¿UN NUEVO ESTATUTO JURÍDIDO PARA EL CIUDADANO?

Juan Antonio Hernández Corchete[1]

I.- LA INTELIGENCIA ARTIFICIAL DETERMINA LA ARQUITECTURA DEL ESPACIO DIGITAL. LA EXISTENCIA VIGILADA Y MEDIDA. II.- LA VIGILANCIA CONSTANTE Y LA CONDICIÓN JURÍDICA DE PERSONA. ¿CONVENIENCIA DE UN DERECHO GENERAL DE LIBERTAD? III.- LOS RIESGOS ESPECÍFICOS DE LA INTELIGENCIA ARTIFICIAL Y LOS DERECHOS PROCEDIMENTALES COMO RESPUESTA COMÚN. IV.- EL USO DE LA INTELIGENCIA ARTIFICIAL POR LAS ADMINISTRACIONES PÚBLICAS. V.- LA DEFINICION DE LOS DERECHOS EN EL ESPACIO DIGITAL. ESPECIAL REFERENCIA A LA PROPUESTA DE REIA.

I.- LA INTELIGENCIA ARTIFICIAL DETERMINA LA ARQUITECTURA DEL ESPACIO DIGITAL. LA EXISTENCIA VIGILADA Y MEDIDA.

La tecnología o tecnologías a las que nos referimos con el nombre de "inteligencia artificial" es un campo en pleno desarrollo, y por ello muy abierto en el sentido de que comprende bajo una misma denominación modalidades que varían sustancialmente unas respecto de otras[2]. Es cierto que ya durante algunos años la inteligencia artificial viene siendo objeto preferente de discusión tanto en el espacio público como entre los investigadores universitarios (incluidos los de las ramas jurídicas). Y que este tiempo asciende a décadas si nos referimos a los tecnólogos y las empresas que han participado activamente en su desarrollo[3]. Sin embargo, los modelos fundacionales, que son aquellos que a partir de predicciones apoyadas en análisis de datos no etiquetados pueden realizar una generalidad de tareas, han saltado al conocimiento público, como una gran novedad sobre cuyas posibilidades concretas aún se sabe bastante poco, hace escasos meses, en la primera parte del año 2023[4]. La propia Propuesta de la Comisión Europea sobre el Reglamento de Inteligencia Artificial que data de 2021 no contemplaba estos modelos

[1] Profesor Titular Derecho Administrativo. Universidad de Vigo.

[2] Por referencia a Estados Unidos, cfr. Waldman, A. E., "Power, Process, and Automated Decision-Making". Fordham Law Review, Vol. 88, 2019. Para la experiencia europea, cfr. Tangi L. y otros: AI Watch European landscape on the use of Artificial Intelli-gence by the Public Sector, JRC Science For Policy Report, Unión Europea, 2022.

[3] Veáse Zuboff, S., "We make them dance: surveillance capitalism, the rise of instrumetarian power, and the threat to human rights", en Human rights in the age of platforms, MIT Press, 2019, pp 3-51, en especial pp 9 y ss donde sitúa en 2001 la decisión de Google de conservar los datos de las búsquedas para prestar, a partir de su combinación con algoritmos, otros servicios económicos.

[4] Sobre los distintos modelos de inteligencia artificial y el rápido desarrollo de los modelos fundacionales, cfr. General-Purpose Artificial Intellegence, documento publicado por el Parlamento Europeo el 30.3.2023 y que se puede consultar en https://www.europarl.europa.eu/thinktank/en/document/EPRS_ATA(2023)745708

más avanzados de inteligencia artificial[5]. Ha sido el Parlamento Europeo el que, en la fijación de su posición en el proceso legislativo aprobada en junio de 2023 y durante las rondas de negociación (trílogos) mantenidos en otoño de 2023, ha reclamado una regulación específica para estas aplicaciones de inteligencia artificial por ser sustancialmente diferentes[6].

El hecho es que lo que denominamos "inteligencia artificial" no es una realidad homogénea y que se necesita, en especial para elaborar una respuesta jurídica adecuada, una labor de delimitación conceptual que precise las características de las distintas modalidades y los riesgos que cada una de ella conlleva para las personas. Una cosa es un algoritmo cuyos criterios de decisión estén predeterminados y se limita a ejecutarlos de un modo automático con las ventajas que ello conlleva en términos de eficiencia y de menor uso de recursos. Y otra muy distinta es un sistema de inteligencia artificial que, a partir de la consideración de un conjunto masivo de datos procedentes de experiencias anteriores, prediga cuáles son los criterios que deben guiar la decisión para conseguir el objetivo perseguido. En este último caso el sistema de inteligencia artificial no solo proporciona agilidad y ahorro, sino que determina autónomamente el criterio que orienta la decisión. No es lo mismo, en segundo lugar, que un modelo de carácter predictivo esté limitado a una tarea muy concreta a que sus posibilidades se extiendan a una generalidad de actividades. Otro factor relevante girará en torno a si la inteligencia artificial se inscribe en una actividad preparatoria de una decisión o, por el contrario, conforma la propia decisión o la condiciona de un modo determinante. Por último, debe tenerse presente la gravedad del impacto que la actividad configurada con inteligencia artificial es susceptible de producir sobre las personas destinatarias, lo que se relaciona no tanto con las características de los modelos de inteligencia artificial sino con los usos que se les den, con el contenido material de las decisiones en que influyan.

De todos estos factores, y también de otros que se vayan identificando a medida que este esfuerzo de construcción jurídica vaya dando sus frutos, depende la afectación de los intereses de los ciudadanos y, por tanto, qué derechos se les reconocen. No parece que quepa proponer que la inteligencia artificial hace nacer un estatuto jurídico del ciudadano completa o sustancialmente idéntico, sino que presentará asimetrías relevantes según las características de cada modelo y el impacto de sus distintos usos. Precisamente este enfoque se desprende también de los trabajos preparatorios del Reglamento de Inteligencia Artificial, pues parte de una clasificación fundada en el riesgo para las personas y en cuanto a los supuestos de alto riesgo exige que los proveedores y/o los usuarios evalúen el concreto impacto que el uso de la inteligencia artificial supondrá para las personas, con el fin de adoptar las medidas de salvaguarda que resulten necesarias. En otras palabras, la delimitación última del estatuto jurídico del ciudadano se realizará casuísticamente y se confiará, al menos en un momento inicial y sin perjuicio de la supervisión posterior por la autoridad pública competente, a la apreciación del proveedor y/o usuario. Esta fórmula de colaboración público-privada que desemboca en una delimitación asimétrica de los derechos del ciudadano tiende a permitir que se aprovechen

[5] La Propuesta de la Comisión Europea se identifica como COM (2021) 206 final, de 21 de abril de 2021.
[6] Sobre la inclusión de los modelos fundacionales en el proceso legislativo del REIA y las posiciones de los colegisladores durante los trílogos, cfr. DIGITALEUROPE, "AI Act trilogues: A vision for future proofing, governance and innovation in Europe", que es un documento publicado el 16 de octubre de 2023.

todos los usos posibles de la inteligencia artificial con tal que se revistan de las medidas necesarias para salvaguardar la posición jurídica de las personas afectadas. Se persigue con ello asegurar un alto nivel de protección de los derechos de los ciudadanos sin por ello obstaculizar innecesariamente los nuevos usos que puedan tener las tecnologías de inteligencia artificial. Se trata de compatibilizar el objetivo irrenunciable de poner en el centro a la persona humana con el fomento de la innovación que genere progreso y riqueza para las sociedades de los países europeos.

Esta perspectiva es uno de los ejes que informan este trabajo. Sin embargo, quiero abordar aquí un enfoque complementario, según el cual resulta desaconsejable tratar de un modo completamente separado las distintas modalidades de inteligencia artificial a los efectos de determinar los derechos de las personas destinatarias. Me detendré en dos factores que sostienen este enfoque complementario.

De un lado, todas las modalidades de inteligencia artificial de carácter predictivo, cualesquiera que sean las características del modelo usado o los impactos de los usos a que se apliquen, tienen en común que operan a partir de la consideración de un conjunto masivo de datos procedentes de experiencias anteriores, a partir de los cuales es posible apreciar correlaciones que saquen a la luz patrones de éxito que repetir y de fracaso que evitar. Es bien conocido que la eclosión de la inteligencia artificial como fenómeno extraordinariamente prometedor hunde sus cimientos en el aumento de la capacidad de computación y en el perfeccionamiento de los algoritmos, pero sobre todo en que actualmente hay enormes conjuntos de datos organizados a partir de los cuáles entrenar los modelos de inteligencia artificial. Cuanto mayor y más variada sea la acumulación de datos más fiable será la predicción que se derive del uso de dichos modelos.

Esta disponibilidad de datos no se produce por generación espontánea, sino que se han adaptado los resortes necesarios con vistas a producir a gran escala esta materia prima que es imprescindible para que funcione este nuevo modelo productivo. Y ello tanto en el espacio digital como en la vida analógica.

Por lo que hace a lo primero, no hace tanto tiempo los datos generados con motivo de la prestación de servicios digitales eran desechados porque no se veía en ellos ningún valor económico. Luego, poco después del año 2000, se cayó en cuenta que todos esos datos permitían, a partir de su análisis, prestar otros servicios y fue entonces cuando se adoptó la configuración del espacio digital que hoy conocemos, caracterizado porque los datos derivados de los servicios en línea se conservan con el objeto de que sirvan para, mediante el uso de técnicas de inteligencia artificial, prestar otros servicios o en general para cumplir otros objetivos. Es más, han proliferado los servicios que se prestan digitalmente con la finalidad principal de que, con ocasión de su uso, se vayan acumulando más datos sobre experiencias anteriores que puedan ser usados para entrenar de un modo eficiente los modelos de inteligencia artificial. Es, por tanto, el funcionamiento de las aplicaciones de inteligencia artificial y la acumulación de datos que actúa como su presupuesto lo que motiva que la sociedad digital, de entre varias posibles arquitecturas, haya quedado organizado de tal modo que la actuación de las personas está en constante vigilancia[7].

[7] Sobre la arquitectura de la sociedad digital, Zuboff, S.,*The age of surveillance capitalim: the fight for a human future at the new frontier of power*, New York, 2019. Y en España, véase Piñar Mañas, J.L., *Derecho e innovación tecnológica. Retos de presente y futuro*, CEU Ediciones, Madrid, 2018.

Ahora bien, la presión por acumular datos que ejerce la operatividad de la inteligencia artificial no se ha conformado con condicionar la arquitectura del espacio digital sino que ha saltado también a la vida analógica. Cada vez son más las actividades que las personas realizan fuera del espacio digital que son medidas por todo tipo de aparatos tecnológicos (asistentes de voz, domótica, wearables etc) y convertidas en datos. Es el objetivo de incrementar y dar variedad a los conjuntos de datos que son utilizados por las aplicaciones de inteligencia artificial lo que presta un incentivo especial a la extensión exponencial de estos dispositivos que designamos bajo el calificativo "Smart".

En conclusión, que la inteligencia artificial con todas sus promesas de eficiencia y progreso se sitúe como eje del modelo productivo como si de una nueva revolución industrial se tratase[8], en sí mismo considerado como un fenómeno general y más allá de que unas manifestaciones generen más riesgos que otras, exige como presupuesto la disponibilidad de datos en cantidad y variedad suficiente, lo que empuja hacia la construcción tanto del espacio digital como de la realidad analógica en clave de vigilancia constante de las actividades que realizan las personas. De aquí se desprende que la definición del estatuto jurídico del ciudadano no debe fijarse únicamente teniendo en cuenta los riesgos concretos que implica cada una de las aplicaciones de inteligencia artificial. Hay un impacto relevante sobre la persona, consistente en su sometimiento a una existencia vigilada, que también ha de ser considerado y que es inherente a la inteligencia artificial como fenómeno global que informa el modelo productivo.

El segundo factor que apuntala este enfoque complementario es que la realidad digital se comporta de un modo fluido y se resiste a fragmentaciones injustificadas. Hay actividades que no se justifican sino por su condición de preparatorias o accesorias de otras actividades de estructura distinta. Abordarlas de un modo completamente separado al efecto de delimitar los derechos de las personas afectadas no es acertado porque deja resquicios para la aparición sistemática de situaciones de falta de tutela. Vale la pena detenerse a explicar esta cuestión con algún detalle.

Las posiciones doctrinales más próximas a la perspectiva empresarial sostienen que las responsabilidades asociadas al uso de la inteligencia artificial deben ser atribuidas entre los distintos actores de un modo adecuado e incluso que resulta conveniente permitir que ellos mismos distribuyan entre sí estas responsabilidades en virtud de mecanismos contractuales. Esta no es una idea nueva. Un enfoque muy similar viene siendo postulado desde hace años por varias instancias para el reparto de las obligaciones de protección de datos entre los responsables de aquellos tratamientos de datos personales que, siendo distintos y sucesivos, se condicionan de algún modo, de suerte que no habría una corresponsabilidad completa frente al titular de los datos, quien no tendría otra opción que exigir sus derechos de uno u otro de los responsables del tratamiento según la fase de tratamiento afectada y todo ello sin perjuicio de que los varios responsables acordasen con efectos externos otra distribución de responsabilidades[9].

El artículo 26 RGPD y sobre todo la jurisprudencia del TJUE han adoptado una posición ecléctica, haciendo jugar el principio de corresponsabilidad completa y al mismo tiempo

[8] Veáse K. Schwab, *La cuarta revolución industrial*, Barcelona 2016.

[9] Grupo del Art. 29, Dictamen 1/2010, sobre los conceptos de "responsable del tratamiento" y de "encargado del tratamiento", pág. 22.

admitiendo una cierta idea de distribución de responsabilidades por fases de tratamiento[10]. La doctrina ha entendido que este enfoque fragmenta injustificadamente una realidad profundamente interconectada y que ello hace peligrar el alto nivel de protección de los derechos de los titulares de los datos que persigue la normativa europea[11].

Este dilema se va a reproducir con igual o mayor intensidad en el ámbito de la inteligencia artificial. Va a ser común que las predicciones a partir de tecnologías de inteligencia artificial sean actividades que se realizan únicamente porque hacen más eficientes las decisiones que se adoptan en el seno de otras actividades. La calificación crediticia de las personas, que es uno de los usos más conocidos de la inteligencia artificial, solo tiene sentido en función de la decisión de la entidad financiera de acceder o no a una concreta solicitud de crédito. Solo en la primera hay inteligencia artificial pero sería desacertado no considerarla en conjunto con la segunda que es la que le da sentido. Esta cuestión ya se ha planteado desde la perspectiva del derecho que el artículo 22 RGPD reconoce a los titulares de los datos personales frente a las decisiones automatizadas. Y probablemente se planteará en lo venidero en cuanto a otros aspectos, como por ejemplo si la gravedad del impacto que deriva de la decisión final debe tenerse en cuenta para delimitar los derechos que el ciudadano tiene respecto de la actividad preparatoria, lo que va a exigir a su vez distinguir entre los casos en que al tiempo de realizar la actividad preparatoria se conoce el tipo de decisiones finales a que va a servir y aquellos en que las predicciones resultantes de la aplicación de inteligencia artificial son susceptibles de ser utilizadas para usos imprevistos.

Por todo ello, el carácter heterogéneo de lo que hoy denominamos inteligencia artificial remite a una delimitación asimétrica del estatuto jurídico del ciudadano, caracterizada por una ponderación de los riesgos en cuya realización tendrá un papel destacado el proveedor y/o usuario del modelo de inteligencia artificial. No obstante, la generalización del uso de la inteligencia artificial como modelo productivo, como fenómeno general e independientemente de sus modalidades, tiene como presupuesto someter a la persona a una existencia vigilada, circunstancia esta que tiene entidad suficiente para justificar una reflexión sobre qué ajustes son aconsejables en el estatuto jurídico del ciudadano.

II.- LA VIGILANCIA CONSTANTE Y LA CONDICIÓN JURÍDICA DE PERSONA. ¿CONVENIENCIA DE UN DERECHO GENERAL DE LIBERTAD?

Los usos de la inteligencia artificial, en la medida que se vayan confirmando los beneficios que promete, se asentarán en todos los sectores de actividad, sean públicos o

[10] Para un análisis del art 26 RGPD y de la jurisprudencia TJUE respectiva, cfr. Hernández Corchete, J.A., "Corresponsables del tratamiento: atribución de obligaciones, responsabilidad civil y régimen sancionador", en *Privacidad en un mundo global*, Valencia, 2023.

[11] de Hert, P., y Papakonstantinou, V, "The new general data protection regulation: Still a sound system for the protection of individuals". *Computer Law and Security Review*, 32, 2016, 179–194. Sostienen en la p. 184 que "the idea of a single data controller that will carry all liability under data protection law while all other parties to the same processing carry less or no responsibility at all is outdated and lacks an understanding of where technology an lifestyles are headed"

privados. Y la innovación tecnológica perfeccionará estos usos, mejorando su acierto y reduciendo sus riesgos, lo que los convertirá en más atractivos y aumentará su utilización. No es de esperar otra cosa que la consolidación de las tecnologías de inteligencia artificial y, en consecuencia, como uno de sus presupuestos de existencia, de la vigilancia constante de las personas, tanto de su comportamiento en el espacio digital como de su actuación en la realidad analógica por medio del llamado "internet de las cosas". Se necesita una reflexión jurídica dirigida a desvelar si esta circunstancia, siempre o al menos cuando se dan ciertas condiciones, priva al individuo del margen de autonomía suficiente que constituye uno de esos "tramos de los que creemos que la humanidad no se puede separar"[12] y cuyo aseguramiento va indisolublemente unido a la protección de su especial dignidad como ser humano. Y, de ser así, qué derechos individuales y/o colectivos podrían coadyuvar al reequilibrio necesario.

No faltan estudios doctrinales que se han ocupado de llamar la atención sobre los peligros que entraña esta situación de vigilancia constante de la persona humana. Son especialmente ilustrativos los del ZUBOFF[13] que, a partir de una descripción precisa de los orígenes y desarrollo de este fenómeno, muestran como el despliegue de esta vigilancia constante no es casual sino premeditada, y que se orienta a verificar una "apropiación de la experiencia" humana con el fin no solo de predecir el comportamiento futuro de la persona sino, mucho más allá, de condicionarlo de un modo beneficioso para el proveedor de servicios. El propósito declarado de estos modelos de negocio, que se asemejan en apoyarse en la vigilancia constante de las personas, es proveer de mejores servicios al vigilado, pero su objetivo verdadero es limitar su autonomía "guiándola" hacía las decisiones que mejor convienen al prestador de los servicios. También proliferan trabajos de comisiones éticas o jurídicas instituidas en el seno de la Unión Europea o de organizaciones supranacionales de nuestro entorno que, a la vista de los peligros para la autonomía de la persona, reclaman que las tecnologías de inteligencia artificial se desarrollen respetando una serie de valores respetuosos de la dignidad de la persona humana[14].

A pesar de estos estudios doctrinales e informes de comisiones oficiales, las personas directamente afectadas no parecen estar particularmente preocupadas por estos peligros. O bien les pasan desapercibidos por completo, o bien no alcanzan a comprender su gravedad y no tienen inconveniente en asumirlos a cambio de disponer de las ventajas que les ofrecen estas innovaciones tecnológicas. Hay consenso en la doctrina especializada en que, como dice PIÑAR MAÑAS, "la sociedad digital y con ella la innovación que la acompaña pueden generar riesgos … que, precisamente por moverse

[12] Rodotá, S., "Del ser humano al posthumano", en *Sociedad Digital y Derecho*, Madrid 2018, p. 89.

[13] Veáse Zuboff, S. op cit. En la p. 13 se puede leer "Surveillance capitalism originates in this act of digital dispossesion, operationalised in the rendition of human experience as behavioral data". Y en la p. 31, como consecuencia, que "under the regime of instrumentarían power, the mental agency and sel-possession of autonomus human action are gradually submerged beneath a new kind of automaticity: a lived routine of stimulus-response-reinforcement that operates outside of awareness an is aggregated as statistical phenomena: the comings and goings of mere organisms".

[14] Para una amplia referencia a las comisiones instituidas en el seno de la Unión Europea, del Consejo de Europa y de la OCDE, cfr. Martire, D., "Intelligenza artificiale e Stato Costituzionale", *Diritto Pubblico*, 2, maggio-agosto 2022, pp. 402 a 407.

en el entorno digital, no siempre son fáciles de detectar"[15]. En esta misma línea, BECK, y entre nosotros VIDA FERNÁNDEZ, argumentan que los riesgos que se manifiestan en el entorno digital tienen una caracterización especial, en el sentido que no tienen consecuencias físicas, lo que les hace difícilmente reconocibles[16].

Aun cabría hacer una distinción. La vigilancia constante que se produce en el espacio digital y que no presenta una realidad física constituye en sí misma una fuente de peligro para el individuo vigilado. No obstante, esa vigilancia constante tiene como objetivo último incidir en el ámbito de intereses de la persona vigilada, la mayoría de los cuales se manifiestan en el espacio físico. Sin embargo, esta repercusión de la vigilancia constante, a pesar de que tiene una clara expresión en el espacio físico, resulta mayormente inadvertida para el individuo vigilado. Hay decisiones de terceros que le afectan y que no llega a conocer. Otras sí alcanza a conocer pero solo de un modo muy remoto las atribuye a la vigilancia constante de la que es objeto. En fin, hay otras decisiones que el individuo vigilado atribuye a su propia elección sin comprender suficientemente cómo es "guiado" a adoptarlas.

En este contexto, en que las tecnologías de inteligencia artificial se harán más presentes a medida que se perfeccionen y que sus indudables peligros conectados con la vigilancia constante de las personas son difícilmente reconocibles, se hace imprescindible una reconstrucción conceptual que identifique con precisión cuáles son los intereses jurídicos del individuo que resultan afectados, cómo se produce esa afectación y qué derechos es necesario reconocerle para asegurar que la dignidad humana prevalezca.

PIÑAR MAÑAS ha postulado la necesidad de que se reconozca el derecho a la identidad digital de la persona[17]. Destaca en su trabajo que el derecho a la identidad, que no ha de ser entendido meramente por referencia a los medios de identificación de la persona sino que alude al control sobre la "proyección social de la propia personalidad"[18], es uno de los elementos conformadores de la dignidad de la persona humana. Y sostiene que el derecho a la identidad así concebido debe respetarse no solo en la realidad física sino también y con la misma importancia en el entorno digital, lo que exige desarrollar técnicas específicas porque "en el entorno digital las posibilidades de conformar desde fuera del

[15] Piñar Mañas, J.L., "Derecho e innovación. Privacidad y otros derechos en la sociedad digital", *El derecho a la protección de datos personales en la sociedad digital* (Coord. Casas Baamonde), p. 43.

[16] Véase Beck, U, "El riesgo de la libertad digital: un reconocimiento demasiado frágil", *Cuadernos del Mediterráneo,* 22, 2015, pp. 311 a 314. Beck, U., conocido por acuñar la expresión "risk society" en los años 90 del pasado siglo, ha enfatizado que la particularidad que entrañan los riesgos digitales es que, al no conllevar catástrofes físicas que se hacen evidentes para la generalidad de los ciudadanos, son difícilmente reconocibles.

Entre nosotros, en el mismo sentido, cfr. Vida Fernández, J., "The risk of digitalization: Transforming government into a Digital Leviathan", *Indiana Journal of Global Legal Studies*, 30, 1 (Winter 2023), pp. 3 y ss. Sostiene, además, que estos riesgos han aumentado con los últimos desarrollos de las tecnologías digitales y que no se justifica que, con tal de favorecer al máximo la innovación, el espacio digital siga desregulado.

[17] Piñar Mañas, J.L., "Identidad y persona en la sociedad digital", en *Sociedad Digital y Derecho*, Madrid 2018, pp. 94 a 111.

[18] Ibídem, p. 96.

propio sujeto su identidad y con ello su personalidad son sin duda mucho más numerosas, y cualitativamente diversas"[19].

Esta propuesta doctrinal parece haber tenido acogida en la Carta de Derechos Digitales adoptada por el Gobierno de España[20]. Su artículo II lleva por rúbrica "Derecho a la identidad en el entorno digital" y en el apartado primero se declara expresamente que "el derecho a la propia identidad es exigible en el entorno digital" y se precisa a continuación que "esta identidad vendrá determinada por el nombre y por los demás elementos que la configuran de acuerdo con el ordenamiento jurídico nacional, europeo e internacional", de donde se desprende que el reconocimiento de este derecho va más allá de la protección del nombre e identificación de la persona.

En definitiva, lo que está en juego es un concepto de privacidad dinámica que asegura a la persona como elemento esencial de su dignidad un espacio autónomo de decisión a partir del cual pueda conformar en libertad sus opciones vitales más básicas y las relaciones que entabla con otros sujetos. Esto que podríamos designar como "derecho general de libertad", que nuestra Constitución no erige en derecho autónomo pero que tiene presente al consagrar el libre desarrollo de la personalidad del individuo en el artículo 10.1 CE, resulta afectado por la arquitectura que sustenta el uso eficaz de las tecnologías de inteligencia artificial al menos de tres maneras: a) la vigilancia constante de la persona puede inhibir comportamientos que de otro modo habrían tenido lugar; b) la información masiva derivada de esa vigilancia constante permite a los terceros que la controlen condicionar las opciones que hacen las personas, reduciendo así su espacio autónomo de decisión y, en fin, c) al generalizarse la previsión de conductas futuras "se corre el riesgo de que la persona pueda ser evaluada por sus propensiones y no por sus acciones"[21].

Algunas sentencias de tribunales europeos han abordado esta cuestión de la relación de las aplicaciones de inteligencia artificial predictiva con el ámbito privado de los ciudadanos[22]. Me quiero referir a dos que terminan en fallo de anulación, con el fin de exponer que los fundamentos en que descansan se relacionan directamente con ese ámbito

[19] Ibídem, p. 101, donde argumenta que "la identidad a que vengo refiriéndome se construye fundamentalmente en el entorno de la realidad física. Pero puede asimismo trasladarse al entorno digital. En éste, en efecto, confluyen elementos que configuran tanto la identidad que cada uno quiere o pretende darse como la que se otorga. Lo que ocurre es que en el entorno digital la heteroformación de la identidad depende de factores que no siempre operan en el mundo físico o lo hacen de un modo muy diverso. Pues en el entorno digital las posibilidades de conformar desde fuera del propio sujeto su identidad y con ello su personalidad son sin duda mucho más numerosas, y cualitativamente diversas".

[20] La Carta de Derechos Digitales es un instrumento de soft law que, elaborada por un grupo de expertos presidido por el profesor De la Quadra-Salcedo, fue adoptada por el Gobierno de España y presentada en julio de 2021 por el Presidente del Gobierno.

[21] Véase Pérez Luño, A.E., "Las generaciones de derechos humanos ante el desafío posthumanista", en *Sociedad Digital y Derecho*, Madrid 2018, p. 152.

[22] Aparte de las sentencias que se reseñan a continuación, es importante la jurisprudencia italiana (cfr. Civitarese Mateucci, S, "Umano troppo umano, Decisioni amministrative automatizzate e principio de legalitá", Diritto Pubblico, 1, 2019, pp. 27 a 34) y francesa (cfr. Rodríguez Pontón, F.J., "El régimen jurídico del uso de los sistemas de inteligencia artificial en el Derecho público francés. Fuentes. respuestas y debates", *Revista General de Derecho Administrativo*, 63, 2023).

privado de la persona que es parte integrante de su dignidad. Se apoyan, como se argumentará, en una especie de "derecho general de libertad"[23].

De un lado, la Sentencia de 5 de febrero de 2020 dictada por la Corte de Distrito de la Haya[24], que analiza y termina prohibiendo un programa que a partir de correlaciones desveladas mediante el uso de tecnologías de inteligencia artificial establecía previsiones sobre qué ciudadanos tenían más probabilidades de incurrir en fraudes relacionados principalmente con ayudas sociales, con el fin último de hacerles objetivo prioritario de las actividades de inspección. El órgano judicial, sin dejar de reconocer que la persecución del fraude fiscal es una finalidad que legítima intervenciones que restrinjan los derechos de los ciudadanos, argumenta que el mecanismo utilizado no contiene las salvaguardas necesarias y que ello determina que constituya una restricción desproporcionadas del derecho a la vida privada protegido por el artículo 8 CEDH.

No es extraño que esta sentencia haya recurrido a este fundamento jurídico, sobre todo si tenemos en cuenta el papel que el TEDH le está haciendo jugar como derecho que permite amparar aquellas exigencias de autonomía individual que, por no ser perceptibles al momento de su adopción en 1950, el Convenio no protege mediante derechos específicos. Es este artículo 8 y la ductilidad de la noción "vida privada" lo que ha convertido al CEDH en un instrumento vivo que ha logrado dar respuesta a los nuevos desafíos para la persona que ha traído consigo la evolución de la sociedad, tanto los retos que se manifiestan en las relaciones personales como los que trae consigo el progreso tecnológico y en especial el que tiene lugar en un cada vez más complejo espacio digital.

Respecto a la primera cuestión, y con motivo de determinar el ámbito de autonomía individual que ampara a las personas transexuales y limita correlativamente las imposiciones que se pueden prever en la leyes del Estado en aras a garantizar otros intereses relevantes, el TEDH ha afirmado que el artículo 8 CEDH "protege el derecho al desarrollo personal y el derecho a establecer y consolidar relaciones con otros seres humanos y con el entorno que le rodea" y que "el Tribunal considera que el concepto de autonomía personal es un importante principio que debe informar la interpretación de las garantías del artículo 8"[25].

En cuanto a lo segundo, la posible utilización de los progresos tecnológicos que conlleve consecuencias perjudiciales para el individuo también se ha reconducido por el TEDH a la noción de "vida privada" del art. 8 CE. Ya en la sentencia S. y Marper c. Reino Unido [GC) de 4 de diciembre de 2008, en relación al uso de bases de datos biométricas para la persecución de los delitos, el TEDH resolvió que "el artículo 8 del Convenio se

[23] Nótese que la expresión "derecho general de libertad" se utiliza aquí de un modo que no es equivalente al "principio general de libertad". Este último alude a la vinculación negativa de la persona a la ley, de modo que forma parte de su libertad de acción todo lo que no esté prohibido en la ley. El derecho general de libertad", por el contrario, aludiría al ámbito esencial de autonomía personal (integrado por distintas manifestaciones) y que como tal es un límite para el legislador.

[24] Cotino Hueso, L. "'SyRI, ¿a quién sanciono?' Garantías frente al uso de inteligencia artificial y decisiones automatizadas en el sector público y la sentencia holandesa de febrero de 2020", *La Ley Privacidad*, Wolters Kluwer nº 4, mayo 2020. El autor expone que la utilización del artículo 8 CEDH como parámetro de análisis puede deberse a las particularidades del sistema holandés, que no contempla control de constitucionalidad de las leyes pero sí admite examinar su compatibilidad con los tratados internacionales.

[25] Véase, en relación a los derechos de la personas transexuales, la STEDH de 10 de marzo de 2015, *Y.Y. c. Turquía*, § 57, y la doctrina a la que se remite.

debilitaría de forma inaceptable si se autorizase a cualquier precio el uso de las técnicas científicas modernas en el sistema judicial penal, sin sopesar cuidadosamente los beneficios que pudieran resultar de un amplio recurso a estas técnicas, de un lado, y los intereses esenciales relacionados con la protección de la vida privada, de otro". Y, en este contexto, el TEDH siempre que se le ha requerido que evalúe la afectación que sufre una persona por las nuevas posibilidades tecnológicas de tratamiento masivo de datos personales lo ha hecho desde la perspectiva de esta concepción amplia del derecho a la "vida privada" protegida por el artículo 8 CEDH[26].

Me ocuparé, en segundo lugar, de la sentencia de 16 de febrero de 2023 del Tribunal Constitucional Federal Alemán[27]. Examina sendas leyes de Hesse y de Hamburgo que articulan el uso de aplicaciones automatizadas que a partir del análisis de conjuntos masivos de datos realizan previsiones orientadas a prevenir delitos y otros peligros para los intereses públicos. La sentencia las declara inconstitucionales porque las salvaguardas que establecen no son suficientes para considerar que el sacrificio de los derechos de las personas afectadas es proporcionado[28].

Lo que me interesa en este momento es poner el foco sobre qué dice la sentencia acerca de cuál es la afectación que deriva del uso de estas aplicaciones de inteligencia artificial y cuál es el derecho de la persona que resulta afectado. El tribunal alemán razona, respecto de lo primero, que la restricción de los derechos de la persona derivan de la procedencia de los datos personales utilizados y, además, de un modo independiente que requiere examen específico de proporcionalidad, del tratamiento automatizado que se hace de esos datos para elaborar previsiones de conducta. En cuanto a lo segundo, la sentencia parte de que resulta afectado el derecho al libre desarrollo de la personalidad (artículo 2.1[29]) en conexión con el carácter intangible de la dignidad humana (artículo 1[30]), con lo que no hace sino reiterar un enfoque ya bien conocido. En efecto, el Tribunal Constitucional Federal Alemán lleva décadas apreciando que en la combinación de los artículos 2.1 y 1 de la Ley Fundamental de Bonn supone el reconocimiento de un derecho general de libertad de la persona y lo ha utilizado para amparar en él aquellas exigencias de su autonomía individual que no estaban de otro modo protegidas, prestando así al texto constitucional alemán el dinamismo y adaptabilidad que son necesarias para los nuevos

[26] Véanse, sin afán de exhaustividad, Satakunnan Markkinapörssi Oy y Satamedia Oy c. Finlandia, 27 junio 2017, § 129 y ss; M.L. y W.W. c. Alemania, nos. 60798/10 and 65599/10, 28 Junio 2018, § 116; y Biancardi c. Italia no. 77419/16, 25 Noviembre 2021, § 48.
[27] 1 BvR 1547/19 y 1 BvR 2634/20.
[28] Véase, para un análisis detallado en español, Cotino Hueso, L, "Una regulación legal y de calidad para los análisis automatizados de datos o con inteligencia artificial. Los altos estándares que exigen el tribunal constitucional alemán y otros tribunales, que no se cumplen ni de lejos en España", *Revista General de Derecho Administrativo*, 63, 2023.
[29] El artículo 2.1 prevé que «Toda persona tiene el derecho al libre desarrollo de su personalidad siempre que no viole los derechos de otros ni atente contra el orden constitucional o la ley moral»
[30] El artículo 1 dispone que «La dignidad humana es intangible. Respetarla y protegerla es obligación de todo poder público»;

desafíos que han ido surgiendo para la identidad de la persona, tanto en la realidad física[31] como en el espacio digital[32].

Se desprende de lo anteriormente expuesto una primera conclusión, el TEDH, con la noción de "vida privada" del artículo 8 CEDH, y el TCFA, con la combinación de los artículos 1 y 2.1 de la Ley Fundamental de Bonn, han encontrado el modo de afirmar un derecho general de libre desarrollo de la personalidad que permite proteger, mediante un análisis fundado en el principio de proporcionalidad, las distintas formas que puede adoptar la autonomía individual como elemento esencial a la dignidad humana, designemos ese espacio de libre decisión como identidad, privacidad o con otra expresión semánticamente parecida. Un derecho como este, por su capacidad de adaptación a las exigencias de autonomía que se puedan presentar, ofrece un estatuto de protección adecuado a la persona respecto de los nuevos retos asociados a los desarrollos de inteligencia artificial que la innovación tecnológica traiga consigo.

La Constitución Española, al menos tal y como se viene interpretando de un modo pacífico, opone algunas aristas serias al reconocimiento de un derecho general de libertad. Por un lado, aunque el libre desarrollo de la personalidad está expresamente contemplado en el artículo 10.1 CE, solo se reconoce como un principio general del Derecho y como tal condiciona la interpretación de las normas y en especial de los derechos afirmados por el constituyente, pero no es fuente a partir de la cual identificar otros ámbitos concretos de libertad a favor de los ciudadanos y frente al legislador[33]. Por otro lado, su art. 10.2 CE contiene un mecanismo de apertura a los tratados internacionales de derechos humanos y la jurisprudencia que los interprete, pero su eficacia se ciñe a la delimitación del contenido y alcance de los derechos fundamentales ya reconocidos por la propia Constitución, negándose que la apertura que representa permita la "creación" de nuevos derechos[34]. Este criterio interpretativo se ha reflejado muy especialmente en las reticencias del Tribunal Constitucional español a incorporar como derechos fundamentales algunas de las dimensiones de autonomía individual que la jurisprudencia

[31] El Tribunal Constitucional Federal Alemán ha tenido ocasión de pronunciarse varias veces acerca de situaciones de limitación de su autonomía de acción impuestas por distintas leyes a las personas transexuales (sentencias 16 de marzo de 1982, 26 de enero de 1993, 6 de diciembre de 2005, la de 18 de julio de 2006, la de 27 de mayo de 2008, de 28 de enero de 2011 y 10 de octubre de 2017). En todos estos casos el marco básico de análisis ha sido el derecho general a la propia personalidad derivado de la conjunción de los arts. 1.1 y 2.1 GG.

[32] Como es conocido, el TCFA afirmó en la sentencia que dictó su Primera Sala el 15 de diciembre de 1983 [BVerfGE 65, 1 (Censo de Población)] el derecho a la autodeterminación informativa, que derivó una vez más de la conexión entre los arts. 2.1 y 1 de la Ley Fundamental de Bonn. De ahí en adelante y hasta la sentencia que comentamos de febrero de 2013 el parámetro constitucional de referencia de los asuntos que implican tratamiento automatizado de datos ha sido constantemente el mismo.

[33] Véase Jiménez Campo, J., "Art 10.1 CE", *Comentarios a la Constitución Española* (Dir. Casas Baamonde y Rodríguez-Piñero), p. 179, donde razona que el art. 10.1 CE contiene verdadero Derecho, si bien que "el Derecho del art. 10.1 CE es, en todo caso, un Derecho de principios" y, añade, "que ante un elenco de principios, cada uno carente, por su generalidad y abstracción, de supuesto de hecho definible es lo que impide presentar a cualquiera de ellos, en ningún sentido significativo, como derecho fundamental".

[34] Véase Saiz Arnaiz, A., "La interpretación de los derechos fundamentales y de los tratados internacionales sobre derechos humanos", en *Comentarios a la Constitución Española* (Dir. Casas Baamonde y Rodríguez-Piñero), pp. 206 y 207.

de Estrasburgo ha desprendido de la amplia noción de "vida privada" ex articulo 8 CEDH[35].

Con esta interpretación de los apartados 1 y 2 del artículo 10 CE lo que se viene a asegurar es la posición del constituyente como la única fuente de derechos fundamentales que se puedan oponer frente al legislador. En efecto, admitir que un sistema constitucional contiene un derecho general de libertad que comprende todas aquellas manifestaciones centrales de la dignidad humana, aunque no esté prevista su protección de un modo expreso en ninguna norma, supone necesariamente aceptar que el órgano jurisdiccional que interprete el alcance de esa cláusula tiene en sus manos el cometido de configurar nuevos ámbitos concretos de libertad como derechos fundamentales que el legislador solo podrá restringir por razones legítimas y con un alcance proporcionado.

Que en España falte el reconocimiento constitucional de un derecho general de libertad se ha venido salvando, por lo que hace a los nuevos retos para la autonomía individual que plantea el tratamiento masivo de datos en el espacio digital, acudiendo a que el artículo 18.4 CE prevé que "la ley limitará el uso de la informática" y desprendiendo de esa previsión tan general un derecho a la protección de datos de carácter personal susceptible de ser interpretado extensivamente a partir de la normativa y jurisprudencia europea que lo desarrolla.

Este derecho a la protección de datos personales ha cumplido con un cierto éxito una función instrumental de garantía de la libertad de las personas en la sociedad digital. El espacio digital y su dinámica de funcionamiento ha traído consigo formas de poder desconocidas previamente que someten a los individuos en maneras sustancialmente diferentes, y frente a ello el ordenamiento jurídico no estaba adecuadamente pertrechado, no contenía derechos eficaces. La primera reacción podríamos calificarla de indirecta o instrumental, pues partiendo del carácter personal de muchos de los datos que están en la base del ejercicio de estos poderes se han dispuesto una serie de principios a los que debe sujetarse su tratamiento y se han otorgado a sus titulares un conjunto de derechos frente a dicho tratamiento, todo ello con la finalidad última de limitar los poderes incontrolados surgidos en el espacio digital y garantizar correlativamente ciertos espacios de autonomía individual a las personas. En definitiva, se imponen límites al modo de tratar los datos personales en el convencimiento de que con ello se aseguran ciertos espacios de libertad a los individuos. Este derecho, sin perjuicio de reconocer abiertamente que ha realizado con cierto éxito la función instrumental que se le había encomendado, no deja de ser una aproximación indirecta respecto de las amenazas para la autonomía individual que entraña el espacio digital. Para afrontar los retos que la inteligencia artificial plantea y planteará en el futuro, como se ha puesto de relieve por algunos autores[36], convendría sumar perspectivas adicionales, que en mi opinión tiene que partir de enfocarlos de un modo directo, en la medida que ello sea posible.

[35] Por ejemplo, SSTC 186/13, de 4 de noviembre, y 99/2019, de 18 de julio

[36] Yuste, R. y De la Quadra-Salcedo, T., "Neurorights and New Charts of Digital Rights: A Dialogue beyond the Limits of the Law", *Indiana Journal of Global Legal Studies*, 30, 1 (Winter 2023), pp. 27 y ss.

III.- LOS RIESGOS ESPECÍFICOS DE LA INTELIGENCIA ARTIFICIAL Y LOS DERECHOS PROCEDIMENTALES COMO RESPUESTA COMÚN

III.1.- La conexión entre inteligencia artificial y decisión individual automatizada

Las tecnologías de inteligencia artificial, además de contribuir a una concreta arquitectura general del espacio digital que condiciona de un modo importante algunas dimensiones de la autonomía individual, conllevan riesgos específicos para las personas que van ligados a que las decisiones que les afectan adquieren una forma automatizada o se ven influidas de un modo determinante por el resultado de un tratamiento automatizado. En esencia, la inteligencia artificial, al menos en su modalidad predictiva, consiste en un tratamiento automatizado de datos que arroja una previsión probable de una conducta futura. Estas operaciones no necesariamente constituyen en sí mismas decisión alguna, pero las pautas que mediante ellas se logra desvelar están destinadas a informar decisiones sobre personas en las que esa precisa conducta es relevante. Aunque el otorgamiento o denegación de un crédito puede ser resuelto por un empleado de la entidad financiera, lo importante es si se limita a trasladar a las circunstancias del caso los patrones identificado previamente mediante el tratamiento automatizado de datos en que consiste la inteligencia artificial.

En definitiva, las decisiones últimas que afectan a las personas, aun en el caso de que sean adoptadas por humanos, están determinadas por la aplicación de inteligencia artificial y en este sentido se puede hablar de decisiones automatizadas. Este concepto relativamente amplio de decisión individual automatizada, que hace bien pocas semanas ha sido respaldado por el TJUE con ocasión de delimitar el alcance de ámbito objetivo del artículo 22 RGPD[37], es el que vamos adoptar en este trabajo.

[37] Wachter S, Mittelstad B, Floridi L, "Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation". *International Data Privacy Law*, 2017, 7, n°2, p. 92, recuerdan que las decisiones automatizadas, mientras en los trabajos preparatorios se definen como las basadas *predominantemente* en el tratamiento automatizado de datos, la redacción final las identifica como las basadas *únicamente* en el tratamiento automatizado, sugiriendo que la mera intervención humana, por intrascendente que fuera, impide la calificación de decisión automatizada. En España, también defiende un entendimiento estricto del alcance objetivo de la noción de decisión automatizada Huergo Lora, A., "Una aproximación a los algoritmos desde el derecho administrativo", en *La regulación de los algoritmos*, Aranzadi, 2020, epígrafe III.6.
Esta cuestión ha sido el centro del litigio en el asunto SCHUFA Holding (Scoring), C-634/21, donde entra en juego una agencia de información comercial, que genera automatizadamente una calificación crediticia (scoring), y una entidad crediticia que con intervención humana accede o rechaza un solicitud de crédito por una persona física. Recogiendo el parecer de las Conclusiones del AG Pikamäe (apartados §§ 44 a 52), la Sentencia TJUE de 7 de diciembre de 2023 afirma que "en circunstancias como las del litigio principal, en las que el valor de probabilidad generado por una agencia de información comercial y comunicado a un banco desempeña un papel determinante en la concesión de un crédito, la generación de dicho valor propiamente dicha debe calificarse como decisión que produce «efectos jurídicos» en un interesado o que «[lo afecta] significativamente de modo similar», en el sentido del artículo 22, apartado 1, del RGPD" (§ 50). El TJUE, separándose de la lectura literal que proponía la doctrina científica citada al principio de esta nota, ha optado por una interpretación relativamente amplia, que busca que el precepto cumpla en la realidad el propósito a que sirve.

III.2.- Los riesgos específicos asociados a las decisiones individuales automatizadas

Las decisiones automatizadas, entendidas en el sentido que hemos expuesto, presentan riesgos para las personas afectadas porque puede adoptarse a partir de patrones erróneos o que, sin serlo, reflejen factores prohibidos por el ordenamiento por ser sospechosos de generar o consolidar supuestos de discriminación[38].

Todos los sistemas de adopción de decisiones son susceptibles de cometer errores. Lo específico de las decisiones automatizadas es que pueden ser groseros y, además, pasar desapercibidos. La primera fuente de estos errores puede estar en los conjuntos de datos con los que se entrena la aplicación de inteligencia artificial. Si estos datos no son relevantes para el objeto de que se trate, o simplemente cuando sean insuficientes por contemplar solo situaciones generales y no otras que presenten especialidades, las pautas identificadas como útiles serán equivocadas y conducirán a decisiones injustificadas. Pero el mayor peligro de errores deriva de que las aplicaciones de inteligencia artificial aprenden a partir de correlaciones que están presentes en los datos analizados y la máquina no distingue si van referidas al objeto que se quiere valorar o, por el contrario, aunque comúnmente están presentes en su entorno son colaterales a él. Un modo de razonar así puede enfocar por mera casualidad patrones de decisión que sean útiles en muchos casos, generando una sensación de confianza, pero que al tratar de generalizarse conduzcan a errores, lo que ocurrirá en aquellos casos en que esos elementos colaterales no estén presentes[39].

El otro peligro específico es que los tratamientos automatizados identifiquen pautas relevantes conectadas con factores que el ordenamiento considera sospechosos de entrañar una discriminación prohibida. La problemática más evidente surge cuando entre los datos utilizados por el modelo de inteligencia artificial se incluyen los pertenecientes a categorías especialmente sensibles y por ello esta práctica está prohibida por la normativa de protección de datos personales, actualmente en el apartado 4 del art. 22 RGPD, que solo admite como excepciones el consentimiento explícito y la persecución de intereses públicos[40]. Esta regulación no elimina por completo este peligro, pues la propia aplicación de inteligencia artificial, al aprender a partir de los datos que analiza, incluso de los que aparentemente no tienen ninguna relación con factores sospechosos de discriminación, realiza inferencias y a través de ellas perpetúa las estructuras de desigualdad que subyacen en los datos analizados. De este modo, los modelos de inteligencia artificial pueden sacar a la luz pautas mediatizadas por los factores de

[38] Sobre la identificación de los ámbitos en que la inteligencia artificial puede generar daños que justifican una regulación específica, Vease Huergo Lora, A, op. cit. Epígrafe III.6. Y también el Considerando 71 del RGPD.

[39] Grieman K. y Early J, "A Risk-based Approach to AI Regulation: System Categorisation and Explainable AI Practices", *Scripted*, 20, 1, Febrero 2023, pp. 63, 64 y 77 ss.

[40] En la doctrina se argumenta con razón que el uso por modelos de inteligencia artificial de categorías de datos especialmente sensibles puede ser necesario precisamente para evitar resultados discriminatorios, por ejemplo para auditar el funcionamiento de dichos modelos. Se discute, sin embargo, si ese uso debe restringirse a los supuestos de excepción de los arts. 9.2 y 22.4 RGPD o conviene una nueva excepción como la prevista en el art. 10.5 de la Propuesta de RUIA. Cfr. Van Bekkum M. y Borgesius, F. "Using sensitive data to prevent discrimination by artificial intelligence: Does the GDPR need a new exception?", *Computer Law & Security Review*, 48, Abril 2023.

discriminación presentes en las experiencias anteriores y postularlas como la mejor ratio de futuras decisiones.

III.3.- Los derechos procedimentales como respuesta

La protección del ciudadano frente a ambos riesgos pasa por poner en su mano derechos de carácter procedimental, dirigidos en especial a permitirle disfrutar de una vía efectiva en la que puedan contestar el contenido de las decisiones automatizadas y lograr que sean revisadas[41]. Al derecho a contestar las decisiones automatizadas se le ha dado en ocasiones un alcance muy recortado, de carácter meramente formal, que se satisfacía con permitir al interesado instar una nueva decisión[42]. En mi opinión, por el contrario, el derecho a contestar la decisión automatizada necesariamente ha de revestirse de una dimensión material, no basta con poder exigir una nueva decisión sino que debe alcanzar a someter a crítica la decisión inicial.

III.3.A.- Derecho a ser informado de la lógica del tratamiento o derecho a obtener una explicación

El punto partida debería ser necesariamente el reconocimiento de un derecho a conocer los motivos que han guiado la decisión automatizada, sin lo cual es absolutamente imposible articular una verdadera contestación. Y aquí lo que se viene resaltando como la caracterización misma de la inteligencia artificial se erige en un obstáculo. Las formas de inteligencia artificial más potentes y que prometen previsiones más seguras, precisamente porque combinan una enorme cantidad de datos y aprenden autónomamente con sus distintos usos y con los nuevos datos que van incorporando de distintas procedencias, son opacas, en el sentido que ni siquiera los tecnólogos que las configuran están en condiciones de explicar el modo en que conforman las previsiones que arrojan. De hecho, la doctrina (incluso la reciente) afirma con rotundidad que el avance en términos de eficacia que llevan aparejados estos modelos de inteligencia artificial tiene el alto precio de la opacidad[43].

La perspectiva que procede adoptar ha de ser, en mi opinión, una diferente. Desde el ángulo que ahora interesa, que no incluye todavía el juego que tiene el principio de legalidad en la actuación de los poderes públicos, lo importante no es poder determinar a priori cómo la aplicación de inteligencia artificial da lugar a la previsión sino tener constancia a posteriori del contenido de la misma y en una medida razonable del peso que atribuye a los distintos factores relevantes. Esta última información es la que coloca al ciudadano afectado por una decisión concreta en posición de poder contestarla,

---

[41] Citron D., "Technological Due Process", 85 *Wash. U. L. Rev*, 85, 2007, p.1256; Selbst A. y Barocas S., "The Intuitive Appeal of Explainable Machines", *Fordham L. Rev*. 87, 2018, pp. 1092–94.

[42] Wachter S, Mittelstad B, Floridi L, op. cit., p. 91, por referencia al art. 12.2.b) de la UK Data Protection Act 1998, mediante la que se transponía al Derecho inglés la Directiva 95/46/CE, cuyo art. 15 ya contenía un régimen específico sobre las decisiones individuales automatizadas.

[43] Martire, D, "Intelligenza artificiale e Stato costituzionale", *Diritto pubblico*, 2/2022, pp. 401-402, concluye que "L'abbandono della programmazione meramente lógico-deduttiva ha richiesto, tuttavia, un prezzo…. Un passo ulteriore verso l'efficienza, al prezzo, tuttavia, di una minore trasparenza".

demostrando que su fundamento no se corresponde con la realidad o se apoya en factores prohibidos por discriminatorios.

La perspectiva *ex ante* de cómo funciona el algoritmo y la *ex post* de cómo ha actuado en un caso concreto distan mucho desde la óptica que nos ocupa, la de los condicionantes que derivan de la caracterización autónoma e impredecible de las formas más complejas de inteligencia artificial. Que a priori sea incontrolable la operatividad de un algoritmo complejo, que aprende autónomamente y está en constante adaptación a una realidad que le proporciona nuevos datos, no significa que, una vez actuado en un caso concreto, la previsión ofrecida y los factores que la determinan sean necesariamente opacos. Basta con imponer la obligación de que los modelos de inteligencia artificial se configuren y se usen bajo ciertas exigencias de trazabilidad para que quede constancia de cada uno de sus usos concretos y con apoyo en esa constancia quepa no solo exponer la decisión sino también explicarla por referencia a los factores principales tomados en consideración y al peso que le ha sido atribuido.

Ciertas obligaciones de trazabilidad de los modelos de inteligencia artificial ya aparecen en la Propuesta de REIA al menos en relación a aquellos que en ella se clasifican de alto riesgo[44]. Y la literatura científica da cuenta de los intentos de hacer explicable los usos de la inteligencia artificial[45]. Cabría esperar, por tanto, que la falta de explicabilidad de los modelos de inteligencia artificial, en el sentido de conocimiento *a posteriori* de los motivos que sustentan cada uno de sus usos concretos, es fundamentalmente un problema de desarrollo de esta tecnología, provocado porque está aún en una fase incipiente de instalación y que irá desapareciendo a medida que se vaya perfeccionando[46], proceso de desarrollo que también estará caracterizado por la exigencia de adaptación a aquellos requerimientos que el Derecho le imponga para que en su uso prevalezca un enfoque antropocéntrico, de suerte que los valores de eficacia y progreso vayan acompañados del respeto de la dignidad de la persona y de los derechos fundamentales que de ella se desprenden[47].

Me he referido a la doble perspectiva *ex ante* y *ex post* de la transparencia del algoritmo para destacar que, mientras la primera está realmente afectada por los problemas de

[44] El art. 12 de la Propuesta de RUIA se refiere a la obligación de que los sistemas de inteligencia artificial de alto riesgo se diseñen con la capacidad de llevar registro de todas sus actividades. Y en este contexto su apartado 2 impone se alcance un "nivel de trazabilidad del funcionamiento del sistema durante todo su ciclo de vida que sea adecuado a su finalidad prevista". Esta trazabilidad va referida en general al funcionamiento del sistema, pero supone reconocer que a posteriori sí es posible conocer cómo actuó el sistema. A posteriori no hay opacidad porque hay constancia de lo ocurrido.

[45] Grieman K. y Early J, op. cit. 77 y ss.

[46] Rivero Ortega, R., "Algoritmos, inteligencia artificial y policía predictiva del Estado vigilante", *Revista General de Derecho Administrativo,* 62, 2023, estima que las cuestiones que están siendo objeto de crítica no son otra cosa que "un efecto secundario de una gran transformación tecnológica y social, no calibrada ni anticipada por los promotores del empleo creciente y exitoso de la inteligencia artificial". Esta falta de previsión "es la que produce perplejidad, pero no debería conducir a la parálisis".

[47] Martire, D. op. cit., 407, se refiere "all'esigenza di condizionare ed ancorare lo sviluppo alla salvaguardia de la dignità umana" y concluye que "la oramai riconosciuta capacità della tecnología de incidire- anche in negativo- sull'essere umano richiede … un intervento condizionante in grado di modificare l'essere in vista del raggiungimento di finalità – anche e sopratutto costituzionale-: un dover-essere …", que cifra en la "questione della eventuale codificazione di nuovi diritti digitali".

En España, Vida Fernández, J., op. cit., p. 7, señala "Digital transformation has so far taken place under the principle of freedom (laissez-faire) letting innovation unfold without limits" y aboga por abandonar por una cierta intervención pública ("Leaving the Digital Laissez-Faire Era")

opacidad, la segunda no, pues los mecanismos de trazabilidad pueden dejar constancia de los factores determinantes de la previsión que arroja. Sin embargo, esta doble perspectiva sirve también a un segundo propósito, que es el de precisar qué información es necesaria para que el ciudadano pueda contestar una decisión automatizada. En la doctrina científica especializada en protección de datos personales se diferencia, en referencia a las decisiones automatizadas, el "derecho a ser informado" (*right to be informed*) del "derecho a una explicación" (*right to an explanation*). El primero opera *ex ante* y atañe a una noticia general y abstracta de la lógica y significación del sistema automatizado. El segundo actúa *ex post* y su objeto son los factores que han sido relevantes para que una aplicación de ese sistema en un caso concreto arroje un cierto resultado predictivo, con consideración expresa de las circunstancias específicas del caso[48]. Evidentemente, solo si el ordenamiento jurídico reconociera este último derecho los ciudadanos estarían en posición de poder contestar de un modo efectivo las decisiones automatizadas que les afecten, aportando razones de porqué tales decisiones son incorrectas o discriminatorias y deben ser sustituidas. Por el contrario, afirmar el derecho a que se comunique cómo funciona abstractamente un sistema automatizado, aun cuando alcance un importante grado de detalle, tendrá su utilidad desde el punto de vista de darle al titular de los datos personales un mayor control sobre ellos, pero siempre será inadecuado a los efectos de permitirle contestar decisiones concretas en que se traduzca el funcionamiento del sistema automatizado, que es la vía más segura de sacar a la luz que revisten errores o que se fundan en criterios discriminatorios[49].

Hasta el momento la regulación más directa de las decisiones automatizadas y de los derechos del ciudadano respecto de ellas se encuentra en la normativa europea de protección de datos personales. Aun cuando se trata de un sector regulatorio muy acotado y ello condiciona el alcance y la interpretación de estos derechos, lo importante es que el Considerando 71 del RGPD, con ocasión de justificar las razones por la que las decisiones automatizadas se sujetan a una regulación específica, las cifra precisamente a que están especialmente predispuestas a incurrir en errores y resultados discriminatorios que dañen los derechos de las personas destinatarias[50]. Como he sugerido ya antes, el derecho a la protección de datos personales tiene un destacado componente instrumental, en el sentido de que actuando sobre los datos personales como elemento imprescindible del espacio digital apunta a crear derechos de los ciudadanos en dicho espacio.

Como es bien sabido, el RGPD prohíbe de partida las decisiones automatizadas. En aquellos casos limitados en que, por excepción a la prohibición general, el RGPD las admite quedan sujetas a reglas especiales de información y acceso previstas en los

[48] Wachter S, Mittelstad B, Floridi L, op. cit., p. 77 y 78.

[49] En la doctrina está extendido un debate diferente, que centra la transparencia de las decisiones automatizadas en el acceso al código fuente del algoritmo. Cfr., Boix Palop A. y Soriano Arnanz A., "Transparencia y control del uso de la inteligencia artificial por las Administraciones Públicas", en *Derecho Público de la Inteligencia Artificial*, Zaragoza, 2023.

[50] Afirma este Considerando 71 que las obligaciones específicas del responsable del tratamiento en relación a las decisiones automatizadas se le imponen "para garantizar, en particular, que se corrigen los factores que introducen inexactitudes en los datos personales y se reduce al máximo el riesgo de error" y para "asegurar los datos personales de forma que …. se impidan, entre otras cosas, efectos discriminatorios en las personas físicas por motivos de raza u origen étnico, opiniones políticas, religión o creencias, afiliación sindical, condición genética o estado de salud u orientación sexual, o que den lugar a medidas que produzcan tal efecto".

artículos 13 a 15 RGPD que, por ser obligaciones del responsable del tratamiento que actúan *ex ante*, no pueden contener un derecho a explicación de las aplicaciones concretas del algoritmo. El artículo 22 RGPD sí alude a la decisión individual resultante del tratamiento automatizado, pero no es rotundo en el reconocimiento de un derecho a la explicación de la aplicación concreta del algoritmo. Se limita a exigir que respecto de las decisiones automatizadas que se amparen en las letras a (necesarias para la celebración o ejecución de un contrato) y c (consentimiento explícito del titular de los datos) de su apartado 2 se adopten salvaguardas adecuadas, "como mínimo el derecho a obtener intervención humana por parte del responsable, a expresar su punto de vista y a impugnar la decisión". Por su parte, el apartado b) del artículo 22.2 exceptúa de la prohibición general las decisiones automatizadas autorizadas por Derecho de la Unión o de un Estado Miembro, exigiendo a estas normas que contemplen "medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado", sin precisar cuáles.

En definitiva, la regulación indeterminada que el RGPD hace de la transparencia *ex post* de las decisiones automatizadas deja notablemente abierta la cuestión. Permite lecturas restrictivas, que apuntan a un cierto pragmatismo del legislador comunitario, que habría eliminado en la redacción final del artículo 22 la referencia explícita al derecho a una explicación para facilitar la innovación tecnológica en materia de inteligencia artificial[51]. También admite que el TJUE, a través de una interpretación que tengan en cuenta la finalidad del precepto (hacer posible la contestación de la decisión) y que traiga en su apoyo un Considerando 71 que es mucho más claro (derecho "a recibir una explicación de la decisión tomada después de tal evaluación") y, afirme que el artículo 22 otorga un verdadero derecho a la explicación de las decisiones automatizadas. Y entretanto, aprovechando esta indefinición y aunque la naturaleza de Reglamento UE del RGPD puede oponerse como un obstáculo, hay algunos legisladores nacionales que han optado por incluir en sus normas de adaptación al RGPD un verdadero derecho a la explicación de las decisiones automatizadas.

Esto último es lo que sucede con el artículo 21 de la ley que adapta el ordenamiento francés al RGPD, por el que se da nueva redacción al artículo 10 de la Ley nº 78-17 de 6

[51] Wachter S, Mittelstad B, Floridi L, op. cit., 78 a 82, dan cuenta de que, mientras el Parlamento Europeo había añadido al texto del que entonces era el art. 20 el "derecho a recibir una explicación de la decisión tomada después de tal evaluación", en el texto definitivo del RGPD ese derecho se eliminó del articulado (no consta en su art. 22) y solo se mantuvo en el Considerando 71. Este argumento les lleva a sostener que el RGPD no consagra un derecho a la explicación a posteriori de decisiones concretas y solo, con apoyo en los arts. 13 a 15, un derecho a ser informado de un modo abstracto y general de la lógica y consecuencias del tratamiento automatizado de que se trate. Una interpretación más garantista, pero solo en lo que se refiere a los arts. 13 a 15, cfr. Selbst, A. y Powles, J., Meaningful Information and the Right to Explanation, *International Data Privacy Law*, vol. 7(4), 233-242 (2017).El AG Pikamäe, acogiendo en cierto modo esta argumentación doctrinal, propuso en sus Conclusiones de 16 de marzo de 2023 relativas al asunto SCHUFA (C-634/21) que las decisiones automatizadas en el sentido del 22 del RGPD solo confieren al titular de los datos un derecho a ser informado de la lógica abstracta y general del tratamiento de datos y no un derecho a una explicación de la decisión concreta que comprenda una referencia adecuada a las circunstancias del titular de los datos afectado por ella.
La Sentencia TJUE recaída en este asunto el día 7 de diciembre de 2023 ha dejado sin resolver esta cuestión, pues se ha quedado en una cuestión previa, la de si en el caso hay una base legal suficiente que permita excepcionar la regla general de prohibición de las decisiones automatizadas.

de enero de 1978[52], que dispone que las decisiones automatizadas que se sostienen en los apartados a) y c) del artículo 22 RGPD solo serán admisibles "a condición que las reglas que informan el tratamiento y las principales características de su aplicación concreta sean comunicadas, salvo los secretos protegidos por la ley, por el responsable del tratamiento al interesado si éste lo solicita"[53]. Ese mismo precepto francés contiene una normativa específica en relación a las decisiones administrativas automatizadas, respecto de las que afirma igualmente el derecho del ciudadano a su explicabilidad. La situación aquí es parcialmente distinta en lo que a este punto interesa. No (re)interpreta el significado del artículo 22.3 RGPD porque se apoya en el apartado b) del artículo 22.2 RGPD, que se configura como una de las denominadas "cláusula de apertura" del RGPD[54], en el sentido que remite al legislador de la Unión o del Estado Miembro un espacio de apreciación en la determinación de los supuestos admisibles de decisiones automatizadas y de las salvaguardas que han de acompañarlas.

A diferencia del caso francés, en España, de un modo seguramente más ortodoxo con la caracterización de Reglamento UE del RGPD, la Ley Orgánica 3/2018 se limita a remitirse al art. 22 RGPD, sin precisar el sentido de los derechos consagrados en el artículo 22.3 RGPD, con lo cual queda abierta la cuestión de si existe o no un derecho a obtener una explicación de las decisiones individuales automatizadas[55]. Como se ha dicho en una nota previa de este mismo trabajo, la STJUE de 7 de diciembre de 2023 (asunto SCHUFA C-634/2021), a pesar de que las Conclusiones de AG Pikamäe afrontaban esta duda y optaban por una interpretación restrictiva, no la ha esclarecido porque se ha quedado en un momento anterior del razonamiento.

Volvamos al artículo 21 de la citada ley francesa ("a condición que …las principales características de su aplicación concreta sean comunicadas .. por el responsable del tratamiento al interesado *si éste lo solicita*[56]") porque es un buen punto de partida para abordar cómo ha de ser otra de las notas de este derecho a obtener una explicación del porqué de una decisión automatizada individual, la relativa a si opera *per se* o solo a petición del interesado. Este derecho tendría más intensidad si consistiera en que todas las decisiones individuales automatizadas se comunicasen al interesado con su respectiva motivación referida específicamente a las circunstancias del caso, sin que sometiera a la condición de éste asumiera la carga de solicitarla.

---

[52] Ley n° 2018-493 de 20 junio de 2018, sobre protección de datos de carácter personal. Esta ley adopta la forma de modificación de la emblemática Ley nº 708-17 de 6 de enero, sobre la informática, los ficheros y las libertades.

[53] Sobre la gestación y sentido de este precepto, Cfr. Tambou, O., "The French adaptation of the GDPR", en *National adaptations of the GDPR*, pp. 56 y 47, accesible en https://blogdroiteuropeen.com/2019/02/27/open-access-book-on-national-adaptations-of-the-gdpr-mc-cullagh-tambou-bourton-eds/

[54] Miscenic, E, y Hoffmann, A-L., "The Role of Opening Clauses in Harmonization of EU Law: Example of the EU's General Data Protection Regulation (GDPR)". *EU and comparative law issues and challenges series (ECLIC)*, 2020, pp. 44-61

[55] Su art. 18 dice que "los derechos relacionados con las decisiones individuales automatizadas, incluida la realización de perfiles, se ejercerán de acuerdo con lo establecido … en [el artículo] 22". El art. 11.2, en el contexto del derecho a recibir información de los tratamientos de datos personales, sí contiene una previsión específica, al disponer que "el afectado deberá ser informado de su derecho a oponerse a la adopción de decisiones individuales automatizadas que produzcan efectos jurídicos sobre él o le afecten significativamente de modo similar".

[56] Cursiva añadida por el autor

Traer aquí el debate doctrinal sobre la naturaleza del artículo 22 RGPD nos ofrece una perspectiva útil de análisis. En efecto, desde tiempos de la Directiva 95/46/CE y provocado por la diferente trasposición que unos y otros Estados Miembros hicieron de su artículo 15, se discute si el régimen especial previsto para las decisiones individuales automatizadas se configura como una prohibición o como un derecho a oponerse que tiene que ser accionado por el interesado (una especie de *opt ut*). Como es una constante en la regulación del espacio digital, la posición más pragmática procede del mundo anglosajón, siendo en este caso la ley inglesa de trasposición la que exigía un requerimiento escrito del interesado para que se desplegaran los derechos previstos en el régimen especial del art. 15[57]. Según este planteamiento, no habría una verdadera prohibición que pesara sobre el responsable del tratamiento, que solo levantaría si realizaba el tratamiento de modo que se satisficiesen a los interesados determinados derechos, sino una especie de derecho de oposición del interesado. Y conforme con él tendría todo el sentido que la legislación francesa que hemos referido sujetase el derecho de los individuos a una explicación de las decisiones automatizadas que les afecten a que ellos la requieran activamente.

Sin embargo, esta posición no ha prevalecido. En 2017, cuando ya se había aprobado el RGPD pero aún no había entrado en vigencia, el Grupo del Art. 29 se inclinó por considerar que el artículo 22 contenía una "*prohibición general* de tomar decisiones individuales basadas únicamente en el tratamiento automatizado" [58]. Y la STJUE de 7 de diciembre de 2023, en la misma línea, afirma que "el artículo 22, apartado 1, del RGPD consagra un «derecho» del interesado a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles. Dicha disposición establece una *prohibición de principio cuya inobservancia no necesita ser invocada proactivamente por el interesado*"[59]. Esta interpretación del artículo 22 RGPD que hace el TJUE avalaría que las decisiones individuales automatizadas solo serían posibles, como excepción a la prohibición de principio, cuando estemos en uno de los supuestos salvados de dicha prohibición y se cumpla con los derechos del interesado previstos como salvaguarda, lo cuales operan per se y no a petición de aquél. Otra cosa es que uno de estos derechos afirmados por el artículo 22 RGPD como salvaguarda de la posición de individuo afectado sea el de obtener una explicación de la decisión concreta por referencia a las circunstancias específicas del caso, cuestión esta que, como se ha expuesto más atrás, es objeto de discusión doctrinal y sobre la que aún no hay jurisprudencia interpretativa del TJUE.

---

[57] Cfr. Veale M., Edwards L., "Clarity, surprises, and further questions in the Article 29 Working Party draft guidance on automated decision-making and profiling", *Computer Law & Security Review*, 34, 2, 2018, p. 400.
El art. 12 de la UK Data Protection Act 1998, bajo la rúbrica "Rights in relation to automated decision-taking", dice en su apartado primero que "An individual is entitled at any time, by notice in writing to any data controller, to require the data controller to ensure that no decision taken by or on behalf of the data controller which significantly affects that individual is based solely on the processing by automatic means of personal data in respect of which that individual is the data subject for the purpose of evaluating matters relating to him such as, for example, his performance at work, his creditworthiness, his reliability or his conduct".

[58] Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679. (WP251rev.01), Adoptadas el 3 de octubre de 2017 y revisadas por última vez y adoptadas el 6 de febrero de 2018, pp. 22 y 25.

[59] Asunto C-634/21 (SCHUFA Holding), § 52. La cursiva es añadida por el autor.

III.3.B.- La intervención humana como derecho de las personas destinatarias de una decisión individual automatizada

El otro elemento procedimental que es susceptible de contribuir de un modo relevante a que las decisiones individuales automatizadas no sean erróneas o discriminatorias es la intervención humana, que puede adoptar muy diversos registros. Uno de ellos aparece bien representado en la supervisión humana que la Propuesta de REIA impone en el artículo 14 respecto de todos los sistemas de inteligencia artificial que reciban la clasificación de alto riesgo. Esta clase de sistemas deberán ser diseñados y desarrollados de tal forma que haya personas físicas que los mantengan bajo su control durante todo el tiempo que estén en uso. Y ese control debe ser significativo, lo que supone que las personas físicas encargadas deben ser capaces de entender correctamente el resultado del funcionamiento del modelo [art. 14.4.c)], de decidir en un caso concreto no usarlo e incluso desechar o sustituir el resultado obtenido [art. 14.4.d)] y, en fin, de detener y poner fin al funcionamiento del modelo [art. 14.4.e)]. Esta supervisión humana que exige este artículo 14 es imprescindible para lograr el objetivo de evitar que los sistemas de inteligencia artificial sean erróneos o discriminatorios, pero no se mueve en el plano de las decisiones individuales automatizadas sino en otro plano, en el del funcionamiento general y abstracto del sistema de inteligencia artificial[60], de un modo similar a como hace el apartado 2.b) del artículo XXV de la Carta de Derechos Digitales cuando prevé que "en el desarrollo y ciclo de vida de los sistemas de inteligencia artificial: .. b) se establecerán condiciones …. De supervisión humana" por contraste con su apartado 3 que dispone que "las personas tienen derecho a solicitar una supervisión e intervención humana".

En el plano de las decisiones individuales en que se traducen las previsiones resultantes de sistemas de inteligencia artificial la intervención humana puede adoptar distintas modalidades. La más radical consiste en que el ordenamiento jurídico señale algunas materias que quedan fuera del juego de los sistemas de inteligencia artificial o que, pudiéndose utilizar, su resultado no puede usarse para adoptar decisiones o acciones cualesquiera que afecten a individuos sin que medie previamente una intervención humana significativa, en el sentido de que valore efectivamente la decisión y quedé constancia de ello. Esta categoría de intervención humana, que vendría a ser lo que PONCE ha denominado "reserva de humanidad"[61], tiene al menos las siguientes manifestaciones: a) el artículo 5 de la Propuesta de RUIA prohíbe por completo una serie de prácticas de inteligencia artificial, de un modo especial las que se orientan a manipular a las personas o a explotar sus vulnerabilidades[62]; b) el artículo 14.5 de la Propuesta de RUIA, dentro de un ámbito muy acotado de materias de entre las que dan lugar a un sistema de alto riesgo, exige que para que cualquier resultado de un sistema de inteligencia artificial se traduzca en una decisión individual previamente debe ser

[60] Nótese la diferencia con el apartado 5 de ese mismo artículo 14 ("For high-risk AI systems referred to in point 1(a) of Annex III, the measures referred to in paragraph 3 shall be such as to ensure that, in addition, *no action or decision is taken by the user* on the basis of the identification resulting from the system *unless this has been verified and confirmed by at least two natural persons*"). La cursiva es añadida por el autor.

[61] Ponce Solé, J., "Reserva de humanidad y supervisión human en la inteligencia artificial", *El Cronista*, nº 100, pp. 63 a 65.

[62] Cohen, T. "Regulating Manipulative Artificial Intelligence", *Scripted*, 20, 1, Febrero 2023, que identifica la manipulación en el carácter oculto de la influencia y en el intento de explotación de vulnerabilidades de la persona.

verificada por dos personas físicas; c) el artículo 22.4 RGPD prohíbe, salvo dos excepciones, que una decisión individual automatizada se adopte sobre la base del tratamiento de categorías especiales de datos. Cabría decir que las decisiones individuales que se deriven de tratamientos de datos sensibles quedan reservados a los humanos, salvo algunas excepciones.

Aun cuando podamos enumerar algunos supuestos de "reserva de humanidad", hay otros muchos en que sí resultan admisibles las decisiones individuales automatizadas, entendiendo esta expresión en el sentido amplio que venimos sosteniendo, como comprensiva no solo de las que están completa o íntegramente automatizadas sino también de las que van seguidas de una intervención humana meramente formal, de suerte que el resultado del modelo de inteligencia artificial es determinante para decisión final. Es en estos supuestos, que son los que conforman el ámbito de aplicación del artículo 22 RGPD, donde principalmente hay que plantearse el alcance de la intervención humana como derecho de los individuos afectados por la decisión. De hecho, el apartado 3 de este artículo 22 señala expresamente a la intervención humana como una de las salvaguardas necesarias para que las decisiones individuales automatizadas se entiendan excepcionadas de la prohibición general. Lo que no precisa es en qué consistirá esta intervención humana, dejando un espacio notable para la construcción doctrinal y jurisprudencial. En mi opinión, la podríamos caracterizar conforme a varias notas: a) es una intervención humana que se produce *a posteriori* de que la decisión automatizada haya sido adoptada; b) necesariamente ha de ser una intervención humana significativa y configurarse en función de proporcionar a las personas una vía efectiva para contestar las decisiones que les afecten; c) al ser una noción abierta debe ser susceptible de modulación según la gravedad de los riesgos.

III.4.- Los derechos procedimentales, aunque siempre sirven a un mismo sustrato, no tienen una forma única y completamente predeterminada en la ley

La inteligencia artificial dista mucho de ser un concepto homogéneo. Hay diversidad de aplicaciones y modelos de inteligencia artificial, también dentro de los que actúan mediante técnicas de aprendizaje autónomo[63]. Además, los usos a que pueden destinarse son innumerables, pudiendo incidir en ámbitos poco transcendentes para la persona y también operar en la satisfacción de sus necesidades básicas. La consecuencia es que la decisión individual automatizada que refleje los errores o factores de discriminación identificados por un sistema de inteligencia artificial resultará más o menos perjudicial dependiendo tanto de la configuración de éste como del uso en que se emplee. Y en consonancia con ello la intensidad con que se conformen los derechos procedimentales que aseguren que dicha decisión se puede contestar de un modo efectivo debe ser directamente proporcional a la gravedad del posible daño. El estatus jurídico de ciudadano, caracterizado por el sustrato común de hacer posible la revisión de la decisión

[63] Cfr. Soriano Arnanz, A., "Decisiones automatizadas: problemas y soluciones jurídicas. Más allá de la protección de datos", *Revista de Derecho Público: Teoría y Método*, Vol. 3, 2021 pp. 85-127, expone la diferencia entre sistemas automáticos y autónomos, señalando que los segundos son los que más riesgos suponen para las personas.

automatizada que actúa como referencia que proporciona un mínimo de seguridad jurídica, ha de adaptarse a las circunstancias. Será en cuanto a sus manifestaciones concretas asimétrico, so pena de revelarse en muchas de sus aplicaciones como desproporcionado, en el sentido de condicionar en exceso la configuración de los sistemas de inteligencia artificial. Esta idea de principio merece ser examinada aquí desde dos perspectivas.

El Parlamento Europeo en una Resolución de 12 de febrero de 2019[64], con el trasfondo de la Comisión Europea trabajando ya en elaborar un proyecto de norma general sobre la inteligencia artificial, expone que "la IA es un concepto que abarca una amplia gama de productos y aplicaciones" y que por ello "debería abordarse con cautela una ley o regulación integral de la IA"65, sugiriendo la posibilidad de que adoptar un enfoque sectorial fuese una solución más adecuada. Además, en la misma resolución, pide a la Comisión que "eval[úe] periódicamente la legislación actual con el fin de garantizar que sea adecuada para su propósito en relación con la IA y … que intente modificar o sustituir las nuevas propuestas cuando no sea así"66, con lo que alude a la incertidumbre ligada al desarrollo tecnológico de los sistemas de inteligencia artificial, que requiere dotar de una cierta flexibilidad a la regulación que se adopte en lugar de predeterminar toda ella en una ley con pretensión de estabilidad[67].

La Propuesta de RUIA de la Comisión presentada en 2021 recoge al menos en una parte muy relevante estos planteamientos. Aunque sigue teniendo por objeto una regulación integral de la inteligencia artificial, dentro de ella coexisten diferentes regímenes de obligaciones dependiendo del riesgo que entrañen los distintos modelos y aplicaciones. Solo respecto de los sistemas que se califican de alto riesgo la norma prevé un importante conjunto de requisitos obligatorios, e incluso respecto de ellos el posible impacto para los derechos fundamentales se remite a una evaluación específica caso por caso que aporta flexibilidad y capacidad de adaptación a la incertidumbre en la determinación de las medidas de salvaguardia. Algo parecido ha sucedido con las reglas que, no estando en la Propuesta de la Comisión Europea, se han incorporado en relación a las formas de inteligencia artificial de uso general, que ante su situación de plena evolución y la consiguiente falta de conocimiento cierto sobre sus riesgos[68] el texto resultante del compromiso político alcanzado entre los colegisladores en diciembre de 2023 ha optado por someter a los modelos de impacto más extenso o de mayor riesgo a evaluaciones específicas[69].

Conforme a este enfoque, no solo las garantías de buen funcionamiento de los sistemas de inteligencia artificial sino también los derechos frente a las decisiones individuales en

[64] Resolución del Parlamento Europeo, de 12 de febrero de 2019, sobre una política industrial global europea en materia de inteligencia artificial y robótica (2018/2088(INI))
[65] Punto 116
[66] Punto 114
[67] Esteve Pardo, J.,"La regulación de riesgos. Gestionar la incertidumbre", *El Cronista*, 96-97, p. 32 y ss, sostiene que Los riesgos tecnológicos y la incertidumbre que les rodea son inapropiados para el establecimiento en la ley de referencias generales y fijas.
[68] DIGITALEUROPE, op. cit, pp. 12 y 13.
[69]Véase el comunicado de prensa del Parlamento Europeo de 9 de diciembre de 2023, accesible en https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai

que su uso se traduzca deberán determinarse en la norma con arreglo a su transcendencia, e incluso remitirse en cuanto a su forma precisa a las evaluaciones de impacto específicas que los proveedores o usuarios realicen. Aparentemente cabría ver un límite en el artículo 22 RGPD, pues establece un régimen general para las decisiones individuales automatizadas dentro del cual enuncia una serie de derechos del titular de los datos personales tratados, pero hay dos elementos que alteran esta conclusión. De un lado, el carácter abierto con que se enuncian los derechos en el artículo 22.3 RGPD, que admite distintas formas más o menos exigentes de realización última. De otro lado, la interpretación sistemática que conllevaría tener en cuenta el principio de responsabilidad proactiva del artículo 24 RGPD, según el cual el responsable del tratamiento debe evaluar cómo afecta éste a los titulares de los datos personales y modular en consecuencia las medidas técnicas y organizativas necesarias, entre ellas por supuestos las que otorguen efectividad a los derechos previstos en el citado artículo 22 en relación a las decisiones individuales automatizadas.

La segunda perspectiva que reclama ahora nuestra atención es la del reparto de responsabilidades entre el proveedor del sistema de inteligencia artificial y el usuario del mismo, en especial en relación a la valoración de riesgos a partir de la que fijar la intensidad de los derechos procedimentales. El modelo de inteligencia artificial que desarrolla un proveedor puede ser destinado a múltiples usos. Esta distancia entre proveedor y usuario se incrementará a medida que se generalicen los llamados modelos básicos (permiten su rápida adaptación a distintas tareas), pudiendo también aparecer sujetos intermedios entre el creador del modelo y el que finalmente lo utiliza para un uso determinado mediante la producción de decisiones que afectan a los individuos. Dado que quien se coloca al principio de la cadena de valor no conoce los usos finales, no puede valorar con conocimiento de causa los riesgos e incorporar en consecuencia una u otra configuración al sistema que construye. Solo el usuario final tiene a la vista la transcendencia de las decisiones individuales que derivan de las pautas identificadas por el sistema de inteligencia artificial y está en situación de hacer una valoración de los perjuicios que ocasionaría dichas pautas contuvieran errores o factores de discriminación. Sin embargo, el usuario final estará condicionado por las características con las que el proveedor haya conformado el modelo de inteligencia artificial.

En definitiva, la obligación de evaluar cómo un uso de inteligencia artificial afectará la salud, la seguridad y los derechos fundamentales y, a partir de ello, dar una u otra forma a los derechos procedimentales del individuo afectado debe ser asumida por el usuario final, sin perjuicio de que el proveedor debe tener la obligación de facilitar esta tarea, tanto mediante el diseño del sistema como asumiendo en relación al usuario obligaciones de transparencia sobre el funcionamiento del sistema. Las evaluaciones del sistema de inteligencia artificial que corresponda realizar al proveedor no eximen al usuario de abordar la identificación de riesgos provenientes del uso a que destine el sistema. Es cierto, no obstante, que el usuario de un sistema de inteligencia artificial orientado a producir decisiones individuales reviste al mismo tiempo la condición de responsable del tratamiento de datos personales, recayendo sobre él las obligaciones de evaluación que conlleva el principio de responsabilidad proactiva[70]. Ambas actividades de evaluación de

[70] Especialmente las que se derivan de los arts. 24 y 35 del Reglamento (UE) 2016/679 (RGPD)

impacto tienen aspectos en común, por lo que sería conveniente que se posibilitase alguna fórmula de cumplimiento conjunto o coordinado[71].

III.5.- El sustrato común de la protección de los ciudadanos frente a los poderes públicos y a las empresas privadas. La generalización de los derechos procedimentales.

La tecnología no solo entraña riesgos cuando la usa la Administración Pública. La presencia de errores o de factores de discriminación en usos que se hagan de sistemas de inteligencia artificial por empresas privadas y a partir de los cuales se adopten decisiones que afecten a las personas también son susceptibles de producir daños en su salud, seguridad o derechos (sean o no fundamentales). Esos perjuicios procedentes de la actuación de sujetos privados pueden revestir grados muy importantes de gravedad, tanto por el objeto como por el sujeto. En cuanto a lo primero, cada día son más los servicios transcendentes que los ciudadanos reciben de empresas privadas en el espacio digital y para cuya prestación se utilizan de un modo creciente tecnologías de inteligencia artificial (por supuesto, seguros, crédito y salud; pero también la información, e incluso la "arena digital" donde realiza la relación entre personas). Por lo que se refiere al aspecto subjetivo, la consideración de las empresas privadas como verdaderos poderes es una realidad que la doctrina reconoce desde hace varias décadas[72], pero es precisamente en el espacio digital donde se está poniendo de manifiesto que las empresas privadas ejercen potentísimas formas de poder, seguramente no de carácter coercitivo como las Administraciones públicas, pero sí el poder igualmente dominante que deriva de controlar la información, los escenarios en que se produce y los cauces por medio de los que adquiere valor[73].

Se sigue de todo lo anterior que los poderes públicos no solo han de hacer uso de las aplicaciones de inteligencia artificial de una manera respetuosa con los derechos de los ciudadanos sino que, en el contexto de una sociedad digital protagonizada por la preeminencia de las empresas privadas que lideran la innovación tecnológica y dominan la información[74], deben implicarse en garantizar que el uso que éstas hagan de los sistemas de inteligencia artificial sea acorde con los intereses jurídicos de las personas afectadas, asegurando que no se vean privadas del ámbito de autonomía personal

[71] El Parlamento Europeo en la posición que adoptó en junio 2023 sobre la Propuesta de RUIA imponía al usuario la obligación de evaluar los riesgos del uso de inteligencia artificial que pretendía hacer. DIGITALEUROPE, op. cit, pp. 11 y 12, lo considera una duplicación e insta a que al menos se eviten los solapamientos con la normativa de protección de datos personales.

[72] Cfr. Doménech Pascual, G., *Derechos fundamentales y riesgos tecnológicos*, Madrid, 2006. Luego de una cita de Nieto ["La nueva misión del Derecho Administrativo consiste .. en defender a los ciudadanos y a la Administración de las eventuales (y reales) agresiones de grupos privados, mucho más poderosos que las Administraciones Públicas", Recensión al libro de Esteve Pardo, *Autorregulación: génesis y efectos*, RAP, 160, 2003, p. 429], añade "Durante mucho tiempo el Derecho público se ha preocupado especialmente por poner límites y frenos a las intervenciones estatales. Ahora la necesidad es obligar al Estado a intervenir, garantizar que no permanezca inactivo frente a las nuevas amenazas para la libertad".

[73] Sobre los poderes privados en el espacio digital y el papel del Derecho, cfr. Vettori, G., "Sui poteri privati. Interazioni e contaminazioni", *Diritto Publico*, 3, 2022, pp. 829 y ss.

[74] Piñar Mañas, J.L.: *Derecho e innovación tecnológica. Retos de presente y futuro*, CEU Ediciones, Madrid, 2018.

indispensable para conducir una vida digna de persona humana y que disponen, en una medida variable que se corresponda con el contenido de las decisiones individuales automatizadas de que se trate, de los derechos procedimentales que sustenten una vía que permita de un modo efectivo contestar aquellas decisiones y lograr su revisión y sustitución por otras.

No solo es que las amenazas que entraña la inteligencia artificial para los ciudadanos se manifiestan tanto cuando es utilizada por la Administración Pública como cuando es usada por una empresa privada. Es además que su expresión presenta un sustrato similar, concretado en el peligro para la autonomía individual y que el contenido de la decisión individual pueda ser erróneo o discriminatorio. Tiene sentido, por tanto, que el ordenamiento jurídico reaccione proporcionando al ciudadano un conjunto de derechos que en su esencia es común a una y otra situación. Estos derechos no son completamente idénticos, dado que la especial posición de la Administración Pública conlleva necesariamente lo que podríamos denominar adiciones o agravantes, pero la base ha de ser común.

El régimen jurídico de las decisiones individuales automatizadas definido por el artículo 22 RGPD parece apuntar en otra dirección[75]. Es un precepto que impone una prohibición general y prevé distintos supuestos de excepción. Dos de ellos, los que se enuncian de un modo preciso y respecto de los que el artículo 22.3 prevé el reconocimiento de una serie de derechos al titular de los datos, no son adecuados para la actuación administrativa. Esta solo puede encauzarse a través del otro supuesto de excepción, el que regula el artículo 22.2.b) y que consiste en una remisión a normas de la Unión Europea o de los Estados Miembros que autoricen este tipo de tratamientos automatizados. Si nos atenemos estrictamente a la letra del precepto, se trata de una remisión completamente abierta que permitiría que una norma de un Estado Miembro autorizase que las Administraciones Públicas adoptasen decisiones individuales automatizadas con sujeción al régimen especial que dicha norma estableciese. La norma autorizatoria sería entonces enteramente libre de configurar los derechos de los ciudadanos de un modo diverso, separado de las facultades básicas previstas en el artículo 22.3 y desarrolladas en el Considerando 71, respecto de estas decisiones automatizadas administrativas[76].

Esta, sin embargo, no es la única interpretación posible. Otra interpretación, que me parece preferente y que ha sido sostenida en la doctrina especializada en materia de protección de datos personales[77], es que la cláusula de apertura recogida en el artículo 22.2.b) ha de entenderse integrada en el sistema del RGPD del que forma parte. Según este entendimiento, la norma nacional que autorizara las decisiones automatizadas administrativas tendría que partir del respeto a los principios del RGPD y acomodarse además al régimen básico que para las decisiones individuales automatizadas deriva del

[75] Cfr., Boix Palop A. y Soriano Arnanz A., p. 272, que destacan que las previsiones del art. 22 RGPD "son normas pensadas más para el tráfico privado que la acción pública".

[76] Wachter S, Mittelstad B, Floridi L, op. cit., parten de este criterio literal para interpretar este precepto, con el objeto de configurar restrictivamente el marco dentro del que el TJUE podría afirmar un derecho a obtener una explicación de la decisión individual tomada a partir de algoritmos. Es decir, para justificar que las normas dictadas al amparo del art. 22.2.b) podían separarse del reconocimiento de facultades que en favor del individuo afectado por la decisión se halla en el art. 22.3.

[77] Véase Palma Ortigosa, A., "Decisiones automatizadas en el RGPD. El uso de los algoritmos en el contexto de la protección de datos", *Revista General de Derecho Administrativo*, 50, 2019.

Considerando 71 y se refleja, por lo que hace a algunas de sus manifestaciones, en el artículo 22.3[78]. Por tanto, la remisión del artículo 22.2.b) no habilitaría para establecer regímenes distintos sino para especificar el fijado por el propio RGPD en función de las exigencias particulares que entrañase el tipo de decisión automatizada que se autoriza, especificación que alcanzaría a dar una forma más o menos intensa a las facultades que el RGPD atribuye a los titulares de los datos y a fijar requerimientos añadidos allí donde la materia o el sujeto lo requiera[79].

Creo que podemos extraer de todo ello un par de conclusiones. La primera es que la inteligencia artificial plantea unas amenazas para los ciudadanos que en su esencia fundamental son similares independientemente de que quien la utilice sea el poder público o las empresas privadas, de modo que tiene sentido que el estatus jurídico del ciudadano frente a la inteligencia artificial tenga un sustrato común. Ello no obsta para que su expresión final adquiera particularidades dependiendo del caso de uso, por lo que se comporta en lo concreto como un estatus jurídico asimétrico. Estas especialidades no son exclusivas de que la decisión individual automatizada de carácter administrativo, dado que las normas que exceptúan la prohibición general al amparo del artículo 22.2.b) también pueden referirse a usos realizados por empresas privadas. La segunda conclusión es que ese sustrato común que caracteriza el estatus jurídico del ciudadano se apoya en la configuración de derechos procedimentales orientados a proporcionar transparencia en los motivos de la decisión y en última instancia hacer posible una vía efectiva mediante la que poder contestar las decisiones individuales automatizadas que sean erróneas o discriminatorias y lograr que sean revisadas. Asistimos, por tanto, a un proceso de generalización de técnicas elaboradas inicialmente para proteger al ciudadano frente al exceso del poder público, que ahora se demuestran también útiles para equilibrar las nuevas relaciones de poder que vinculan a los ciudadanos con los nuevos poderes privados que señorean el espacio digital.

[78] La STJUE de 7 de diciembre de 2023 (Asunto C-634/21, SCHUFA) ha venido a respaldar expresamente esta interpretación. Señala en los apartados 65 y 66 que el art. 22.2.b) impone que las normas que autoricen ciertos tipos de decisiones individuales automatizadas definan medidas de salvaguarda de los intereses de los afectados por ellas, precisando que esas medidas han corresponderse con las reglas previstas en el Considerando 71. Y añade en los apartados 67 y 68 que la cláusula de apertura del art. 22.2.b) no abre la puerta a que las normas que se dicten a su amparo autoricen tratamientos de datos que no se sometan a los principios consagrados en los arts. 5 y 6 RGPD.

[79] A este planteamiento responde el art. 21 de la ley francesa que acomoda ese ordenamiento nacional al RGPD (Loi n° 2018-493 de 20 junio de 2018), que dentro de su apartado I dice: "2° Des décisions administratives individuelles prises dans le respect de l'article L. 311-3-1 et du chapitre Ier du titre Ier du livre IV du code des relations entre le public et l'administration, à condition que le traitement ne porte pas sur des données mentionnées au I de l'article 8 de la présente loi. Ces décisions comportent, à peine de nullité, la mention explicite prévue à l'article L. 311-3-1 du code des relations entre le public et l'administration. Pour ces décisions, le responsable de traitement s'assure de la maîtrise du traitement algorithmique et de ses évolutions afin de pouvoir expliquer, en détail et sous une forme intelligible, à la personne concernée la manière dont le traitement a été mis en œuvre à son égard"

IV.- EL USO DE LA INTELIGENCIA ARTIFICIAL POR LA ADMINISTRACIÓN

Las tecnologías de inteligencia artificial tienen una utilidad general y se han extendido también al ámbito de las actuaciones de las Administraciones Públicas, incluidas aquellas que se traducen en decisiones que afectan directamente a los ciudadanos. En el apartado anterior hemos mostrado cómo en relación con su uso los ciudadanos disfrutan de un estatuto jurídico básico, que deriva de los derechos previstos en la normativa de protección de datos respecto de las decisiones individuales automatizadas. Adicionalmente, la especial posición de las Administraciones Públicas en relación a los ciudadanos, caracterizada por las exigencias de transparencia pública, la participación de los ciudadanos en las decisiones públicas y las implicaciones asociadas al principio de legalidad, determina que los ciudadanos vean reconocidas facultades complementarias.

IV.1.- El concepto amplio de decisión automatizada administrativa

Antes que nada debe precisarse el ámbito objetivo respecto del que se atribuyen derechos a los ciudadanos. Se entiende por actuación administrativa automatizada, según el artículo 41.1 la Ley 40/2015, "cualquier acto o actuación realizada *íntegramente* a través de medios electrónicos por una Administración Pública en el marco de un procedimiento administrativo *y en la que no haya intervenido de forma directa un empleado público"* (anteriormente artículo 39 de la Ley 11/2007). Se trata de un concepto de actividad administrativa automatizada muy restrictivo, que asumido en su estricta literalidad deja fuera toda actuación administrativa en que la decisión final corra a cargo de una persona humana. Este concepto limitado no es exclusivo de la norma española. La legislación alemana de procedimiento administrativo[80] se refiere a los actos administrativos que se dicten "enteramente por medio automáticos" y la ley francesa de adaptación al RGPD[81] alude a las decisiones "que se basen únicamente en un tratamiento automatizado de datos personales"[82]. Esta cercanía entre los citados preceptos, no solo de significado sino incluso de dicción, no es una coincidencia. Deriva de que estos ordenamientos debían acomodarse al artículo 15 de la Directiva 95/46/CE que, al reglamentar las decisiones individuales automatizadas desde el ángulo de la protección de datos personales, las define como aquellas "que se basen únicamente en un tratamiento automatizado de datos personales", precepto que en 2016 pasó con la misma literalidad al artículo 22 RGPD, en vigor desde mayo de 2018.

Este concepto tan reducido de actividad administrativa automatizada no se corresponde con la realidad de las actuaciones administrativas que repercuten en los

[80] Art. 35 a) VwVfg, que está en vigor desde 2017, afirma que "Ein Verwaltungsakt kann *vollständig* durch automatische Einrichtungen erlassen werden, sofern dies durch Rechtsvorschrift zugelassen ist und weder ein Ermessen noch ein Beurteilungsspielraum besteht".

[81] El art. 21 de la Loi n° 2018-493, citada supra, dispone "Aucune décision produisant des effets juridiques à l'égard d'une personne ou l'affectant de manière significative ne peut être prise *sur le seul fondement d'un traitement automatisé de données* à caractère personnel, y compris le profilage, à l'exception" y entre las excepciones enumera "Des décisions administratives individuelles" cuando se cumple ciertos requisitos.

[82] En general, sobre la habilitación de la automatización en las leyes europeas de procedimiento administrativo, véase Mir Puigpelat, O. "La automatización y el uso de algoritmos en inteligencia artificial en Derecho administrativo comparado", *Revista General de Derecho Administrativo*, 63, 2023, p. 6.

ciudadanos. En estos casos el uso de la inteligencia artificial tiene por objeto, sobre todo cuando estamos hablando de sistemas que operan con cierta autonomía, influir de una manera determinante en la decisión final. Lo ordinario es que el empleado público, aun cuando esté en su mano separarse, siga con un cierto automatismo el criterio resultante del tratamiento automatizado. Aunque haya presencia humana en la decisión final, se trata de una intervención meramente formal. La categoría de actividad preparatoria como distinta a decisión final, aunque sigue siendo útil, ha de comprender un ámbito más reducido. No abarcará todos aquellos supuestos en que la decisión final sea adoptada por un empleado público, sino solo aquellos en que se demuestre que su intervención no es meramente formal, lo que exigirá que se separe regularmente (o al menos no solo ocasionalmente) del criterio del sistema automatizado. Quizá fuera suficiente que la persona física estuviera sujeta a examinar dicho criterio de un modo agravado y a dejar constancia del resultado, pero en tal caso será imprescindible que se definan los estándares que deberían estar presentes en tal procedimiento. Además, mantener una concepción reducida de actividad administrativa automatizada como la que se desprende del tenor literal del artículo 41.1 Ley 40/2015 convertiría el régimen especial que lleva asociado en fácilmente eludible. Bastaría con situar en el momento de la decisión final a un funcionario público, aunque su papel no fuese más allá que trasladar automáticamente a aquélla el criterio que emanase del tratamiento automatizado previo.

Razonamientos de esta clase se habían hecho valer en documentos de soft law y en la doctrina[83] y ahora el TJUE les ha conferido carta de naturaleza[84], al establecer que el régimen especial contenido en el artículo 22 RGPD es también aplicable a los tratamientos automatizados que realizan las agencia de información comercial aunque la repercusión en el ciudadano solo se produzca con la decisión de la entidad financiera de conceder o no un crédito. Esta delimitación más amplia del concepto de decisión individual automatizada ex artículo 22 RGPD incide directamente en las normas nacionales que antes examinamos, las cuales deberán ser reinterpretadas[85]. Donde se lee "acto o actuación realizada *íntegramente* a través de medios electrónicos" (ley española), actos administrativos que se dicten "*enteramente* por medio automáticos" (ley alemana) o decisiones "que se basen *únicamente* en un tratamiento automatizado" (ley francesa) hay que hacer el esfuerzo de entender comprendidos también los casos en que el tratamiento automatizado vaya seguido de una intervención humana sea meramente formal.

La fuerza de obligar del concepto amplio de decisión automatizada que conforma esta sentencia afecta al artículo 41.1 Ley 40/2015 solo en la medida que se entienda que sirve para adaptar el Derecho español a la normativa de protección de datos personales. Pero la generalidad de sus argumentos (no conectan con los objetivos específicos de la protección de datos personales) y las ventajas de la homogeneidad coadyuvan en favor de extender este concepto amplio de decisión automatizada más allá de ese ámbito. Ello no es obstáculo tampoco para que las decisiones íntegramente automatizadas constituyan, de

[83] El Grupo del Art. 29, en sus Directrices sobre decisiones individuales automatizadas, más arribas citadas, señala en la pág. 23 que "para ser considerada como participación humana, el responsable del tratamiento debe garantizar que cualquier supervisión de la decisión sea significativa, en vez de ser únicamente un gesto simbólico". En la doctrina, Civitarese Mateucci, S, op. cit., pp. 23 y 26.

[84] STJUE de 7 de diciembre de 2023 (Asunto C-634/21, SCHUFA), § 50.

[85] Sobre la lectura estricta que se viene haciendo (y que deberá modularse) cfr. para el caso alemán Schneider, J.P. y Enderlein, F., "Sistemas automatizados de toma de decisiones en el Derecho administrativo alemán", *Revista General de Derecho Administrativo*, 63, 2023 *Revista General de Derecho Administrativo*, 63, 2023 ("vacío normativo en relación con la toma humana de decisiones apoyada por las tecnologías de la información basadas en datos").

entre todas las decisiones automatizadas, un supuesto al que el legislador anude garantías específicas adicionales.

IV.2.- El régimen jurídico de las actuaciones administrativas automatizadas previsto en el artículo 41 Ley 40/2015.

Las actuaciones administrativas automatizadas, definidas en el sentido amplio que hemos expuesto, se sujetan al régimen especial previsto en el artículo 41.2 Ley 40/2015. Régimen especial que se limita a exigir que se establezca "previamente el órgano u órganos competentes, según los casos, para la definición de las especificaciones, programación, mantenimiento, supervisión y control de calidad y, en su caso, auditoría del sistema de información y de su código fuente. Asimismo, se indicará el órgano que debe ser considerado responsable a efectos de impugnación". Hay diversidad de opiniones en la doctrina[86] acerca de si este precepto es suficiente para satisfacer las exigencias que se desprenden del principio de legalidad en materia de atribución de potestades a las Administraciones Públicas. La problemática desde este punto de vista, a mi juicio, no está tanto en que se trate de una habilitación genérica, que se refiere a todas las actuaciones automatizadas administrativas sin acepción de materia, sino a que el régimen general que dispone quizá sea insuficiente por reducirse a prever que la designación de los órganos competentes se realice previamente.

En lo que no cabe la menor duda es que este precepto no contiene el estatuto jurídico del ciudadano en relación a las decisiones administrativas automatizadas. El art 35 a) VwVfg, aunque tampoco enuncia los derechos que asisten a los ciudadanos afectados cuando son objeto de este tipo de actuaciones administrativas, dista mucho de la disposición española. El precepto alemán exige que cada acto administrativo automatizado "sea permitido por una norma" y será esta base normativa específica la encargada de establecer los derechos de los ciudadanos. El legislador alemán ha entendido que la norma que, en aplicación del artículo 22.2.b) RGPD, exceptúe la prohibición general de decisiones individuales automatizadas ex art. 22.1 ha de ser específica. Solo teniendo presente la materia en la que va a incidir la decisión automatizada cabe precisar cuáles son las salvaguardas necesarias de los derechos del titular de los datos personales y en qué grado de intensidad hay que aplicarlas. Eso sí, esa norma (*Rechtsvorschrift*) no tiene porqué ser una ley[87]. La legislación francesa, a diferencia de la alemana, ha optado por incorporar a la ley por la que adapta su ordenamiento al RGPD una cláusula que exceptúa en general de la prohibición del artículo 22 RGPD a todas las decisiones administrativas automatizadas, siempre que respeten varios derechos de los afectados: a) que no se usen en estos tratamientos automatizados categorías especiales de datos; b) que se informe a las personas que las decisiones que les afectan son decisiones automatizadas; y c) que el sujeto público responsable del tratamiento mantenga el control del sistema algorítmico con el fin de informar en detalle y de forma inteligible a las personas de la manera en que les afectan sus aplicaciones concretas.

[86] Gamero Casado, E., "Sistemas automatizados de toma de decisiones en el Derecho administrativo español", *Revista General de Derecho Administrativo*, 63, 2023 *Revista General de Derecho Administrativo*, 63, 2023; Civitarese Mateucci, pp. 34 a 37.
[87] Cfr. Huergo Lora, A., op. cit., nota 62.

Lo que más interesa ahora no es tanto las diferencias entre las legislaciones alemanas y francesa. Lo relevante es que, a pesar de las distintas fórmulas, en ambos casos se aprueban normas que no solo habilitan la actuación administrativa automatizada sino que también se ocupan de atribuir una serie de derechos a los ciudadanos afectados por ellas. La legislación española, por el contrario, no hace esto último. Ni lo hace la norma que regula el funcionamiento de las Administraciones Públicas (en especial no lo hace el artículo 41 Ley 40/2015) ni tampoco la ley mediante la que se adapta el ordenamiento jurídico español al RGPD (su artículo 18 se limita a remitir al art. 22 RGPD). Este déficit tiene consecuencias graves. El artículo 41.1 Ley 40/2015, en la medida que no cumple con uno de los requisitos exigidos por el artículo 22.2.b), el de prever los derechos que tendrán los ciudadanos frente a las decisiones administrativas automatizadas, no desplaza la prohibición general establecida por el artículo 22 RGPD. En definitiva, las decisiones administrativas automatizadas estarían incursas en dicha prohibición general, resultado que pugna con la creciente utilización que las Administraciones Públicas hacen de estas tecnologías[88].

No creo que sea una solución sostener que los derechos de los ciudadanos frente a las decisiones administrativas automatizadas no necesitan una atención propia porque ya se aplican los derechos que en general están previstos por la legislación de procedimiento administrativo. Y ello por dos razones. La primera porque el artículo 22.2.b) exige que la norma que exceptúe cierto tratamiento automatizado de la prohibición general considere de un modo específico, atendiendo a la materia y a que se trata de una decisión automatizada, la extensión y la intensidad de los contrapesos. Estos vienen definidos de un modo genérico en el Considerando 71 del RGPD [que no es vinculante, pero la STJUE de 7 de diciembre de 2023 ha utilizado como parámetro de interpretación del artículo 22.2.b)], pero su extensión e intensidad concreta deben fijarse en la norma que autorice la excepción. La segunda razón es que en el elenco de derechos de los ciudadanos previstos en la legislación de procedimiento administrativo no están algunas de las salvaguardas previstas en el Considerando 71 del RGPD[89].

Podría argumentarse que, autorizado por la norma española como supuesto de excepción la adopción de decisiones administrativas automatizadas, el silencio en esa norma conlleva la aplicación del conjunto de derechos contemplados en el artículo 22 RGPD y en el Considerando 71. Desde un punto de vista material podría ser una lectura satisfactoria porque al fin y al cabo el régimen de derechos que atribuya a los ciudadanos la norma de excepción tiene respetar ese esquema general. Pero puede encontrar obstáculos formales, como que el artículo 22.2.b) RGPD requiere una consideración específica del caso concreto en la norma de excepción, aparte de que ese esquema general requiere precisión para ser aplicado. La predisposición del TJUE a superar estos obstáculos formales es difícil de predecir.

En definitiva, en rigor solo se abren dos vías para que las decisiones administrativas automatizadas queden fuera de la prohibición general del artículo 22 RGPD. Cabría, por un lado, incorporar a una norma que se refiriese a todas ellas, como por ejemplo el artículo 41 Ley 40/2015 o una norma similar que se integrase en la normativa de protección de datos personales, un elenco de derechos de los ciudadanos en la línea de los regulados en el Considerando 71, precisando su extensión y grado de intensidad. En este mismo precepto se podrían señalar estándares para identificar cuándo la intervención humana en

[88] Esta situación no es exclusiva de la norma española. Un análisis de Derecho comparado puede verse en Mir Puigpelat, O, op cit, p. 6.
[89] Sobre la discusión de esta cuestión en Alemania, Schneider, J.P. y Enderlein, F., op. cit., pp. 16 y 17.

la toma de la decisión final es significativa, quedando excluida en tal caso la aplicación de la normativa propia de las decisiones administrativas automatizadas. La otra opción consistiría en que la previsión de los derechos de los ciudadanos fuese abordada caso por caso en aquella normativa sectorial que autorice actuaciones administrativas automatizadas, pudiendo hacerlo a través de normas de carácter reglamentario, pues al fin y al cabo consiste en la precisión de un régimen ya diseñado con cierto detalle en el RGPD.

IV.3.- La especial posición de la Administración Pública y el estatuto jurídico de los ciudadanos en relación al uso de los algoritmos.

El régimen de derechos previsto en el RGPD y al que obligatoriamente debe ajustarse el ordenamiento español cubre en parte los intereses jurídicos de los ciudadanos ante el uso de algoritmos por las Administraciones Públicas. Aunque su propósito sea uno distinto, dar mayor control al titular de los datos personales sobre cómo se utilizan por los responsables del tratamiento, producen de un modo mediato las garantías que los ciudadanos necesitan frente a la actividad de aquéllas, lo que se ve favorecido por la cláusula de apertura que contiene el artículo 22.2.b) RGPD, mediante la cual el régimen de protección de datos personales puede ajustarse al supuesto específico de la acción pública o incluso de las distintas actuaciones públicas. En la medida que este efecto mediato se produzca, no tendría sentido duplicar la protección asignando nuevos derechos so pena de generar solapamientos ineficaces de regímenes jurídicos y la consiguiente falta de seguridad jurídica.

Sin embargo, esta normativa de protección de datos personales es inadecuada para responder a algunas de las necesidades que plantea el uso de algoritmos por las Administraciones Públicas. Tiene, de un lado, un cariz subjetivista orientado a proteger a los titulares de los datos personales cuando frente a la actuación pública también tienen derechos otros ciudadanos interesados distintos de aquellos y, en ocasiones, los ciudadanos en general. De otro, las Administraciones Públicas en el ejercicio de su poder se tienen que sujetar a principios característicos como los de transparencia, participación y legalidad.

El principio de transparencia de los poderes públicos implica que su gestión debe estar abierta al conocimiento de todos los ciudadanos, como una forma de ejercer una suerte de participación y control democrático. Los ciudadanos, y no solo los particularmente afectados, tienen derecho a conocer cuándo las Administraciones Públicas se sirven de tratamientos automatizados y cuáles son sus características principales. Una técnica de publicidad activa en este contexto puede consistir en establecer registros públicos[90] de los sistemas automatizados que usa cada entidad pública, con una referencia a las principales características de su funcionamiento y de las consecuencias que entraña. Podría pensarse que es una suerte de generalización de la obligación de informar sobre los sistemas automatizados que usa que tiene el responsable del tratamiento en favor de los titulares de datos personales. Pero no es exactamente así, pues el acento es otro. No se trata ya de ilustrar al titular sobre cómo se usan sus datos personales y cómo le va a afectar ese uso, sino de permitir que el público conozca los elementos principales del ejercicio del poder que se hace mediante sistemas automatizados. Se busca evitar que el poder se oculte

[90] Boix Palop, A. y Soriano Arnanz, op. cit., 275, con referencia a las propuestas existentes.

detrás del algoritmo. Este diferente acento ha de reflejarse en el contenido que se publicita a través del registro.

Aparte del contenido es importante definir qué sistemas automatizados han de inscribirse. El artículo 16.1.l) de la Ley 1/2022, de transparencia y buen gobierno de la Comunidad Valenciana, prevé que se publique "la relación de sistemas algorítmicos o de inteligencia artificial que tengan impacto en los procedimientos administrativos o la prestación de los servicios públicos …" y la doctrina[91] subraya que esa relación comprende todos ellos, también los que tengan un impacto solo indirecto, resaltando que con ello su ámbito excede del que previsto en el artículo 41.1 Ley 40/2015. En este trabajo se viene sosteniendo que, yendo más allá de su dicción literal, el concepto actividad administrativa automatizada ex art. 41.1 abarca también los tratamientos que sean determinantes para la decisión final adoptada por un humano. Entiendo que la ley valenciana, al fijar qué sistemas es obligatorio publicar, acude a un concepto aún más amplio dentro del que hay que incluir igualmente los tratamientos automatizados que no son determinantes de la decisión final porque van seguidos de una intervención humana significativa. Desde la óptica de la transparencia, que persigue hacer del conocimiento público la actividad administrativa, todos los sistemas algorítmicos que usen las entidades públicas resultan afectados, lo que no obsta para que otros bienes jurídicos justifiquen excepciones a la obligación de dar publicidad o al menos modulaciones[92].

La transparencia de las Administraciones Públicas se logra también con el ejercicio por los ciudadanos del derecho de acceso a la información pública. Aquí seguiré como guía lo sucedido en el caso de la Fundación Civio en relación al algoritmo usado por el Ministerio de Transición Ecológica llamado Bosco[93]. En síntesis, el subsidio eléctrico denominado "bono social", cuyas condiciones de aplicación se fijan en detalle en las normas sectoriales, se otorga o deniega mediante el referido algoritmo, que debe obviamente ajustarse estrictamente a los criterios normativos de adjudicación. Ante la aparición de resultados inesperados, la fundación solicitó en vía administrativa (ante el Ministerio y el Consejo de Transparencia) y judicial el código fuente del algoritmo Bosco[94]. Lo que me interesa destacar es que este ejercicio del derecho de acceso no es en rigor un ejercicio de transparencia, pues las razones de la decisión administrativa son conocidas en tanto que prefiguradas en la ley. Es más propiamente un mecanismo de control, que persigue cerciorarse que el algoritmo cumple con las reglas establecidas normativamente. Los algoritmos que incorporan criterios ya definidos en las leyes no requieren transparencia sino control. Y el momento en que el control se despliega con mayor garantía es cuando la actuación pública ha desembocado en una resolución concreta, que no será válida si no responde a los parámetros legales por mucho que sea el fruto de la aplicación de un algoritmo. El esfuerzo ha de dirigirse a configurar la actuación administrativa automatizada de forma que la resolución final pueda ser objeto de este control. Habrá que imponer que los sistemas automatizados usados para ejercer el poder público se configuren específicamente para permitir dejar constancia reforzada de los

[91] Boix Palop, A. y Soriano Arnanz, op. cit., 276.

[92] Huergo Lora, A., op. cit., apartado IV.4.7, señala que quedan fuera de la publicidad activa "los algoritmos que solo sirven para seleccionar objetivos" y "los algoritmos no predictivos", pero pareciera que estas afirmaciones no se corresponden tanto a los deberes generales de transparencia cuanto a los derechos de audiencia dentro de un procedimiento concreto.

[93] Un análisis muy ilustrativo en Fuertes, M., "Reflexiones ante la acelerada automatización de actuaciones administrativas", Revista jurídica de Asturias nº45/2022, pp. 118-124.

[94] La Sentencia del Juzgado Central de lo Contencioso-Administrativo nº 8 de 30 de diciembre de 2021, que se puede leer en la página web de la Fundación Civio, rechazó el recurso por entender que deben primar razones de seguridad. Este asunto fue recurrido en apelación,

factores considerados y del peso atribuido a cada uno de ellos. Digo constancia reforzada porque no basta con que muestre los motivos de la decisión. Se requiere además que a partir de ella sea posible verificar que esos motivos se corresponden con los prefijados en la ley y conocer qué peso se ha otorgado a los distintos elementos de juicio, en especial a los aportados por los interesados. Los algoritmos administrativos deben permitir al usuario este dominio agravado[95] al objeto de que los afectados por su uso puedan contestarlo y el resto de ciudadanos u organizaciones satisfacer su derecho a la transparencia.

Distinto es el caso de los algoritmos que, en lugar de aplicar las reglas predeterminadas en las normas, identifican a partir del análisis de datos los criterios que rijan la actuación administrativa. Estos algoritmos son opacos a priori, en el sentido que funcionan con cierta autonomía y por ello ni siquiera es predecible su resultado. Pero no lo son a posteriori, siempre que la configuración del algoritmo deje constancia de su funcionamiento. La transparencia de estos sistemas automatizados debe asegurarse ex post.

Este tipo de algoritmos se adecúan mejor al ejercicio de potestades discrecionales. El principio de legalidad impone que la ley atribuya potestades a la Administración y le señale parámetros de desempeño, pero admite que la propia ley le confiera ciertos espacios libres de condicionamiento, en los cuales puede elegir entre distintas opciones válidas en Derecho. Desde el punto de vista del principio de legalidad nada obsta a que el criterio elegido venga determinado usando sistemas automatizados, incluso cuando éstos expresen directamente la decisión o influyan en ella de un modo determinante. Y además, sobre todo cuando se trata de combinar muchos factores en la decisión, puede ser muy aconsejable por razones de eficacia, dada las capacidades que las tecnologías de inteligencia artificial están demostrando para identificar patrones útiles que de otro modo resultan inadvertidos.

No obstante, al utilizar algoritmos autónomos en estos casos deben tenerse presentes otras perspectivas ligadas a los derechos de los ciudadanos. La primera es que el ejercicio de las potestades discrecionales conlleva derechos específicos de los ciudadanos, como el de participar aportando su visión de cómo ha de ser la actuación administrativa y el derecho a que la Administración motive reforzadamente porqué escogió una determinada solución. Estos derechos no deben sufrir detrimento porque la decisión discrecional se adopte mediante un sistema automatizado o con apoyo en él. Estos sistemas deben configurarse oportunamente para recoger - y considerar de forma principal - las alegaciones de los ciudadanos, debiendo dejarse constancia de ello. Y de otro lado han de poner especial hincapié en que su funcionamiento y resultado sea transparente ex post, estableciéndose estándares que se correspondan con la motivación reforzada propia de las decisiones administrativas discrecionales.

La segunda es que, aunque las decisiones discrecionales como género son susceptibles de ser adoptadas mediante algoritmos, se pueden establecer excepciones. Es conocido que el artículo 35a VwVfg prohíbe que se dicten actos administrativos enteramente automatizados cuando "exista discrecionalidad o margen de apreciación". Sin embargo,

[95] En España no hay ninguna norma que exija que los algoritmos administrativos tengan una configuración determinada, permitiendo un cierto "dominio agravado" sobre los mismos. El art. 21 de la Loi n° 2018-493, citada *supra*, sí lo hace. Requiere, respecto de las decisiones administrativas automatizadas, que el responsable del tratamiento garantice "un control sobre el tratamiento algorítmico y sus desarrollos". Y no cualquier control, sino el que le coloque en grado de informar, en detalle y de un modo inteligible, al afectado de la manera en que el tratamiento le afecte.

este criterio, que ha sido acogido en la doctrina española si bien que recalcando que solo opera respecto de tratamientos enteramente automatizados[96], no tiene reflejo en la legislación española ni tampoco en otras de nuestro entorno. La legislación francesa, a pesar de disponer de un precepto específico respecto de las decisiones administrativas "basadas exclusivamente en tratamientos automatizados", no prevé tampoco una prohibición de este tipo.

Otra cosa sería que algunas decisiones discrecionales por razón de versar sobre las materias más sensibles para las personas quedasen reservadas a los humanos y excluidas del uso de sistemas automatizados. Ello exigiría una decisión legislativa expresa que así lo previese que podría apoyarse en cuanto a la determinación de estos supuestos excepcionales en la doctrina científica que resalta que son los humanos y no las máquinas quienes están dotados de empatía, justificando en esta condición exclusiva que sean los primeros los únicos que pueden adoptar ciertas decisiones[97].

V.- LA DEFINICION DE LOS DERECHOS EN EL ESPACIO DIGITAL. ESPECIAL REFERENCIA A LA PROPUESTA DE REIA.

El devenir de la sociedad digital está siendo una historia de innovación vertiginosa. Cada pocos años aparecen nuevas tecnologías digitales sobre cuya base prestar servicios que antes no eran posibles o safisfacer necesidades preexistentes con mayor eficacia. Desde algunos ángulos se argumenta que este proceso se beneficia de la situación de desregulación que caracteriza al espacio digital. La heteroimposición de reglas, aunque persiga el efecto beneficioso de conformar una sociedad digital en la que las personas tengan derechos, lastrará la innovación y en un mundo globalizado como el actual empujará la inversión y el progreso hacía otros países donde prevalezca la desregulación. En fin, una constante en el desarrollo de la sociedad digital es la pugna entre innovación y regulación.

Estados Unidos ha venido afrontando este dilema privilegiando la innovación y reduciendo al mínimo la regulación impuesta desde los poderes públicos, en una suerte de "digital Laissez-Faire"[98]. En materia de protección de datos personales la Federación aún no ha promulgado una ley general que establezca principios obligatorios y atribuya derechos a los titulares de los datos[99]. Estas cuestiones se dejan a la autorregulación de las empresas responsables de los tratamientos. Otra fórmula de desregulación afecta a las empresas que prestan servicios de intermediación en internet, pues desde la Digital

[96] Huergo Lora, A., "Administraciones Públicas e inteligencia artificial: ¿más o menos discrecionalidad?", Cronista, 96-97, pp. 93 a 95, que destaca la diferencia entre utilizar la inteligencia artificial y hacerlo a través de un tratamiento enteramente automatizado.

[97] Ponce Solé, J, op. cit. 60 a 63. No obstante, este autor hace en la p. 63 un entendimiento amplio del concepto "empatía" ("La empatía está ínsita en el principio y derecho de buena administración" y "esa falta de empatía de las máquinas es la que las hace inadecuadas para adoptar decisiones automatizadas con margen de valoración que afectan a seres humanos") que pareciera hacer necesario que todas las decisiones discrecionales que "afectan a seres humanos" quedaran afectadas por esta reserva de humanidad.

[98] Vida Fernández, J., op. cit. P. 7

[99] Algunos Estados sí lo han hecho, pero muy recientemente. El ejemplo más importante es la California Consumer Privacy Act of 2018, que ha entrado en vigor el 1 de enero de 2023.

Millenium Copyright Act de 1998[100] disfrutan de una exención de responsabilidad siempre que acrediten no conocer los contenidos que sus usuario distribuyen en internet gracias a sus servicios de intermediación. Esta exención de responsabilidad o "safe harbor" fue asumida en Europa en el artículo 15 de la Directiva de comercio electrónico[101], que en parte resulta modificado por el Reglamento UE de Servicios Digitales aprobado en 2022[102], mediante el que se exigen a estos intermediarios comportamientos activos dirigidos a proteger los derechos fundamentales de las personas en el espacio digital que ellos controlan.

El enfoque de la Unión Europea en la regulación del espacio digital ha sido notablemente diferente. Se ha caracterizado por una combinación equilibrada de regulación pública y autorregulación privada. El poder público ha impuesto una serie de principios básicos y ha reconocido un conjunto de derechos a los ciudadanos, pero la precisión última del régimen jurídico aplicable se ha remitido a la autorregulación de los sujetos privados que actúan en el espacio digital. Se da paso a la autorregulación, dice PIÑAR MAÑAS, "por un lado para permitir la adaptación de la regulación a la realidad del caso concreto, sin la rigidez de la heteroregulación sometida a rigurosos procesos de elaboración y aprobación de la norma; por otro para dar entrada a la participación responsable de los diferentes actores en la regulación del entorno en que operan"[103]

El ejemplo paradigmático nos viene dado por el RGPD, que ha incorporado en su artículo 24 el principio de responsabilidad proactiva como eje maestro de su regulación. Los principios esenciales a que ha de someterse todo tratamiento de datos personales se establecen de un modo obligatorio en el artículo 5 y ello se hace identificándolos mediante reglas muy abiertas (como tratamiento lícito, leal y transparente, minimización de datos, exactitud de los datos y otras similares). El amplio margen que resta se confía a los responsables del tratamiento de los datos, que en virtud del artículo 24 deberán evaluar la naturaleza y riesgos del tratamiento concreto que realizan y, adaptando los principios enunciados con carácter general en la norma, disponer las medidas organizativas y técnicas que sean necesarias para atender a ellos. En definitiva, aunque todos los titulares de datos personales tienen derecho a que éstos sean tratados conforme a los principios de lealtad, licitud y transparencia (y otros más específicos previstos en el artículo 5), la extensión precisa de sus derechos deberá buscarse en la autorregulación que cada responsable de tratamiento haga a la vista de su especial situación. Este principio de responsabilidad proactiva combina este amplio margen de maniobra en la definición del régimen jurídico que se concede al responsable del tratamiento con una obligación correlativa de un alcance igualmente extenso, según la que le resultaría exigible no solo

[100] Digital Millennium Copyright Act (DMCA), Pub. L. No. 105-304, Sección 512

[101] Directiva 2000/31/CE del Parlamento Europeo y del Consejo, de 8 de junio de 2000, relativa a determinados aspectos jurídicos de los servicios de la sociedad de la información, en particular el comercio electrónico en el mercado interior (Directiva sobre el comercio electrónico).

[102] Reglamento (UE) 2022/2065 del Parlamento Europeo y del Consejo, de 19 de octubre de 2022, relativo a un mercado único de servicios digitales y por el que se modifica la directiva 2000/31/CE.

[103] Piñar Mañas, J.L., *Derecho e innovación tecnológica. Retos de presente y futuro*, CEU Ediciones, Madrid, 2018, p. 19.

que adoptase medidas que resulten suficientes sino también que pueda demostrarlo cuando la autoridad de control se lo reclame[104].

La pugna entre, por un lado, la desregulación que reclama el mercado para innovar sin ataduras y, por otro, la regulación que asegure que la construcción que se hace del espacio digital mantiene a la persona en el centro y protege sus intereses a través del reconocimiento de derechos vive un nuevo episodio en el terreno de la inteligencia artificial. La normativa europea que está en proceso de elaboración vuelve a confiar en los sujetos privados que actúen en este ámbito, particularmente en los proveedores de los sistemas de inteligencia artificial, para que evalúen los riesgos que conllevan los sistemas de los que son responsables. Así se desprende del artículo 9 de la Propuesta de la Comisión Europea, que como uno de los requisitos que se exigen a los sistemas de IA de alto riesgo, ordena que "se establecerá, implantará, documentará y mantendrá un sistema de gestión de riesgos". A ello se une que, como parte del compromiso político alcanzado por los colegisladores a principios el 9 de diciembre de 2023, se incorporará al texto legislativo, de un lado, la obligación de los usuarios de someter los sistemas de IA de alto riesgo a una evaluación de impacto sobre los derechos humanos y, de otro, la obligación de evaluar los riesgos de los sistemas de IA de uso general antes de que se pongan en funcionamiento.

No obstante, se aprecian algunas importantes diferencias entre el modelo regulatorio del RGPD y el que ahora se recoge en el proyecto de normativa sobre IA. La primera es que el art. 24 RGPD era una regla general, aplicable a todos los responsables de tratamiento independientemente de su naturaleza y de los eventuales peligros que conllevase. Por el contrario, la normativa de inteligencia artificial remite a esta fórmula de autorregulación únicamente respecto de un conjunto de sistemas de IA, los de alto riesgo. Aunque también es cierto que los demás, los que no son de alto riesgo, o bien están prohibidos o bien no están sometidos apenas a regulación alguna, tampoco a supuestos de heteroregulación salvo requerimientos muy puntuales.

La segunda diferencia es quizá más importante. La legislación en proyecto regula en los artículos 10 a 15 los requisitos esenciales que deben respetar los sistemas de IA de alto riesgo y lo hace con bastante precisión, con mucha más que la utilizada por el artículo 5 RGPD para definir los principios a los que somete a todo tratamiento de datos personales. Es más esos preceptos remiten a su vez a la concreción posterior de esos requisitos esenciales mediante la elaboración y aprobación de especificaciones técnicas. La regulación detallada que se contiene en los artículos 10 a 15 del Reglamento en tramitación y la mayor concreción adicional que provenga de las especificaciones técnicas que se aprueben en el futuro agotarán gran parte de la materia, de suerte que poco espacio restará para la autorregulación[105]. No obstante, no debe perderse de vista que la definición de las especificaciones técnicas participan de un modo principal los organismos de normalización, que son entidades privadas en las que tienen presencia

[104] Una descripción muy elocuente del funcionamiento del principio de responsabilidad proactiva y su centralidad en el régimen previsto por el RGPD, cfr. Piñar Mañas, J.L., "Protección de datos. Las claves de un derecho fundamental imprescindible", *Cronista*, 88-89, p.13.

[105] De Gregorio, G. y Dunn, P., "The european risk-based approaches: connecting constitutional dots in the digital age", *CMLR*, 59, pp. 488 a 493.

relevante los empresas que protagonizan la producción del producto o la prestación del servicio objeto de la normalización[106].

Con todo la diferencia más relevante es la tercera porque se separa de la segunda dimensión del principio de responsabilidad proactiva, la que vincula al responsable del tratamiento a demostrar que las medidas que adoptó eran las que exigía la naturaleza y riesgos del tratamiento concreto que realizaba. Nótese que el artículo 19 del Reglamento IA en tramitación dispone que "los proveedores de sistemas de IA de alto riesgo se asegurarán de que sus sistemas sean sometidos al procedimiento de evaluación de la conformidad oportuno, de conformidad con el artículo 43, antes de su introducción en el mercado o puesta en servicio". Esta "evaluación de la conformidad" al igual que la remisión a la concreción de los requisitos legales mediante especificaciones técnicas son elementos característicos de lo que se conoce como regulación armonizadora de "nuevo enfoque". Esta técnica regulatoria, que inicialmente se aplicaba a la seguridad de los productos y ahora se ha extendido a los servicios, incluidos los que se ubican en el espacio digital, consiste en que la legislación europea se limita a señalar unos requisitos esenciales para cuyo cumplimiento se establecen las correspondientes especificaciones técnicas. La normalización (definición de las especificaciones técnicas) puede realizarse mediante distintos cauces, entre los que destaca la elaboración por los organismos europeos de normalización de "normas técnicas armonizadas", las cuales a pesar de su origen privado producen un efecto jurídico-público consistente en que los productos y servicios que se ajusten a ellas gozan de una presunción de conformidad con los requisitos previstos en la legislación armonizadora de nuevo enfoque[107].

La diferencia es sustancial con el principio de responsabilidad proactiva. Este confía al responsable del tratamiento decidir qué medidas son necesarias dada la naturaleza y riesgos concretos que se asocian a su tratamiento y correlativamente le somete a la obligación de demostrar qué dichas medidas eran las adecuadas. En el esquema de la legislación armonizadora de nuevo enfoque, dado que basta con producir el producto o prestar el servicio con arreglo a las "normas técnicas armonizadas" para que se presuma la conformidad con los requisitos legales, el proveedor o usuario de un sistema de IA de alto riesgo no tendrá la obligación de demostrar que las medidas que adoptó eran las estrictamente necesarias. Simplemente se presume que lo eran.

Ahora bien, el esquema de la legislación armonizadora "nuevo enfoque", que implica precisión en las exigencias que hay que cumplir y presunción de conformidad, está pensado y ha demostrado su utilidad para afrontar riesgos para la salud o para la seguridad. El otro enfoque, el de la responsabilidad proactiva, que combina mayor margen de apreciación con la obligación de demostrar el cumplimiento, se ha experimentado en relación a riesgos para los derechos fundamentales, en especial el de protección de datos pero también otros asociados al espacio digital[108].

Los sistemas de inteligencia artificial tienen una dimensión de producto o servicio que es susceptible de poner en riesgo la salud o la seguridad (e.j. coches autónomos). Pero tienen

[106] Véase Álvarez García, V, Las normas técnicas armonizadas (una peculiar fuente del Derecho europeo), Madrid 2020.
[107] Álvarez García, V., y Tahiri Moreno, J., "La regulación de la inteligencia artificial en Europa a través de la técnica armonizadora del nuevo enfoque", *Revista General de Derecho Administrativo*, 63, 2023
[108] De Gregorio, G. y Dunn, P., op. cit., pp. 485 y 486.

otra de afectación a los derechos fundamentales de las personas, que en el compromiso político de los colegisladores de 9 de diciembre de 2023 se refleja en la imposición de que los sistemas de IA de alto riesgo se sujeten a una evaluación de impacto en los derechos fundamentales. No está claro que esta segunda vertiente pueda ser abordada con éxito con el esquema de la legislación armonizada de "nuevo enfoque" y no solo porque los comités de los organismos de normalización se integran solo por expertos en aspectos técnicos del producto o servicio, lo que obviamente podría cambiar con la incorporación de expertos en derechos fundamentales. Quizá la materia regulada no se presta a un grado de predeterminación normativa tal (por lo que tiene de situación subjetiva) como para poder asociarle a su cumplimiento una presunción de conformidad. O quizá simplemente la materia derechos fundamentales no se presta a una regla de presunción de conformidad.