Estado de Minnesota, EEUU, y exigió hablar con un responsable. A su hija adolescente le habían enviado una carta celebrando un embarazo y ofreciéndole cupones para ropa y accesorios para bebés, y él quería saber por qué la tienda la alentaba a quedar embarazada.

Días después, cuando el gerente lo llamó para disculparse, el hombre a su vez se disculpó insistentemente con él y le informó que su hija había admitido estar embarazada. **Ella no le había revelado esta información a Target, sino que la empresa la había averiguado analizando su historial de compras con un software destinado a predecir la probabilidad de que una mujer estuviera embarazada.** Encontraron, por ejemplo, que alrededor de la semana 16 compraban cremas sin perfume y durante las primeras 20 un conjunto de vitaminas.

La adolescente había comprado varios productos de un lista de 25 diferentes que, analizados juntos, le permitieron a Target “adivinar” en qué trimestre estaba y estimar su fecha de parto para poder enviarle cupones cada vez que estuviera a punto de hacer una compra nueva.

Este ejemplo es citado frecuentemente por académicos y expertos. Ver, por ejemplo, Holder, A. E., “What We Don’t Know They Know: What To Do About Inferences in European and California Data Protection Law”, 2020, p.1. disponible en https://doi.org/10.15779/Z38MP4VP1V, p. 1, y Blanke, J., Op. Cit., p. 81.

Vemos en este caso cómo el resultado de una inferencia afectó desde la privacidad hasta la autonomía y potencialmente la integridad física de la adolescente, sin que ella haya brindado información sobre su situación ni su consentimiento para realizar este análisis y tomar acciones (de marketing, en este caso) en base al resultado. Veremos en la sección 3.i más ejemplos para dimensionar su impacto en la vida cotidiana.

Además, al ser creadas por terceros a través de métodos más allá de la mera observación, **las inferencias resultan un tipo de información subjetiva y no verificable**[12] .

Por ejemplo las evaluaciones de los bancos para aprobar préstamos, las de las compañías de seguros al fijar la prima para cada asegurado o incluso en el mundo laboral, cuando se evalúa el mérito para un ascenso[13], son deducciones hechas por terceros donde se infieren opiniones subjetivas o características que no son simplemente "observables" de los datos que poseen. En otras palabras, este tipo de decisiones que toman los bancos, seguros o empleadores no se basan en datos objetivos brindados por cada individuo, sino que se "crean" en base a aquellos que sí brindó o que se obtuvieron de su contexto, dando como resultado información subjetiva. De hecho, fueron calificados como "opiniones" en tanto se trata de un tipo de "datos personales que constituyen una aseveración sobre una persona, construídos sobre un marco interpretativo para producir nuevos hechos probables sobre esa persona."[14] .

A la vez, la mayoría de las veces se tratará de afirmaciones difíciles o imposibles de verificar. Por ejemplo, es común que los seguros de autos fijen su prima según cuán "peligroso" puede ser un conductor, aunque esta evaluación no se le comunica a los clientes. Se infiere, por ejemplo, de las horas de sueño registradas por distintas aplicaciones. Si bien este tipo de datos permite construir o arribar a otros, lo cierto es que serán imposibles de verificar: ¿cómo asegurar que va a ocurrir?. De hecho, estudios recientes muestran que el impacto de pocas horas de sueño en el desempeño profesional no está demostrado por falta de mediciones objetivas de ambas variables[15]. Las decisiones tomadas sobre esta base, sin embargo, sí tendrán efectos certeros sobre los usuarios: siguiendo el ejemplo, la prima del seguro sí subirá o bajará. ¿Cómo podría un usuario discutir el valor de su póliza de seguros basado en una premisa falaz, si ni siquiera conoce que se basó en ella? Veremos en la sección 5 por qué esto requiere una regulación especial de manera urgente.

Estas características se vuelven aún más riesgosas en tanto las inferencias se usan para crear perfiles (tal como veremos más adelante) y para luego extraer otra información sobre éstos. En otras palabras, las inferencias se toman en general como de parte permanente del perfil de una persona o de un grupo del que una persona forma parte, sin aclarar que no es información "cruda" sino construida, y por lo tanto dándola por cierta, como si se tratara de hechos[16] .

[12] Wachter, S. and Mittelstadt, B., Op. Cit., p. 27, con cita al Grupo de Trabajo del art. 29, "Dictamen 4/2007 sobre el concepto de datos personales", 01248/07/ES WP 136, p. 6. Disponible en https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_es.pdf.
[13] Grupo de Trabajo del art. 29, "Dictamen 4/2007 sobre el concepto de datos personales", 01248/07/ES WP 136, p. 6. Disponible en https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_es.pdf
[14] Hallinan, D. and Borgesius, F. Z., . "'Opinions Can Be Incorrect (in Our Opinion)! On Data Protection Law's Accuracy Principle", International Data Privacy Law 10, no. 1 (February), 2010, p.1, disponible en https://doi.org/10.1093/idpl/ipz025.
[15] Park and Arian, et al., "Online Mobile App Usage as an Indicator of Sleep Behavior and Job Performance", en WWW '21: Proceedings of the Web Conference 2021, abril de 2021, junio del 2021, p.2. Disponible en https://arxiv.org/abs/2102.12523.
[16] Ibídem, p. 85.Ibídem, p. 85.

De hecho, las inferencias se basan en procesos estadísticos (la posibilidad de que si alguien cumple con determinados parámetros entonces adopte ciertos comportamientos, tenga ciertos gustos, etc.), otro punto problemático para destacar es que, cuando se trata de procesos inferenciales automatizados, no está asegurada la causalidad entre el dato "crudo" que se tiene del individuo y el que se infiere: no es lo mismo correlación que causalidad[17] . Dicho de otro modo, si bien es cierto que las dos variables pueden estar relacionadas, las herramientas actuales fallan al detectar si eso sucede porque son causa-consecuencia o simplemente por casualidad.

Es fácil encontrar ejemplos de hechos que estadísticamente se correlacionan (evolucionan proporcionalmente, por caso), pero que desde una perspectiva lógica humana podemos advertir fácilmente que no hay causalidad entre ello: existen aumentos y descensos proporcionales (una correlación muy cercana) entre el número de doctorados en matemática concedidos en EE.UU. y el uranio almacenado en las plantas nucleares de ese país entre 1996-2008, así como entre la edad de la ganadora anual de Miss América y las muertes por vapor y objetos calientes en un período similar, o entre personas ahogadas por caerse en una pileta y películas en las que apareció Nicolas Cage, etc[18]. Más allá de lo grotesco, estos ejemplos sirven para advertir que, si bien es cierto que estadísticamente hay una correlación, no podemos de ninguna manera afirmar que uno sea consecuencia de otro: claramente más uranio no implica más doctores, ganadoras de América más jóvenes no prevendrán las muertes por vapor y si Nicolas Cage deja de actuar no se acabarán los ahogamientos.

Entonces, y como profundizaremos más adelante, la toma de decisiones basadas en inferencias es problemática en varios sentidos:

1. Puede afectar derechos básicos como la privacidad, reputación, igualdad y no discriminación, identidad, autodeterminación informacional, libertad de expresión, los derechos de consumidor, entre otros.
2. Se trata de información subjetiva y no verificable.
3. Suelen quedar fuera de la órbita de acceso y control por parte de sus titulares, en tanto no están claros normativamente los derechos que traen aparejados.
4. En tanto se trata de un proceso crecientemente mediado por inteligencia artificial y análisis de Big Data, la relación entre acciones que uno lleva a cabo y las percepciones o deducciones que se pueden generar:
   - a. No siempre se basa en una causalidad lógica (a pesar de que exista una correlación entre las variables).
   - b. Se vuelve cada vez menos intuitiva y más difícil de imaginar o suponer por parte de los titulares.

[17] Ibídem, p. 85.Ibídem, p. 85.
[18] Para más ejemplos de correlaciones no causales ver https://www.tylervigen.com/spurious-correlations

Todo esto lleva a crear oportunidades para la toma de decisiones y elaboración de perfilamientos discriminatorios, sesgados e invasivos [19], procesos que revisaremos más adelante.

Sin embargo, si bien se suele hacer foco en el hecho de que se arribe a ellos a través de procesos automatizados, lo cierto es que **lo más problemático no es el método a través del que se construyen sino para qué se usan, qué decisiones fundamentan**. Las inferencias implican usar data existente para generar nueva información o predicciones sobre una persona, por lo que **deberíamos enfocarnos más en la legalidad y salvaguardas del proceso de inferencia que en el hecho de que ésta sea automatizada** [20] .

Como veremos a lo largo de este informe, el marco normativo nacional e internacional de protección de datos personales no parece ser suficiente para proteger efectivamente de los nuevos riesgos que suponen los procesos de inferencias. De hecho, su estatus jurídico es disputado en la academia y jurisprudencia interpretativa de los instrumentos de protección clásicos de datos personales.

Es que hoy como individuos tenemos muy poco control sobre cómo se usan nuestros datos personales para hacer inferencias sobre nosotros: los datos inferidos no estarían reconocidos como tales en las normas (salvo en algunas excepciones que veremos), y por lo tanto no parece ser clara su protección, los derechos y garantías que traen aparejados, la capacidad o mecanismos de control sobre su uso y producción, ni los remedios para cuestionar decisiones tomadas en base a ellos. Si bien "uno esperaría que información de este tipo, que es la que dirige la economía moderna -y en la que se invierte tanto para desarrollar-, estuviera estrictamente regulada (...) la regulación de estas inferencias permanece, por el momento, imperfecta, incompleta e incierta."[21] .

Vimos entonces cómo los datos inferidos hacen referencia a información relativa a un individuo identificado o identificable. En esa medida, **la información inferida también será personal.**

[19] Wachter, S. and Mittelstadt, B., Op. Cit., p. 4., con cita a Brent Daniel Mittelstadt, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter & Luciano Floridi, The Ethics of Algorithms: Mapping the Debate, BIG DATA & SOC'Y, July–Dec. 2016, pp. 1–2.
[20] Solove, D. J., "The Limitations of Privacy Rights", en 98 Notre Dame Law Review 975 (2023), disponible en https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4024790
[21] Para más información sobre la regulación de distintos sistemas regionales de DDHH sobre el derecho a la privacidad, ver Maqueo Ramírez, M. S., Moreno González, J. y Recio Gayo, M, "Protección de datos personales, privacidad y vida privada: la inquietante búsqueda de un equilibrio global necesario", Rev. Derecho (Valdivia) vol.30 no.1, 2017. Disponible en http://dx.doi.org/10.4067/S0718-09502017000100004

ii. Los datos inferidos como datos personales

Veamos un ejemplo:

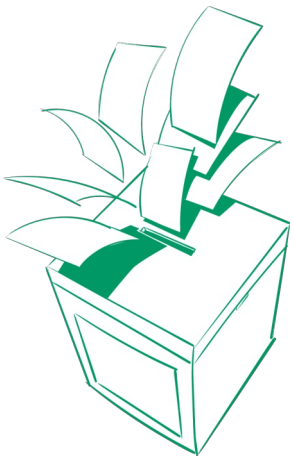
Un partido político hace inferencias sobre las probables características de personas viviendo en un determinado distrito electoral. El partido asocia esta información con los nombres y direcciones postales que constan en el padrón electoral. Categoriza a las personas y les asigna un porcentaje de probabilidad (“puntaje”) de apoyo al partido.

Estos son datos personales. Esta información se relaciona con individuos identificables, en tanto las características inferidas, las categorías y los puntajes están asociados a sus nombres y direcciones.

Un partido político hace inferencias sobre las probables características de personas viviendo en determinados distritos electorales. El partido recibe la información de cada distrito como un conjunto y no intenta asociarla a nombres o direcciones individuales. Categoriza los distritos y asigna un porcentaje de probabilidad (“puntaje”) de apoyo al partido en cada zona.

Estos datos no son personales. La información no se relaciona con individuos identificables en tanto las características inferidas, categorías y puntajes se asocian a zonas generales.

Ejemplo extraído de la web oficial de la Information Commissioner's Office, un organismo independiente del Reino Unido que trabaja por la defensa de los derechos de la información. Disponible en https://ico.org.uk/for-organisations/guide-to-data-protection/key-dp-themes/guidance-for-the-use-of-personal-



Además de tratarse de información que permite identificar a un individuo en particular, sabemos que éste no suele tener conocimiento de las inferencias realizadas sobre él o sobre los conjuntos de los que se los considera parte, y que son la base de decisiones que lo afectarán directamente. Por su naturaleza, vulneran tanto la **privacidad** como la **autodeterminación informacional**, y numerosos derechos personalísimos que se desprenden de ellos: desde la intimidad hasta la identidad, reputación, libertad de expresión, derecho a la salud, a un trabajo digno, a la educación y derechos de consumidor, entre otros.

Es que, indefectiblemente, la **privacidad** no sólo es un derecho en sí misma, sino que funciona además como condición para el ejercicio de otros. Sobre el derecho a la vida privada, la Corte Interamericana de Derechos Humanos[22] tiene dicho que:

"La protección a la vida privada abarca una serie de factores relacionados con la dignidad del individuo, incluyendo por ejemplo la capacidad para desarrollar la propia personalidad y aspiraciones, determinar su propia identidad y definir sus propias relaciones personales. (...) La efectividad del ejercicio del derecho a la vida privada es decisiva para la posibilidad de ejercer la autonomía personal sobre el futuro curso de eventos relevantes para la calidad de vida de la persona. La vida privada incluye la forma en que el individuo se ve a sí mismo y cómo decide proyectarse hacia los demás, y es una condición indispensable para el libre desarrollo de la personalidad" [23] .

Es fundamental, en este sentido, contar con herramientas jurídicas que aseguren su protección y promuevan su vigencia de forma amplia.

Por otro lado, la **autodeterminación informacional,** entendida como el derecho de cada persona a decidir por sí misma sobre la publicidad, divulgación y uso de sus datos personales, también es condición para el ejercicio de otros derechos y por lo tanto debe ser estrictamente garantizada. Es que se trata de "(...) reconocer a las personas una serie de facultades jurídicas que se les atribuyen precisamente para enfrentar las extralimitaciones (...) y que puedan evitar que de su mal uso lesionen bienes o derechos constitucionales como la intimidad y los derechos conexos."[24] .

Este derecho fue reconocido como tal por primera vez por el Tribunal Constitucional Federal alemán (Bundesverfassungsgericht) en 1983 en una sentencia sobre el censo[25] , basándose en la dignidad humana y el derecho a desarrollar libremente la personalidad.

[22] Para más información sobre la regulación de distintos sistemas regionales de DDHH sobre el derecho a la privacidad, ver Maqueo Ramírez, M. S., Moreno González, J. y Recio Gayo, M, "Protección de datos personales, privacidad y vida privada: la inquietante búsqueda de un equilibrio global necesario", Rev. Derecho (Valdivia) vol.30 no.1, 2017. Disponible en http://dx.doi.org/10.4067/S0718-09502017000100004

[23]Corte IDH, Caso Artavia Murillo y Otros vs. Costa Rica, Sentencia de Fondo (Excepciones Preliminares, Fondo, Reparaciones y Costas), 2012, párr. 143.

[24] Castillo Córdova, L., Comentarios al Código Procesal Constitucional, tomo II, Ed. Palestra, Perú, 2006, p. 975, citado en Orrego, A. C., "Una aproximación al contenido constitucional del derecho de autodeterminación informativa en el ordenamiento jurídico peruano", en Anuario de Derecho Constitucional Latinoamericano, año XIX, Bogotá, 2013, p. 317. disponible en https://biblioteca.corteidh.or.cr/tablas/r32202.pdf

[25] BVerfG, fallo del Primer Senado de 15 de diciembre de 1983, 1-BvR-209/83, § 1-215. Disponible en inglés en https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/EN/1983/12/rs19831215_1bvr020983en.html. Traducciones propias.

La entiende como “el derecho del individuo, derivado de la autodeterminación, a decidir por sí mismo cuándo y dentro de qué límites se revelan hechos y situaciones de la vida personal.” [26] y agrega que debe contar con medidas específicas de protección, en la medida en que la elaboración automática de datos permite que la información se funda “(...) con otras colecciones de datos en un perfil de personalidad parcial o ampliamente definido, sin que el interesado pueda controlar suficientemente su exactitud y su utilización.”[27] Menciona el peligro de que sus acciones se vean cohibidas por el temor de que trasciendan[28] y concluye que “El derecho a la autodeterminación informacional no sería compatible con un orden social y su ordenamiento jurídico subyacente en el que los ciudadanos ya no puedan saber quién, qué, cuándo y con qué motivo sabe algo sobre ellos.” [29]

En función de esto, **se vuelve clara la necesidad de entender a los datos inferidos como datos personales**, en tanto su naturaleza y posibles afectaciones son los mismos, y por lo tanto garantizarles la misma protección jurídica: es que, en tanto revelan información sobre una persona identificada o identificable, devienen personales y corresponde en consecuencia que rija su misma protección.

Dicho de otro modo, para asegurarles a los titulares de los datos inferidos derechos amplios, sin restricciones (para protegerlos, discutirlos, acceder a ellos, etc), es necesario considerarlos jurídicamente como datos personales.

Veremos que los instrumentos internacionales más modernos tienden a incluirlos de forma explícita (aunque aún son poco claros o insuficientes al respecto) y que sus interpretaciones más protectoras de derechos promueven esta línea.

[26] BVerfG, cit., § 144
[27] Ibíd., § 145
[28] Ibíd., § 145, 146: “Esto no sólo perjudicaría las posibilidades de desarrollo del individuo, sino también el bien común, porque la autodeterminación es una condición funcional elemental de una comunidad democrática libre basada en la capacidad de sus ciudadanos para actuar y participar. ”. Esta autocensura también es conocida como “chilling effect”. Ver sección 3.
[29] Ibíd., § 146.

Sobre este punto, es frecuente escuchar posturas contrarias a la idea de considerar a las inferencias como datos personales y por ende otorgarles igual estatus jurídico.

En primer lugar, sostienen que no deben serlo en tanto no se trata de datos aportados por su titular. Sin embargo, es posible sostener que con más razón deberían entonces protegerse: el hecho de que no sea éste quien los aporte en nada obsta a que se trate de información sensible, personal, que deba quedar bajo el mismo (o incluso mayor) paraguas de protección que los del mismo tipo que sí brinda.

Por otro lado, alegan que usualmente consisten en presunciones probabilísticas y no conocimiento probado ni, en muchos casos, pasible de ser probado (por ejemplo, cuando se categoriza a alguien como "conductor peligroso"). En este punto, es importante destacar que cualquier información o evaluación sobre una persona, sea verificable o no, correcta o incorrecta, puede tener consecuencias graves para el titular. De hecho, es posible que aquellos incorrectos sean aún más gravosos y discriminatorios por lo que, con más razón, deberían quedar comprendidos en el espectro de protección de los datos personales[30]. Al tratarse de decisiones muchas veces automatizadas, se presenta una exigencia mayor y más precisa de rendición de cuentas sobre cómo suceden estos procesos que tienen como resultado información invasiva y no verificable.

Al mismo tiempo, es necesario garantizar herramientas para entenderlas y discutirlas: que estas inferencias no estén probadas o no sean verídicas no sólo no impide que deban ser consideradas como datos personales (según lo entiende gran parte de la doctrina y la jurisprudencia[31]), sino que incluso "las normas de protección de datos prevén la posibilidad de que la información sea incorrecta y confieren al interesado el derecho de acceder a esa información y de refutarla a través de los medios apropiados"[32]. Es necesario entenderlas dentro del paraguas de protección de datos personales para asegurar el ejercicio de derechos de los titulares.

Por último, es muy frecuente el argumento de que, como el algoritmo mediante el cual se procesan los datos originarios y se construyen las inferencias se encuentra bajo secreto comercial por constituir propiedad intelectual de las empresas, éstas no pueden someterse al mismo régimen de acceso y transparencia que los datos personales. Indefectiblemente, los derechos personalísimos -como la privacidad-[33] priman sobre aquellos patrimoniales -como la propiedad intelectual-.

[30] Kröger, J. L., Op. Cit., p. 2.
[31] Wachter, S. and Mittelstadt, B., Op. Cit., p. 28 y sus referencias.
[32] Grupo de Trabajo del art. 29, "Dictamen 4/2007 sobre el concepto de datos personales", Op. Cit., p. 7.
[33] Los derechos personalísimos se encuentran reconocidos en el Código Civil y Comercial de la Nación, capítulo 3 (art. 51 y ss). La doctrina los definió como "(...) una inconfundible categoría de derechos subjetivos esenciales, que pertenecen a la persona por su sola condición humana y que se encuentran respecto de ella en una relación de íntima conexión, casi orgánica e integral." (Rivera, J. C., Instituciones de Derecho Civil. Parte General, Bs. As., Ed. Abeledo Perrot, 2010, p. 704).

Sobre este punto respondió acabadamente el Procurador General de California, sosteniendo que no había recibido "(...) ningún ejemplo concreto de situaciones donde las inferencias fueran en sí mismas secretos comerciales, o donde la divulgación de inferencias expondría secretos comerciales. Mientras que el algoritmo que usa una empresa para derivar sus inferencias sí podría serlo, el CCPA sólo requiere la divulgación de productos suyos individualizados, no el algoritmo en sí". Y concluyó que "inferencias generadas internamente que una empresa posee sobre un consumidor **es información personal** según la entiende el CCPA, y debe divulgarse a pedido del consumidor. Una empresa que retiene inferencias con el argumento de que son secretos comerciales protegidos tiene la carga de probar que en efecto lo son bajo la ley vigente"[34].

Entonces, si asimilamos los datos inferidos a personales -en tanto refieren a una persona humana identificada o identificable-, se abre un nuevo escenario. Deberían aplicarle, naturalmente, los mismos derechos que a los personales: de mínima, aunque no exclusivamente, los clásicos de acceso, rectificación, cancelación y oposición ("ARCO" por sus siglas). Pero es indispensable crear las condiciones necesarias para poder ejercerlos: asegurar mecanismos de rendición de cuentas sobre cómo se producen, establecer rigurosamente quién debe probar su exactitud (para dar lugar a eventual rectificación), clarificar las reglas del consentimiento, si el alcance de los derechos que los amparan es exactamente el mismo que a los datos personales (vimos, por ejemplo, que el RGPD los excluye de la portabilidad), etc. Todas estas cuestiones demuestran la urgencia de regularlos, cuestión sobre la que nos explayaremos en la sección 5 de este informe.

Por eso, más allá de la discusión conceptual sobre su naturaleza jurídica, lo cierto es que es fundamental proteger la información personal que surge de las inferencias. Como sostienen Wachter y Mittelstadt, "Es crucial señalar, sin embargo, que la cuestión sobre si las inferencias son o no datos personales no es la más importante. El problema subyacente es mucho más profundo y se relaciona con la tensión sobre si los individuos tienen derechos, control y recursos sobre cómo son vistos por otros"[35]. Veremos a lo largo de este trabajo cómo estas cuestiones se ven obstaculizadas en el caso de los datos inferidos con los marcos regulatorios actuales.

Comienza el capítulo 2

Construcción y uso de los datos inferidos

[34] Cal. Op. Att'y. Gen. No. 20-303 (Cal. A.G.), pp. 14 y 15.
[35] Wachter, S. and Mittelstadt, B., Op. Cit., p. 6. Traducción propia.

CAPÍTULO 2

Construcción y uso de los datos inferidos

i. ¿Cómo se construyen?

Como mencionamos al principio, la masividad y alcance de las tecnologías permiten la recolección masiva de tipos de datos y en cantidades que antes eran impensables. Hoy en día muchos de los objetos que usamos cotidianamente tienen la posibilidad de recolectar datos sobre nosotros: desde el chip del celular que permite saber la ubicación, hasta la balanza que registra nuestro peso o los cepillos de dientes eléctricos que pueden almacenar cuántas veces por día nos los lavamos o por cuánto tiempo. También en el mundo digital dejamos nuestros datos indirectamente e incluso a veces sin notarlo: la mayoría de los sitios y plataformas desarrollaron formas de seguir a los usuarios a través de internet y observar sus hábitos en línea (un ejemplo conocido es a través de las cookies), para poder agregar estos datos a sus registros[36].

De esto se desprenden dos consecuencias: por un lado, en la actualidad es simple y frecuente para las empresas conseguir información sobre usuarios sin ninguna acción de parte de ellos. Por otro, en tanto la tecnología lo habilita y a la vez el mercado lo demanda, se desarrollaron técnicas más efectivas de procesamiento masivo de datos, que permitieron que se complejicen y sofistiquen las inferencias sobre usuarios, en tanto cada vez es más fácil y frecuente asignar puntajes o crear perfiles. Es que estos procesos habilitan a las empresas a "extraer nuevos conocimientos y perspectivas o crear nuevas formas de valor, en formas que cambian mercados, organizaciones, la relación entre ciudadanos y gobiernos, y más." [37]. En el caso de la publicidad, por ejemplo, las plataformas prometen habilitar una llegada más precisa a los anunciantes para volverse más competitivos en su sector del mercado[38].

ii. ¿Para qué se usan?

Como dijimos, las inferencias funcionan principalmente como base para procesos conocidos como ***profiling*** (generalmente traducido como elaboración de perfiles o perfilamiento) y ***scoring*** (asignaciones de puntaje), cuyos resultados son usados para dirigir publicidad, fijar precios diferenciales para cada consumidor, tomar decisiones sobre clientes, estudiantes, trabajadores, etc.

[36] Para más información sobre este punto y otros ejemplos, ver Holder, A. E., Op. Cit. pp. 7 y 8.
[37] Viktor Mayer-Schönberger & Kenneth Cukier, Big Data: A Revolution That Will Transform How We Live, Work, And Think (2013), p. 6., referenciado en Holder, A., E., Op. Cit., p. 8.
[38] Holder, A. E., Op. Cit., pp. 6-9.

En cuanto al *profiling*, se trata de un proceso usado frecuentemente para hacer predicciones sobre personas: **se usa información de varias fuentes para inferir algo sobre un individuo basándose en las cualidades de otros que estadísticamente parecen asimilarse.** Así crea nueva información sobre ellas, que no fue provista directamente por ellas.

Podemos definirlo como "toda forma de tratamiento automatizado de datos personales consistente en utilizar datos personales para evaluar determinados aspectos personales de una persona física, en particular para analizar o predecir aspectos relativos al rendimiento profesional, situación económica, salud, preferencias personales, intereses, fiabilidad, comportamiento, ubicación o movimientos de dicha persona física"[39]. Es decir que se configura con tres elementos: 1) ser una forma automatizada de procesamiento, 2) realizarse con datos personales y 3) tener como objetivo evaluar aspectos personales de un individuo. En función de este último punto, vemos que no cualquier clasificación de personas basada en características conocidas (por ej. edad, sexo, altura) implica necesariamente un *profiling*, sino que dependerá de su propósito: por ejemplo, una empresa puede clasificar a sus clientes según su género con fines estadísticos, pero sin hacer predicciones o sacar conclusiones de cada uno en particular y por lo tanto no tratarse de un profiling de ellos[40]. Por último, es importante destacar que suele ser un proceso poco transparente, del que las personas aludidas no suelen enterarse: "El proceso de elaboración de perfiles suele ser invisible para el interesado. (...) Las personas tienen distintos niveles de comprensión y les puede resultar difícil entender las complejas técnicas de los procesos de elaboración de perfiles y decisiones automatizadas."[41].

El siguiente es un clásico ejemplo de *profiling*:

> Un comercio podría listar los clientes que compraron artículos de camping durante el año pasado, identificar que eran todos hombres con un cierto código postal y un determinado límite en sus tarjetas de crédito, y usar esta información para agregar a clientes con condiciones similares en una base de datos más grande que se llame "Clientes interesados en comprar artículos de camping".

Ejemplo extraído de "Data Brokers: A Call For Transparency and Accountability: A Report of the Federal Trade Commission", mayo 2014, disponible en https://www.ftc.gov/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014, referido en Holder, A. E., Op. Cit.
Para más ejemplos sobre perfilamiento ver Privacy International, Examples of Data Points Used in Profiling (2017), disponible en https://privacyinternational.org/sites/default/files/2018-04/data%20points%20used%20in%20tracking_0.pdf.?q=controllers. Traducción propia.

[39] RGPD, art. 4.4).
[40] Grupo de Trabajo del Art. 29, Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679, 17/EN, WP251rev.01, art. II.A. Disponible en español en https://ec.europa.eu/newsroom/article29/items/612053/en
[41] Grupo de Trabajo del Art. 29, Guidelines on Automated Individual Decision-Making (...), art. III.A.I

scoring

Por otro lado, hablamos de *scoring* cuando se trata del uso de inferencias no para armar perfiles de personas pero sí para ranquear ciertos rasgos y comportamientos.

> Las campañas políticas ordenan a los espectadores de televisión según la posibilidad de que voten o no al candidato según qué vieron y cuándo; algunas empresas convierten sus análisis de datos e intereses de clientes en "puntajes" de marketing que los clasifican según cuán posible es que respondan a cierta publicidad o que hagan una compra, cómo es su presencia en la web o cuánto influencian a otros, u otras métricas, etc.

Ejemplos presentados en Holder, A. E., Op. Cit., p. 6.

decisiones automatizadas

En este punto es importante introducir la noción de "decisiones automatizadas" con el objetivo de evitar confusiones: se trata de "la habilidad de tomar decisiones a través de medios tecnológicos sin intervención humana" y puede basarse en cualquier tipo de datos, tanto entregados directamente por las personas involucradas (ej. contestando un cuestionario), observados sobre ellas (ej. su ubicación a través de una aplicación) o derivados/inferidos de otros que se tienen sobre ellas. Vemos entonces que puede implicar o no un proceso de profiling o de scoring, dependiendo de cómo y para qué se use la data producida. A la vez, éstos últimos pueden ser o no una decisión automatizada según si la evaluación de cada persona se hace a través de este tipo de tecnologías. En otras palabras, no son excluyentes pero tampoco se dan necesariamente en simultáneo[42].

Veamos un ejemplo:

[42] Grupo de Trabajo del Art. 29, Guidelines on Automated Individual Decision-Making (...), art. II.B.

> “La imposición de multas por exceso de velocidad únicamente sobre la base de las pruebas de los radares de velocidad es un proceso de decisiones automatizadas que no implica necesariamente la elaboración de perfiles. Sin embargo, puede convertirse en una decisión basada en la elaboración de perfiles si los hábitos de conducción de la persona se supervisan a lo largo del tiempo y, por ejemplo, la cuantía de la multa impuesta es el resultado de una evaluación que implique otros factores, como si el exceso de velocidad es un caso de reincidencia o si el conductor ha cometido otras infracciones de tráfico recientemente.”

Grupo de Trabajo del Art. 29, Guidelines on Automated Individual Decision-Making (...), art. II.B.

Como mencionamos en la introducción, **lo más problemático no es el método a través del que se construyen los datos inferidos (automatizado, manual, etc) sino para qué se usan, qué decisiones fundamentan y el poco control de parte de los titulares.** Todas estas técnicas representan nuevas oportunidades para **decisiones discriminatorias, sesgadas e invasivas de la privacidad**, tomadas en base a análisis inferidos[43].

En resumen, entonces, se utilizan datos personales como insumo para alimentar un algoritmo que según ciertas indicaciones derivará en un determinado resultado, lo cual tiene varios problemas: hay que tener mucho cuidado con cómo/con qué se alimenta ese algoritmo, para evitar sesgos, desviaciones o resultados discriminatorios; en función de esa automatización se producen nuevas informaciones (personales, privadas) sobre nosotros, muchas veces sin una correlación verdaderamente causal; se nos perfila e incluye en “grupos” o “listas” de los que no estamos al tanto, muchas veces no queremos ser parte y es muy difícil apartarse; en base a esa pertenencia se toman decisiones con impacto relevante y directo sobre nosotros (veremos ejemplos a continuación); y no parecen existir mecanismos efectivos de reclamo.

Comienza el capítulo 3

Relevancia e Impacto en la vida de las personas

[43] Wachter, S. and Mittelstadt, B., Op. Cit., p. 13.

CAPÍTULO 3

Relevancia e impacto en la vida cotidiana

Sabemos entonces que las inferencias pueden ser muy invasivas: desde el embarazo inferido por parte de la cadena de comercios Target en EEUU que presentamos al inicio, hasta el caso de investigadores que lograron inferir el nivel de satisfacción de usuarios haciendo búsquedas en internet según su interacción con el mouse o el sistema de scoring de alcance masivo para el acceso a créditos sociales en China[44].

Además, pueden ser hechas sin que se enteren las personas en cuestión y son muchas veces contraintuitivas, poco "imaginables": esto lleva a afectaciones de los derechos a la privacidad, identidad, reputación y autodeterminación informacional. Más aún, el tipo de invasión a la privacidad va más allá de lo que tradicionalmente se concebía: pueden implicar amenazas mediatas o casi inmediatas a la integridad personal, a la autonomía, a derechos civiles e incluso a perspectivas financieras de miembros de poblaciones vulnerables[45] (veremos más adelante algunos ejemplos).

Es que, frente a la incertidumbre de qué podría pasar con su información, las personas pueden alterar su comportamiento al usar estos tipos de tecnologías. Esta autocensura, que se conoce como *chilling effect* (literalmente, "efecto de congelamiento"), socava la **libertad de expresión y la autodeterminación** sobre las acciones, elecciones, preferencias y demás de cada persona, y termina afectando su propia identidad, en tanto cambia las conductas que adoptaría en función de lo que se podría llegar a derivar de ellas[46]. Por ejemplo, "(...) quien se siente inseguro de si en todo momento se registran cualesquiera comportamientos divergentes y se catalogan, utilizan o transmiten permanentemente a título de información, procurará no llamar la atención con esa clase de comportamiento. Quien sepa de antemano que su participación, por ejemplo, en una reunión o en unas iniciativa cívica va a ser registrada por las autoridades y que podrán derivarse riesgos para él, por este motivo renunciará presumiblemente a lo que supone un ejercicio de los correspondientes derechos fundamentales"[47].

[44] Para más información sobre estos casos, ver ibídem p. 16 y sus referencias.
[45] Holder, A. E., Op. Cit., p. 28.
[46] En lo que concierne a la UE, se pronunció sobre el tema el Tribunal Europeo de DDHH. Para más ifnormación ver Council of Europe, Case Law of the European Court of Human Rights Concerning the Protection of Personal Data, T-PD(2017)23 (2017), https://rm.coe.int/case-law-on-data-protection/1680766992 [https://perma.cc/H4F2-9WVZ], citado por Wachter, S. and Mittelstadt, B., Op. Cit., p. 19.
[47] Tribunal Constitucional Federal Aleman, sentencia de 15 de diciembre de 1983 (Ref. 1 BvR 209/83), Fondo - Ley del Censo.

Existe otra característica problemática propia de la información generada en estos procesos: no sólo no suele ser verificable, los titulares no suelen enterarse y potencialmente afecta numerosos derechos, sino que trata además de **registros "persistentes"**. A diferencia de lo que sucedía anteriormente, en la actualidad el análisis que se puede realizar sobre una persona o un grupo de personas se mantiene almacenado indefinidamente y sirve de base para futuros análisis, por ejemplo de *profiling* o *scoring*, e incluso por parte de terceros. Se abren así nuevos escenarios de afectación a derechos como la reputación y la identidad. Este tipo de vulneraciones está contemplada en relación con los datos personales (a través del "derecho al olvido") pero, en tanto no existe una regulación de los datos inferidos, tampoco es tampoco es clara su protección o derechos aplicables en estos casos.

Además, como mencionamos, **no existen aún mecanismos que permitan a los usuarios afectados o interesados conocer esta información, acceder a ella, discutirla, pedir su rectificación o eliminación, etc.** A la vez, no rige tampoco obligación legal para los responsables del tratamiento de revelar o justificar el tipo de análisis que realizan y en base al cual toman decisiones. "Actualmente, los individuos no tienen asegurado conocer procesos de tomas de decisión potencialmente problemáticos, y en general carecen de marcos legales para examinarlos"[48].

Si bien los riesgos fueron contemplados por la academia y la doctrina, **en muchos casos todavía no se plasmaron en cambios normativos.** Como veremos más adelante, la legislación vigente (europea, principalmente), se enfoca más en la protección de la información "primaria" que luego será procesada, que de aquella que surge del procesamiento. Este enfoque no es suficiente en pleno auge del Big Data: en términos de afectación de derechos y mecanismos de protección, los que conllevan más riesgos son los que resultan del proceso de análisis y tratamiento, que además son fundamento de decisiones que nos impactan cotidianamente.

i. Algunos ejemplos

Como mencionamos, el problema de los análisis inferidos, especialmente de grandes volúmenes de información -más allá de la técnica con la que se realicen- es que traen aparejado el riesgo de afectaciones preocupantes sobre individuos y grupos. Es que se utilizan para tomar decisiones que van desde publicidad dirigida hasta empleabilidad, riesgos asegurables, cuestiones crediticias y políticas de Estado. Veremos a continuación algunos ejemplos con la intención de ilustrar el abanico de consecuencias en la vida cotidiana[49].

[48] Wachter, S. and Mittelstadt, B., Op. Cit., p. 21. Traducción propia.
[49] Extraídos de ibídem, p. 13 y ss, así como de sus referencias citadas.

Las principales **plataformas** son tal vez el caso más conocido: Facebook, por ejemplo, puede inferir la orientación sexual y otras características protegidas (raza, opiniones políticas, etc), así como emociones (desde tristeza y ansiedad hasta intentos de suicidio inminentes), usar ese conocimiento construido para hacer **publicidad** dirigida [50]. Similar capacidad tiene Twitter. También terceros pueden usar sus datos para crear inferencias propias, por ejemplo el nivel o postura de alguien sobre el derecho al aborto. De hecho, la información disponible en redes muchas permite inferir otros datos que expresamente no fueron compartidos; un caso conocido es la posibilidad de deducir el número de seguridad social de ciudadanos de EEUU[51].

Pero los efectos son múltiples, no tienen que ver únicamente con las ventas: el hecho de que una plataforma pueda inferir características sensibles sobre una persona y usarla, por ejemplo, para hacer publicidad, tiene más consecuencias que el *marketing*.

Tal es el caso de David, un joven británico que fue echado de su casa en 2012 cuando sus padres se enteraron que era homosexual por la publicidad que aparecía en su Facebook: si bien él no había posteado nada al respecto ni estaba en ningún grupo o comunidad gay online, la plataforma ubicaba publicidad de este tema en su perfil sin notificarlo y sin su consentimiento.

Como afirman estudios al respecto, "Si bien uno podría argumentar que los algoritmos de Facebook sólo estaban haciendo su tarea -rastrear a David para dirigirle publicidad precisa-, lo que se cuestiona acá es la falta de sensibilidad en el envío de esa publicidad a usuarios que no indicaron su sexualidad".

Al mismo tiempo, tampoco sería correcto argumentar que David fue ingenuo o poco cuidadoso con sus configuraciones de privacidad, en tanto él había elegido no revelar sus preferencias sexuales en su perfil e igual los algoritmos lo identificaron como homosexual. De hecho, relata que intentó convencer a Facebook de lo contrario para mantenerse "en el closet" con respecto a sus padres. "Esto revela los modos arbitrarios en que los algoritmos de Facebook clasifican los datos, porque deberían haber respondido a los intentos de David de cambiar las publicidades que recibía. En este caso, los datos producidos se correspondían con la vida que vivía, a pesar de sus intentos de dirigir el algoritmo hacia otro lado".

Holder, A. E., Op. Cit., p. 29 y Kenneth C. Werbin, Mark Lipton & Matthew J. Bowman, The Contextual Integrity of the Closet: Privacy, Data Mining and Outing Facebook's Algorithmic Logics, 2(1) QUEER STUD. IN MEDIA & POPULAR CULTURE (2017), p. 37.

[50] Para más información y ejemplos sobre publicidad dirigida y marketing ver Holder, A. E., Op. Cit., p. 9.
[51] Privacy International, Op. Cit., p.8.

En función de este caso, imaginémonos qué ocurre con procesos similares cuyo resultado es este tipo de publicidad, por ejemplo, en países donde la homosexualidad está penalizada[56].

Otro caso de público conocimiento es el de Cambridge Analytica[52]: En el contexto de las elecciones presidenciales de Estados Unidos de 2016, la compañía recolectó datos de Facebook de 87 millones de personas, incluyendo entre otras cuestiones qué páginas visitaban y les gustaban, para así armar un perfil de cada uno e inferir su alineación política. En función de esto, se mostraba determinado tipo de publicidad a quienes, por ejemplo, no habían decidido a quién votar o tenían menor convicción, con el fin de incidir en su comportamiento electoral.

Por otro lado, estos datos no sólo se usan para dirigir publicidad: son parte de los fundamentos detrás de gran parte de las decisiones que toman sobre nosotros empresas, organizaciones e incluso el Estado. Los datos útiles para inferencias también son recolectados por los propios bancos, comercios físicos, fabricantes y otros tipos de organizaciones, así como también por agencias estatales[53]. De hecho, existen compañías ("*data brokers*") dedicadas específicamente a recolectar información personal de consumidores y revenderla o extenderla a empresas u organizaciones con las que seguramente el usuario no supo ni tuvo la intención de compartir. Además, los *data brokers* suelen hacer también sus propias inferencias y elaborar perfiles en base a eso (veremos de qué se trata esta práctica a continuación), que también venden a otros.

Por ejemplo, las **aseguradoras** usan datos de las redes sociales para inferir cuestiones de salud de los usuarios y en base a eso determinar la prima de su seguro: pueden inferir desde estados depresión (a través del uso de Twitter y Facebook), brotes de gripe y otras infecciones (a través de Google), predecir Parkinson o Alzheimer (a través de las búsquedas en motores de Microsoft), otras cuestiones de salud a través de patrones en el habla al comunicarse con Alexa (Amazon), etc. **Así, "es esperable que 'preventivamente' una compañía de seguros de salud sepa que una persona sufre de una condición de salud crónica, en función del historial de compras del usuario que muestra que adquiere determinado medicamento todos los meses, y su historial de búsqueda que muestra que investigó sobre remedios caseros para ciertos síntomas"[54].**

El Estado, por su parte, también ha basado históricamente muchas de sus decisiones en el análisis de múltiples sets de información para inferir o predecir datos sobre una persona: desde áreas de salud para determinar y predecir enfermedades, hasta agencias de seguridad cruzando datos de distintas fuentes para establecer el riesgo que presenta un pasajero y en función de eso a qué controles lo somete[55].

[56] Para más información sobre este supuesto y casos concretos, ver Holder, A. E., Op. Cit., p. 30.

[52] Para información general sobre este caso ver, por ejemplo, "5 claves para entender el escándalo de Cambridge Analytica que hizo que Facebook perdiera US$37.000 millones en un día", BBC, 21/03/2018, disponible en https://www.bbc.com/mundo/noticias-43472797.

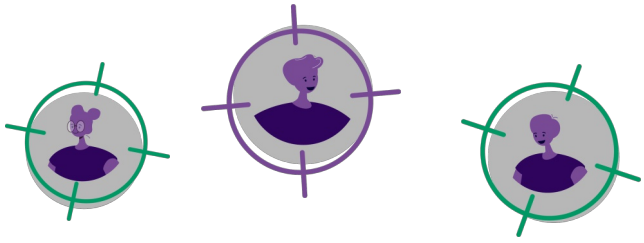
[53] Holder, A. E., Op. Cit., p. 9.

[54] Ibídem, p. 9. Traducción propia.

[55] Ibídem, p. 2 y Dixon, P. y Gellman, R., World Privacy F., The Scoring Of America: How Secret Consumer Scores Threaten Your Privacy And Your Future 16 (2014), p. 62 y 76, citados en el mismo documento.

Un **caso argentino** que tomó público conocimiento fue el anuncio de la implementación[57] en la provincia de Salta de una aplicación desarrollada por Microsoft que permitiría **predecir embarazos adolescentes**: "Con la tecnología vos podés prever cinco o seis años antes, con nombre, apellido y domicilio, cuál es la niña, futura adolescente, que está en un 86 por ciento predestinada a tener un embarazo adolescente", dijo el entonces Gobernador [58]. Frente a este anuncio, el Laboratorio de Inteligencia Artificial Aplicada UBA-CONICET[59] estudió la metodología del sistema de inteligencia artificial utilizado y encontró "serios errores técnicos y conceptuales, que ponen en duda los resultados reportados (...) y que comprometen el empleo de la herramienta generada, en una cuestión tan sensible como el embarazo adolescente." [60]. Destacó tres problemas principales: 1) los resultados estaban sobredimensionados por la metodología usada para construirlos; 2) como los embarazos adolescentes son un tema sensible y muchas veces confidencial, los datos relativos a ellos suelen estar sesgados o incompletos (por ejemplo, es posible que representen más a un sector de la sociedad que a otro), por lo que el sistema ya se alimentaba sesgadamente; y 3) los datos de base fueron inadecuados con respecto al objetivo del sistema (predecir futuros embarazos, no acertar embarazos en curso).

De lo expuesto podemos advertir tres núcleos problemáticos: por un lado, que tanto los problemas metodológicos como los datos poco confiables "plantean el riesgo de llevar a tomar medidas incorrectas a los responsables de políticas públicas", tal como concluyó el LIAA. Por otro lado, el uso discriminatorio y afectación directa a derechos que implica este tipo de usos de las inferencias (imaginemos todos los impactos que puede tener en la vida de una niña que sea objeto de esta predicción; desde su vida familiar, educativa, laboral hasta el derecho a elegir sobre su propia vida, su integridad física y psicológica, etc). Finalmente, vemos cómo la información construida en base a datos personales (sensibles, en este caso) también resulta ser personal y necesita contar con la misma protección. De ninguna manera podría, en un ejemplo así, argumentarse que la posibilidad de embarazo inferida de los datos recolectados no representa información personal de la adolescente en cuestión. **Así, existen problemas tanto en la construcción como en el uso y protección de la información.**



[57] Según la información pública disponible al momento de producción de este informe, la aplicación aún no está implementada.
[58]https://icc.fcen.uba.ar/investigadores-encuentran-graves-errores-en-el-sistema-de-prediccion-de-embarazos-adolescentes/
[59] El LIAA pertenece al Departamento de Computación, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires y al Instituto de Investigación en Ciencias de la Computación, CONICET- UBA.
[60] LIAA, "Sobre la predicción automática de embarazos adolescentes", julio de 2018. Disponible en https://liaa.dc.uba.ar/es/sobre-la-prediccion-automatica-de-embarazos-adolescentes/

En una línea similar, es interesante mencionar que las inferencias son muchas veces violatorias de derechos -estigmatizantes, discriminatorias- en sí mismas: por ejemplo Google Images aprendió a asociar determinados nombres propios con determinadas etnias. El proceso de inferencia no es ilegal en sí mismo, pero sí se vuelve preocupantemente discriminatorio en el momento en que se incluyen estas características en perfiles sobre los usuarios. Recordemos que además éstos pueden ser compartidos con data brokers que a su vez los venderán a otras compañías que tomen acciones en función de esto, como dirigir o, peor aún, no dirigir, cierta publicidad a determinados usuarios. El proceso que busca dirigir publicidad de forma excluyente (en función de distintas categorías, no únicamente la etnia) es frecuente no sólo en *marketing* sino también al ofrecer préstamos crediticios, tarjetas de crédito, búsquedas laborales, etc., y termina perpetuando en muchos casos la desigualdad económica, en tanto afecta desproporcionadamente a comunidades históricamente vulneradas. Al mismo tiempo, los usuarios no tienen cómo saber que estas publicidades les están siendo ocultadas por ciertos rasgos, y por lo tanto no están al tanto de la violación que implica para sus derechos civiles (igualdad, no discriminación)[61]. Así, inferencias que son de por sí estigmatizantes darán lugar a decisiones del mismo tipo, por lo tanto violatorias de derechos. De hecho, es posible afirmar que más que corregir los sesgos, estos modelos tienden a reforzarlos[62].

Una vez más, para tomar dimensión de la gravedad de que este tipo de decisiones se base en inferencias, recordemos que, en la mayoría de los casos, no se trata de información verificable, que es subjetiva, que los titulares no suelen estar al tanto de su existencia, que no hay previstos mecanismos de protección y control suficientes, que tienen un sentido de permanencia y que potencialmente afectan numerosos derechos.

Comienza el capítulo 4

Antecedentes internacionales sobre regulación de datos inferidos

[61] Holder, A. E., Op. Cit., pp. 30-31.
[62] Blanke, J., Op. Cit., p. 84 y sus referencias.

CAPÍTULO 4

Antecedentes internacionales sobre regulación de datos inferidos

Si bien no existe al día de hoy una regulación sobre la protección de los datos inferidos como tal que sirva de modelo o ejemplo a seguir, sí hay discusiones o propuestas interesantes en curso en otros sistemas jurídicos, que pueden servir como indicadores a la hora de diseñar una norma local.

i. California

En primer lugar, la legislación que parece ser más avanzada es la de California, EEUU. En el Consumer Privacy Act (CCPA)[63] del año 2018 se incorporó específicamente a los datos inferidos como parte de su definición de información personal[64].

Por un lado, entiende a los datos personales como "toda aquella información que identifica a, se relaciona con, describe y/o puede razonablemente ser asociada directa o indirectamente con un consumidor o un hogar" y agrega una lista de tipos de información que pertenecen a esta categoría siempre que cumplan con estos requisitos[65]. Por otro lado, tal como mencionamos en el apartado 1, también provee una definición de los datos inferidos como "la derivación de información, datos, presunciones o conclusiones a partir de hechos, evidencia u otra fuente de información o datos"[66]. Pero da un paso más allá y regula, además, cómo deben ser tratados en términos de protección de derechos: determina que deben ser considerados datos personales -es decir, gozar de su misma protección- en la medida en que se trate de inferencias derivadas de cualquier información categorizada como "datos personales" (según la definición anterior) para crear "un perfil sobre un consumidor que refleje sus preferencias, características, tendencias psicológicas, predisposiciones, comportamiento, actitudes, inteligencia, habilidades y aptitudes".[67]

En 2022 el Procurador General de California emitió una opinión[68] sobre la interpretación de este tipo de datos en el marco del CCPA, al ser consultado si el acceso de acceso que prevé esta norma es aplicable a consumidores para conocer datos creados por empresas sobre ellos.

[63] California Consumer Privacy Act (CCPA), Op. Cit.
[64] Kröger, J. L., Op. Cit., p. 1.
[65] California Consumer Privacy Act (CCPA), Op. Cit., art. (v)(1).
[66] Ibidem, art. (r)
[67] Ibídem, art. (v)(1)(k). Traducción propia.
[68] Cal. Op. Att'y. Gen. No. 20-303 (Cal. A.G.), disponible en https://oag.ca.gov/system/files/opinions/pdfs/20-303.pdf

Si bien resolvió que debe evaluarse caso a caso, sí estableció que los datos inferidos deben ser tratados como "información personal" en caso de que se cumplan dos requisitos: 1) la inferencia debe haber sido generada a partir de categorías específicas de "información personal" reconocidas en el CCPA (art. v.1., citado más arriba), y 2) haberse realizado con el propósito de crear un perfil sobre un consumidor que refleje sus preferencias, características, predisposiciones, comportamiento, etc. En otras palabras, se trata de una equiparación absoluta de los datos inferidos a los personales, en tanto amplía el espectro de protección y los derechos de acceso.

Como vimos, el CCPA está destinado a mediar la relación entre consumidores y empresas. Si bien es un gran avance y resulta fundamental que esté regulada, vale aclarar que también existe la producción de datos inferidos y su uso por parte de otro tipo de entidades, cuyas decisiones no necesariamente son hacia consumidores (por ejemplo, organismos estatales). A la hora de legislar sobre esta materia, es importante contemplar a los individuos como consumidores pero también tener en cuenta todas las otras facetas de su vida que pueden verse afectadas por medidas tomadas en base a datos inferidos.

ii. Australia

Es interesante recuperar el caso de este país porque también menciona expresamente las inferencias. En sus Principios de Privacidad[69] (2014), un instrumento específico que complementa la ley general de protección de la privacidad de 1988, Australia incluye expresamente a las inferencias que permiten identificar a personas, sean o no correctas o precisas, al definir qué se entiende como información personal. De hecho, la guía oficial[70] sobre analítica de *Big Data* a la luz de estos principios, lo ejemplifica así: "Una organización infiere información sobre un individuo en base a sus actividades en línea, como gustos y preferencias en función de compras en línea, según el historial de búsqueda y transacciones realizadas. Incluso si la información inferida es incorrecta, sigue siendo información personal".

De esta forma, en la medida en que refieran a individuos identificados o identificables, las inferencias serán consideradas datos personales y por lo tanto le aplicará el mismo régimen legal de protección.

[69] Disponible en https://www.oaic.gov.au/privacy/australian-privacy-principles. Para más información ver Bensal, D., "Scope And Analysis Of Inferred Data: Application And Implications", diciembre 2021, disponible en https://tclf.in/2021/12/27/scope-and-analysis-of-inferred-data-application-and-implications/#The_Australian_Privacy_Principles

[70] Disponible en https://www.oaic.gov.au/privacy/privacy-guidance-for-organisations-and-government-agencies/more-guidance/guide-to-data-analytics-and-the-australian-privacy-principles

iii. China

Si bien China no tiene una normativa específica sobre los datos inferidos ni menciones expresas a su estatus jurídico, sí ofrece una definición muy extensiva de información personal en su Ley de Protección de Datos Personales[71]: "(...) todo tipo de información, registrada por medios electrónicos u otros, relacionada con personas naturales identificadas o identificables, excluyendo la información después de un proceso de anonimización. El manejo de información personal incluye la recolección, el almacenamiento, el uso, el procesamiento, la transmisión, la provisión, la revelación, el borrado, etc. de información personal.".

Vemos que no limita la información personal a aquella provista por el titular pasiva o activamente, sino que incluye también en esta categoría a la que el responsable de datos o terceros puedan construir mediante un proceso de inferencia. En otras palabras, bajo esta definición parece estar claro que los datos inferidos son datos personales. De hecho, muchas de las garantías que establece para la información personal parecen aplicarse directamente a las inferencias. Por ejemplo, requiere que su recolección se limite "al alcance más reducido para realizar el propósito de manejo" y que su procesamiento sea a través de los métodos menos invasivos posibles en los derechos e intereses del individuo (art. 4). Más adelante nos adentraremos en la necesidad de regular las inferencias expresamente pero es interesante destacar las regulaciones que no las excluyen y que incluso prevén garantías frente a los efectos propios de estos procesos.

iv. Europa: Reglamento General de Datos Personales, Grupo de Trabajo del Art. 29 y Tribunal de Justicia de la Unión Europea

A continuación nos referiremos a la situación normativa en la Unión Europea. Es particularmente importante en tanto el Reglamento General de Protección de Datos (en adelante RGPD o GDPR, por sus siglas en inglés), su principal instrumento jurídico de regulación y protección de datos personales, es tomado muchas veces como modelo a nivel global. Sin embargo veremos que tiene importantes lagunas y la regulación de los datos inferidos permite interpretaciones ambiguas.

Partimos de la base de que el concepto de datos personales a **nivel europeo** es amplio, de manera que podría interpretarse de forma que incluya inferencias, predicciones y presunciones que tienen impacto en o hacen referencia a un individuo; si se considerarán como datos personales, las personas tendrían asegurados derechos bajo el marco de protección de datos personales.

[71] Personal Information Protection Law (PILP), Art. 4. Disponible en inglés en https://digichina.stanford.edu/work/translation-personal-information-protection-law-of-the-peoples-republic-of-china-effective-nov-1-2021/

. Sin embargo, los datos inferidos como tales no se definen expresamente en el RGPD. Su estatus jurídico se disputa en la academia y está marcado por inconsistencias y contradicciones entre el Grupo de Trabajo del art. 29[72] y el Tribunal de Justicia de la UE[73]. Como mencionamos antes, en general la protección se limita a las condiciones de recolección y procesamiento de los datos personales, pero no a lo que se puede derivar de ellos. Son pocos y/o débiles los mecanismos de control para conocer cómo se elaboran y usan sus producidos, aún sabiendo que causan riesgos para los individuos[74].

En primer lugar, existen distintos tipos de datos personales tanto dentro del **RGPD** como en las interpretaciones realizadas por el Grupo de Trabajo del art. 29, como veremos más adelante[75]. Mientras que el art. 4 define datos personales como "toda información sobre una persona física identificada o identificable", el 9.1. incorpora la diferencia entre datos normales/no sensibles y aquellos pertenecientes a "categorías especiales" (por ejemplo orientación sexual, orígenes étnicos o raciales, opiniones políticas, creencias religiosas o filosóficas, pertenencia sindical o cuestiones de salud), que implican mayores restricciones en su procesamiento. **En este sentido, si las inferencias se consideran datos personales, esta distinción entre datos sensibles y no sensibles y su respectivo estándar de protección también sería aplicable[76].**

Sin embargo, la situación de los datos inferidos en sí misma no resulta tan clara y queda a constante interpretación de los tribunales. No están regulados expresamente e, incluso, como el concepto de "datos personales" no está definido allí de manera exhaustiva, deja lagunas sobre si se incluyen como tales aquellos inferidos de información personal[77]. Por ejemplo, al referirse a la portabilidad de los datos aclara que los interesados tendrán "derecho a recibir los datos personales que le incumban, que haya facilitado a un responsable del tratamiento, (...)" (art. 20.1), es decir que lo acota a aquellos entregados expresamente y no inferidos por aquel, marcando una diferencia entre datos personales y datos inferidos.

En principio, para que este tipo de inferencias constituya "datos personales" en los términos del RGPD, deberían darse tres condiciones, que responden al test que el Reglamento establece para evaluar qué data puede ser considerada como personal[78]: ser 1) el contenido, propósito o resultado de procesamiento de datos que 2) se relacione directa o indirectamente con 3) una persona física[79] identificada o identificable. De esta forma, datos no-personales pueden devenir personales si se asocian con una persona identificable: por ejemplo, el valor de una casa puede considerarse información personal en la medida en que se usa para evaluar los impuestos que pagan sus propietarios.

[72] El Grupo de Trabajo del artículo 29 ("Art. 29 Working Group", por el art. 29 de la Directiva europea de protección de datos, 95/46/CE) fue un grupo asesor europeo independiente conformado por representantes de las autoridades de protección de datos de cada Estado de la UE, del Supervisor Europeo de Protección de Datos y de la Comisión Europea. Su función era atender cuestiones relacionadas con la protección de la privacidad y los datos personales. Funcionó hasta mayo de 2018, cuando entró en vigencia el Reglamento General de Protección de Datos. Si bien no era vinculante para los Estados, sí fueron y siguen siendo relevantes sus interpretaciones en términos de estándares.
[73] Wachter, S. and Mittelstadt, B., Op. Cit., p. 5.
[74] Ibídem, Op. Cit., p. 21.
[75] Ibídem, Op. Cit., p. 23.
[76] Más adelante nos referiremos a una opinión reciente del TJUE sobre esta cuestión.
[77] Madge, R. Five loopholes in the GDPR. 2018, p. 109.
[78] Corte, L. D. "Scoping Personal Data: Towards a Nuanced Interpretation of the Material Scope of EU Data Protecton Law". In: European Journal of Law and Technology 10.1 (2019).
[79] Kröger, J. L., Op. Cit., p. 2

Sin embargo, es tan confuso el límite establecido por el Reglamento que termina dependiendo en gran medida del contexto y la discrecionalidad de quien lo interpreta, por lo que se evalúa caso a caso[80]. Tal como afirma la doctrina: "Dada la vaguedad de las definiciones del RGPD, el resultado de estas evaluaciones [para determinar si las inferencias constituyen datos personales] -y, por lo tanto, el efecto protectorio de la ley frente a interferencias intrusivas- depende fuertemente en las interpretaciones y la jurisprudencia"[81].

Es más protectoria la interpretación del **Grupo de Trabajo** del art. 29. En primer lugar, sostuvo que "datos personales" son todos aquellos creados a través de cualquier procesamiento de datos cuyo contenido, propósito o resultado se relaciona con una persona directa o indirectamente identificable[82], pero aclaró que los datos pueden ser clasificados como personales incluso cuando no describen a una persona identificable, o aunque no vayan a ser usados para tomar una decisión sobre esa persona, siempre que tengan el potencial de impactar los intereses y derechos de una persona identificable[83]. En consecuencia, "datos inferidos que no incluyan un sujeto identificado, pero que tengan el potencial de afectar su vida, son considerados datos personales en la postura del Grupo de Trabajo"[84].

Por otro lado, y más puntualmente sobre inferencias, en sus documentos hace una diferencia entre datos provistos y observados frente a aquellos derivados e inferidos[85]. Los primeros refieren a cualquier dato que el titular haya provisto al responsable del tratamiento directamente (ej. dirección de correo electrónico) o indirectamente (o pasivamente, se trata de los observados; ej. ubicación, caligrafía, forma de hablar).

En cambio, los derivados (ej. país de residencia a partir del código postal u otros datos a partir de los clicks en internet) e inferidos (ej. puntaje para crédito, resultado de un chequeo de salud) no son provistos por él ni pasiva ni activamente, sino que son creados por el responsable o terceros a partir de datos provistos por el titular y, en algunos casos, otra información de contexto. Si las inferencias se consideran como datos personales, claramente entrarían en este segundo tipo[86].

De hecho, más allá de las definiciones que se pueden construir a partir de sus documentos sobre la regulación de los datos inferidos en sí mismos, es importante destacar que el Grupo también determinó la aplicabilidad de los típicos derechos de privacidad que asisten a los datos personales (acceso, información, etc) a los datos derivados o inferidos. Esto sería imposible en caso de no considerarlos personales.

[80] Corte, L. D. "Scoping Personal Data: Towards a Nuanced Interpretation of the Material Scope of EU Data Protecton Law". In: European Journal of Law and Technology 10.1 (2019).
[81] Kröger, J. L., Op. Cit., p. 2
[82] Article 29 Data Prot. Working Party, Guidelines on the Right to Data Portability, 16/EN, WP242rev.01, at 9–11 (Dec. 13, 2016), https://ec.europa.eu/newsroom/document.cfm ?doc_id=44099. Traducción propia.
[83] Holder, A. E., Op. Cit., p. 13. El autor explica que mientras que el Grupo de Trabajo no dio ejemplos específicos, la lectura del texto permite entender que información inferida, anonimizada y agregada puede ser considerada "datos personales" si puede ser re-identificada.
[84] Ibídem, 13
[85] Art. 29 Data Prot. Working Party, Op. Cit., pp. 9–11.
[86] Holder, A. E., Op. Cit., p. 16

Si bien estos documentos no son vinculantes para los Estados, sí influencian a tomadores de decisiones, entre ellos el **Tribunal de Justicia de la Unión Europea** (TJUE). Este Tribunal históricamente fue más restrictivo que el Grupo de Trabajo antes mencionado, considerando las inferencias como datos personales sólo en la medida en que los sujetos tengan un interés en corregirlos, borrarlos o impedir su procesamiento[87].

Explica la doctrina que "En la jurisprudencia vigente, el TJUE restringió sistemáticamente la aplicabilidad de normas de protección de datos para evaluar la legitimidad de datos personales sometidos a procesamientos o para rectificarlos, bloquearlos o borrarlos. También ha dejado en claro que la ley de protección de datos no pretende garantizar la precisión de las decisiones y los procesos de toma de decisiones que involucran datos personales, o hacer que estos procesos sean completamente transparentes."[88]

Sin embargo, su opinión más reciente al respecto[89] fue bastante protectoria: allí se discutía el procesamiento de datos nominales de allegados al demandante (por ejemplo de su pareja) que permitían deducir información sobre su propia orientación o vida sexual, a raíz de un caso en Lituania.

El Tribunal sostuvo que cualquier información inferida a través de una operación intelectual que incluye comparación o deducción ("*by means of an intellectual operation involving comparison or deduction*") sobre categorías especialmente protegidas constituye datos personales especialmente protegidos: cualquier dato inferido que revele información sobre categorías especiales (orientación sexual en este caso, pero aplicable a las demás) debe considerarse dentro de la categoría de datos especiales sujetos a las limitaciones que impone el RGPD para el procesamiento en su art. 9 (párr. 185).

De hecho agrega que debe hacerse una interpretación amplia de los términos "categorías especiales de datos personales" y "datos sensibles" para "asegurar un alto nivel de protección de las libertades y de los derechos fundamentales de las personas físicas, en particular, de su intimidad, en relación con el tratamiento de los datos personales que las afectan" (párr. 125). De esta manera, si bien se mantiene restrictivo en cuanto a que no asimila por completo el tratamiento de cualquier dato inferido con datos personales, sí amplía su protección en el caso de que pertenezcan a categorías especiales.

Pero la práctica parece estar más alejada: un informe de Privacy International[90] muestra que "muchas empresas (...) parecen estar trabajando bajo la presunción de que datos derivados, inferidos y predecidos no cuentan como personales, aún si están asociados a identificadores únicos o usados para identificar a individuos".

[87] Holder, A. E., Op. Cit., p. 16
[88] Wachter, S. and Mittelstadt, B., Op. Cit., p. 2, traducción propia.
[89] European Court of Justice, Judgment of the Court (Grand Chamber) of 1 August 2022, case C-184/20 "OT v Vyriausioji tarnybinės etikos komisija". Disponible en https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A62020CJ0184.
[90] Wachter, S. y Mittelstadt, B. (Op. Cit) los consideran "'economy class' personal data" por el tratamiento que reciben.

En los hechos, los individuos tienen muy poco control sobre cómo se usan sus datos personales para hacer inferencias sobre ellos. En el RGPD son tratados como datos personales "de segunda clase"[91] en función de los derechos, control y recursos que los individuos tienen como ellos, desde los derechos clásicos (conocer, rectificar, eliminar, objetar o portabilidad) hasta los remedios para cuestionar decisiones tomadas en base a ellos. Y el panorama no parece ser alentador: las nuevas iniciativas regulatorias sobre protección de la privacidad en Europa no clarifican la situación jurídica de las inferencias sino, al revés, limitan los derechos individuales sobre información personal con el objetivo de permitir análisis de grandes volúmenes de datos (*Big Data analytics*) y minería de datos (*data mining*). De esta forma, parece debilitarse aún más la garantía de derechos frente a datos inferidos.

Finalmente, vale mencionar que el Reino Unido tiene su propia regulación luego del Brexit, aunque en su mayor parte se adecúa al RGPD. Lo interesante a destacar es que no restringe la definición de "datos personales" a información fáctica sobre un individuo: incluye opiniones e inferencias en la medida en que se relacionen con un individuo y permitan que sea identificado directa o indirectamente[92].

Vemos de este modo que las regulaciones internacionales reconocen la importancia de las inferencias, aunque lo hacen hasta cierto punto y mantienen lagunas legales[93]. Sin embargo, son bastante claras al incluir las inferencias bajo el paraguas de protección y derechos que asisten a los datos personales, en la medida en que se relacionan con y/o permiten identificar -de forma directa o indirecta- a un individuo. En algunos casos existen interpretaciones más restrictivas, como la del TJUE que considera que datos inferidos basados en categorías especialmente protegidas constituyen datos personales especialmente protegidos, de alguna u otra forma los admiten dentro del conjunto.

Comienza el capítulo 5

Sobre la necesidad de regularlos. Propuesta de integración de la protección legal de los datos inferidos para una reforma de la ley de protección de datos personales

[91] Wachter, S. y Mittelstadt, B. (Op. Cit) los consideran "'economy class' personal data" por el tratamiento que reciben
[92] https://ico.org.uk/for-organisations/guide-to-data-protection/key-dp-themes/guidance-for-the-use-of-personal-data-in-political-campaigning-1/personal-data/
[93] Holder, A. E., Op. Cit., p.4

CAPÍTULO 5

Sobre la necesidad de regularlos. Propuesta de integración de la protección legal de los datos inferidos para una reforma de la ley de protección de datos personales

i. Sobre la necesidad de regularlos expresamente

Del desarrollo que hicimos hasta acá podemos observar que no parece haber hoy una legislación modelo que incluya expresamente a los datos inferidos como tales en todas sus dimensiones y le otorgue suficientes derechos a los titulares, en forma análoga a lo que sucede con los datos personales. Tampoco está suficientemente regulado, con una perspectiva de derechos humanos, el proceso mediante el cual se construyen.

La preocupación por la falta de regulación y protección de las inferencias en sí mismas fue expresada por distintos actores durante los últimos años. **Expertos de la academia sostienen**, por ejemplo, que "en un mundo de *Big Data*, en general lo que debe ser escrutado no es la exactitud de los *datos crudos* sino de las *inferencias* derivadas de ellos"[94]. También que este tipo de análisis de datos "(...) produce inferencias invasivas, impredecibles y contraintuitivas que amenazan estos componentes [refiere a protección de datos, derecho a la identidad, a presentarse a sí mismo, a la reputación y a la autonomía] de la privacidad. En respuesta, los interesados requieren tener **mayor control sobre cuándo, cómo y en qué condiciones están siendo evaluados por sistemas automatizados**"[95].

[94] Omer Tene & Jules Polonetsky, Big Data for All: Privacy and User Control in the Age of Analytics, 11 NW. J. TECH. & INTELL. PROP. 239, 270 (2013), citado en Wachter, S. and Mittelstadt, B., Op. Cit., p. 18. Traducción propia, énfasis propio de la versión original.
[95] Wachter, S. and Mittelstadt, B., Op. Cit., p. 125.

Además, recalcan la necesidad de promover marcos regulatorios adecuados especialmente teniendo en cuenta que las inferencias incluyen predicciones algorítmicas que resultan imposibles de verificar en tanto, naturalmente, hablan del futuro[96]: "¿Cuán probable es que determinada persona cometa un delito en el futuro? ¿Que sea productiva y honesta en su trabajo? ¿Que salde sus deudas a tiempo? Estos son algunos de los asuntos en que se hacen predicciones sobre personas. Las predicciones algorítmicas prevalecen en finanzas, educación, empleo y aseguradoras- y continúan extendiéndose a otras áreas críticas de la vida personal"[97]. Para estos supuestos no alcanza de ninguna forma la protección actual de la privacidad, en tanto los derechos que habilita tienen que ver con cuestiones del pasado o del presente (acceder, corregir, etc), pero no contemplan información sobre lo que aún no sucedió. Sin embargo, esa información inferida se utiliza como base para tomar decisiones del presente, por lo que es urgente ofrecerle salvaguardas.

En el mismo sentido, **organismos especializados** como el Grupo de Trabajo del art. 29 y el Supervisor Europeo de Protección de Datos[98] afirmaron que "**usualmente, lo sensible no es la información recolectada en sí misma** sino las inferencias que se derivan de ella y la forma en que este proceso se realiza. Eso puede ser motivo de preocupación."[99]. También organismos de la sociedad civil vienen insistiendo a nivel internacional en la urgencia de contemplarlos normativamente[100].

En otras palabras, necesitamos asegurar garantías sobre este tipo de datos.

La época en la que vivimos requiere proteger toda aquella información que se produzca en base a datos nuestros, teniendo en cuenta que estos procesos son cada vez más masivos y automatizados, y son la base de los modelos de negocios de la actualidad. Como sostuvimos a lo largo de este documento, **es clara la necesidad de entender a los datos inferidos como personales** dado que, en tanto permiten identificar a un individuo se trata de información personal: la naturaleza jurídica es la misma, por lo que les cabrá el mismo andamiaje jurídico de protección. Es que, tanto por la forma en que se construyen como por sus posibles afectaciones, es **fundamental contemplar sobre ellos derechos y garantías**. Puesto que los datos inferidos son personales, cualquier suposición o agrupamiento que una empresa, el Estado o la organización que sea haga sobre una persona debería ser accesible y caberle el mismo régimen de protección de aquellos que recolectan directamente.

[96] Ver sección 1, sobre la imposibilidad de verificar las inferencias.
[97] Matsumi, H. and Solove, D. J., "Prediction Society: Algorithms and the Problems of Forecasting the Future", mayo 2023 (borrador) , p. 5, disponible en https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4453869. Traducción propia.
[98] Organismo de la Unión Europea dedicado a "garantizar que, a la hora de tratar datos personales, las instituciones y organismos de la UE respeten el derecho a la intimidad de los ciudadanos". Para más información ver https://european-union.europa.eu/institutions-law-budget/institutions-and-bodies/search-all-eu-institutions-and-bodies/edps_es
[99] Article 29 Data Prot. Working Party, Opinion 03/2013 on Purpose Limitation, at 47, 00569/13/EN, WP203 (Apr. 2, 2013), https://ec.europa.eu/justice/article-29/documentation/opinion- recommendation/files/2013/wp203_en.pdf, citado en Wachter, S. and Mittelstadt, B., Op. Cit., p. 18. Traducción propia.
[100] Ver Wachter, S. and Mittelstadt, B., Op. Cit., p. 19, especialmente nota al pie nº 59

ii. El actual proyecto de reforma de ley en Argentina. Lineamientos mínimos y propuestas para una regulación

A la fecha de elaboración de este informe está discutiéndose un proyecto de actualización de la Ley n° 25.326 de Protección de Datos Personales elaborado por la Agencia de Acceso a la Información Pública, donde se inserta la Dirección Nacional de Protección de Datos Personales. Previamente fue sometido a consulta abierta como anteproyecto, proceso en el que participaron organizaciones de la sociedad civil y expertos en la materia[101].

La ley es del año 2000 y, en su versión actual, quedó anacrónica e insuficiente para gran parte de las problemáticas que afectan o pueden afectar hoy en día a los datos personales. Quedó rezagada frente al avance de la industria tecnológica y los nuevos negocios. Vimos cómo fue insuficiente para responder a casos recientes de público conocimiento[102] y vemos también cómo avanzan legislaciones de otros países, cuestiones que vuelven urgente su actualización. Pero, además, ésta es necesaria por las tecnologías que nos rodean inevitablemente (a quienes no necesariamente brindamos consentimiento expreso o libre para el tratamiento de nuestros datos, como mencionamos antes) y su gran capacidad de procesamiento de datos surgidos de nuestras interacciones para realizar nuevas inferencias que, en definitiva, afectarán futuras interacciones que realicemos[103].

También deviene indispensable prever una actualización de la norma en tanto Argentina es parte del Convenio 108+, ratificado recientemente por ley[104]. Si bien se trata de un Protocolo modificatorio del convenio original que no está aún vigente, pronto lo estará y de allí surgen con total claridad derechos exigibles frente a datos inferidos (por ejemplo, su art. 9 establece el derecho a recibir una explicación suficiente sobre procesamientos de datos suyos). De acuerdo al derecho internacional de los derechos humanos y cómo éste se incorpora en nuestro ordenamiento jurídico interno[105], nos rigen plenamente las disposiciones de tratados internacionales -como será el caso del Convenio cuando entre en vigencia- y las normas nacionales deben ser armonizarse en ese sentido.

[101] Proyecto disponible en https://www.argentina.gob.ar/sites/default/files/2018/10/proyecto_leypdp2023.pdf. Para más información sobre el proceso que derivó en el proyecto ver https://www.argentina.gob.ar/aaip/datospersonales/proyecto-ley-datos-personales. A la fecha de cierre de este informe no se ha elevado al Congreso Nacional.

[102] Algunos casos resonantes al momento de producción de este informe son, por ejemplo, la venta de datos del Registro Nacional de las Personas en la deep web, el uso de cámaras con tecnología de reconocimiento facial en la Ciudad Autónoma de Buenos Aires (discutido judicialmente), la exclusión de activistas de la cumbre de la OMC en Argentina en 2017, entre otros.

[103] Para más información sobre la necesidad de sancionar una nueva ley y algunas cuestiones mínimas que debería incluir, ver conferencia "Protección de datos personales en Argentina: Fortalezas y debilidades de la reforma" (27/03/2023, Facultad de Derecho de la Universidad de Buenos Aires). Disponible en https://www.youtube.com/watch?v=vgGv8MrIwp0

[104] Ley 27.699 de noviembre de 2022.

[105] Ver, entre otros, arts. 31 y 75.22 de la Constitución Nacional, y Fallos 315:1492 - CSJN, Ekmekdjian, Miguel Angel c/ Sofovich, Gerardo y otros. s/ Recurso de hecho, sentencia del 7 de julio de 1992.

Entonces, volviendo a los problemas que reviste la **ley actual** con respecto a lo que nos interesa en este documento, vemos por ejemplo que **no contempla cómo se construye la información que hay sobre nosotros** en poder de distintas agencias y empresas, **cómo podemos acceder a ella** ni **qué derechos nos amparan sobre su contenido**. Partiendo de que hoy en día el mayor problema con respecto a los datos no es su recolección sino su inferencia, esta es una de las cuestiones centrales que debería estar incorporada.

Sin embargo, el proyecto parece ir en sentido contrario en algunos aspectos. A modo de ejemplo, explícitamente excluye del derecho a la portabilidad a los datos inferidos, derivados, creados, generados u obtenidos a partir de un análisis hecho por el responsable del tratamiento (art. 32 in fine). Como mencionamos anteriormente, las inferencias realizadas sobre personas condicionan tanto decisiones futuras sobre ellas como sus propias actuaciones. Dejar al titular de este tipo de datos sin un mecanismo que permita obtener una copia de los datos objeto del tratamiento implica violaciones a otros derechos.

Por otro lado, el texto enumera ciertas excepciones al ejercicio de los derechos allí establecidos (art. 35) e incluye, entre otras, la salvaguarda de "la seguridad pública, la defensa de la Nación, la protección de la salud pública, y las libertades de terceros y en resguardo del interés público". Si bien aclara que la restricción deberá hacerse por ley (o resolución judicial o administrativa cuando se trate de afectaciones a investigaciones en curso) y siempre que sea necesario, adecuado y proporcional, lo cierto es que no define taxativamente de qué se tratan estos supuestos. Sabemos que, en el caso del Estado, es frecuente que se ampare en este tipo de excepciones para restringir derechos (el acceso, principalmente) sin justificación suficiente y que en la práctica es muy difícil discutirlo[106]. Sabemos también que los datos en general, y los inferidos en particular, se usan para tomar decisiones de gestión pública (tal como vimos anteriormente, cuestiones de salud -la aplicación para predecir embarazos adolescentes, por ejemplo-, seguridad, etc). Habilitar entonces un régimen de excepciones con supuestos tan vagos cuando partimos de estas premisas, implica un alto nivel de desprotección para el titular y de desnaturalización del derecho de acceso y sus derivados.

Por último, nos interesa destacar que el proyecto limita el derecho a conocer las decisiones automatizadas cuando su explicación "afecte derechos intelectuales del Responsable del tratamiento" (art. 27.i), cuestión a la que nos referimos en sección 1, al considerar por qué las inferencias deben gozar de la misma protección que los datos personales. Tal como sostuvimos allí, los derechos personalísimos priman sobre un mero derecho patrimonial como la propiedad intelectual.

[106] Sobre este punto ver Litvachky, P., Trovato, M., et al, "El secreto. La seguridad nacional como coartada para un Estado sin controles", en Centro de Estudios Legales y Sociales (CELS), Derechos Humanos en Argentina. Informe 2019., Ed. Siglo XXI, Buenos Aires, 2019. Disponible en https://www.cels.org.ar/web/capitulos/el-secreto-la-seguridad-nacional-como-coartada-para-un-estado-sin-controles/

Al mismo tiempo, conocer la lógica detrás de estas decisiones no implica violar el secreto comercial o el derecho de propiedad intelectual relativos al algorimo que se usa, sino brindar una explicación clara y razonable el proceso que pueda permitir al titular cómo se arribó al dato inferido en cuestión. Si bien el proyecto prevé el "derecho a no ser objeto de una decisión basada única o parcialmente en el tratamiento automatizado de datos, incluida la elaboración de perfiles e inferencias", lo contempla únicamente en el caso de que le produzca al titular "efectos jurídicos perniciosos, lo afecte significativamente de forma negativa o tengan efectos discriminatorios.". Esto es un problema porque los efectos no necesariamente serán inminentes ni directos y por lo tanto gran parte de los casos de decisiones basadas en inferencias quedarán excluidas.

Una vez más, insistimos en que toda información inferida que versa sobre una persona identificada o identificable constituye un dato personal, más allá de sus efectos directos.

Entonces, si los datos inferidos implican información personal, de ninguna forma pueden quedar por fuera de una nueva legislación sobre datos personales. **De mínima, es necesario que una ley en esta materia establezca, con perspectiva de derechos humanos, tres cuestiones**[107]: en primer lugar, explicitar que **los datos inferidos son datos personales** y que, por lo tanto, quedan comprendidos en las previsiones de la ley y **sujetos a todos los derechos y garantías** que ésta otorga independientemente de que supongan un daño o perjuicio al titular. Esto incluye conocer, acceder, modificar y eliminar los datos, con explicación suficiente de su significado y de los procesos mediante los cuales fueron construidos.

En segundo lugar, **establecer un límite a las inferencias** para evitar tratamientos abusivos, tanto con relación a los procesos inferenciales mediante los que se producen como a los posibles usos de la información.

En consecuencia, por último, es fundamental que se asiente **la prohibición de llevar a cabo decisiones automáticas, inferencias o identificaciones probabilísticas producto de procesos que no puedan ser explicados** en los términos descriptos.

Si intentamos referenciar experiencias internacionales con respecto a este último punto, lo cierto es que la academia no es unánime al interpretar si el derecho a recibir una explicación apropiada está consagrado para los datos inferidos. Mientras algunos consideran que se desprende del propio RGDP[108], otros lo ponen en duda o directamente interpretan que no, y proponen entonces reformas legislativas para incorporarlo.

[107] Nos remitimos en este punto al aporte que realizamos en el marco de la consulta pública sobre el anteproyecto en cuestión, en noviembre de 2022. Disponible en https://www.vialibre.org.ar/aporte-ley-datos/.
[108] Ver, entre otros, Goodman, B. and Flaxman, S. "European Union regulations on algorithmic decision-making and a 'right to explanation'", AI Magazine 38, 2017; Selbst, A. and Powles J. "Meaningful Information and the Right to Explanation", Int'l. Data Privacy L. 7, 2017; Kaminski, M. "The Right to Explanation, Explained", Berkeley Tech. L. J. 35 2019.

Entre estos últimos se encuentran Watcher y Mittelstadt[109]. Es interesante traer su propuesta para asegurar la correcta explicación o rendición de cuentas: además de los derechos clásicos de los datos personales (los denominados ARCO)[110], invitan a aplicar en el caso de las inferencias un mecanismo extra de transparencia y control. Se refieren a **establecer no sólo un control "ex-post"**, es decir luego de producida la inferencia (transparencia sobre cómo se construyó, qué información se utilizó, si se utilizó para perfilamientos o *scoring*, qué tipo de decisiones motivó, etc), **sino también una instancia "ex-ante", donde el responsable del tratamiento deba justificar la razonabilidad de realizar esa inferencia.** Sostienen que, debería expedirse sobre por qué esa información es una base "aceptable" sobre la que hacer inferencias, por qué esas inferencias serían aceptables y normativas para después tomar decisiones y si los métodos de procesamiento para inferir son exactos y confiables. Así, podría pensarse en un requisito de estándares mínimos para la toma de decisiones, que no implicaría hacer público todo el proceso sino conocer los pasos de ese proceso de decisión para entender si los derechos fueron respetados.

En una línea parecida, Kröger propone la publicación por parte de los responsables del tratamiento de información comprehensiva sobre todo tipo de inferencia que estén realizando o intentando realizar y que, si implican un alto riesgo, se contemple la prohibición legal de realizar determinados tipos o de usarlas para determinados propósitos[111]. Agrega que, como la complejidad de los sistemas modernos de procesamiento hace aún más difícil la discusión por la transparencia y protección de este tipo de datos, hacen falta regulaciones que "pongan el foco en este problema, por ejemplo estableciendo límites legales estrictos a usos de información éticamente indefendibles y mejorando la aplicación de las normas y la rendición de cuentas a través de políticas de transparencia y mayor protección a los denunciantes (*whistleblowers*)"[112].

Por último, es crucial mencionar la necesidad de una autoridad de aplicación fuerte, robusta, con autonomía funcional y presupuesto suficiente para desempeñarse de manera adecuada en el contralor y aplicación de la ley que se sancione, de acuerdo con los estándares internacionales en la materia y los principios generales de derechos humanos.

[109] Wachter, S. and Mittelstadt, B., Op. Cit., pp. 2 y 11. A lo largo de esta publicación, los autores proponen la necesidad de un derecho "a una inferencia razonable" y establecen estos mecanismos de control ex-ante y ex-post para evaluar y garantizar su cumplimiento, como parte de un conjunto más amplio de requisitos y salvaguardas. También sostienen la necesidad de legislar sobre la explicabilidad, entre otros, Edwards, L. and Veale, M."'Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for", Duke L. & Tech. Rev. 16, 2017; y Malgieri, G. and Comandé, G. "Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation", Int'l. Data Privacy L. 7, 2017.
[110] Lo mismo sostiene Kröger en Kröger, J. L., Op. Cit., p. 3.
[111] Ibídem, pp. 3-4.
[112] Ibídem, p. 4. Traducción propia.

Esto es especialmente importante por dos motivos: en primer lugar, porque teniendo en cuenta que al hablar de protección de datos inferidos estamos abordando los principales modelos de control estatal y de negocios. De hecho, "existe una ironía detrás de escena que necesita ser examinada abiertamente: en qué medida información altamente personal que la gente elige no revelar incluso a su familia, amigos o colegas termina siendo compartida con completos extraños. Estos extraños participan en el ecosistema -generalmente opaco- de tecnologías y *data brokers* donde las empresas tienen un motivo económico para compartir data en una escala sin precedentes"[113].

Y, en segundo lugar, porque el derecho a la privacidad, bien jurídico a proteger, tiene también una dimensión colectiva[114] (tal como mencionamos cuando nos referimos al consentimiento, ver sección 1 y nota al pie 9) y por lo tanto deben pensarse remedios y mecanismos de salvaguarda que la contemplen. Es que estamos rodeados de tecnologías (plataformas digitales, pero también electrodomésticos inteligentes, por ejemplo) cuyo modelo de negocio se basa en los datos masivos de usuarios y, principalmente, de grupos o colectivos que permitan derivar patrones para luego inferir quiénes pertenecen a ellos. Si bien la privacidad se entendió tradicionalmente como la protección de la autonomía y dignidad personal, hoy se proyecta también sobre grupos de personas. Es que "(...) al ser objeto de tratamiento estadístico los datos agregados de diversos individuos que comparten unas características determinadas, ya sean estas demográficas, de intereses o relativas al comportamiento, de nada sirve que uno o pocos individuos decidan oponerse al tratamiento de su información: mientras siga siendo posible adscribirlos al grupo estudiado, las consecuencias sobre su privacidad seguirán siendo las mismas."[115]. Por estos dos motivos, hace falta robustecer la defensa de los derechos de forma tal que no dependan de la acumulación de reclamos individuales frente a la posición dominante de quienes producen estos datos[116].

Comienza el capítulo 6

Conclusión →

[113] Cohen, K., "Location, health, and other sensitive information: FTC committed to fully enforcing the law against illegal use and sharing of highly sensitive data", Federal Trade Commission, 2022. Disponible en https://www.ftc.gov/business-guidance/blog/2022/07/location-health-and-other-sensitive-information-ftc-committed-fully-enforcing-law-against-illegal. Traducción propia.
[114] Wachter, S. and Mittelstadt, B., Op. Cit., p. 20.
[115] Barbudo, C. F., Op. Cit., p. 74.
[116] Algunos expertos proponen incluso pensar en mecanismos colectivos de reclamo. Ver conferencia "Fortalezas y debilidades de la reforma" (27/03/2023, Facultad de Derecho de la Universidad de Buenos Aires). Disponible en https://www.youtube.com/watch?v=vgGv8MrIwp0

CAPÍTULO 6

Conclusión

Vimos que las inferencias, sin ninguna regulación que las limite, plantean importantes desafíos para la protección de los derechos humanos, especialmente en relación a la privacidad y la no discriminación.

A lo largo de este documento describimos y ejemplificamos distintos aspectos problemáticos de las inferencias y su impacto en nuestra vida cotidiana. Desde vulneraciones a la intimidad hasta al derecho a la salud, trabajo digno, libertad de expresión, entre otros. **La lista continúa porque las derivaciones de la falta de privacidad, autodeterminación e igualdad son muchas.**

Los aspectos preocupantes surgen tanto de sus condiciones de producción -mencionamos los problemas metodológicos, la falta de verificabilidad y de consentimiento, la presunción errónea de causalidad, entre otros- como del uso que se le da. La elaboración de perfiles o scorings, que de por sí es problemática, se vuelve aún más gravosa cuando se basa en este tipo de datos. Lo mismo con las decisiones automatizadas. No nos interesó en este informe discutir la automatización en sí misma, sino poner de relieve la urgencia de regularla con miras a proteger a los titulares de la información que de allí surja. Recordemos que suelen tener sesgos y desviaciones que se potencian con el tiempo y su resultado no es más que una profundización de situaciones discriminatorias e invasivas.

En la sociedad actual no parece verosímil pretender que los datos inferenciales dejen de existir como tales, es decir que dejen de elaborarse o de utilizarse para distintas acciones. Es fundamental, en cambio, promover regulaciones con perspectiva de derechos humanos que prevengan afectaciones a derechos, que aseguren garantías para todos los individuos junto con mecanismos efectivos para hacerlas valer y que limiten la arbitrariedad absoluta con la que hoy se manejan.

Esto implica, de mínima, reconocer que se trata de datos personales e incluirlos explícitamente bajo su paraguas legal, en tanto contienen información relativa a individuos identificados o identificables.

Al mismo tiempo, es necesario generar conciencia sobre su producción y uso entre la ciudadanía. Consideramos estas cuestiones indispensable especialmente viendo hacia dónde tienden las discusiones internacionales y las lagunas interpretativas que aún hoy se presentan en aquellas que suelen ser consideradas “de avanzada”.

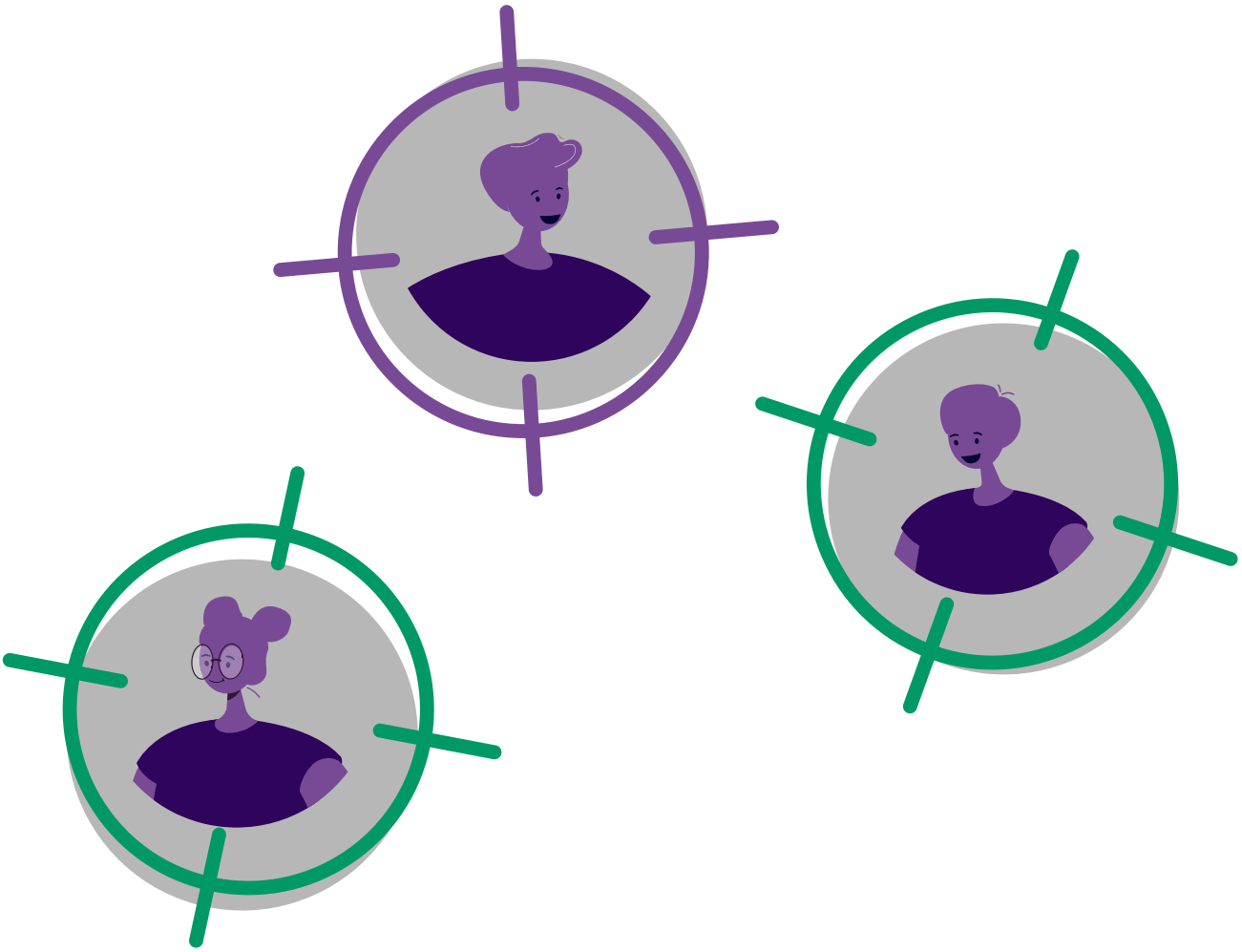
Es que, de lo contrario, el modelo actual seguirá perpetuándose e incluso, como vimos, tendiendo hacia la desprotección con el objeto de profundizar un modelo de negocios. El impacto es directo, muchas veces sin que siquiera lo sepamos y a través de asociaciones poco intuitivas que nos costaría imaginar. Mencionamos ejemplos de decisiones tomadas en base a inferencias realmente gravosas para las personas involucradas.

El universo de las inferencias es complejo y tiene tantas aristas como aspectos problemáticos para abordar. Sin embargo, como mencionamos al principio, este informe no pretende ser exhaustivo con respecto a ellas sino ilustrar la urgencia de contemplar legalmente su existencia. Pretendemos destacar la importancia de las decisiones que fundamentan, así como algunos de los problemas más frecuentes durante su elaboración, que llevan a resultados muchas veces erróneos. Es que, como afirman los expertos: "Las inferencias en base a datos personales devinieron más peligrosas para la privacidad individual que la recolección y almacenamiento masivo de datos en sí mismos"[117].

Es urgente enfrentar el escenario y construir -o reivindicar- derechos y garantías sobre este tipo de datos. Este informe pretende ser un aporte para esta discusión.

[117] Blanke, J., Op. Cit., p. 81, traducción propia, y Grupo de Trabajo del Art. 29, ver nota al pie 97.

Lo más importante no es cómo se construyen los datos inferidos sino qué decisiones fundamentan.



www.vialibre.org.ar
Twitter: @FViaLibre
Instagram: @fvialibre
Mastodon: @fvialibre@rebel.ar