BIG DATA: RIESGOS Y SOLUCIONES DESDE EL DERECHO Y DESDE LOS PRINCIPIOS

FRANCISCO JAVIER MARTÍN JIMÉNEZ

SUMARIO

Fecha recepción: 13/09/2022
Fecha aceptación: 03/10/2023

BIG DATA: RIESGOS Y SOLUCIONES DESDE EL DERECHO Y DESDE LOS PRINCIPIOS

FRANCISCO JAVIER MARTÍN JIMÉNEZ[1]

Universidad Pontificia de Salamanca

I. INTRODUCCIÓN

La transformación digital ha creado nuevas expectativas en el ámbito empresarial y, también, en la actividad pública. Esa transformación digital es el resultado de la incorporación de nuevas herramientas tecnológicas. El Big Data (BD) es una más pero con una potencialidad abrumadora. Un ejemplo, a caballo entre las realidades pública y privada lo muestra la economía colaborativa. La posibilidad de compartir el uso de un coche es un ejemplo de economía colaborativa, con evidentes ventajas para las partes: disminución de su infrautilización, más plazas de aparcamiento libres, mejoras en la contaminación acústica o atmosférica, menos accidentes, menos gasto energético, más calles peatonales y un entorno más agradable para el ciudadano. Es cierto que pueden competir con otros sectores ya asentados, para los que estas economías de colaboración suponen una merma en sus ingresos, pero llevar el coche lleno en cualquier recorrido siempre es más eficiente que llevarlo semivacío. Compartir coche es una actividad que se ha venido realizando desde que el coche existe, pero es con las nuevas tecnologías cuando podemos contactar de forma rápida con nuevos interesados que antes resultaba inimaginable.

El Reglamento de Gobernanza de Datos[2] reconoce que, a lo largo de la última década, los datos son elementos centrales de la transformación tecnológica y que la innovación basada en los datos reportará enormes beneficios tanto a los ciudadanos

[1] Profesor asociado de la Universidad Pontificia de Salamanca. Facultad de Informática. Calle de la Compañía, 5, 37004, Salamanca. Email: fjmartinji@upsa.es. ORCID: https://orcid.org/0000-0003-3736-5110

[2] Reglamento (UE) 2022/868 del Parlamento Europeo y del Consejo de 30 de mayo de 2022 relativo a la gobernanza europea de datos y por el que se modifica el Reglamento (UE) 2018/1724. Considerando 2.

de la Unión Europea[3] como a la economía, por ejemplo, mejorando y personalizando la medicina.

Frente al BD se han alzado voces sin darse cuenta de que son muchos sus beneficios[4]. Que existan éstos, no quita que deban determinarse normas que protejan los derechos que puedan verse afectados, muchas veces los derechos y libertades fundamentales. Es aquí donde puede existir la discrepancia: se quiere más autorregulación y más auto control o se prefiere más heterorregulación o control externo.

La eficacia real de las medidas adoptadas no sólo dependerá de la decisión adoptada sino del derecho supranacional en el cual esté inmerso ese territorio, teniendo en cuenta, además, que el tratamiento del sector público o sector privado serán diferentes[5]. El marco del primero se mueve en aras del interés público y está más fuertemente limitado por el principio de legalidad. El ámbito privado permite mayores cuotas de autocontrol.

Pero se atisba que las nuevas tecnologías precisan, por su variabilidad innovadora, normas o principios que se adapten a las nuevas realidades, tarea que no parece fácil. Pero si al Derecho o la Ética les cuesta adaptarse a las nuevas realidades, de la misma forma tenemos que asumir que el humano tiene que acostumbrarse a trabajar con las máquinas y con la forma de uso de los algoritmos utilizados. Esta relación le obligará a conocer su funcionamiento.

II. BD, CONCEPTO Y CONTENIDO

1. Concepto

La expresión BD, «datos masivos», podemos encontrarla en el trabajo de Schönberger y Cukier[6] cuyos autores la relacionan con la astronomía y la genética, que experimentaron por primera vez esa explosión en la década de 2000. En español se utiliza la expresión inglesa «BD». También se habla de «macrodatos».

---

[3] Debemos contextualizarlo en el marco de una «Estrategia Europea de Datos» a que se refiere la Comunicación de 19 de febrero de 2020 sobre una Estrategia Europea de Datos. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A52020DC0066

[4] Tene, Omer; Polonetsky, Jules. Judged by the tin man: Individual rights in the age of big data. *J. on Telecomm. & High Tech. L.*, 2013, vol. 11, p. 351.

[5] Becerra, Jairo, et al. Marco teórico y aproximación jurídica al BD: algoritmos, inteligencia artificial y transformación digital. Becerra, J., Cotino-Hueso, L., León, IP, Sánchez-Acevedo, ME, Torres-Ávila, J., & Velandia-Vega, J. (2018). *Derecho y BD.* Bogotá: Editorial Universidad Católica de Colombia, 2018, p. 47.

[6] Mayer-Schönberger, Viktor y Cukier, Kenneth *BD. La revolución de los datos masivo*s. Edición original: *BD. A Revolution That Will Transform How We Live, Work, and Think,* 2013. Publicado por acuerdo especial con la editorial Houghton Mifflin Harcourt. Turner Publicaciones S.L., 2013, p. 7.

El Reglamento de Gobernanza de Datos[7], en su artículo 2, definiciones, apartado 1 indica que «datos», es toda «representación digital de actos, hechos o información, así como su recopilación, incluso como grabación sonora, visual o audiovisual». Esta definición amplifica el concepto clásico de datos adaptándolo a las nuevas realidades. Sin duda, el dato aporta información, esto es, descubre algo desconocido para su receptor. Su amplitud puede ser enorme, tomado el concepto en su sentido más extenso.

Por su parte, el Parlamento Europeo afirma[8] que: «el concepto de macrodato se refiere a la recopilación, análisis y acumulación constante de grandes cantidades de datos, incluidos datos personales, procedentes de diferentes fuentes y que son objeto de un tratamiento automatizado mediante algoritmos informáticos y avanzadas técnicas». Se utilizan «tanto datos almacenados como datos transmitidos en flujo continuo, con el fin de generar correlaciones, tendencias y patrones (analítica de macrodatos)».

Los datos se clasifican en estructurados, semiestructurados y no estructurados[9], siendo estos últimos los de mayor dificultad de tratamiento. El origen de estos datos es diverso[10], pudiendo distinguir los que proceden de las transacciones o del uso de internet y redes sociales, los que generan las máquinas a través de sensores o medidores, registros de facturación, información biométrica, geolocalización, seguridad y sanidad[11], etc.

De forma sintética, podemos referirnos al BD como la capacidad de manejar grandes cantidades de datos de diferente índole, ponerlos en relación, y así obtener información con valor añadido[12]. De otra forma, BD permite a una organización la captura y análisis de datos de cualquier procedencia que aporten información relevante para la toma de decisiones. Junto al concepto referido ha nacido uno nuevo: *Data mining*, o minería de datos, como una tecnología que parte de la información disponible en sus bases de datos, para «predecir futuras tendencias y comportamientos, permitiendo

---

[7] Reglamento (UE) 2022/868 del Parlamento Europeo y del Consejo de 30 de mayo de 2022 relativo a la gobernanza europea de datos y por el que se modifica el Reglamento (UE) 2018/1724 https://www.boe.es/buscar/doc.php?id=DOUE-L-2022-80835 Entra en vigor el día 24 de septiembre de 2023 (artículo 38)

[8] Resolución del Parlamento Europeo, de 14 de marzo de 2017, sobre las implicaciones de los macrodatos en los derechos fundamentales: privacidad, protección de datos, no discriminación, seguridad y aplicación de la ley (2016/2225(INI))

[9] Paredes-Moreno, Antonio. Big Data: Estado de la cuestión. *International Journal of Information Systems and Software Engineering for Big Companies*, 2015, vol. 2, no 1, p. 49.

[10] Paredes-Moreno, Antonio. Big Data..., Ob. cit. p. 41

[11] Cotino Hueso, Lorenzo. «BD e inteligencia artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales». *Dilemata*, núm. 24, 2017, pp. 131-150 Ob. cit., p. 133

[12] Martín Lineros, Eduard. «El trinomio dato-información-conocimiento» en VV.AA.: Manual sobre utilidades del BD para bienes públicos. Entimema, Madrid, 2017, pp. 38 y 44.

efectuar una toma de decisiones proactiva y conducida por un conocimiento extraído de la información (*knowledge-driven*)»[13].

El Big Data es una tecnología cuyo objeto se centra en resolver la problemática de tratar cantidades ingentes de datos de tal forma que su resultado satisfaga el esfuerzo de su procesamiento y que permite gestionar los datos a gran escala, obtener ventajas o beneficios, sean económicos o no.

3. CONTENIDO DEL BD

No hay duda de que el volumen de datos generados crece sin medida y exponencialmente[14]. La sociedad asiste a la eclosión de la era del BD con la generación interminable de colecciones de datos estructurados y no estructurados en tiempo real o diferido.

Reconoce el propio Parlamento Europeo que el sector de los macrodatos crece a un ritmo del 40 % anual, y que esto supone un avance siete veces más rápido que el sector de las tecnologías de la información[15].

En el ámbito de la Unión Europea se habla ya de la necesidad de «un espacio común europeo de datos», que se describe como un mercado interior de datos en el que estos pudieran utilizarse independientemente de su ubicación física de almacenamiento en la Unión de conformidad con el Derecho aplicable»[16].

Esta tecnología es una realidad gracias a otros avances[17]: un menor coste en las posibilidades de almacenaje, el incremento de la velocidad de los procesadores, la mejor transmisión y administración confiable de los datos.

Se explican los atributos de los datos resumidos en cinco «V»[18]: volumen, variedad, velocidad, veracidad[19]. y valor. El volumen, determinado por una ingente cantidad de datos que precisan adecuado almacenamiento y procesamiento; la variedad de los datos utilizados, estructurados, semiestructurados y no estructurados, de diversa procedencia; la velocidad que debe acercarse al «tiempo real», a la inmediatez y la veracidad o calidad de los datos. Se puede añadir la visualización[20], o facilidad de acceso y lectura de los datos como complemento al de velocidad.

---

[13] Secades, Vidal Alonso. Big Data: la eclosión de los datos. *Cuadernos Salmantinos de Filosofía*, 2015, vol. 42, no 1.p. 319 y 320.

[14] Martínez Martínez, Silvia; Lara Navarra, Pablo. «El BD transforma la interpretación de los medios sociales». *El profesional de la información*, v. 23, n. 6, 2015. noviembre-diciembre, pp. 575-581.

[15] Resolución del Parlamento Europeo, de 14 de marzo de 2017. Norma citada.

[16] Comunicación de 19 de febrero de 2020 sobre una Estrategia Europea de Datos Bruselas, 19.2.2020 COM(2020) 66 final.

[17] Puyol Moreno, Javier (2014). «Una aproximación a BD». *Revista de Derecho UNED*, 14, p. 485

[18] Gartner. *Emerging Market Analysis*: IT. Mexico, 2012 and beyond Gartner.

[19] Puyol Moreno, Javier. «Una aproximación a BD», op. cit., p. 486

[20] Rivero, Alfonso José López. Tratamiento estadístico de big data: un cambio de paradigma tecnológico en la utilización de la información. *Cuadernos salmantinos de filosofía*, 2015, no 42, p. 335.

Y, por último, el valor, como colofón de los anteriores: los resultados deben ser relevantes y valiosos para la organización, que en el ámbito privado se dirige a los aspectos de eficiencia económica, pero en el ámbito público tiene como finalidad la satisfacción del interés público.

La veracidad (nivel de fiabilidad o de calidad de los datos) va a predisponer el valor que pueda ser obtenido. Si los datos tienen inconsistencias, son superfluos o es demasiado grande el número de atributos, el resultado del análisis se verá afectado[21]. Recientemente, se habla del término «Smart Data» para referirnos a aquellos datos que son exactos, procesables y ágiles (*accurate*, actionable y *agile*, en inglés). El objetivo es obtener datos de calidad y que sean fácilmente escalables[22].

En definitiva, los datos masivos o macrodatos deben ser captados, gestionados y tratados para obtener valor. Por lo tanto, el BD «no sólo debe referirse a las herramientas y procedimientos utilizados para manipular y analizar grandes cantidades de datos»[23], sino también resulta necesario un acercamiento al pensamiento computacional y a la investigación[24].

Más allá de los datos y los procesos, el BD plantea nuevas formas del ejercicio del poder, económico, social, político. Detrás de aquellos debemos percibir una intencionalidad, quizás nuevas ideologías o el rediseño de las anteriores, una nueva forma de entender las cosas. Del lado de la economía, las decisiones empresariales ya no pueden ser ajenas al BD. Desde el sector público, no podrá excusar la incorporación de nuevas formas de administrar.

El poder estatal tendrá que reformularse sin poder olvidar la velocidad, volumen, variedad, veracidad y valor de los datos manejados y que van dirigidos a perfilar nuevas estructuras de control en todas las materias imaginables: la justicia, la actividad policial y militar, la prestación social, la normativa regulatoria, la protección y garantía de los derechos, etc.

El problema se suscita cuando las aplicaciones son proporcionadas por las empresas y son utilizadas por el Estado, sin saber exactamente qué hace y cómo[25]. De esta forma no está el control en la aplicación de estas nuevas tecnologías sino en la transparencia de las mismas, aquella que nos permite conocer qué hacemos realmente y cómo lo hacernos. Esta transparencia se convierte en la base central del control.

---

[21] Gil González, Elena. *BD, privacidad y protección de datos*. Madrid: Agencia Estatal Boletín Oficial del Estado, 2016, pp. 23 y 24.

[22] García, Salvador, Ramírez-Gallego, Sergio, Luengo, Julián, Herrer, Francisco «BD: Preprocesamiento y calidad de datos». *Novatica* núm. 237 julio-octubre 2016, p.17.

[23] Boyd, Danah y Crawford, Kate (2011): *Six Provocations for BD*, A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society, Acceso SSRN.

[24] Burkholder, L, ed. (1992) *Philosophy and the Computer*, Boulder, San Francisco, and Oxford.

[25] Becerra, Jairo, et al. Marco teórico y aproximación jurídica al BD: algoritmos, inteligencia artificial y transformación digital. Ob. cit. p. 78

Si antes el sector público creaba instrumentos de control de los ciudadanos, ahora es el sector privado el que ofrece sus servicios para tal fin al sector público[26], incluso para el caso de que las funciones desempeñadas sean de gran trascendencia y vinculación con el interés público como las administrativas, judiciales o policiales. Los propios investigadores y académicos deben colaborar con el sector privado si no quieren quedar fuera del acceso a los conocimientos[27]

Pero, por otra parte, reconozcamos la importancia de la inteligencia artificial (IA) o el BD en el control del sector público. Estas tecnologías pueden poner de manifiesto la desviación de poder o decisiones donde se aplique una discrecionalidad indebida o simplemente errónea. El ahorro en tiempo y costes puede ser notorio. También se reduce la arbitrariedad y el trato de favor o discriminatorio. Se gana en coherencia y homogeneidad[28], en seguridad jurídica y previsibilidad[29].

Tampoco podemos obviar la relación íntima entre el BD y la IA[30]. El primero obtiene información relevante y la IA trata de resolver problemas y tomar decisiones a partir de una enorme cantidad de datos que se tratan aplicando algoritmos y patrones de razonamiento que pretende acercarse al razonamiento de los humanos. En esta dirección, el Parlamento Europeo indica que los macrodatos nutren «dispositivos de IA»[31].

La necesidad de procesar y extraer conocimiento valioso de tal inmensidad de datos supone un gran desafío para la sociedad. El gran objetivo del BD es la obtención de valor, esto es, los beneficios que surgen del tratamiento de los grandes volúmenes de datos que se generan día a día.

Su campo de actuación es amplísimo: el BD puede, también, transformar la información sobre datos no texto (localización, fotografías, imágenes y ficheros de audio, etc.). A esto se le llama por la comunidad científica datificación[32] (o *datafication* en inglés). Incluso, nuestras amistades o nuestros gustos son transformados en datos procesables[33]. El BD impulsará la creación de nuevas empresas dedicadas a capturar

---

[26] Joh, Elizabeth E. The undue influence of surveillance technology companies in policing. *NYUL Rev. Online*, 2017, vol. 92, p. 19.

[27] Pasquale, Frank A. Privacy, autonomy, and internet platforms. *Privacy in the Modern Age: The Search for Solutions*, 2015, p. 165-174.

[28] Trazegnies Granda, Fernando. ¿Seguirán existiendo jueces en el futuro?: el razonamiento judicial y la inteligencia artificial. Ob. Cit. p. 112-131

[29] Zeleznikow, John. Using web-based legal decision support systems to improve access to justice. Information & Communications Technology Law, 2002, vol. 11, no 1, p. 15-33.

[30] Laltour, B.»Tarde's idea of quantification», in The Social After Gabriel Tarde: Debates and Assessments, (ed M. Candea), London: Routledge, 2009, pp. 145-162.

[31] Cotino Hueso, Lorenzo. «BD e inteligencia artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales». Dilemata, núm. 24, 2017, p. 132.

[32] Kenneth Neil Cukier y Viktor Mayer-Schöenberger. «The Rise of BD. How It's Changing the Way We Think About the World». Foreign Affairs Vol. 92, núm. 3, 2013.

[33] Gil González, Elena. *BD, privacidad y protección de datos*. AEPD. Madrid: Agencia Estatal Boletín Oficial del Estado, 2016, p. 18.

y analizar datos sobre productos y servicios, proveedores y clientes en todo tipo de actividad[34].

Aparecen nuevos conceptos asociados a estas tecnologías, como las ciudades inteligentes[35], que proporcionan información actualizadas en tiempo real, por ejemplo, del tráfico o de otros servicios administrativos, y que sirven en gran medida a mejoras indudables en la ciudadanía. Las posibilidades son inmensas y el futuro es prometedor.

Por último, tenemos en cuenta los actores y los que no lo son. Tene y Polonetsky[36] describen tres clases de personas que surgen en el entorno del BD: los que generan los datos (con o sin su consentimiento), los que recogen y administran los datos y aquellos que tienen la capacidad para analizarlos. Falta añadir, con Lerman[37], los que no aportan, que se quedan en la periferia, y sus preferencias o necesidades no son tenidas en cuenta para bien o para mal.

III. RIESGOS DEL BD Y CONFLICTOS DE DERECHOS

1. Riesgos

Sabemos que el BD tiene que ver con un volumen ingente de datos, cualquiera que sea su procedencia, recopilados, almacenados, gestionados e interpretados[38].

Esos datos, en buena parte son datos personales, muchos de ellos se corresponden con categorías especiales de datos[39], en los que el consentimiento necesario debe ser reforzado. Esta circunstancia nos pone sobre aviso ante la posibilidad de producirse lesiones al derecho a la intimidad y a la protección de datos personales. Cotino Hueso[40] apunta a una mayor atención jurídica, desde el respeto a la dignidad y los derechos fundamentales, especialmente los derechos de la personalidad y la vida privada.

Conocemos que es posible la manipulación del estado de ánimo o la inducción a realizar comportamientos con una personalización masiva de noticias más o menos

---

[34] Gil González, Elena. *BD, privacidad y protección de datos*. Ob. cit., p. 32.

[35] Townsend, Anthony M. *Smart cities: Big data, civic hackers, and the quest for a new utopia*. WW Norton & Company, 2013, p. 93.

[36] Tene, Omer; Polonetsky, Jules. Judged by the tin man: Individual rights in the age of big data. *J. on Telecomm. & High Tech. L.*, 2013, vol. 11, p. 351.

[37] Lerman, Jonas. Big data and its exclusions. *Stan. L. Rev. Online*, 2013, vol. 66..

[38] Bello-Orgaz, Gema; Jung, Jason J.; Camacho, David. Social big data: Recent achievements and new challenges. *Information Fusion*, 2016, vol. 28, p. 45-59

[39] Art. 9 del RGPD

[40] Cotino Hueso, Lorenzo. «Riesgos e impactos del BD, la inteligencia artificial y la robótica: enfoques, modelos y principios de la respuesta del derecho». *Revista general de Derecho administrativo*, 2019, núm. 50, p. 6.

positivas[41]. No es lo mismo verse influenciado por los argumentos presentadas desde la pluralidad y la veracidad a hacerse con el uso de técnicas oscuras que inducen al individuo a actuar de forma dirigida. Es muy diferente, influir en la libre toma de decisiones utilizando argumentos, desde una información veraz, que inducir, mediante mecanismos opacos a la adopción de comportamientos o decisiones más o menos inconscientes[42].

Como reconoce el Parlamento Europeo[43], el tratamiento de los datos «podría dar lugar a algoritmos sesgados, correlaciones falsas, errores, una subestimación de las repercusiones éticas, sociales y legales, el riesgo de utilización de los datos con fines discriminatorios o fraudulentos y la marginación del papel de los seres humanos en esos procesos, lo que puede traducirse en procedimientos deficientes de toma de decisiones con repercusiones negativas en las vidas y oportunidades de los ciudadanos, en particular los grupos marginalizados». Y pone de manifiesto algo que no debe olvidarse: la concentración de grandes conjuntos de datos generados por las nuevas tecnologías ofrece información fundamental para las grandes empresas. Este hecho podría crear monopolios y prácticas abusivas que modifiquen las reglas del juego de la competencia en los mercados.

La utilización masiva de datos choca con la protección de estos. Una mayor circulación de los datos provocará mayores problemas de seguridad jurídica. La tecnología avanza tan rápidamente que no parece posible desarrollar fácilmente mecanismos jurídicos que la controlen[44].

Para el Supervisor Europeo de Protección de Datos[45] de datos de la Unión Europea los riesgos se encuentran en la falta de transparencia, que permita saber cómo se procesan los datos y qué utilidad puede tener para el responsable del tratamiento, en el desequilibrio entre las personas que proporcionan los datos y las empresas que los tratan, en tanto el resultado del tratamiento puede empeorar las situación de las primeras frente a las segundas, que tendrán información valiosa para utilizar en su beneficio (por ejemplo, gustos y disposición a pagar). Y en la ausencia de legitimación en tanto ésta debe ir inexorablemente unida a un tratamiento con fines determinados, explícitos e informados.

No olvidemos, además, la existencia de ingentes cantidades de datos no estructurados recopilados por las empresas durante años con sistemas de gobernanza de datos

---

[41] Solove, Daniel. Facebook's psych experiment: Consent, privacy, and manipulation. *Huffington Post*, 2014..

[42] Martín Miralles, Ramón. «BD vs Small low», Congrés IDP 2013 Butlletí Kdades: Butlletí electrònic de tecnologia, auditoria i seguretat de la informació, Núm. 24, 2013.

[43] Resolución de 14 de marzo de 2017.

[44] Ortega Giménez, Alfonso y Gonzalo Doménech, Juan José. «Nuevo marco jurídico en materia de protección de datos de carácter personal en la Unión Europea». *Revista de la Facultad de Derecho*, núm. 44, ene.-jun. 2018, pp. 31 a 35.

[45] Supervisor Europeo de Protección de Datos / Opinion 7/2015 Meeting the challenges of Big Data / 19 de noviembre de 2015

que seguramente tienen una falta de adecuación a la normativa vigente de protección de datos.

Algunos autores manifiestan la necesidad de mantener un especial cuidado con las llamadas tecnologías «*captology*», esto es las «*Computers As Persuasive Technologies*». Se trata de tecnologías informáticas creadas para modificar nuestras intenciones y comportamiento, como, por ejemplo, las psicografías, que llevan a cabo segmentación y análisis de datos de la personalidad[46].

Igualmente se constatan riesgos sobre la protección y garantía de los derechos con la utilización de programas que simulan el comportamiento humano (*bots*) o con la presencia de noticias falsas (*fake news*), que pueden afectar al derecho a la información en las sociedades democráticas.

Un ejemplo de esa preocupación lo pone de manifiesto el Consejo Nacional de Ética de Alemania que señala que el BD representa un reto importante para el sistema legal y, en concreto, para el Derecho Constitucional[47].

Estos riesgos se acrecientan si la información es sensible, por ejemplo, la relacionada con la salud[48]. La información sobre el estado de salud puede afectar al derecho a la intimidad, pero también a otros aspectos de relevancia como el acceso al empleo, al crédito o al aseguramiento[49].

Gil González[50] menciona tres riesgos del BD: el de la obtención de informaciones erróneas, la toma de decisiones automatizadas sin un sesgo humano y el ataque a la privacidad de las personas. Las dos primeras, de hecho, también pueden suponer un ataque a privacidad si, consecuencia de ellas se crea un desvirtuado o erróneo perfil personal.

Ante esta situación, la normativa no se encuentra adaptada al nuevo entorno tecnológico. En este sentido, indica Gil González[51] los siguientes problemas: el principio de «minimización de datos» no se cumple, confía demasiado en el consentimiento informado del individuo, la anonimización ha demostrado tener limitaciones y el aumento del riesgo relacionado con la toma de decisiones de forma automática.

Los problemas añadidos serían los siguientes[52]:

---

[46] Samuel, Alexandra. Psychographics are just as important for marketers as demographics. *Harvard business review*, 2016, vol. 3, no 11.

[47] Ethikrat, Deutscher. «BD and Health–Data Sovereignty as the Shaping of Informational Freedom». *Opinion. Executive Summary & Recommendations*., vol. 10, 2017. P. 14.

[48] Chan, S., «Bioethics in the BD era: health care and beyond», *Rev Bio y Der*., vol. 41, 2017, p. 10.

[49] Martínez, Ricard Martínez. «BD, investigación en salud y protección de datos personales…, ob. cit., p. 236.

[50] Gil González, Elena. *BD, privacidad y protección de datos*. Ob.cigt, p. 32.

[51] Rubinstein, Ira S. «BD: The 070747End Of Privacy Or A New Beginning?». *International Privacy Law,* Vol. 3, n.º 2. 2013; Cate, Fred H. «The failure of Fair Information Practice Principles». Consumer Protection In The Age Of Information Economy. 2006.

[52] Gil González, Elena. «BD, Privacidad y Protección de Datos». Ob. cit., pp. 52-53.

1ª) El principio de «minimización de datos» no se cumple en la práctica, teniendo en cuenta, además, que este principio se enfrenta al mismo concepto de BD que consiste en la recogida masiva de datos.

2ª) Para recopilar y tratar sus datos de carácter personal, se confía en exceso en la utilización del consentimiento informado del individuo cuando, de hecho, carece de una información completa. En la práctica, numerosos autores[53] indican que el consentimiento del tratamiento aparece como el perdedor ante la fuerza arrolladora del BD. Es difícil creer que el consentimiento para el tratamiento de datos masivos sea tal, cuando se consiente sobre un proceso desconocido y con un resultado imprevisible para el interesado.

3ª) La anonimización tiene limitaciones. La forma de reidentificar a los sujetos después de la anonimización es cada vez más sencilla y en ella podemos incluir tanto los datos personales como los datos no personales. Por lo tanto, aquella no es una fórmula fuerte para proteger el derecho al tratamiento de datos, en tanto la tecnología, como el BD, constituye una potente herramienta que facilita esa reidentificación. No obstante, a pesar de que la anonimización absoluta no existe, Elliot[54] indica que, partiendo del concepto de insignificancia (*negligibility*), podemos determinar que si el riesgo de reidentificación es muy pequeño o estadísticamente cercano al cero, podemos considerar que la anonimización utilizada es suficiente.

4ª) El BD puede, además, incrementar el riesgo en relación a la toma de decisiones de forma automática. Si la elaboración de un perfil personal supone una valoración desfavorable puede tener consecuencias muy perjudiciales en la vida cotidiana o, como indica Garriga Domínguez[55], «su discriminación en otros actos de su vida». Es por ello que la elaboración de perfiles puede suponer una lesión en los derechos de los interesados en tanto se recogen y se tratan gran número de datos personales y de fuentes diversas[56].

---

[53] Becerra, Jairo, et al. Marco teórico y aproximación jurídica al BD: algoritmos, inteligencia artificial y transformación digital Ob. cit., p 58. Rubinstein, Ira. Big data: The end of privacy or a new beginning? *International Data Privacy Law (2013 Forthcoming), NYU School of Law, Public Law Research Paper*, 2012, no 12-56. Martínez, Ricard. Ética y privacidad de los datos. *Ponencia presentada en Big Data: de la investigación científica a la gestión empresarial, Fundación Ramón Areces*, 2014. Heeger, Eva. Controlling Your Online Profile: Reality or an Illusion? *A Research into Informed Consent as a Mechanism to Regulate Commercial Profiling (August 17, 2015)*, 2015.

[54] Elliot, Mark. *To be or not to be (anonymous)? Anonymity in the age of big and open data*. [Conferencia] Computers, Privacy & Data Protection on the Move (CPDP). 2015.

[55] Garriga Domínguez, Ana, *Nuevos retos para la protección de datos personales: en la era del BD y de la computación ubicua.* Dykinson, Madrid, 2016, p. 67.

[56] Bravo, Álvaro Sánchez. *Derechos humanos y protección de datos personales en el siglo XXI. Homenaje a Cinta Castillo Jiménez*. Punto Rojo Libros, 2018, p. 18.

El principio de cautela o precaución, indicado por la Comisión Europea en 2000[57] implicaría abandonar una acción si no existe consenso entre los expertos sobre la bondad de esa acción, esto es, si puede causar un daño a las personas o al entorno.

Por último, otro de los retos ineludibles del BD va a consistir en explicar cómo se lleva a cabo el tratamiento de datos[58]. La falta de transparencia en el diseño de los algoritmos supone uno de los grandes problemas de la utilización de BD y la IA: lo que se ha denominado «oscuridad por diseño» (*obscurity by design*)[59]. Sobre la necesidad de una aplicación del tratamiento de datos desde el diseño y por el defecto, la Agencia Española de Protección de Datos (AEPD) ha elaborado unas guías específicas[60] que describen las medidas que deben contener el tratamiento, desde la óptica del órgano público, para adaptarse al Reglamento General de Protección de Datos (RGPD). También son de interés las Directrices 4/2019 relativas al artículo 25 Protección de datos desde el diseño y por defecto[61]

2. CONFLICTO DE DERECHOS

Cada vez más, las organizaciones públicas o privadas de todo tipo (empresas, organismos públicos, ONGs, partidos políticos, la comunidad científica, etc.) están aprovechando la analítica de esta cantidad ingente de datos para mejorar sus

---

[57] Comisión Europea. (2000). Comunicación de la Comisión sobre el recurso al principio de precaución, COM/2000/0001 final https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52000DC0001&from=ES

[58] Martínez, Ricard Martínez. «BD, investigación en salud y protección de datos personales. ¿Un falso debate?». *Revista valenciana d'estudis autonòmics*, núm. 62. 2017. 276-277

[59] Hartzog, Woodrow; Stutzman, Frederic. Obscurity by design. Wash. L. Rev., 2013, vol. 88, p. 385. En https://digitalcommons.law.uw.edu/cgi/viewcontent.cgi?article=4758&context=wlr

[60] Entre otras: Guía de Protección de Datos por Defecto. https://www.aepd.es/es/documento/guia-proteccion-datos-por-defecto.pdf

Guía de Privacidad desde el Diseño. https://www.aepd.es/es/documento/guia-privacidad-desde-diseno.pdf

Orientaciones y garantías en los procedimientos de anonimización de datos personales. https://www.aepd.es/es/documento/guia-orientaciones-procedimientos-anonimizacion.pdf

Código de buenas prácticas en protección de datos para proyectos Big Data. https://www.aepd.es/es/documento/guia-codigo-de-buenas-practicas-proyectos-de-big-data.pdf

10 malentendidos relacionados con la anonimización. https://www.aepd.es/es/documento/10-malentendidos-anonimizacion.pdf

Requisitos para Auditorías de Tratamientos que incluyan IA. https://www.aepd.es/es/documento/requisitos-auditorias-tratamientos-incluyan-ia.pdf

Y la norma técnica: La ***K***-anonimidad como medida de la privacidad. https://www.aepd.es/sites/default/files/2019-09/nota-tecnica-kanonimidad.pdf

[61] En este sentido es reseñable el apartado 3: Aplicación de los principios de protección de datos en el tratamiento de datos personales utilizando la protección de datos desde el diseño y por defecto. https://edpb.europa.eu/system/files/2021-04/edpb_guidelines_201904_dataprotection_by_design_and_by_default_v2.0_es.pdf

resultados, económicos, de beneficio social o de otro tipo. Es reconocido que el potencial del BD es enorme: «asistencia sanitaria, la lucha contra el cambio climático, la reducción del consumo energético, la mejora de la seguridad en el transporte y la posibilidad de establecer ciudades inteligentes, la mejora de las condiciones laborales y a la detección y la lucha contra el fraude»[62].

Vistos los indudables beneficios para la sociedad propiciados por las nuevas tecnologías, nos vamos a encontrar, necesariamente, con un conflicto de derechos e intereses.

Empecemos por un ejemplo esclarecedor: en el ámbito de la medicina y de la salud las aplicaciones de BD tienen una enorme importancia dado que la investigación utiliza de forma habitual el método cuantitativo, al correlacionar miles de procesos asistenciales, sin olvidar que dichos resultados deben ser después contextualizados[63]. Los resultados de estas investigaciones son, objetivamente, beneficiosos para la sociedad. Hay razones de interés público que permiten limitar el derecho a no compartir los datos personales cuando la salud de terceros (y, por ende, la propia) puede depender de ellos.

El conflicto de derechos se resuelve más fácilmente cuando relacionamos la explotación masiva de los datos de una persona, o de miles, con la posibilidad de salvar la vida de otros tantos. La balanza se inclina, en estas situaciones, del lado de la aplicación masiva de datos médicos personales frente al derecho de oposición individual a su tratamiento. En este punto debemos poner en valor la existencia de un interés público esencial, recogido en el artículo 9.2 g) del RGPD[64], en tanto la vida de miles de personas, caso de la recientemente vivida pandemia por COVID-19, va más allá del mero interés público, pues podría estar en juego la propia supervivencia de poblaciones o grupos de población enteros. La Disposición adicional decimoséptima de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales extiende expresamente el ámbito del referido artículo del RGPD a los tratamientos de datos relacionados con la salud y de datos genéticos regulados en distintas leyes con ese contenido, entre ellas, la Ley 14/1986, de 25 de abril, General de Sanidad, La Ley 41/2002, de 14 de noviembre, básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica o la Ley 14/2007, de 3 de julio, de Investigación biomédica.

---

[62] Resolución del Parlamento Europeo, de 14 de marzo de 2017.

[63] De Montalvo Jääskeläinen, Federico. «Una reflexión desde la teoría de los derechos fundamentales sobre el uso secundario de los datos de salud en el marco del BD». *Revista de Derecho Político*, 2019, vol. 1, núm. 106, p. 47.

[64] g) el tratamiento es necesario por razones de un interés público esencial, sobre la base del Derecho de la Unión o de los Estados miembros, que debe ser proporcional al objetivo perseguido, respetar en lo esencial el derecho a la protección de datos y establecer medidas adecuadas y específicas para proteger los intereses y derechos fundamentales del interesado;

Como indica De Montalvo[65], el derecho a la protección de la salud está reconocido como un derecho constitucional, pero también como un derecho social, bajo un modelo asistencial universal y «no se sostiene ni ética ni legalmente que los ciudadanos no acepten que sus datos puedan ser empleados en beneficio de la salud de terceros». A lo que hay que añadir que lo anterior no quita que pueda garantizarse la confidencialidad de los datos con mecanismos como la anonimización, la pseudonimización o, en casos extremos de crisis sanitaria como una pandemia, sin esta última[66].

Esta argumentación que sirve para la salud puede extrapolarse a otras actividades en las que interés público pueda ser relevante[67].

Este es el caso del derecho a la libertad de expresión cuando la utilización de algoritmos permite censurar de forma automatizada los contenidos fruto de aquel derecho, y, por lo tanto, evitar su propagación[68]. Debe preservarse el contenido esencial del primero ante las injerencias y abusos que, amparados en la opacidad del algoritmo utilizado, tengan sus efectos nocivos[69].

De la misma forma, en la otra cara de la moneda, cabe preguntarse si los contenidos generados de forma automática por estas tecnologías están protegidos por el derecho a la libertad de expresión[70] y, en caso afirmativo, qué problemas pueden suscitarse en el caso de que proceda exigir responsabilidades por el uso indebido de ese derecho.

El uso automatizado en el ámbito judicial o administrativo tiene muchos años de existencia, desde los años 70, y el avance desde aquel momento ha sido notorio[71]: de ayudar a la toma de decisiones sencillas a preparar propuestas sobre asuntos complejos que pueda ser objeto de validación y ratificación por el ser humano[72]. De futuro,

---

[65] De Montalvo Jääskeläinen, Federico. «Una reflexión desde la teoría de los derechos fundamentales sobre el uso secundario de los datos de salud en el marco del BD». Ob. cit., pp. 66-67.

[66] El informe del Comité de Bioética de España sobre investigación con datos de salud y muestras biológicas en pandemia Covid 19, de a 28 de abril de 2020, (p. 22) indica que «excepcionalmente se puede permitir el uso secundario de los datos personales de salud con fines de investigación sin seudonimizarlos, cuando no exista otro medio de obtener un conocimiento relevante que permita alcanzar un resultado decisivo en la lucha contra la pandemia provocada por el SARS-CoV-2»

[67] Sancho López, Marina. «Internet, BD y nuevas tecnologías: repercusiones y respuestas del ordenamiento jurídico». *CEFD Cuadernos Electrónicos de Filosofía del Derecho*. 2019. p. 318.

[68] Narayanan, Arvind; Zevenbergen, Bendert. No encore for encore? ethical questions for web-based censorship measurement. *Ethical Questions for Web-Based Censorship Measurement (September 24, 2015)*, 2015.

[69] Krotoszynski JR, Ronald J. Reconciling Privacy and Speech in the Era of Big Data. *William & Mary Law Review*, vol. 56, p. 1279.

[70] Massaro, Toni M.; Norton, Helen. Siri-ously? Free speech rights and artificial intelligence. *Nw. UL Rev.*, 2015, vol. 110, p. 1169. Se muestran de acuerdo.

[71] De Trazegnies Granda, Fernando. ¿Seguirán existiendo jueces en el futuro?: el razonamiento judicial y la inteligencia artificial. *Ius et Veritas*, 2013, no 47, p. 123

[72] Coglianese, Cary; Lehr, David. Regulating by robot: Administrative decision making in the machine-learning era. *Geo. LJ*, 2016, vol. 105, p. 1147. Alarie, Benjamin; Niblett, Anthony; Yoon,

no existirá ámbito administrativo que no sea susceptible de implementar aplicaciones de BD o IA, de una menor o mayor complejidad.

Al final, con Cotino Hueso[73] se trata de cambiar el enfoque. El autor aboga por «trabajar con una dimensión colectiva de los derechos», en tanto entiende que «el daño individual producido por el BD y la IA puede ser imperceptible para el derecho fundamental desde la perspectiva individual», pero se hace visible en tanto afecta a la sociedad en su conjunto. Pero sin olvidar el derecho al respeto a la dignidad de las personas y al libre desarrollo de la personalidad[74].

IV. SOLUCIONES

1. Soluciones desde el Derecho

La aportación más importante, cuando de derechos fundamentales se refiere, viene de la mano de la CE y desde la base sólida que de la interpretación de la misma conforma la Jurisprudencia del TC[75], cuya doctrina es aplicada por los tribunales ordinarios ante cualquier afrenta al derecho a la intimidad y al derecho a la protección de datos personales. En ambos, sabemos, deberá respetarse su contenido esencial. Recordemos, respecto a éste último, que ha sido definido por el TC[76], como «el poder de disposición y de control sobre los datos personales que faculta a la persona para decidir cuáles de estos datos proporcionar a un tercero, sea el Estado o un particular, o cuáles puede este tercero recabar, permitiendo también al individuo saber quién posee esos datos personales y para qué, pudiendo oponerse a esa posesión o su uso». Su carácter de derecho fundamental le otorga determinadas características, como la de ser irrenunciable y el hecho de prevalecer sobre otros derechos no fundamentales.

Los autores[77] coinciden que entre el régimen jurídico de protección de datos personales y el BD o la IE se está produciendo un choque de trenes, un enfrentamiento de supervivencia para el primero cuando nos referimos al tratamiento de megadatos

Albert H. Law in the Future. *University of Toronto Law Journal*, 2016, vol. 66, no 4, p. 423-428.

[73] Cotino Hueso, Lorenzo. «BD e inteligencia artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales». Ob. cit., p. 137

[74] art. 10. 1 CE

[75] Sin olvidar la influencia que sobre el propio TC ejerce el TJUE cuando se trata de interpretar normas de Derecho Europeo y, en especial, el Convenio para la protección de las personas con respecto al tratamiento automatizado de datos de carácter personal, hecho en Estrasburgo el 28 de enero de 1981. https://www.boe.es/buscar/doc.php?id=BOE-A-1985-23447

[76] STS 76/2019, de 22 de mayo de 2019 y STC 292/2000, de 30 de noviembre de 2000.

[77] Crawford, Kate; Schultz, Jason. Big data and due process: Toward a framework to redress predictive privacy harms. *BCL Rev.*, 2014, vol. 55, p. 93. Rubinstein, Ira. Big data: The end of privacy or a new beginning?. *International Data Privacy Law (2013 Forthcoming), NYU School of Law, Public Law Research Paper*, 2012, no 12-56. Tene, Omer; Polonetsky, Jules. Judged by the tin man: Individual rights in the age of big data. *J. on Telecomm. & High Tech. L.*, 2013, vol. 11, p. 351.

o datos masivos. Apuntan al análisis predictivo y la categorización como las grandes amenazas. No obstante, se puede argumentar en contra que los resultados de esa categorización no son el fruto de unos datos individuales sino de millones de datos, muchos de ellos no atribuibles a personas identificables, que se nutren, entre otras fuentes, de las redes sociales[78]. Si los datos son anónimos o han pasado a ser anónimos, ya no sería aplicable el régimen jurídico del derecho a la protección de datos personales. No obstante, es difícil garantizar que los datos anonimizados puedan perder el anonimato en el futuro por lo que es preciso adoptar medidas suplementarias[79]. De lo que se trata es de evitar la reidentificación mediante técnicas[80] que desvinculen un dato con una persona. También hay dificultades cuando la información se obtiene de una enorme cantidad de fuentes y los datos han podido ser fragmentados previa o simultáneamente al tratamiento.

Es preciso mejorar las condiciones para el intercambio de datos mediante la creación de un marco armonizado en los Estados miembros. Con ese objeto el Reglamento de Gobernanza de Datos tiene como objetivo generar confianza entre los particulares y las empresas en relación con el acceso a los datos, su control, intercambio, utilización y reutilización, especialmente mediante el establecimiento de mecanismos adecuados[81].

Para generar confianza, los mecanismos de reutilización de los datos dentro de la Unión y, con mayor motivo, con terceros países, máxime si se trata de datos sensibles como los sanitarios, el transporte, la energía, el medio ambiente y las finanzas, deberán extremarse las medidas.

El Reglamento de Gobernanza de Datos[82] indica que existen técnicas suficientes para garantizar la privacidad en el afán de analizar datos personales como puede ser la anonimización, la privacidad diferencial, la generalización, la supresión y la aleatorización o el uso de datos sintéticos».

El Parlamento Europeo, en esa línea señala que el tratamiento de los datos personales debe, necesariamente realizarse cumpliendo los fundamentos jurídicos recogidos en el artículo 6 del Reglamento (UE) 2016/679»[83]. Pero también afirma, que el derecho que subyace de la anterior norma solo se aplica «al tratamiento de estos datos correlacionados cuando un individuo es realmente identificable». Los datos

---

[78] Martín Miralles, Ramón. «BD vs Small low», Congrés IDP 2013 Butlletí Kdades: Butlletí electrònic de tecnologia, auditoria i seguretat de la informació, Núm.. 24, 2013, pp. 7-8.

[79] AEPD e ISMS, 2017,. Código de Buenas Prácticas en Protección de Datos para Proyectos de BD. p. 31

[80] Véase el Dictamen 05/2014 sobre técnicas de anonimización del Grupo de Trabajo del artículo 29 de 10 de abril de 2014. https://www.aepd.es/es/documento/wp216-es.pdf

[81] Reglamento (UE) 2022/868 del Parlamento Europeo y del Consejo de 30 de mayo de 2022 relativo a la gobernanza europea de datos y por el que se modifica el Reglamento (UE) 2018/1724 (Reglamento de Gobernanza de Datos). Considerando 5

[82] Reglamento (UE) 2022/868 del Parlamento Europeo y del Consejo de 30 de mayo de 2022. Considerando 7

[83] Resolución del Parlamento Europeo, de 14 de marzo de 2017. Norma citada

anónimos no tienen ningún problema y son una fuente limpia para el uso de las tecnologías de BD siempre y cuando no sirvan para obtener consecuencias que afecten a los derechos de las personas. Luego, ni en estos extremos, la solución es fácil. En todo caso, la aplicación de la seudonimización, la anonimización o el cifrado de los datos personales deben ir unidos necesariamente al establecimiento de garantías que eviten la reidentificación.

En este sentido, la regulación existente en protección de datos precisa de un desarrollo porque las nuevas tecnologías como el BD comportan problemas complejos[84]. No basta un sólido conjunto de principios, sino que será necesario un mayor desarrollo en relación con el BD[85] o la IA[86]. Podemos encontrarnos ante el consentimiento sobre un tratamiento para una finalidad y el BD nos lleva a otra u otras finalidades, por ejemplo, de una finalidad de consumo a otra de salud. De esta forma, sin tratar datos especialmente sensibles como los relativos a la salud, la orientación sexual, la opción política, la religión, el historial delictivo o sancionador, pueden inferirse de otros datos. La geolocalización, las interacciones en redes sociales, el análisis del vocabulario empleado, las relaciones en las redes sociales, etc., pueden llevar a describir aquellos datos sensibles referidos a una persona determinada. La inferencia no es preciso que parta de datos actuales o actualizados, sino que puede proceder de datos históricos.

Se trata, pues, de garantizar los derechos de información, acceso, rectificación, supresión, oposición, portabilidad, no ser objeto de decisiones individuales automatizadas, conservación y limitación que conforman las facultades del derecho a la protección de datos personales, bien delimitados en el Reglamento General de

[84] Lo anterior no es obstáculo para reconocer iniciativas que aportan soluciones como la iniciativa que supuso el Premio a la proactividad y buenas prácticas de la AEPD: Proyecto. Healthdata. Guía Legal y Repositorio para fomentar la compartición de Datos de Salud, de24 noviembre 2020, una iniciativa que tiene por objetivo el fomentar la compartición de datos de salud anonimizados con fines de investigación. https://www.aepd.es/es/documento/premio-rgpd-mod-a-2020-fundacion-29.pdf

[85] Martínez, Ricard. Ética y privacidad de los datos. *Ponencia presentada en Big Data: de la investigación científica a la gestión empresarial, Fundación Ramón Areces*, 2014.

[86] En esta línea de desarrollo se encuentran la Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de Inteligencia Artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión de 21.4.2021 https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A52021PC0206&lang1=ES&from=EN&lang3=choose&lang2=choose&_csrf=a024d78b-015e-4ed2-9ad5-e7af111f0280

la Propuesta de Reglamento de Parlamento Europeo y del Consejo sobre el espacio europeo de datos sanitarios de 3.5.2022

https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0197

o la Propuesta de Reglamento del Parlamento Europeo y del Consejo sobre normas armonizadas para un acceso justo a los datos y su utilización (Ley de Datos) de 23.2.2022

https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=COM%3A2022%3A68%3AFIN

Protección de Datos [87](RGPD). Por lo tanto, es esta norma una aportación muy relevante a la solución de los problemas que presenta el BD.

En el caso del derecho de acceso, Martín Miralles[88] propone la idea de un «derecho de acceso amplificado» que no se limita a los datos propios de un tratamiento utilizados por un responsable, comunicados o no por el interesado, sino también a los tratamientos concretos utilizados y la información obtenida a partir de ellos.

Una de las soluciones que aporta la tecnología es aplicar la política Privacidad desde el diseño (*Privacy by design*) que contiene una serie de principios relatados por Cavoukian[89] que deben respetarse en el diseño de las aplicaciones informáticas. La tecnología debe ser neutral[90], de tal forma que no disminuya el ejercicio y la garantía de los derechos de los interesados. La pretensión final se resume en que esta tecnología sea respetuosa con la privacidad (*Privacy Enhancing Technologies* – PETs) y asegure que el titular de los datos de carácter personal tenga un control sobre sus datos, evitando el acceso no consentido a los mismos[91].

Por ello, los fabricantes de las máquinas, de las aplicaciones y demás herramientas tecnológicas deben diseñar sistemas que aseguren el nivel de privacidad exigido por el ordenamiento. Se trata de evitar los riesgos, pero también de prevenirlos incorporando, por ejemplo, técnicas preventivas de evaluación de impacto. En esta línea podemos mencionar a Selbst[92] que quiere introducir técnicas como «*discrimination impact assessments*» de una forma similar a lo que puede ser una declaración de impacto ambiental o Edwards[93] que aboga por una ética en el diseño que perfile una «evaluación del impacto social». Rubistein[94], en esta línea, elabora las características el modelo a seguir:

> «1. El individuo es el centro del sistema de recolección, gestión y uso de los datos.
> 2. El individuo decide qué información revelar, de forma selectiva.

---

[87] REGLAMENTO (UE) 2016/679 DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 27 de abril de 2016

[88] Martín Miralles, Ramón. «BD vs Small low», Congrés IDP 2013 Butlletí Kdades: Butlletí electrònic de tecnologia, auditoria i seguretat de la informació, Núm. 24, 2013, pp. 7-8.

[89] Cavoukian, Ann. *Privacy by Design - The 7 Foundational Principles*

[h]ttps://www.ipc.on.ca/wp-content/uploads/resources/7foundationalprinciples.pdf

[90] Navas Navarro, Susana. «Computación en la nube: BD y protección de datos personales (Cloud Computing: BD and Personal Data Protection)». *InDret,* vol. 4. 2015. P. 42.

[91] Llácer Matacás, M. Rosa. «La autodeterminación informativa en la sociedad de vigilancia: ubiquitous computing». En Protección de datos personales en la sociedad de la información y la vigilancia. Wolters Kluwer, 2011, pp. 88-92.

[92] Selbst, A. D. (2017). «Disparate Impact in BD Policing». *Georgia Law Review*. p. 50 Acceso en: https://ssrn.com/abstract=2819182 http://dx.doi.org/10.2139/ssrn.2819182

[93] Edwards, Lilian; Mcauley, Derek; Diver, Laurence. 2016. *From Privacy Impact Assessment to Social Impact Assessment*. Conference: 2016 IEEE Security and Privacy Workshops (SPW), 2016. pp. 53-57

[94] Rubinstein, Ira S. «BD: The End Of Privacy Or A New Beginning?» *International Privacy Law,* Vol. 3, núm. 2, 2013.

3. Se establece un control sobre los fines para los que se usan los datos, así como sobre su duración, a través de contratos.

4. El individuo tiene mecanismos para ejercer sus derechos forma abierta y flexible.

5. El sistema tiene mecanismos de autenticación de la identidad de los sujetos que accedan al sistema, mecanismos de seguridad y la posibilidad fácil para el interesado de la portabilidad de datos».

Así, el responsable del tratamiento de datos personales tiene que cumplir las exigencias normativas y debe ser capaz de demostrarlo, esto es, tiene una responsabilidad proactiva[95].

El RGPD[96] apuesta por mecanismos proactivos, que prevengan ante los riesgos más que reaccionen ante ellos y que pueden ser muy útiles en relación al BG[97]. Así, el RGPD determina en su artículo 25 la necesidad de aplicar la protección de datos desde el diseño y por defecto. Esto supone que esa protección se integre en las aplicaciones de tal forma que se aplique en el tratamiento en el inicio y, después, durante todo su ciclo de vida. Igualmente, en el artículo 35 se exige la evaluación de impacto de protección de datos.

Como recuerda Martínez[98], el principio de responsabilidad proactiva introduce una forma de actuar que, por ejemplo, exige un esfuerzo de adaptación permanente de los diseños de las aplicaciones al cumplimiento de la norma. Y, precisamente, los usos del BD y la IA parecen destinados a una obligatoria evaluación de impacto[99], por cuanto pueden implicar la elaboración de perfiles y porque los resultados del tratamiento pueden producir decisiones que afecten al individuo significativamente[100], como señala el art. 35. 3º RGPD.

El RGPD exige aplicar una serie de principios: proporcionalidad, necesidad, limitación del propósito, seguridad, minimización y conservación de los datos, no discriminación y consentimiento informado. Estos principios son de difícil aplicación

[95] González, Pedro Alberto. «Responsabilidad proactiva en los tratamientos masivos de datos», en *Dilemata*, Núm. 24, 2017, p. 120 y 125 y ss.

Acceso en https://www.dilemata.net/revista/index.php/dilemata/article/view/412000103

[96] Art. 5 2. del RGPD El responsable del tratamiento será responsable del cumplimiento de lo dispuesto en el apartado 1 y capaz de demostrarlo («responsabilidad proactiva»).

[97] ENISA. *Privacy by design in BD: An overview of privacy enhancing technologies in the era of BD analytics* European Union Agency For Network And Information Security. 2015. Acceso en https://www.enisa.europa.eu/publications/big-data-protection

[98] Martínez, Ricard Martínez. «Cuestiones de ética jurídica al abordar proyectos de BD. El contexto del Reglamento general de protección de datos, *Dilemata*, Núm. 24, 2017, p. 160.

[99] Cotino Hueso, Lorenzo. «BD e inteligencia artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales». Dilemata, núm. 24, 2017, p. 139.

[100] Cotino Hueso, Lorenzo. «Riesgos e impactos del BD, la inteligencia artificial y la robótica: enfoques, modelos y principios de la respuesta del derecho». Ob. cit., p. 29.

ante el procesamiento de los datos masivo que se propone. Por ello, señala Lecuona[101] «es preciso un análisis interdisciplinar y un debate social para establecer los límites de actuación y las pautas éticas» para la toma de decisiones automatizadas.

La necesidad de prestar el consentimiento como elemento básico en el derecho de protección de datos personales, se muestra insuficiente, primero por las excepciones legales en base al interés público, luego por la excepción que puede ampararse en base al interés legítimo. Así, acertadamente, el consentimiento se transforma en papel mojado ante los nuevos avances tecnológicos y su mayor complejidad. No sería, en puridad, factible el consentimiento pues con el BD o la IA, el resultado de las decisiones automatizadas es imprevisible a priori[102].

Serán precisas, ineludiblemente, la prevención en el diseño y por defecto y la utilización, siempre que sea precisa, la evaluación de impacto de protección de datos, con el RGPD como referencia[103].

Se aportan nuevas soluciones: el «*due process*» puede ser una garantía de transparencia de la utilización de los algoritmos y de sus resultados en relación con el uso de tecnologías como el BD e IA[104]. Esa transparencia debe permitir el acceso a esa información en el caso de que sea precisa para la defensa de los derechos afectados, caso de la intimidad o de la protección de datos personales, y, con relación al BD, el acceso a la prueba de la causa del perjuicio que pueda ser alegada en caso de controversia[105].

Es necesario, para el caso de decisiones que parten del uso de estas tecnologías, que el interesado pueda conocer cómo se han tomado las decisiones que le afectan mediante la utilización de datos masivos. El derecho de acceso, como parte del derecho a la protección de datos, reconocido en nuestro derecho en el RGPD es una herramienta adecuada pero insuficientemente desarrollada desde el plano jurídico para afrontar los riesgos del BD[106]. Más aún, se podría hablar de nuevos derechos como el derecho a la criptografía[107], cuyo ejercicio evitaría la utilización de los datos personales, previamente encriptados, para su no uso.

---

[101] Leucona, Itziar. «Evaluación de los aspectos metodológicos, éticos, legales y sociales de proyectos de investigación en salud con datos masivos (BD)». *Gaceta Sanitaria Elsevier*. Noviembre-Diciembre. 2018, p. 1.

[102] Martínez, Ricard. Ética y privacidad de los datos. *Ponencia presentada en Big Data: de la investigación científica a la gestión empresarial, Fundación Ramón Areces*, 2014.

[103] Becerra, Jairo, et al. Marco teórico y aproximación jurídica al BD: algoritmos, inteligencia artificial y transformación digital. Ob. cit.. p 55.

[104] Balkin, Jack M. 2016 Sidley Austin Distinguished Lecture on Big Data Law and Policy: The Three Laws of Robotics in the Age of Big Data. *Ohio St. LJ*, 2017, vol. 78, p. 1217.

[105] Crawford, K. y Schultz, J. (2014). BD and due process: Toward a framework to redress predictive privacy harms. Boston College Law Review, 55(93), 13-36.

[106] Becerra, Jairo, et al. Marco teórico y aproximación jurídica al BD: algoritmos, inteligencia artificial y transformación digital. Ob. cit. p 54

[107] Becerra, Jairo, et al. Marco teórico y aproximación jurídica al BD: algoritmos, inteligencia artificial y transformación digital. Ob. cit., p 47. Lo enlaza el autor con el derecho a dejar a la persona sola (Warren y Brandeis, 1890), The right to privacy. En http://www.cs.cornell. edu/~shmat/courses/cs5436/warren-brandeis.pdf)

2. La normatividad en camino

El Comité Europeo de Protección de Datos («CEPD») y el Supervisor Europeo de Protección de Datos («SEPD») remarcan, en relación con la existencia de un Espacio Europeo de Datos Sanitarios (EEDS), la interacción oportuna y necesaria de la Ley de Gobernanza de Datos, la Ley de Datos y la Ley de Inteligencia Artificial [108].

A) La Propuesta de Reglamento del Parlamento Europeo y del Consejo sobre normas armonizadas para un acceso justo a los datos y su utilización (Ley de Datos)[109].

Para la Comisión Europea, una utilización adecuada va a permitir una transición de la economía actual a una economía digital. La CE pone de manifiesto que muchos de los datos no llegan a utilizarse o están en poder de unos pocos, desaprovechándose el indudable valor y potencial de los mismos. Es, pues, un reto y a la vez una obligación, desde los principios europeos de respeto a los derechos, liberar de obstáculos el desarrollo de la utilización masiva de datos que permitan un uso pleno, ordenado y garantista de los datos que sea accesible para que todos puedan beneficiarse de las oportunidades y una mejor distribución del valor que de su uso surjan, en un ámbito de economía de datos sostenible en Europa.

En esta línea su presidenta, Ursula von der Leyen, se debe equilibrar el flujo y el uso de los datos, pero también mantener la privacidad de los datos, su protección y la ética de su uso[110].

Se concreta en el objetivo de «garantizar la equidad en la asignación del valor de los datos entre los agentes de la economía de los datos y de fomentar el acceso a los datos y su utilización[111]". En definitiva:

Facilitar el acceso a los datos y su utilización por parte de los consumidores y las empresas, teniendo en cuenta que la obtención y el tratamiento de datos tienen un coste que debe ser resarcido en tanto generan valor. Los contratos de intercambio deben ser seguros y su distribución equitativa, sin que quepan las cláusulas abusivas y respetando los mecanismos de incentivos y remuneraciones que determine normalmente el mercado.

---

[108] CEPD-SEPD. Dictamen conjunto 3/2022 del CEPD y el SEPD sobre la propuesta de Reglamento del Espacio Europeo de Datos Sanitarios (EEDS). (2022). Apartado 3. 3. Párrafos 28,29 y 30.

[109] https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=COM%3A2022%3A68%3AFIN

[110] Von der Leyen, Úrsula. «Una Unión que se esfuerza por lograr más resultados — Mi agenda para Europa, Orientaciones políticas para la próxima Comisión Europea 2019-2024», 16 de julio de 2019.

[111] Propuesta de REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSEJO sobre normas armonizadas para un acceso justo a los datos y su utilización (Ley de Datos). Pág. 3.

Garantizar el intercambio obligatorio de datos entre empresas y las entidades públicas competentes en el caso de emergencias afectadas por intereses públicos,

Facilitar el acceso a servicios de tratamiento de datos competitivos e interoperables en la nube eliminando los obstáculos al intercambio de datos en el espacio común europeo.

Así, al adquirir un producto conectado al internet de las cosas, los datos generados pueden servir al proveedor de aquel, para su mejora, y al cliente para una mejor toma de decisiones. Por ejemplo, los agricultores tendrán a su disposición los datos en tiempo real sobre el clima, la temperatura, la humedad, etc., que permitan optimizar la utilización de los recursos y mejorar el rendimiento.

B) Espacio Europeo de Datos Sanitarios (EEDS)[112]

Los datos sanitarios tienen una doble vertiente, por una parte, afectan de forma radical a elementos esenciales de la intimidad de las personas y, por otra, son necesarios para una mejora de la salud de las personas en su conjunto y por lo ello, afectados de un incuestionable interés público. Este espacio pretende ayudar garantizar el control de los datos personales de salud, pero también, mejorar la investigación, la práctica sanitaria y la elaboración de directrices de política sanitaria. Para ello debería implementarse un mercado único de historiales médicos electrónicos que puedan servir a aquellos fines[113].

C) Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de Inteligencia Artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión[114].

En ella se perfilan como objetivos específicos[115]:

- garantizar la seguridad de los productos y la protección de los derechos fundamentales y valores de la Unión;
- garantizar la seguridad jurídica que permita la inversión en IA;
- facilitar el desarrollo de un mercado único europeo en materia de IA

Esta norma pretende diferenciar niveles de riesgo para atacar las posibles intromisiones en los derechos de las personas de forma adecuada.

---

[112] Proposal for a regulation - The European Health Data Space. https://health.ec.europa.eu/publications/proposal-regulation-european-health-data-space_en

[113] En este punto es de interés el Dictamen conjunto 3/2022 del CEPD y el SEPD sobre la propuesta de Reglamento del Espacio Europeo de Datos Sanitarios (EEDS) https://edpb.europa.eu/system/files/2023-04/edpb_edps_jointopinion_202203_europeanhealthdataspace_es.pdf

[114] https://eur-lex.europa.eu/legal-content/ES/TXT/DOC/?uri=CELEX:52021PC0206

[115] Página 3 de la Propuesta

Se entienden de riesgo inaceptable[116] la manipulación, especialmente en colectivos vulnerables, como los menores o las personas con discapacidad, la clasificación de las personas que pudieran lugar a un trato perjudicial o desfavorable o discriminatorio y los sistemas de identificación biométrico salvo que sirva a la persecución de los delitos previa autorización judicial.

Se entienden de alto riesgo[117], aquellos que pudieran afectar a la seguridad el sistema de IA está destinado a ser utilizado como componente de seguridad de productos indicados en el anexo II, y a los derechos fundamentales en ocho materias concretas contenidas en el anexo III[118] o que conlleven un riesgo para los derechos fundamentales, cuya gravedad y probabilidad sean a las materias de ese anexo III. Para estos, será precisa una autorización previa a su comercialización y, después, un seguimiento específico a lo largo de su vida útil. Esas materias abarcan la identificación biométrica y categorización de personas físicas, la gestión y explotación de infraestructuras críticas, la educación, el empleo, el acceso a los servicios privados esenciales, la migración, el asilo y el control de fronteras, la asistencia en la interpretación jurídica.

3. Las recomendaciones de la AEPD (soft law)

La AEPD recomienda, con sus guías y herramientas, a los responsables y encargados del tratamiento de datos personales un mejor cumplimiento de las normas de protección.

Así el Código de buenas prácticas en protección de datos para proyectos de Big Data[119] muestra recomendaciones al respecto. Desde las empresas, como suministradores de bienes y servicios, se argumenta la necesidad de tener información de sus clientes para mejorar esas prestaciones y satisfacer así sus necesidades presentes y latentes. Se habla así de interés legítimo cuando no es un interés del titular de los datos sino de terceros.

El Grupo de Trabajo del Artículo 29, en su Dictamen 06/2014[120], mostraba la necesidad de balancear el interés de terceros para saber en qué medida podrían

---

116 Art. 5 de la propuesta de Reglamento

117 Artículos 6 y 7 de la propuesta de Reglamento

118 1. Identificación biométrica y categorización de personas físicas:
2. Gestión y funcionamiento de infraestructuras esenciales:
3. Educación y formación profesional:
4. Empleo, gestión de los trabajadores y acceso al autoempleo:
5. Acceso y disfrute de servicios públicos y privados esenciales y sus beneficios:
6. Asuntos relacionados con la aplicación de la ley:
7. Gestión de la migración, el asilo y el control fronterizo:
8. Administración de justicia y procesos democráticos

119 Código de Buenas Prácticas en Protección de Datos para Proyectos Big Data. https://www.aepd.es/es/documento/guia-codigo-de-buenas-practicas-proyectos-de-big-data.pdf

120 https://www.aepd.es/sites/default/files/2019-12/wp217_es_interes_legitimo.pdf

anteponerse al interés del ciudadano generador del dato. Deberá, inicialmente, existir un interés del responsable o del tercero que lo alegue, deberá explicarse el impacto que el tratamiento pueda tener en la esfera de derechos del interesado, en sus expectativas o posibles beneficios, en la naturaleza de los datos y el desequilibrio entre las partes. Y el interés legítimo se transforma, en algunos casos más clara y otras no tanto, en un interés público. Así podemos hablar de un interés legítimo en la libertad de información y expresión, en la prevención del fraude, en la seguridad, en el avance científico desde la investigación o en la estadística. También para los mercados y la publicidad.

Uno de los elementos que más preocupa en el uso de BD es la utilización de los datos para finalidades diferentes para las que se recogieron. En ese caso, el artículo 6.4 del RGPD indica que, para que pueda ser compatible, deberá explicarse cualquier relación entre los fines para los cuales se hayan recogido los datos personales y los fines del tratamiento ulterior previsto, la relación entre los interesados y el responsable del tratamiento, la naturaleza de los datos personales, las consecuencias para los interesados y la existencia de garantías adecuadas que permitan la anonimización o la seudonimización[121].

La AEPD en la referida Guía pone sobre la mesa aspectos tales como el origen de los datos, la transparencia de la información, la calidad y conservación de los datos y la problemática de las decisiones automatizadas individualizadas.

En el primer aspecto, los datos pueden tener una procedencia endógena y exógena. En la primera el responsable asume todo el proceso. En la segunda, la obtención de datos de fuentes externas incrementa sus problemas pues no ha contralado todos los aspectos que permiten conocer la pulcritud en la obtención de aquellos datos[122].

La Guía[123] apunta la necesidad de filtros y controles ulteriores. La integración de datos de orígenes o fuentes diversas, dificulta aún más el problema y anima a la utilización de técnicas de depuración que garanticen también la fiabilidad además de la legitimidad. En este sentido, se indica la necesidad de utilizar los metadatos o información secundaria de datos.

La AEPD se hace eco de la Resolución sobre BD de la 36ª Conferencia Internacional de Autoridades de Protección de Datos y Privacidad[124] donde se propone asumir como básico el principio de transparencia para que el consentimiento parta de una información veraz sobre los fines y propósitos del tratamiento. Y recomienda[125] que un proyecto Big Data debe determinar, con carácter previo, las clases de datos

[121] En ese sentido el Grupo de Trabajo del artículo 29 https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf

[122] El Reglamento de Gobernanza de Datos tiene, entre sus objetivos principales, dar confianza en caso de reutilización de los datos.

[123] Código de Buenas Prácticas en Protección de Datos para Proyectos Big Data. ob, cit. pag. 14

[124] http://globalprivacyassembly.org/wp-content/uploads/2015/02/Resolution-Big-Data.pdf

[125] Código de Buenas Prácticas en Protección de Datos para Proyectos Big Data. ob, cit. pag. 16-17

y la finalidad del tratamiento para escoger los aquellos que sean «adecuados, pertinentes y no excesivos» y los estrictamente necesarios para los propósitos perseguidos. Igualmente deben establecerse protocolos de verificación periódica de los datos para constatar que siguen siendo válidos y mecanismos de su conservación y seguridad ante accesos indeseados.

Por su parte, la Guía de Privacidad desde el Diseño[126] recomienda el uso de los patrones de diseño de la privacidad como soluciones válidas para resolver problemas comunes y repetibles ante situaciones similares. Igualmente propone las *Privacy Enhancing Technologies* o PETs que define como un «conjunto organizado y coherente de soluciones TIC que reducen los riesgos que afectan a la privacidad». En ambos casos, se mencionan distintos repositorio o catálogos que pueden ser utilizados. En el caso de las PETs, *European Union Agency for Cybersecurity* (ENISA) ha publicado varios informes[127] sobre la evolución de herramientas PET para comparar su nivel de madurez y eficacia.

También la AEPD ha desarrollado una guía práctica para la aplicación de la protección de datos por defecto, en aplicación del artículo 25 RGPD y siguiendo la correspondiente guía del CEPD[128]. La clave de su aplicación se encuentra, esencialmente, en «la restricción por defecto», de tal forma que, en caso de uso, el acceso a las funcionalidades básicas se aplica el principio de minimización, sin necesidad de ninguna actuación del responsable o del usuario y de la forma más restrictiva posible[129].

De relevancia en esta materia se encuentra la compartición de datos entendidos como la divulgación de datos a terceros ajenos a la organización con el fin de lograr un propósito específico[130]. Recordemos que el Reglamento de Gobernanza de datos tiene como objetivos, entre otros, establecer las condiciones para la reutilización de determinadas categorías de datos que obren en poder del sector público, un marco de supervisión para la prestación de servicios de intermediación de datos y para la cesión de datos con fines altruistas[131]. En este papel ENISA[132], recomienda la Ingeniería de Datos como conjunto de conocimientos que van a respaldar la selección, el despliegue y la configuración de medidas técnicas y organizativas adecuadas para satisfacer los principios específicos de la protección de datos.

---

126 https://www.aepd.es/es/documento/guia-privacidad-desde-diseno.pdf

127 En esta materia ENISA's PETs Maturity Assessment Repository de 31 de enero de 2019. https://www.enisa.europa.eu/publications/enisa2019s-pets-maturity-assessment-repository

128 Guidelines 4/2019 on Article 25 Data Protection by Design and by Default. https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-42019-article-25-data-protection-design-and_en

129 AEPD. Guía de Protección de Datos por Defecto. Página 20. https://www.aepd.es/es/documento/guia-proteccion-datos-por-defecto.pdf

130 Ingeniería de compartición de datos personales. Casos de uso y tecnologías emergentes. Enero 2023 Pág. 6 https://www.aepd.es/es/documento/enisa-ingenieria-comparticion-datos-personales.pdf

131 Art. 1 del Reglamento (UE) 2022/868, ob. cit.

132 Ingeniería de la Protección de Datos. De la Teoría a la Práctica. Enero 2022. Pág. 8. https://www.aepd.es/es/documento/enisa-ingenieria-de-la-proteccion-de-datos.pdf

4. La ayuda de los principios

El éxito de un adecuado tratamiento de datos deberá tener en cuenta tanto del establecimiento de salvaguardias eficaces de los derechos de los ciudadanos plenamente conformes con el RGPD como en el establecimiento de una gobernanza de datos basada en principios sólidos[133].

Los principios comunes (éticos y los no tan incardinados estrictamente en la ética) deben estar asentados en el Derecho, no basta un mero pronunciamiento doctrinal. Esta es una cuestión fundamental que presupone un acercamiento del concepto de principio ético al de principio general del Derecho. Ese asentamiento no precisa de una norma de obligado cumplimiento, sino que pueden estar imbricados en propuestas o documentos que pretenden su adhesión voluntaria, léase guías o códigos de conducta.

La norma, para que tengan virtualidad y efectividad, debe establecer quiénes son las personas que deben evaluar, inspeccionar o garantizar la conformidad con esos principios éticos durante toda la vida útil de las aplicaciones, incluidos algoritmos y datos utilizados o producidos, y el aseguramiento de la responsabilidad si se producen perjuicios. En todo caso, debe reconocerse que estos principios ya están imbricados, con enfoques diferentes, en los textos constitucionales y otros de gran fuerza normativa en infinidad de normas de Derecho europeo y nacional, por lo que tendríamos oportunidad de hacerlos valer, en muchas ocasiones, si fuera preciso.

El Parlamento europeo[134] reconoce que la realización de «evaluaciones de impacto previas y auditorias deben tener en cuenta las preocupaciones de carácter ético, a fin de evaluar el carácter integrador, la exactitud y la calidad de los datos y de garantizar que las personas a las que se refieren las decisiones y/o los agentes que intervienen en los procesos de toma de decisiones son capaces de entender e impugnar la recopilación o el análisis, los patrones y las correlaciones y evitar cualesquiera efectos perjudiciales para determinados grupos de personas».

A) De los principios éticos

Se reclama la necesidad de ofrecer, en el tratamiento de datos personales (con mayor hincapié en los datos sensibles, como pueden ser los sanitarios) «garantías suficientes de una gestión ética, responsable y legítima anclada en los valores de la UE incluido el respeto de los derechos fundamentales»[135].

[133] CEPD-SEPD Dictamen conjunto 3/2022 del CEPD y el SEPD sobre la propuesta de Reglamento del Espacio Europeo de Datos Sanitarios (EEDS). (2022). Ob. cit. Considerando 12

[134] Resolución del Parlamento Europeo, de 14 de marzo de 2017. Norma citada.

[135] CEPD-SEPD Dictamen conjunto 3/2022 del CEPD y el SEPD sobre la propuesta de Reglamento del Espacio Europeo de Datos Sanitarios (EEDS). (2022). Ob. cit. Párrafo 12

Estos principios, en nuestro ámbito europeo, pueden corresponderse con los siguientes.

El principio de respeto de la dignidad humana, la autonomía y la autodeterminación de la persona que evite que su uso conlleve un sometimiento directo o indirecto de la persona, que conculque la autonomía psicológica desde la manipulación, el engaño o la vigilancia injustificada. El BD no debe ser utilizado para socavar la dignidad humana, ni para obtener información dirigida a la manipulación de las personas, eliminando o reduciendo su libre albedrío.

El principio de respeto a la diversidad, a la no discriminación a la equidad y el fortalecimiento del bienestar social y medioambiental. Con la dignidad humana viajan un conjunto de derechos adquiridos dentro de una sociedad de bienestar. El BD no debe servir para afectar negativamente a los principios de no discriminación, especialmente sobre los marginados o en situación de vulnerabilidad, o por razón de sexo u orientación sexual, ideológica, raza, discapacidad, de igualdad de oportunidades, de mejora de la salud y la educación, de protección de la infancia y de los derechos de los trabajadores. El BD deber contribuir a la igualdad de acceso a la tecnología y a una efectiva distribución equitativa de beneficios y oportunidades para todos los ciudadanos.

La defensa del medio ambiente nos lleva al impulso de la sostenibilidad. El desarrollo, el despliegue y el uso de estas tecnologías deben contribuir a la «transición verde», que intente minimizar los daños que puedan sufrir los ecosistemas, por ejemplo, como consecuencia de la extracción de materiales o las emisiones de gases nocivos para el medio ambiente en la generación y consumo de energía. Los procesos de BD (de minería de datos) deben ser amigables con el medio ambiente.

El principio de intervención humana, del control democrático y de la recuperación del control humano. Esto es, debe garantizarse la intervención humana en los procesos y en el control de estos. Además, ante los riesgos posibles del uso de BD, es fundamental también salvaguardar los pilares de nuestra sociedad democrática basada en el Estado de Derecho, la independencia de los medios de comunicación y la garantía de acceso libre a una información objetiva y que preserve una mayor cohesión social. Supone la eliminación de los sesgos[136] que, utilizados de forma automatizada o no, creen formas de discriminación, en particular cuando nos referimos a grupos de personas homogéneos o con características similares. El BD, si parte de datos erróneos o de baja calidad, puede crear o reforzar sesgos en la línea equivocada. El control democrático se establece con la posibilidad de legislar, estableciendo normas que encaucen el uso

[136] Lazcoz Moratinos, Guillermo. Análisis de la propuesta de reglamento sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas. *IUS ET SCIENTIA*, 2020, vol. 6, no 2, p. 26-41: propone la siguiente definición de sesgo: basado en el texto de Hildebrandt, en Hildebrandt, Mireille. The issue of bias: the framing powers of ML. *Machine Learning and Society: Impact, Trust, Transparency, MIT Press forthcoming*, 2020, propone la siguiente definición: l) «sesgo», toda desviación o representación estadística que reconfigura la distribución de los bienes, servicios, riesgos y oportunidades de una o varias personas físicas o jurídicas.

y desarrollo de estas tecnologías. También deben ser el fruto de la participación de la sociedad a través de los cauces de participación democrática

El principio de solidez, precisión y fiabilidad. Solidez para evitar las vulneraciones de la seguridad, los ciberataques y los usos indebidos de los datos personales. Precisión que socave las posibles fugas de datos, las noticias falsas y la desinformación. Ambas respetuosas con la privacidad de los datos, dentro del acervo jurídico de la Unión en esta materia. Fiabilidad, basada en la calidad de los datos empleados, sin sesgos que dañen los datos acumulados y con unos algoritmos limpios, sin estrategias ocultas.

El principio de transparencia y explicabilidad de las tecnologías. El BD se sirve de los algoritmos que podemos definir como operaciones o indicaciones lógicas que permiten llegar a un resultado a partir de la entrada de información. Conocer cuál ha sido el proceso utilizado para tratar los datos y obtener las inferencias resultantes resulta trascendental para garantizar el derecho a la información de los interesados. Esta información debe ser suministrada de manera comprensible, accesible cuando sea solicitada, normalizada, completa. Debe contener el razonamiento utilizado, los resultados o repercusiones posibles. Debe mostrar las formas de verificación, impugnación y, en su caso, corrección y revisión del sistema. Para ello será preciso mitigar la fuerza de la propiedad intelectual para evitar que, amparándose en ella, pueda encubrirse decisiones opacas malintencionadas.

Se podría hablar de la transparencia algorítmica[137] y su falta resultará ser el enemigo a batir si queremos asegurar que los usos del BD se queden del lado del Derecho. Con Véliz[138], los algoritmos son meras herramientas y los ataques a la ética son responsabilidad de los que los crean o los usan: que no pretendan echar la culpa a los algoritmos de tratamientos de datos personales poco éticos.

Deben conocerse, también, las actividades de mantenimiento de los sistemas, cómo se reparan o mejoran, por ejemplo, en particular, mediante las actualizaciones para corregir sus vulnerabilidades.

El principio de privacidad y de protección de datos. En la medida que el BD se nutre de datos, muchos de ellos personales, son aplicables todos los principios ya establecidos para el tratamiento de los datos personales recogidos fundamentalmente en el artículo 5 del RGPD. Estos son los principios de licitud, lealtad y transparencia, de limitación de la finalidad, de minimización de datos, de exactitud, de limitación del plazo de conservación, de integridad y de confidencialidad. Ni los robots físicos, ni aquellos que operan a través las redes deben recopilar, tratar o transferir datos sin consentimiento. Más aún, no deben influir en el desarrollo personal o en las

[137] Resolución del Parlamento Europeo de 14 de marzo de 2017.

[138] Véliz, Carissa. *Moral zombies: why algorithms are not moral agents. AI & SOCIETY*, 2021, vol. 36, no 2, p. 496.
https://link.springer.com/content/pdf/10.1007/s00146-021-01189-x.pdf?pdf=button%20sticky

relaciones interpersonales, libres de vigilancia. Podemos hablar de un nuevo derecho: el derecho a no ser perfilado, medido, analizado, aconsejado o provocado[139].

El principio de rendición de cuentas. Se debe garantizar a las personas, como un derecho, las vías de recurso o revisión imparciales, accesibles, asequibles y efectivas, efectuada por personas físicas, tanto en el sector público como en el privado. Deben desarrollarse soluciones sólidas que asignen responsabilidades de forma clara y justa desde una legislación vinculante eficiente. Esta redención de cuentas no debe ser sólo formal, sino material, que culmine, en su caso, con la reparación del daño. La rendición de cuentas debe garantizarse a lo largo de las distintas etapas de la toma de decisiones para evaluar no solo la representatividad, la exactitud y la calidad de los datos sino también la idoneidad de cada una de las decisiones que vayan a adoptarse sobre la base de dicha rendición[140].

El principio de prevención de daños y de responsabilidad. La necesidad de prevenir daños debe impulsar y desarrollar procesos de verificación, validación y control de los sistemas de BD que atribuyan certificados o etiquetados de buenas prácticas. Igualmente, la elaboración de códigos deontológico o de buena conducta a los que se deban o puedan adherir las partes implicadas y que pongan en valor la protección de los principios de dignidad humana, integridad, libertad, privacidad, no discriminación y, en general, los derechos humanos fundamentales. Pero si la prevención no ha producido sus efectos, se derivará la responsabilidad a los causantes que deberán, en su caso, asumir las consecuencias resarcitorias e indemnizatorias. Esa responsabilidad también incorpora el deber ser creados de tal forma que su diseño sea respetuoso con los valores y los derechos fundamentales Participo con Ramón Fernández[141] que no es acertado atribuir la responsabilidad civil a las aplicaciones informáticas, ni dotarlos de semiderechos que servirían para crear pantallas cuyo objetivo evidente sería sortear la responsabilidad civil. Detrás de un robot, un programa informático, un producto, está la mente creadora del ser humano.

El principio de proporcionalidad y justificación de la necesidad de limitación de derechos. En la utilización de medidas que sean restrictivas de derechos por causa debidamente justificada en el interés público, deberán mitigarse, por ejemplo, los efectos adversos sobre el derecho a la intimidad, a la protección de datos y a la no discriminación. Es el caso de la utilización de tecnologías de reconocimiento facial que debe ser restringido en el tiempo y debe estar sujeta a control judicial para evitar los abusos como la vigilancia masiva en actuaciones policiales predictivas. El BD puede necesitar, por razones de interés público, limitar algunos derechos, por ejemplo, en el curso de investigaciones que tengan su causa en la comisión de delitos.

[139] Grupo Europeo sobre Ética de la Ciencia y las Nuevas Tecnologías. Declaración sobre Inteligencia artificial, robótica y sistemas «autónomos». Marzo 2018. Pina 17.

[140] Resolución del Parlamento Europeo, de 14 de marzo de 2017.

[141] Ramón Fernández, Francisca. Robótica, inteligencia artificial y seguridad: ¿Cómo encajar la responsabilidad civil? *La Ley (Online)*, 2019, no 9365, p. 8.

Todos estos principios éticos deben ser normativizados para una mejor aplicación que prevea sanciones para el caso de su incumplimiento. En este sentido, Lazcoz Moratinos[142], expresa que es necesario un régimen sancionador si queremos un efectivo cumplimiento de los principios éticos y que resulta imprescindible su existencia en sectores de alto riesgo.

Nos posicionamos proponiendo que todo proyecto de BD debería incorporar aquellos principios éticos considerados fundamentales. Se trata de hacer el bien y no hacer el mal, esto es, conjugar el interés del sector privado o del sector público con la garantía de los derechos fundamentales de las personas dentro de una sociedad cada vez mejor y más justa[143]. Adoptar el enfoque de derechos humanos implica tenerlo en cuenta en todos los aspectos de las políticas y toma de decisión. En esta línea, Fernández Aller[144], propone que debe promoverse la transparencia para determinar el impacto de las tecnologías en los derechos humanos, y debe incorporarse un marco regulatorio que muestre las obligaciones del sector público y del sector privado, de acuerdo con los Principios de Empresa y Derechos Humanos de Naciones Unidas[145] asegurando el respeto del principio de no discriminación.

Por último, no debemos olvidar que un planteamiento que tenga en cuenta a la ética, debe integrase tanto en las propuestas como en las soluciones, y deben incorporarse de forma efectiva o, al menos, deben ser reconocibles a través del Derecho[146].

Debemos, no obstante, poner en valor propuestas de indudable calado como el documento elaborado por el Grupo de expertos de alto nivel sobre inteligencia artificial[147] denominado: «Directrices éticas para una IA fiable», de 8 de abril de 2019. Concluye afirmando[148] que una IA fiable debe ser lícita y cumplir todas las leyes y reglamentos aplicables, debe ser ética, esto es, respetuosa de los principios y valores éticos, y además robusta, que evite los daños tanto los intencionales como los accidentales.

Y no podemos ser ajenos a la estrategia diseñada en el marco de los esfuerzos de la UE en buscar mecanismos que garanticen el desarrollo, el despliegue, el uso y la

---

142 Lazcoz Moratinos, Guillermo. Análisis de la propuesta de reglamento sobre los principios éticos para el desarrollo, el despliegue y el uso de la inteligencia artificial, la robótica y las tecnologías conexas. Ob. cit., p. 40.

143 Martínez, Ricard. «Cuestiones de ética jurídica al abordar proyectos de Big Data. El contexto del Reglamento general de protección de datos», *Dilemata*, 2017, no 24, p. 161

144 Fernández Aller, María Celia. «Salud digital, salud global y ética. Una mirada desde el enfoque de derechos humanos». *Revista Diecisiete*, 2020, núm. 3, pp. 92-93.

145 https://www.ohchr.org/documents/publications/guidingprinciplesbusinesshr_sp.pdf

146 Cotino Hueso, Lorenzo. «BD e inteligencia artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales». Ob. cit.,, p. 136.

147 El Grupo de expertos de alto nivel sobre IA es un grupo de expertos independientes constituido por la Comisión Europea en junio de 2018.

148 Grupo de expertos de alto nivel sobre inteligencia artificial. Directrices éticas para una IA fiable. Abril 2019. Comisión Europea. Bruselas. P. 46. (139). file:///C:/Users/Usuario/Downloads/directrices%20%C3%A9ticas%20para%20una%20ia%20fiable-KK0219841ESN%20(1).pdf

gestión de la IA dentro del respeto a los derechos fundamentales valores y libertades recogidos en los Tratados de la Unión[149].

B) De los principios de privacidad desde el diseño y por defecto

El concepto protección de datos desde el diseño y por defecto ya está recogido implícitamente en el artículo 25 del RGPD por el que el responsable del tratamiento aplicará las medidas técnicas y organizativas apropiadas para salvaguardar los derechos y libertades de las personas físicas tanto en el diseño del proyecto de BD como por defecto.

Desde este planeamiento podemos reseñar los siguientes 7 principios básicos[150]:

1. Proactivo no reactivo; preventivo no correctivo. Debemos anticiparnos y no tener solo mecanismos para responder cuando el daño suceda. Podemos hablar de «prevenir antes que curar».
2. La privacidad como configuración por Defecto. Los datos personales deben estar protegidos automáticamente, sin necesidad de una actuación previa o ulterior del cliente o del proveedor. La privacidad permanece intacta desde el inicio.
3. La privacidad embebida en el diseño. Los diseños de BD están implementados en los procedimientos del responsable y en las propias estructuras tecnológicas, como un componente esencial integrado en las mismas.
4. Funcionalidad completa. Todos ganan, los titulares de derechos y los de intereses legítimos tienen cabida sin que unos existan en detrimento de los otros, y viceversa.
5. Seguridad completa del ciclo vital de los datos, desde la recogida hasta su destrucción.
6. Visibilidad y transparencia. Los interesados pueden, en cualquier momento, acceder a los datos, los procesos y la tecnología utilizada.
7. Respeto a la privacidad del usuario los desarrolladores del sistema tienen siempre presente el interés de las personas.

[149] Por ejemplo y en este sentido la Resolución del Parlamento Europeo, de 20 de enero de 2021, sobre inteligencia artificial: cuestiones de interpretación y de aplicación del Derecho internacional en la medida en que la UE se ve afectada en los ámbitos de los usos civil y militar, así como de la autoridad del Estado fuera del ámbito de la justicia penal. https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52021IP0009&from=HR

[150] Seguimos a la AEPD. Código de Buenas Prácticas en Protección de Datos para Proyectos Big Data. ob, cit. pag. 20. De forma similiar, en la Guía de privacidad desde el diseño. AEPD. https://www.aepd.es/es/documento/guia-privacidad-desde-diseno.pdf

C) El principio de responsabilidad proactiva (*accountability*)

Para la AEPD[151] consistente en «el reconocimiento, asunción de responsabilidad y actitud transparente sobre los impactos de las políticas, decisiones, acciones, productos y desempeño asociados a una organización», con independencia de la existencia de normas de obligado cumplimiento sobre la materia.

En nuestro ámbito, supondría la implantación de medidas, de garantía y cumplimiento de los principios y obligaciones en materia de protección de datos, y el establecimiento de mecanismos internos y externos para evaluar su fiabilidad y demostrar su efectividad cuando así se requiera.

El Grupo de Trabajo del Artículo 29 concreta la aplicación práctica del principio de *accountability*[152], en estar en disposición de demostrar por parte de los responsables que han adoptado las medidas adecuadas, teniendo en cuenta las circunstancias que rodean al tratamiento. Y esto puede probarse mediante:

- El establecimiento de políticas escritas y vinculantes de protección de datos que contendrán procedimientos internos de revisión y control previos y simultáneos a las operaciones de tratamiento de datos.
- El nombramiento de un delegado de protección de datos y la constitución de un mecanismo interno de resolución de quejas de los interesados.
- Una formación adecuada en protección de datos a las personas involucradas: personal, desarrolladores y responsables de los procesos.
- El establecimiento de procedimientos internos de gestión de fallos de seguridad y la realización de evaluaciones de impacto, en su caso.
- La aplicación de auditorías internas o externas

V. CONCLUSIONES

El BD está demostrando ser una herramienta tecnológica fundamental en el desarrollo de múltiples actividades del ser humano que producen un mayor bienestar para la Humanidad. Podríamos decir que todas las actividades en las que exista información reflejada en datos (entendida en un sentido amplio, cualquiera que sea el soporte, texto, localización, fotografías, imágenes y ficheros de audio, etc.) son susceptibles de utilizar esta tecnología para obtener conclusiones predictivas y, por ello, posibilitar un mayor número de aciertos en las decisiones a tomar. La Humanidad, incontestablemente se rige a lo largo de la Historia por aciertos y errores. Cuantos más sean los primeros, menores los segundos, mejor para todos y para las generaciones presentes y futuras.

[151] AEPD. Código de buenas prácticas proyectos de big data. Pág. 21-22

[152] GT29. Opinion 3/2010 on the principle of accountability. Pág. 38-39. https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2010/wp173_en.pdf

Sin embargo, el BD no está exento de problemas en su aplicación que pueda zaherir derechos y libertades individuales y, también, a colectivos determinados. Esta tecnología se diseña con intrincados procesos para los profanos en la materia que pueden ocultar los verdaderos fines, intencionados o no, de su uso. La discriminación arbitraria de personas y colectivos es uno de sus grandes riesgos. También lo es aquello de lo que fundamentalmente se nutre: los datos, muchos de ellos, personales. Es en este punto donde ni siquiera el consentimiento consciente es una buena salvaguardia pues siendo afirmado se asiente sobre el desconocimiento de la propia tecnología y sus efectos directos e inmediatos, indirectos o mediatos, de presente y de futuro. Tampoco el principio de minimización, contrario en esencia al concepto de BD, ayuda mucho. La anonimización como solución pierde fuerza ante las posibilidades de reversión o reidentificación. Por último, las dificultades de descifrar los opacos algortimos de los procesos automáticos nos aportan malos augurios. No podemos olvidar que un sesgo malintencionado, un algoritmo en la sombra, puede provocar efectos perniciosos de difícil resarcimiento en una sociedad globalizada gracias a esas mismas tecnologías.

Es en el Derecho y en la aplicación de los principios dónde podemos proteger aquella esfera individual y colectiva que puede verse dañada.

No cabe duda de que el Derecho, como cualquier cambio social, necesita de una adaptación constante ante las nuevas situaciones. Desde el Derecho se reconoce la necesidad de implementar nuevas y más afinadas normas. Y ya están en curso propuestas de Reglamentos específicos del BD y de la IA que van a aportar una estructura de garantías aún más sólida

El RGPD ha supuesto un hito indiscutible de protección en el tratamiento de datos personales, pero debe afinar en el uso del BD. Constatamos que el consentimiento, allí regulado, queda difuminado ante la fuerza abrumadora y de efectos todavía impredecibles del BD. Otros derechos constitucionalmente protegidos también pueden verse afectados y no hay, todavía, mecanismos suficientemente lúcidos para atajar los peligros, como la libertad de expresión o información, o la no discriminación cualquiera que sea el motivo de ésta.

Las recomendaciones de los organismos supervisores, también cuentan. Desde la AEPD, el CEPD o el SEPD se emiten documentos que marcan un camino de interpretación y aplicación, que sirven también para esos propósitos garantistas.

Y los principios que, por su capacidad intrínseca de abstracción o deductiva, pueden abordar situaciones particulares diferenciadas. Unos principios encajan en la ética y otros más en la realidad práctica del manejo de los datos personales. Tenemos a nuestro favor que muchos de estos principios se encuentran incorporados en normas de nuestro Derecho positivo, europeo y nacional, y que están disponibles para ser incluidos, así mismo, en códigos de conducta voluntariamente asumibles. Se recomiendan ambas, por una parte, interesarse en la incorporación o mejor asentamiento de los principios éticos en nuestro Derecho Positivo y, por otra, invitar a los personajes actores de esta tecnología a asumir voluntariamente los principios reconocidos.

Title
Big Data: risks and solutions from Law and principles.

Summary:

Resumen:

El avance de algunas tecnologías, especialmente las relacionadas con la comunicación (redes sociales) y la utilización generalizada de dispositivos electrónicos capaces de transmitir información, algunos de ellos sin intervención de las personas, ha generado ingentes cantidades de datos que, si son tratados, pueden socavar derechos fundamentales y sus garantías reconocidas en nuestro entorno europeo. Este hecho no ha pasado desapercibido y genera una gran preocupación en la medida en que estos tratamientos masivos de datos no tienen un adecuado marco jurídico que responda a los nuevos riesgos que conllevan. Desde el Derecho, en sentido amplio, sin excluir las recomendaciones de los órganos administrativos supervisores, desde los principios éticos y desde el peculiar manejo de los datos personales, se están aportando diferentes soluciones. El análisis de riesgos y sus posibles soluciones marcan la línea de este trabajo.

Abstract:

The advance of some technologies, especially those related to communication (social media) and the widespread use of electronic devices capable of transmitting information, some of them without the intervention of individuals, has generated huge amounts of data that, if processed, can undermine fundamental rights and their guarantees recognized in our European environment. This fact has not gone unnoticed and generates a great concern insofar as these massive data processing operations do not have an adequate legal framework that responds to the new risks they entail.
From the Law, in a broad sense, including the recommendations of the supervisory administrative bodies, from the ethical principles and from the peculiar handling of personal data, different solutions are being provided. Risk analysis and possible solutions mark the line of this work.

Palabras clave:

Big Data; Derecho; principios éticos; Unión Europea.

Keywords:

Big Data; Law; soft law; principles; European Union