

INTELIGENCIA ARTIFICIAL DESDE EL DISEÑO. RETOS Y ESTRATEGIAS PARA EL CUMPLIMIENTO NORMATIVO

Ricard Martínez Martínez*

Resumen

En este artículo se analizan las repercusiones de la inteligencia artificial en el derecho y los derechos fundamentales. Frente a visiones negativas, reactivas o catastrofistas se propone un enfoque basado en el cumplimiento normativo desde el diseño inspirado en la metodología de la protección de datos desde el diseño y por defecto.

Palabras clave: inteligencia artificial; protección de datos desde el diseño y por defecto; diseño del cumplimiento normativo; ética; derechos humanos; derechos fundamentales; derecho constitucional.

DESIGNING ARTIFICIAL INTELLIGENCE. CHALLENGES AND STRATEGIES FOR ACHIEVING REGULATORY COMPLIANCE

Abstract

This article analyses the repercussions of artificial intelligence in terms of law and fundamental rights. Amidst negative, reactive and alarmist outlooks, we propose an approach inspired by the data protection by design and default methodology, in which emphasis is placed on regulatory compliance from the very design of such technologies.

Key words: artificial intelligence; data protection by design and default; design for regulatory compliance; ethics; human rights; fundamental rights; constitutional law.

* Ricard Martínez Martínez, director de la Cátedra de Privacidad y Transformación Digital. Profesor ayudante doctor en el Departamento de Derecho Constitucional, Ciencia Política y de la Administración de la Universidad de Valencia. Facultad de Derecho, avinguda dels Tarongers, s/n, 46002 València. ricard.martinez@uv.es, [@ricardmm](https://twitter.com/ricardmm).

Artículo recibido el 01.05.2019. Evaluación ciega: 24.05.2019 y 24.05.2019. Fecha de aceptación de la versión final: 26.05.2019.

Citación recomendada: Martínez Martínez, Ricard. (2019). Inteligencia artificial desde el diseño. Retos y estrategias para el cumplimiento normativo. *Revista Catalana de Dret Públic*, (58). 64-81. <https://doi.org/10.2436/rcdp.i58.2019.3317>

Sumario

- 1 ¿Utopía o distopía?: una introducción a la problemática de la inteligencia artificial
- 2 Los procesos de innovación tecnológica y su impacto en nuestro mundo
- 3 Un impacto global en el derecho y los derechos
 - 3.1 ¿Están en juego los derechos fundamentales?
 - 3.2 El impacto en el derecho
- 4 De la ética de los derechos humanos al derecho desde el diseño
 - 4.1 Una ética de los derechos humanos
 - 4.2 Una estrategia inmediata
 - 4.3 El marco jurídico de la protección de datos como modelo inspirador
 - 4.4 El cumplimiento normativo desde el diseño
 - 4.4.1 Formar ética y jurídicamente al equipo
 - 4.4.2 Identificar los requerimientos del proyecto
 - 4.4.3 Integrar las normas en el diseño y documentar el cumplimiento
 - 4.4.4 De la programación a la comercialización
 - 4.4.5 Un ciclo que se mantiene durante toda la vida de una IA
- 5 Una breve conclusión

Bibliografía

1 ¿Utopía o distopía?: una introducción a la problemática de la inteligencia artificial

En el primer capítulo de *Vida 3.0* (Tegmark, 2018: 13-39) se propone una utopía distópica que pone de manifiesto las contradicciones que enfrenta el análisis jurídico de la inteligencia artificial (IA).

La división de investigación de una gran empresa consigue desarrollar una IA de propósito general. Mediante la construcción de una red neuronal, el equipo de investigadores logra un objetivo aparentemente inalcanzable y debe adoptar una serie de decisiones trascendentales.¹ La primera, mantener en secreto el hallazgo y comenzar a trabajar aceleradamente en su desarrollo. A partir de aquí, los resultados que ofrece la IA cambian el mundo desde una ética de la utopía científica. Se trata de erradicar las guerras, liberar al hombre de tareas mecánicas o promover la justicia social. Para ello, “primero hay que ganar dinero”. Así que el primer criterio que guía la máquina resulta bien conocido: obtener ingresos a partir de negocios que permitan un proceso rápido de acumulación de capital en la industria del entretenimiento —redes sociales, producción audiovisual, etc.—, en un modelo que nos resulta, sin duda, familiar.

Una vez conseguido el objetivo, la máquina decide una estrategia de compra y control de medios de comunicación y de inversión en la promoción de políticos alineados con valores como la paz, la libertad, la igualdad y la garantía de los derechos fundamentales. El resultado final permite instaurar un gobierno democrático mundial y el nacimiento de una nueva era para una humanidad en paz en un mundo sostenible pero tutelado por la máquina.

Parece muy evidente cómo respondería la población a una encuesta en la que se preguntase algo del tipo: “¿Desea usted un mundo en paz, ecológicamente sostenible, en el que las máquinas trabajen por usted?”. Y, sin embargo, la propuesta de Tegmark resulta, sin duda, inquietante. En primer lugar, porque supone asumir nuestra inferioridad frente a la máquina. Si algún día logramos crear una inteligencia de propósito general, una inteligencia “fuerte”, superaría de inmediato las barreras de su programación. Podría mejorar su programa, multiplicaría de modo exponencial su capacidad de computación y diseñaría mejoras en *hardware* y *software*.² Exploraría nuevos modelos como la computación cuántica o las herramientas que ofrece la bioinformática basada en el ADN. Ello supondría de inmediato sobrepasar las condiciones de la ley de Moore³ e iniciar una nueva era para la tecnología.

Ahora bien, en la uto-distopía de Tegmark, lo que resulta inquietante es cómo podría alcanzarse el bien común para los humanos, pero sin los humanos. Sería un nuevo despotismo ilustrado en el que las decisiones corresponderían a una oligarquía de unos pocos científicos y técnicos de los que la propia máquina podría independizarse. Y, aunque se lograra una mejora de la condición humana regida por valores humanistas y democráticos, ¿podríamos afirmar que se sustenta en una concepción de la dignidad humana y la libertad, coherente con nuestra visión de los derechos fundamentales?

2 Los procesos de innovación tecnológica y su impacto en nuestro mundo

En reiteradas ocasiones se ha puesto en cuestión el modelo actual de innovación técnica (Martínez, 2019a). En este sentido, la uto-distopía narrada se ajusta fielmente al modelo que viene rigiendo internet desde inicios de este siglo, e incluso antes. En la era de la informática, tanto la investigación científica como la evolución tecnológica han experimentado un proceso de progresiva aceleración regida por criterios de celeridad, eficiencia y rentabilidad. Es un modelo de innovación disruptiva que obliga a los directivos de las empresas a resolver distintos dilemas estratégicos (Christensen, 1997: 225 y ss.) ante decisiones de inversión que pueden cambiar por completo mercados enteros. Se trata de procesos de cambio que requieren de una agilidad desconocida. Ya no bastan las herramientas de la planificación tradicional, los procesos de decisión deben ser muy ágiles y exigen un nuevo perfil de directivo altamente creativo, incluso dispuesto a

1 Este objetivo se ha manifestado muy esquivo desde que se formulase en la Conferencia de Darmouth en 1956 y presupondría que una máquina tendría la capacidad de operar y razonar como un ser humano pero dotada de enormes capacidades computacionales (McCarthy, Minsky, Rochester y Shannon, 1955).

2 Es un escenario de carencia de control perfectamente previsible. Véase Jiménez (28 de julio de 2017).

3 En 1965 Gordon Moore afirmó en la revista *Electronics* que el número de transistores por pulgada en circuitos integrados y la densidad de los datos se doblarían aproximadamente cada 18 meses. Esta ley ha sido puesta a prueba y puede ser desbordada si se consigue desarrollar la llamada “computación cuántica”.

“desaprender” (Marcet, 2018: 55-72). Esto, en el contexto de la inteligencia artificial, puede generar sesgos poco aceptables (O’Neil, 2016: 3).

Y en este contexto se formuló uno de los grandes, y por ende más peligrosos, axiomas del mundo de internet, el que afirma que, para innovar, “hay que romper cosas” (Zuckerberg citado por Blodget, 2009).⁴ Este modo de hacer las cosas ha conformado nuestro mundo en los últimos decenios. La regla podría formularse de un modo sencillo: “Si tienes una idea, si esta idea es negocio, si se puede hacer, no lo dudes ‘hazla’”. Es fundamental llegar primeros, ocupar el nicho de mercado y conseguir un número significativo de usuarios a cualquier precio. No se discute aquí el valor de la innovación ni su contribución al desarrollo humano; muy al contrario, lamentarse y demonizar la tecnología constituiría un grave error.

Sin embargo, los procesos basados en romper cosas no consideran más regla jurídica que la de entender que lo que no se halla prohibido se entiende expresamente permitido. La experiencia de los últimos decenios muestra un modelo que persigue como objetivo esencial maximizar el beneficio con independencia del daño que se pudiera causar. En la práctica, los procesos de acumulación de capital que acompañan a las empresas que alcanzan el éxito han permitido abordar el pago de sanciones e indemnizaciones, con un impacto relativamente bajo en relación con los beneficios alcanzados, hasta el punto de que de algún modo se incorporan a la estructura de costes (Martínez, 2016).

Los procesos de innovación disruptiva definen puntos de inflexión a partir de los cuales se producen cambios que revolucionan nuestro modo de hacer las cosas e incluso de entender el mundo.⁵ Se apuntan la desaparición del modelo laboral debido a la explosión de servicios gratuitos o de bajo coste en internet, la migración de prestaciones laborales a plataformas de intermediación y un nuevo modelo de profesional cualificado y adaptable, en el que, en el futuro, ser un asalariado será sinónimo de incompetencia y en el que el currículum académico universitario resulta insuficiente (Mallard, 2018: 91-95, 176-179). La empresa tradicional debe reinventarse, cambiar y adaptarse a una nueva normalidad en la que el proceso productivo y el modelo de negocio cambian de un modo radical (Bradley, 2016).⁶ La consultora McKinsey destaca cómo el análisis masivo de datos ha permitido la implantación de herramientas de inteligencia artificial o la favorecerá en un futuro inmediato, lo cual constituye un reto empresarial y macroeconómico estratégico (McKinsey, 2017: 6).⁷

La IA implica un proceso de innovación y cambio permanente que reorientará por completo los procesos de gestión y funcionamiento de las empresas en múltiples dimensiones (McKinsey, 2018). El proceso de digitalización creciente de los sectores público y privado, y las capacidades de analizar los datos mediante herramientas de *machine learning* gracias a las posibilidades de almacenamiento y proceso que ofrecen los entornos de *cloud*, favorecen la migración a un modelo de decisiones basadas en datos. La IA aporta aquí todo su valor, ya sea como herramienta de apoyo a la decisión humana asistida, ya sea como proceso automático que opera ofreciendo directamente servicios. En su *briefing note* de mayo de 2017, McKinsey identifica las posibilidades que el estado actual de la industria ofrece para el avance de la disrupción. En este sentido, resulta particularmente interesante como incluso situaciones de ineficiencia se conciben como oportunidades de mejora, y también lo es el amplio espacio sectorial susceptible de crecer de la mano de la IA.

4 La frase exacta fue: “Move fast and break things. Unless you are breaking stuff, you are not moving fast enough” (Zuckerberg citado por Blodget, 2009).

5 Dobbs, Manyika y Woetzel, editores de McKinsey, apuntaban en 2015 hasta una docena de tecnologías o negocios disruptivos: 1) genómica; 2) ingeniería de nuevos materiales; 3) almacenamiento de energía; 4) nuevos modos de exploración y explotación de hidrocarburos; 5) energías renovables; 6) robótica; 7) vehículos autónomos; 8) impresoras 3D; 9) internet móvil; 10) Internet de los Objetos; 11) *cloud*, computación en la nube; y 12) automatización de los procesos basados en conocimiento. (Dobbs, Manyika y Woetzel, 2015: 40). Cada año distintas organizaciones publican este tipo de listas. Prácticamente en todas, la IA aparece como una de las tecnologías disruptivas y, a la vez, como instrumento indispensable para el desarrollo de otras.

6 Bradley analiza estos cambios proyectándolos sobre un tipo de empresa, los medios de comunicación, que son, sin duda, el paradigma de los cambios radicales que las tecnologías de la información y las comunicaciones imponen a la economía.

7 “The application of AI and the automation of activities can enable productivity growth and other benefits not just for businesses, but also for entire economies. [...] At a macroeconomic level, based on our scenario modeling, we estimate automation alone could raise productivity growth on a global basis by 0.8 to 1.4 percent annually. In short, businesses and the economy need the productivity boost from automation”.

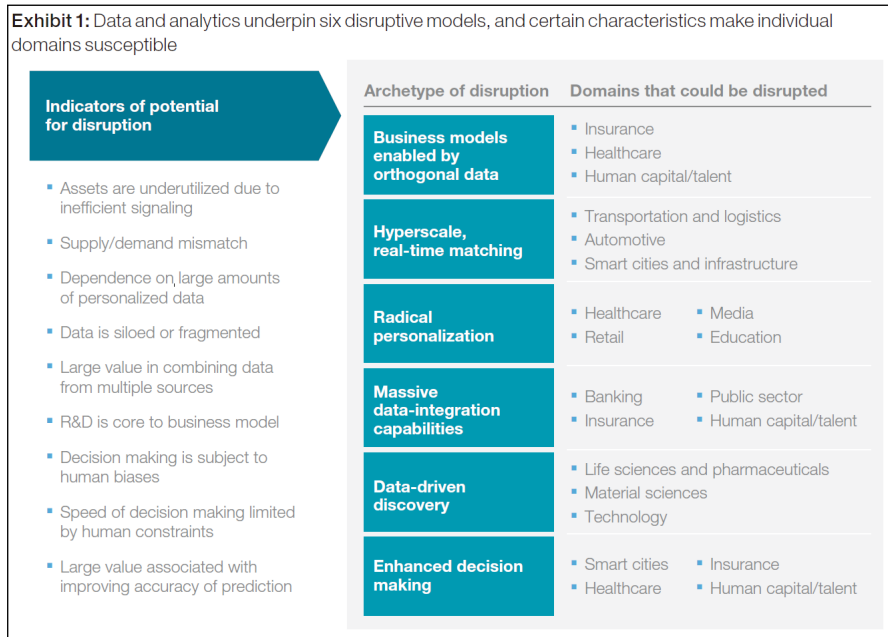


Figura 1. Fuente McKinsey Global Institute (McKinsey, 2017: 2).

El referido informe afirma que economías de países como Estados Unidos, Francia o Alemania explotan a día de hoy el 18, 12 y 10 % del potencial que ofrece la digitalización, mientras que el flujo internacional de datos se ha multiplicado un 45 % desde 2005. El rango de actividades humanas susceptibles de ser gobernadas desde los datos y con procesos automatizados crecerá día a día y transformará profundamente la educación, la economía, la salud o el empleo. Las previsiones de McKinsey señalan que el volumen de los puestos de trabajo susceptibles de una completa automatización todavía puede ser relativamente bajo. Sin embargo, el 60% de los puestos admite rangos de automatización de hasta un 30 % de la actividad.

En un estudio de mayor calado (McKinsey, 2018: 30, 41-43) se profundiza en un análisis que ofrece unas perspectivas significativas de crecimiento y cambio del modelo económico acompañadas de un alto riesgo de brechas entre países y personas. Estas asimetrías derivan de: 1) las capacidades de investigación e innovación de los países —y aquí el liderazgo pertenece a Estados Unidos y China—; 2) de la fortaleza de la economía y su preparación para un mundo interconectado —que caracteriza a Francia, Alemania, Reino Unido, Suecia o Finlandia junto a países como Corea del Sur o Japón—; 3) de países que se encuentran en el umbral de adopción y desarrollo de esta tecnología; y 4) países, como España, por debajo del umbral. Desde el punto de vista de las personas, se apunta a la existencia de brechas significativas entre trabajadores cuya formación se oriente a las competencias necesarias para esta nueva economía, que podrán evolucionar en sus puestos a medida que estos se automaticen, o acceder a los nuevos nichos de empleo, y aquellos que deberán aspirar a empleos residuales en términos de tecnificación o a depender de subsidios de renta mínima vital garantizada.

3 Un impacto global en el derecho y los derechos

3.1 ¿Están en juego los derechos fundamentales?

Al realizar un repaso de las tecnologías o procesos afectados por la IA, podríamos apreciar riesgos y oportunidades (Martínez, 2019a: 264-267). El grupo de trabajo sobre “Inteligencia Artificial y Desarrollo Humano” creado por la Cátedra de Privacidad y Transformación Digital Microsoft-Universitat de Valencia, ha identificado algunos de ellos en sus sesiones.

▪ Riesgo de discriminación digital

La brecha digital puede verse acrecentada con motivo de la IA. La asimetría entre países en una tecnología destinada a cambiar el modelo económico global puede acrecentar todavía más las diferencias norte-sur, y

entre países avanzados en términos digitales y países no avanzados. Esta brecha se proyecta sobre el entero sistema de derechos fundamentales en la medida en que puede ser un elemento generador de dependencia económica, pobreza, exclusión digital y/o exclusión de la economía basada en datos.

▪ Riesgo de discriminación laboral

La desaparición de los puestos susceptibles de ser asumidos por máquinas (Ford, 2016: 359-365) plantea obviamente la expulsión del mercado de trabajo de millones de personas y puede plantear amenazas a la estabilidad social y, con ello, a nuestro sistema de libertades (Mercader, 2018: 619-620). Es un lugar común vaticinar la desaparición de empleos como el del conductor, cajero de supermercado o asistente telefónico, pero también otros de naturaleza administrativa. Así, la IA, ya sea porque es capaz de automatizar o asumir un porcentaje significativo de tareas, ya sea porque se erige en una herramienta esencial de soporte al ejercicio profesional, puede suponer la exclusión del mercado de trabajo de aquellas personas que sean incapaces de adaptar sus conocimientos y competencias a las exigencias de la interacción con la IA.

▪ El impacto en el derecho a la educación

Este es un riesgo estructural que puede afectar a países enteros. Deriva de la inexistencia, en los planes de formación, de herramientas orientadas al pensamiento computacional y de las habilidades de los estudiantes necesarias para el desarrollo, uso y aplicación de la IA. Por otra parte, un sistema educativo anticuado, con profesores con escasas competencias digitales opera como barrera para la introducción de la IA como soporte al proceso educativo. El desarrollo de estrategias adecuadas sobre la IA en la educación y en la formación continuada constituyen oportunidades para la generación de talento. En aquellos países en los que esto no suceda, el riesgo de discriminación laboral y exclusión social se multiplicará de modo exponencial.

▪ El sesgo como riesgo

Los procesos de decisión basados en la IA dependen de procesos de analítica masiva de datos basados en algoritmos. Es bien sabido que el llamado “*machine learning*” puede funcionar de acuerdo con un propósito previamente definido, o bien sin más finalidad que la de establecer inferencias y correlaciones (López y Messeguer, 2017). En ambos casos depende ineludiblemente de la calidad de los datos y del funcionamiento material del algoritmo. Si este presenta algún tipo de sesgo, puede originar discriminación. Podemos identificar impactos del más variado tipo, a partir de ejemplos realmente existentes:

- a) En procesos de selección de personal o de evaluación de la calidad de la prestación laboral puede producir efectos discriminatorios; por ejemplo, fomentar despidos injustos.
- b) En el ámbito de la seguridad pública se ha demostrado la existencia de programas de asistencia que orientan el desarrollo de la actividad policial a la detención de minorías.
- c) Se ha denunciado que programas de asistencia en procesos para obtener libertades condicionales en Estados Unidos operan sesgadamente favoreciendo la concesión de este derecho a personas de raza blanca.
- d) En el plano del derecho a la tutela judicial efectiva, están apareciendo, en el mercado, programas que son capaces de establecer el porcentaje de éxito que alcanzará una determinada demanda ante un determinado tribunal. Desde el punto de vista de la garantía del juez ordinario predeterminado por la ley, y de la independencia de criterio del juzgador en el ejercicio de la función jurisdiccional, resulta preocupante pensar que existan máquinas capaces de predecir un determinado resultado.

Podríamos hacer crecer la lista de riesgos. En el llamado “marketing neuroemocional” (Byung-Chul, 2014: 61, 74-75; Bauman y Lyon, 2013: 15), la manipulación de las preferencias (Pariser, 2017) ha puesto en alerta a las democracias occidentales tras el caso Cambridge Analytica y ha mediatizado la interpretación de las opciones del legislador español al modificar la LOREG (Martínez, 2019b).⁸ Las posibilidades que ofrece la IA en el campo del *scoring* bancario, o la determinación de riesgos actuariales, apuntan a un mundo en el que

⁸ En efecto, la Ley Orgánica de Régimen Electoral General ha sido modificada por la disposición final tercera de la Ley Orgánica 3/2018, introduciéndose un nuevo artículo 58 bis, cuyo primer inciso señala que “La recopilación de datos personales relativos a las opiniones políticas de las personas que lleven a cabo los partidos políticos en el marco de sus actividades electorales se encontrará amparada en el interés público únicamente cuando se ofrezcan garantías adecuadas”.

el riesgo de discriminación por razones laborales, educativas o de salud podría estar muy presente. Es más, en un escenario teórico podría acabar discriminándose a una persona por su código postal.⁹

Sin embargo, también podemos encontrar oportunidades para los derechos. Así, por ejemplo, los avances que está experimentando la investigación biomédica pueden mejorar significativamente la calidad de la atención al paciente. En el ámbito laboral, la robótica avanzada y las herramientas de IA pueden operar como un factor de liberación de los trabajos repetitivos y de eliminación de los riesgos laborales. En la educación, el uso adecuado de estas tecnologías puede operar incrementando el nivel de calidad de nuestro sistema educativo. Resulta inabordable categorizar todas las repercusiones de la IA en los derechos. Sin embargo, a partir de los antecedentes expuestos y de la literatura disponible no resulta descabellado realizar un pequeño ejercicio para identificar impactos significativos en nuestro catálogo de derechos fundamentales, que podríamos clasificar, de modo un tanto maniqueo, como positivos o negativos. No se trata de un ejercicio sistemático, ni están todos los derechos ni se aprecian todos los impactos. Es, más bien, un pequeño experimento mental que el lector también podría hacer añadiendo supuestos o cambiando de columna algunos de ellos.

Derecho a la vida y derecho a la protección de la salud.	+	<ul style="list-style-type: none"> ▪ Medicina asistida por inteligencia artificial. ▪ Desarrollo de nuevos materiales (biónica, biomecánica): un humano robotizado. ▪ Alargamiento de la vida humana, la llamada “singularidad”.
	-	<ul style="list-style-type: none"> ▪ Discriminación por razón de la salud.
Libertad ideológica y participación política	+	<ul style="list-style-type: none"> ▪ Desarrollo de herramientas de participación política y democracia participativa. ▪ Analítica de datos en la prevención de la corrupción política.
	-	<ul style="list-style-type: none"> ▪ Segmentación ideológica de la población. ▪ Manipulación de las preferencias electorales. ▪ Control poblacional y de la conducta.
Libertad y seguridad	+	<ul style="list-style-type: none"> ▪ Herramientas de gestión automática de riesgos y catástrofes. ▪ Identificación y prevención temprana de actos de delictivos.
	-	<ul style="list-style-type: none"> ▪ Discriminación social y demográfica en la acción de las Fuerzas y Cuerpos de Seguridad.
Vida privada	+	<ul style="list-style-type: none"> ▪ Herramientas de gestión del perfil de privacidad. ▪ Gestores capaces de identificar políticas de privacidad y asistir en las decisiones.
	-	<ul style="list-style-type: none"> ▪ <i>Profiling</i>. ▪ Manipulación emocional y de las preferencias.
Libertad de circulación	+	<ul style="list-style-type: none"> ▪ Robotización de la gestión de la circulación del tráfico y la seguridad vial.
	-	<ul style="list-style-type: none"> ▪ El Estado decide nuestros desplazamientos.
Derecho a la información y libertad de expresión	+	<ul style="list-style-type: none"> ▪ Herramientas de análisis, síntesis, localización y lectura automatizada.
	-	<ul style="list-style-type: none"> ▪ Inteligencias sintéticas redactan las noticias.

⁹ Las variables socioeconómicas, demográficas, climáticas o de contaminación geográficamente vinculadas ofrecen potentísimos modelos de análisis. Por ejemplo, una compañía aseguradora podría cruzar datos relativos al clima, la contaminación y los hábitos de vida de los códigos postales en los que nació, se educó y vive un cliente, para singularizar el riesgo. Se pasaría así de un riesgo mutualizado a otro singularizado, en el que el origen social y la residencia concedieran ventajas a las capas más favorecidas de la sociedad.

Tutela judicial efectiva	+	<ul style="list-style-type: none"> ▪ La analítica de datos permite el análisis de documentación en procesos complejos (delincuencia económica o corrupción). ▪ Los sistemas de IA dotan de mayor robustez al tráfico jurídico y evitan conflictos. ▪ Sistemas robotizados de resolución de conflictos.
	-	<ul style="list-style-type: none"> ▪ Los sistemas recomiendan “jueces” y ofrecen ventajas competitivas a demandantes o demandados.
Garantías en el proceso penal	+	<ul style="list-style-type: none"> ▪ Mayor capacidad para la generación de evidencias digitales.
	-	<ul style="list-style-type: none"> ▪ Vuelta a sistemas de adopción de medidas cautelares basadas en el perfil del acusado o demandado. ▪ Sesgo en sistemas automatizados de asistencia decisonal en concesión de beneficios penitenciarios.
Derecho a la educación	+	<ul style="list-style-type: none"> ▪ Nuevos modelos de aprendizaje. ▪ Capacidad de generar talento y oportunidades. ▪ Herramientas de asistencia a la orientación vocacional.
	-	<ul style="list-style-type: none"> ▪ Discriminación por carencias formativas o tecnológicas.
Sostenimiento de los gastos públicos	+	<ul style="list-style-type: none"> ▪ Modelos decisonales basados en datos. ▪ Analítica de datos en la gestión del gasto público. ▪ Analítica orientada a la prevención y detección de la prevención y el fraude fiscal.
	-	<ul style="list-style-type: none"> ▪ Limitaciones o impactos en la vida privada de funcionarios y políticos objeto de análisis.
Propiedad privada	+	<ul style="list-style-type: none"> ▪ Aparición de una nueva economía. Nuevos nichos de negocio y empresa. ▪ Gobernanza de la empresa. ▪ Política económica basada en analítica de datos.
	-	<ul style="list-style-type: none"> ▪ Discriminación en el acceso al crédito. ▪ Monopolios empresariales de la información.
Derecho al trabajo	+	<ul style="list-style-type: none"> ▪ Nuevos perfiles y puestos de trabajo. ▪ Eliminación de puestos con riesgo.
	-	<ul style="list-style-type: none"> ▪ Incremento significativo del paro. ▪ Discriminación en el acceso al trabajo.

3.2 El impacto en el derecho

Desde los años 70, la reflexión sobre el impacto de la informática en los derechos fundamentales ha sido una constante. Y la historia parece esforzarse en demostrarnos que no se pueden poner puertas al campo (Trías, 1992).

La experiencia regulatoria adquirida en este ámbito prefigura de modo muy claro cuáles puedan ser los impactos de la IA en los ordenamientos nacionales, y en el internacional. En los epígrafes precedentes hemos podido apreciar como los procesos que han guiado el desarrollo de las tecnologías de la información se han caracterizado por no considerar el cumplimiento normativo como un elemento previo para el diseño, y parecen haber sido ajenos, más allá de lo imprescindible, a consideraciones de orden ético o jurídico. Cualquier bien o valor jurídico podía ser sacrificado en el altar de la burbuja tecnológica. Se desarrolló un

modelo de negocio de carácter transnacional que ha buscado los entornos más flexibles de regulación y más laxos en materias como la protección de datos. Paralelamente, se ha invocado, precisamente, el criterio de un establecimiento geográfico previamente diseñado, como argumento de exclusión del control administrativo y jurisdiccional.

Por último, se ha construido un discurso, en el que se ha tratado de demonizar al derecho como una herramienta que frena la evolución tecnológica y el desarrollo económico. Desde este punto de vista, se le atribuye un efecto paralizante de la tecnología, a la que se atribuye un carácter intrínsecamente benigno. Adicionalmente, se ha señalado la incompatibilidad de los tiempos del legislador con los de la tecnología, considerando, por ello, que el desarrollo e implementación de estos nuevos modelos podía realizarse al margen del derecho básicamente por dos razones. La primera, porque los ordenamientos jurídicos son anticuados, pensados para un mundo analógico, y, por tanto, inaplicables al mundo digital. En segundo lugar, por considerar que la tecnología no puede esperar al legislador.

Sin embargo, cuando el legislador trata de realizar su trabajo, los argumentos previamente señalados evolucionan en un doble sentido. Primero, profundizando en la idea de que cualquier regulación de las tecnologías de la información puede ser en sí mismo negativa, y el mercado debe autorregularse. Por otra parte, desplegando una significativa tarea de *lobby*, lo cual no deja de ser un interés perfectamente legítimo, que materialmente retrasa ese proceso legislativo al que precisamente se acusa de lento.

4 De la ética de los derechos humanos al derecho desde el diseño

El escenario al que se enfrenta el derecho se encuentra cargado de incertezas. No es casual que el impacto de la IA en el derecho privado haya planteado ya una primera reflexión al Parlamento Europeo.¹⁰ En este sentido, basta con explorar intuitivamente el ámbito del derecho de daños para apreciar la complejidad de los retos jurídicos que plantea esta tecnología. Así, por ejemplo, definir qué responsabilidad pueda derivarse de un error en el funcionamiento de una IA puede plantear teorías para todos los gustos. La más sencilla, por obvia y más pragmática, consideraría que el diseñador de la inteligencia artificial y/o la empresa que provee el servicio deberían responder en cualquier caso de los daños causados por su criatura.

Sin embargo, el mundo jurídico no se encuentra exento de teorías que están explorando la naturaleza de la personalidad jurídica de un robot, sea este o no una máquina corpórea. Y tratan de identificar la presencia de una suerte de *tertium genus*. Como en el caso de los animales, se trataría de sujetos de derecho *sui generis*. Y cabría plantearse si existe una responsabilidad distinta de la propia máquina cuando sus decisiones dependen de un juicio que requiera un análisis de su entorno y un proceso de toma de decisiones autónoma.

En este sentido, la existencia de inteligencias artificiales que interactúen con las personas puede plantear cuestiones jurídicas adicionales. En un epígrafe anterior señalábamos como el sesgo podía determinar errores de funcionamiento de la IA que causan discriminación. Pero este sesgo no solo se origina en la programación del algoritmo, en muchas ocasiones depende de los datos con los que se alimenta el sistema cuando aprende. De este modo, una IA que tuviese por finalidad interactuar con terceros prestándoles servicios de atención a través de conversaciones podría ser xenófoba si aprende en un entorno hostil a los inmigrantes o si infiere que las elecciones de la empresa que la usa privilegian a los nacionales. En estos casos, la dificultad de establecer la responsabilidad gana en intensidad; y distintos autores plantean diferenciar entre la responsabilidad de la máquina considerada como sujeto de derecho, la responsabilidad del usuario y la responsabilidad del proveedor. Y, sin embargo, corremos el riesgo de que este tipo de argumentos sirvan como escudo fácil para eludir todo tipo de responsabilidad.

4.1 Una ética de los derechos humanos

Es precisamente aquí donde la investigación debe aportar valor innovando el ordenamiento. La evolución del derecho, la aparición de nuevas figuras jurídicas, de nuevos sujetos y operadores, va a ser fruto de un esfuerzo de análisis y teorización que se va a alargar, sin ningún género de dudas, durante años. Pero, mientras esto sucede, necesitamos ser capaces de superar las barreras a las que nos enfrenta la filosofía propia

¹⁰ Parlamento Europeo (2017).

de un derecho de la Unión Europea que sistemáticamente trata de codificarlo todo. Es un modelo reactivo, extraordinariamente lento y no exento de errores. El primero de ellos, como reiteradamente se ha señalado, consiste en una aproximación a la tecnología centrada en cada avance concreto. Primero regulamos “la informática”, después nos ocupamos de internet, de las redes sociales, del Internet de los Objetos... Y con cada nueva tecnología es necesario regular de modo específico. Y ello tiene un inconveniente manifiesto: la lentitud.

El proceso legislativo debe ser reflexivo y buscar la conciliación de todos los intereses en presencia. Este proceso es, en sí mismo, de una lentitud que se compadece muy mal con la aceleración exponencial de la tecnología. No se discute que, pasado un tiempo prudencial desde la implantación de la tecnología sea razonable dictar disposiciones específicas. Pero no puede ser la única estrategia que debemos aplicar. Sin embargo, tampoco es posible el modelo descrito más arriba, en el que la ausencia de una regulación específica nos permite “romper cosas”.

En este sentido, el Grupo de Expertos de la Unión Europea para la Inteligencia Artificial¹¹ ha recorrido un camino que ya habían abierto reguladores europeos como la Comisión Nacional de Informática y Libertades francesa, y el Supervisor Europeo de Protección de Datos, y ha identificado un elemento esencial para la ética de la IA. La ética de la inteligencia artificial debe ser una ética de la dignidad humana centrada en la garantía de los derechos fundamentales. Este enfoque nos permite una aproximación general al fenómeno de la IA, capaz de establecer una primera barrera jurídica al desarrollo de la tecnología, y funcional al modelo constitucional y democrático en el que debería desarrollarse. Permite situar la dignidad del ser humano en el centro y considerar una IA que no responda a criterios de mera eficiencia económica, sino que se centre en la función social de la inteligencia artificial y en el empleo de los datos para el bien común (Nadella, 2017: 146, 193). Por otra parte, este enfoque puede operar como un elemento de aplicación territorial del derecho en cada Estado y, a la vez, dinamizar el consenso de la comunidad internacional respecto del marco regulador de la IA.

Desde el punto de vista en nuestro país, el constituyente previó la necesidad de ordenar el uso de las tecnologías de la información en el artículo 18.4 de la Constitución española. La jurisprudencia y la doctrina, como se ha criticado reiteradamente,¹² habían condenado al precepto a acoger el derecho fundamental a la protección de datos, cercenando las capacidades que el precepto ofrece. El correcto entendimiento de la norma obligaría a prescindir de su formulación negativa, que ordena al legislador a limitar los usos de la informática. Este precepto debe ser interpretado de acuerdo con el contexto histórico en el que vivimos. Hoy el precepto debe interpretarse en el sentido de atribuir al legislador la tarea de ordenar el adecuado uso de la informática. Y ordenar no solo implica limitar, sino que muchas ocasiones podría llegar a significar una labor de impulso. La regulación de “la informática” debe hacerse no solo desde la óptica del artículo 18 de la Constitución, sino considerando que se ordena el desarrollo normativo de cualquier derecho fundamental, de cualquier principio y valor constitucional.¹³

4.2 Una estrategia inmediata

Como se acaba de señalar, la tecnología avanza y no podemos esperar. Ello obliga, sin perjuicio de regulaciones futuras, a adoptar decisiones estratégicas que se encuentren guiadas por un conjunto de principios que se consideran particularmente relevantes.

- Asegurar un desarrollo de la tecnología que tenga en cuenta los derechos fundamentales no solo como límite, sino también como objetivo.

11 European Commission’s High-Level Expert Group on Artificial Intelligence. (2019). *Ethics guidelines for trustworthy AI*.

12 Martínez, 2004.

13 Este ha sido, sin duda, el enfoque de la redacción del título X de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y Garantía de los Derechos Digitales. Un enfoque funcional que permite salvar la discusión respecto de la necesidad de reformar la Constitución española para incluir nuevos derechos en el catálogo, y agiliza el proceso no solo de atribución de garantías o límites frente a usos inadecuados, sino también para el impulso de políticas públicas en ámbitos tan fundamentales como la reducción de la desigualdad y el fomento de la educación digital.

- Garantizar que los procesos de desarrollo de la tecnología tengan en cuenta el ordenamiento jurídico preexistente; siendo capaces de identificar los principios, los valores y las reglas que deban ser aplicadas a la tecnología sin perjuicio de su carácter previo e incluso analógico.
- Asegurar que el abordaje jurídico de la inteligencia artificial logre una aproximación equilibrada a la tecnología, en lugar de enfoques puramente reactivos centrados en el miedo, la prohibición y la limitación.
- Ser capaces de entender que la aproximación jurídica al desarrollo de la IA exige un enfoque coherente con los propios procesos de diseño. Si en su día Lawrence Lessig afirmó que el programador dicta las leyes del programa al escribir el código (Lessig, 2001), no es menos cierto que el jurista debe aproximarse a los procesos de innovación tecnológica con una metodología iterativa, ajustada a la realidad y apropiada a su naturaleza.

El legislador europeo debe ser consciente de cuáles son los valores en juego. Y se equivocaría completamente si se centrara en una estrategia reactiva que tuviera por único objetivo establecer los límites al uso y desarrollo de la inteligencia artificial. La IA no está exenta de riesgos; pero, al mismo tiempo, promete cambios sustanciales y mejoras de las condiciones colectivas e individuales, y definirá una nueva economía para la entera humanidad (Schwab, 2016: 15-16).

Un legislador lento, medroso, y limitado condenaría a la Unión Europea a convertirse en el cliente de tecnologías que desarrollen otros. Se ha visto en epígrafes anteriores como el esfuerzo de innovación en inteligencia artificial está siendo liderado por Estados Unidos y China. China no puede ser ejemplo, en la medida en que la ausencia absoluta de garantía de los derechos ofrece un espacio sin límites. Sin embargo, debemos ser capaces de ser competitivos con Estados Unidos desde la perspectiva del modelo europeo de garantía de los derechos fundamentales. Y ello exige ser capaces de definir marcos legales que concilien el desarrollo y la innovación con la garantía de los derechos.

Es cierto que la UE ha empezado a andar con ciertos resultados como los citados trabajos del Parlamento Europeo y del Grupo de Expertos en IA, a la vez que ha dictado un reglamento que fomentará el uso de datos no personales,¹⁴ y se plantea la revisión de la directiva sobre reutilización de datos del sector público. Pero no lo es menos que la implementación de una estrategia competitiva en IA en I+D+i,¹⁵ y el crecimiento de los sectores públicos y privados en esta materia, no pueden esperar. Y ello obliga a los operadores jurídicos a identificar metodologías y soluciones prácticas.

4.3 El marco jurídico de la protección de datos como modelo inspirador

Para conseguir los objetivos enunciados en el epígrafe anterior, puede resultar extraordinariamente útil la experiencia adquirida en el marco de la protección de datos. Ciertamente, se podrá objetar que no se ha caracterizado precisamente por ser de una legislación favorecedora de la innovación. Sin embargo, tal vez sea hora de reconocer que ha sido la aproximación realizada por las autoridades de protección de datos la que ha operado como un freno al desarrollo de iniciativas en nuestro país. Así, el problema deriva no tanto de las carencias de la legislación como de un enfoque por parte de la autoridad, que desconoce, con particular y preocupante regularidad, cuál es la realidad material a la que está aplicando la norma. La experiencia adquirida ha determinado, sin duda, la incorporación expresa al Reglamento General de Protección de Datos (RGPD)¹⁶ de procedimientos que pueden resultar estratégicos para un desarrollo de la inteligencia artificial respetuoso con el derecho y con los derechos. Nos referimos, al análisis de riesgos, a la evaluación del impacto en la protección de datos y a la protección de datos desde el diseño y por defecto. Por otra parte, no solo se trata de una metodología funcional, sino incluso obligatoria. El despliegue de las herramientas de inteligencia artificial puede repercutir significativamente sobre la esfera de derechos de las personas; y, para ello, requerirá del tratamiento de grandes volúmenes de datos de carácter personal.

¹⁴ Reglamento (UE) 2018/1807 del Parlamento Europeo y del Consejo, de 14 de noviembre de 2018, relativo a un marco para la libre circulación de datos no personales en la Unión Europea. DOUE. Serie L 303/59 de 28 de noviembre de 2018.

¹⁵ Ministerio de Ciencia, Innovación y Universidades. (2019). *Estrategia Española de I+D+i en Inteligencia Artificial*. Madrid: Secretaría General Técnica del Ministerio de Ciencia, Innovación y Universidades.

¹⁶ Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE. DOUE. Serie L 119/1 de 4 de mayo de 2016.

Adicionalmente, pueden utilizarse algunos de los institutos jurídicos presentes en el artículo 5 del RGPD que, partiendo de la licitud en el uso de los datos, determinan un enfoque basado en la proporcionalidad —entendida como exigencia de ajustar el volumen y tipología de los datos objeto de tratamiento a las finalidades y necesidades del mismo—, la calidad y veracidad de los datos. Por otra parte, el principio de transparencia puede aportar un requisito adicional para la inteligencia artificial: la capacidad de ser fiscalizada o fiscalizable. Y, finalmente, en el marco de la garantía de los derechos, la capacidad del sujeto de ejercer un control sobre los tratamientos, y oponerse a los mismos, puede ser inspiradora para definir estatutos de garantía de los derechos de las personas respecto de la utilización de la IA.

4.4 El cumplimiento normativo desde el diseño

Veamos, siquiera brevemente, cómo puede repercutir el empleo de las metodologías de protección de datos en el desarrollo de proyectos de IA para aseguramiento del cumplimiento normativo y de la garantía de los derechos. Para ello, tomaremos como referencia una metodología de desarrollo de *software* recomendada por la autoridad noruega de protección de datos (Datatilsynet, 2018).

Esta autoridad propone procesos de diseño de *software* con fases diferenciadas:

- Formación.
- Toma de requerimientos.
- Diseño de la aplicación.
- Programación del código.
- Prueba o verificación.
- Puesta en producción.
- Mantenimiento.

Se trata de un procedimiento circular, o ciclo de Deming, orientado no solo a garantizar un diseño previo adecuado, sino también a mantener el estándar de cumplimiento normativo como compromiso de permanente actualización.

4.4.1 Formar ética y jurídicamente al equipo

Esta aproximación metodológica parte de un empoderamiento previo de las personas implicadas. Para los proyectos de IA, esto implica que el equipo debe contar con una base de formación que incorpore los valores éticos humanistas y de garantía de los derechos fundamentales. No se trata de convertir a los desarrolladores, ni a los gestores en expertos juristas. Sencillamente obliga a que en el conjunto de la organización exista una cultura de garantía de los derechos y de respeto del derecho, que oriente las decisiones. Y comporta incorporar al proyecto, desde el inicio, a expertos en materias como la ética y el derecho.

4.4.2 Identificar los requerimientos del proyecto

Desde esta fase temprana deberán tenerse en cuenta los principios jurídicos aplicables. Y esto implica un enfoque de cumplimiento normativo desde el diseño, que tiene significativas consecuencias de orden práctico. En primer lugar, el desarrollo de un proyecto de IA deberá ponerse en el contexto del sector del ordenamiento en el que se va a desenvolver, y deberá considerar la regulación vigente. En ningún caso servirá considerar que el ordenamiento preexistente es analógico y no resulta de aplicación a una nueva tecnología.

En cualquier caso, como se ha defendido reiteradamente, existe un principio jurídico fundamental que nunca puede ser obviado. Un proyecto de innovación tecnológica no puede diseñarse con la conciencia de que puede producir daños a las personas (Martínez, 2019a: 270). La IA no puede desarrollarse en detrimento de los derechos de los individuos, o minusvalorando principios o valores constitucionales esenciales. Así pues, el cumplimiento normativo desde el diseño incorpora una primera regla: desde el momento cero de gestación

de un proyecto de IA debe realizarse un análisis exhaustivo del derecho aplicable al mismo, desde una aproximación abierta a considerar todas las posibilidades que pudieran ordenar su adecuado funcionamiento.

En una fase tributaria de la anterior es necesario desarrollar un análisis de los riesgos e implicaciones de la tecnología que se está desarrollando. Se trata de una metodología plenamente asentada en relación con las normas de calidad en el desarrollo del producto, y con la prevención de riesgos para la seguridad de los sistemas y para las personas que eventualmente los van a utilizar. Por tanto, se trata de una metodología conocida con reglas muy específicas y con criterios de aplicación particularmente precisos. Se trata de establecer cuáles pueden ser las vulnerabilidades intrínsecas al proyecto que se desarrolla, y las amenazas a las que se enfrenta. En la metodología de análisis de riesgos debemos establecer una relación entre la probabilidad de que esas amenazas o vulnerabilidades se materialicen en el mundo real y su impacto al afectar, en lo que aquí interesa, al cumplimiento normativo, causar daños a las personas y, muy especialmente, vulnerar sus derechos fundamentales. Esta relación nos ofrece una medida del riesgo.¹⁷

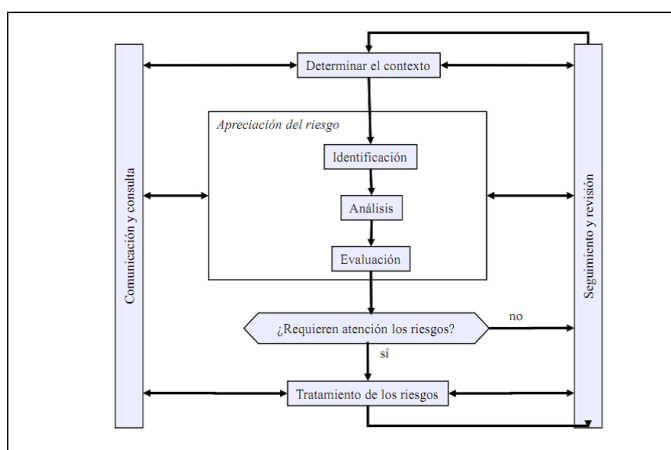


Figura 2. Norma UNE 71504:2008. Metodología de análisis y gestión de riesgos para los sistemas de información. AENOR

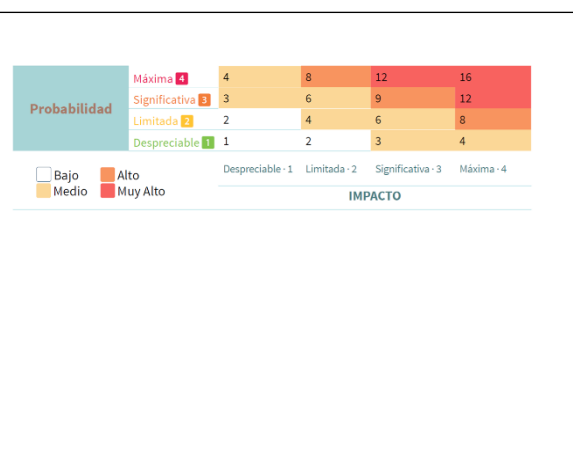


Figura 3. Fuente: Agencia Española de Protección de Datos

Una vez establecido el riesgo, procede adoptar medidas que lo eviten o lo minimicen hasta el punto de hacerlo tolerable. Desde el punto de vista del cumplimiento normativo, ello implica que se tendrá la capacidad de identificar *a priori* las situaciones en las que exista un riesgo de incumplimiento del ordenamiento o de vulneración de los derechos fundamentales, y se adoptarán las medidas adecuadas para evitar que no se produzca.

Aplicando la racionalidad propia del análisis de riesgos, el jurista no podrá exigir desarrollar un escenario de riesgo cero salvo en aquellos casos en los que el perjuicio para la garantía de los derechos de las personas resultase inasumible. Y precisamente para estas situaciones de riesgo grave para los derechos, el artículo 35 del RGPD propone lo que podríamos definir como un análisis de riesgos agravado: la evaluación del impacto en la protección de datos.

No vamos a exponer aquí en detalle en qué consiste esta metodología que ha sido sistematizada por las autoridades de protección de datos en distintas guías. Sin embargo, sí que resulta particularmente relevante subrayar que de las conclusiones de la evaluación de impacto derivarán decisiones estratégicas para

¹⁷ Veamos un ejemplo muy sencillo que puede ayudar a entender lo que aquí se propone. Cualquier usuario de redes sociales ha podido apreciar como el etiquetado de fotografías ha cambiado con el tiempo. En principio, se podía etiquetar a un usuario en fotografías sin ningún tipo de restricción. Sin embargo, hoy en día, todas las redes sociales permiten al usuario bloquear a los terceros que puedan etiquetarlo. Se dirá que esta es una consecuencia lógica de la evolución normativa en materia de protección de datos. Pero es una afirmación completamente falsa. Para cualquier conocedor de nuestro ordenamiento jurídico es una obviedad que el tratamiento de imágenes sin consentimiento se encuentra regulado en el artículo 18.1 CE y en la Ley Orgánica 1/1982, de 5 de mayo, de protección civil del derecho al honor, a la intimidad personal y familiar y a la propia imagen. Y no solo aquí, también en las normas internacionales en materia de derechos humanos, siendo tutelado en la jurisprudencia del Tribunal Supremo de Estados Unidos, del Tribunal Europeo de Derechos Humanos y de los tribunales nacionales. Esto significa que, si los desarrolladores de redes sociales hubieran tenido en cuenta la metodología que aquí se defiende, este tipo de funcionalidad debería haber estado presente desde la versión beta de sus productos.

el desarrollo del producto, que nos permitirán establecer cuándo este es sencillamente inviable, o cómo podemos modularlo para que ofrezca resultados útiles.¹⁸

4.4.3 Integrar las normas en el diseño y documentar el cumplimiento

Superada la fase de toma de requerimientos y de análisis de riesgos, debemos considerar cómo los principios y valores del artículo 5 del RGPD pueden inspirar de modo muy preciso el desarrollo de las inteligencias artificiales.¹⁹ En este sentido, el punto de partida ineludible para cualquier proyecto, y no solo en el marco de la inteligencia artificial, es el del respeto al principio de licitud. Por definición, aquello que es ilícito no puede ser objeto de desarrollo, no es posible programar un modelo de IA si se es consciente de estar infringiendo el ordenamiento.

Por otra parte, la lealtad y transparencia de los procesos decisionales basados en la IA plantea también la necesidad de desarrollar una estrategia de documentación del proyecto, que facilite lo que se ha venido a denominar “la transparencia del algoritmo”. Por otra parte, los principios ordenados a garantizar que el volumen de datos que se van a utilizar es adecuado, su origen legítimo y la información confiable resultan estratégicos para asegurar un adecuado funcionamiento de la IA. Por último, el valor de la “seguridad”, informática, técnica y de funcionamiento, debería considerarse como un requisito indispensable.

4.4.4 De la programación a la comercialización

Las subsiguientes fases de gestación de un proyecto tecnológico suelen discurrir a través de un proceso ordenado y, a la vez, creativo. Se produce lo que en la jerga se denomina “iteración”. Si bien cada fase de desarrollo incluye un proceso de codificación que genera paquetes entregables, su propia dinámica está muy abierta a la innovación.

Así, durante la programación del código y del algoritmo, el programador verifica sus posibilidades y problemas, y ello lo obliga a cambiar su enfoque en muchas ocasiones. Por otra parte, no es inusual que se aprecien nuevas potencialidades. Es más, en procesos vinculados a la analítica de datos, la posibilidad de enfrentarse al hallazgo casual puede ser muy alta. Por otra parte, se hace imprescindible someter el desarrollo a procesos de prueba o verificación que aseguren su adecuado funcionamiento.

Ello obliga a disponer de un soporte jurídico permanente que asegure el cumplimiento de la norma como un objetivo intrínseco al desarrollo y con una doble función. En primer lugar, no solo el producto final, sino cada una de las tareas que contribuyen a su gestación debe cumplir el derecho. No parece razonable considerar adecuado un resultado cuyo proceso de diseño haya vulnerado normas. En segundo lugar, es necesario asegurar, en la puesta en producción y la comercialización, un resultado adecuado desde un punto de vista jurídico.

¹⁸ En todo caso, es muy probable que este tipo de análisis se encuentre presente en cualquier proyecto de desarrollo de IA que obligue a tratar datos personales. Por ejemplo, en el ámbito del *marketing* mediante perfilado de los usuarios, el *scoring* bancario o en seguros.

¹⁹ Este dispone que:

“1. Los datos personales serán:

- a) tratados de manera lícita, leal y transparente en relación con el interesado (‘licitud, lealtad y transparencia’);
- b) recogidos con fines determinados, explícitos y legítimos, y no serán tratados ulteriormente de manera incompatible con dichos fines; de acuerdo con el artículo 89, apartado 1, el tratamiento ulterior de los datos personales con fines de archivo en interés público, fines de investigación científica e histórica o fines estadísticos no se considerará incompatible con los fines iniciales (‘limitación de la finalidad’);
- c) adecuados, pertinentes y limitados a lo necesario en relación con los fines para los que son tratados (‘minimización de datos’);
- d) exactos y, si fuera necesario, actualizados; se adoptarán todas las medidas razonables para que se supriman o rectifiquen sin dilación los datos personales que sean inexactos con respecto a los fines para los que se tratan (‘exactitud’);

[...]

f) tratados de tal manera que se garantice una seguridad adecuada de los datos personales, incluida la protección contra el tratamiento no autorizado o ilícito y contra su pérdida, destrucción o daño accidental, mediante la aplicación de medidas técnicas u organizativas apropiadas (‘integridad y confidencialidad’).

2. El responsable del tratamiento será responsable del cumplimiento de lo dispuesto en el apartado 1 y capaz de demostrarlo (‘responsabilidad proactiva’).”

4.4.5 Un ciclo que se mantiene durante toda la vida de una IA

Ningún proceso de cumplimiento normativo vinculado a la tecnología puede ser estático. La propia naturaleza de la IA hace de este principio una ley. Debemos entender que nos situamos en una fase embrionaria de esta tecnología, en la que hemos aprendido que los procesos de análisis de datos que la sustentan son muy sensibles al sesgo. Por otra parte, comenzamos a confiar procesos básicos de decisión a esta tecnología, con consecuencias jurídicas o materiales para sus destinatarios. Un cumplimiento normativo adecuado obliga a un estado de permanente seguimiento y actualización que se despliega en varios niveles:

- 1) Aprender del propio funcionamiento de la tecnología. Los resultados obtenidos, los errores de funcionamiento, las incidencias..., cualquier elemento verificado o verificable debería ser indexado y estudiado también por el soporte jurídico. Y no únicamente para prevenir posibles conflictos y responsabilidades, sino, sobre todo, para mejorar las condiciones de cumplimiento normativo.
- 2) Profundizar en el diseño de cumplimiento normativo proponiendo mejoras cuando resulte necesario.
- 3) Acompañar cada fase o evolución del producto.

5 Una breve conclusión

La IA plantea, al derecho y a los juristas, un reto descomunal. No sería exagerado afirmar que podríamos estar asistiendo a un cambio de paradigma que puede revolucionar pilares esenciales de nuestra tradición, empezando por el modo de entender conceptos clave como la personalidad jurídica o la dignidad humana. Resultaría imprudente, pues, negar que los cambios que augura la inteligencia artificial plantean retos inimaginables que deben ser abordados cuanto antes. Pero lo sería tal vez mucho más negar cualquier virtualidad al marco normativo vigente. Debemos aplicar las herramientas de las que disponemos.

En este sentido, este trabajo se ha planteado con un objetivo modesto: ofrecer una bitácora para un abordaje inmediato de los proyectos de IA. Y, para ello, necesitamos ser capaces de integrar el desarrollo de la IA con una interpretación evolutiva del marco normativo exenta de prejuicios. En este sentido, se exigen al jurista ciertos esfuerzos necesarios e imprescindibles.

El primero de ellos, asumir las dinámicas propias de la gestación de un proyecto de naturaleza tecnológica en sus distintas fases. Y, por tanto, adaptar su metodología para hacerla funcional al modelo de gestión de los proyectos de desarrollo de las tecnologías de la información y las comunicaciones. No se trata de convertirse en ingenieros, sino de ser capaces de trabajar de modo funcional a sus metodologías.

El segundo, ser capaces de descender a la realidad material. El derecho aplicado a la tecnología exige un esfuerzo adicional de comprensión. No se trata de algo distinto de lo que se exige a un buen profesional en cualquier sector: conocer los hechos. Sin embargo, en este ámbito, la intensidad de la exigencia es muy alta y obliga a una interacción abierta en la que la receptividad y comprensión son determinantes.

En tercer lugar, es necesaria una concepción abierta y dinámica del cumplimiento normativo. Los proyectos de IA van a obligar a desarrollar un enorme esfuerzo de interpretación sistemática del ordenamiento desde sus principios y valores constitucionales al ordenamiento sectorial, de lo local a lo transnacional. Pero esta apertura no solo es material, también lo es metodológica. El esfuerzo de soporte jurídico debe ser permanente, nunca finaliza, acompaña a la IA durante todo su ciclo de vida.

Y esta tarea se despliega en un contexto normativo emergente y con reglas en construcción. Aunque ello no implica un cheque en blanco ni a la industria ni a los investigadores. El marco de los derechos fundamentales y de los principios y valores constitucionales constituye una barrera infranqueable. Del mismo modo, el argumento de la inaplicabilidad de un derecho “análogo” resulta inaceptable. El programa normativo que incorpora nuestro sistema legal define unos objetivos y principios susceptibles de ser identificados y aplicados al conjunto de las relaciones humanas y de sus desarrollos. Considerar lo contrario atentaría contra toda razón.

Precisamente por ello, los profesionales del derecho vienen claramente obligados a incorporar nuevas metodologías de soporte al desarrollo de la tecnología. Estamos obligados a bajar del pedestal distante desde el que solemos pontificar, descender al suelo, llenarnos las manos de barro y contribuir a construir la sociedad de la transformación digital orientada al bien común desde los valores que incorpora la garantía de la dignidad del ser humano y los derechos que le son inherentes.

Bibliografía

- Bauman, Zygmunt y Lyon, David. (2013). *Vigilancia líquida*. Barcelona: Austral.
- Benanti, Paolo. (2019). La dignidad de la persona en la era de Máquina Sapiens [Apunte de blog]. *Fundación para la Investigación del Derecho y la Empresa*. Consultado el 1 de mayo de 2019 desde <http://bit.ly/2K6Fmx0>
- Blodget, Henry (1 de octubre de 2009). [Mark Zuckerberg on innovation](#). *Business Insider*. Consultado el 1 de mayo de 2019.
- Bradley, Chris y O'Toole, Clayton. (2016). An incumbent's guide to digital disruption. *McKinsey Quarterly*. Consultado el 1 de mayo de 2019 desde <https://mck.co/2IlxhC0>
- Bughin, Jacques; Seong, Jeongmin; Manyika, James; Chui, Michael y Joshi, Raoul. (2018). [Notes from the AI frontier. Modeling the impact of AI on the world economy](#). McKinsey Global Institute.
- Byung-Chul, Han. (2014). *Psicopolítica*. Barcelona: Herder.
- Christensen, Clayton M. (1997). *The innovator's dilemma. The revolutionary book that will change the way you do business*. Nueva York: Harper Collins.
- Commission Nationale de l'Informatique et des Libertés. (2017). *Comment permettre à l'homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle*. Consultado el 1 de mayo de 2019 desde <http://bit.ly/2VayB3O>
- Datatilsynet. (2018). *Software development with data protection by design and by default*. Consultado el 1 de mayo de 2019 desde <https://bit.ly/2QPzfA3>
- Dobbs, Richard; Manyika, James y Woetzel, Jonathan. (2015). *No ordinary disruption*. Nueva York: Public Affairs.
- European Commission's High-Level Expert Group on Artificial Intelligence. (2019). [Ethics guidelines for trustworthy AI](#).
- Ford, Martin. (2016). *El ascenso de los robots: la amenaza de un futuro sin empleo*. Ciudad de México: Paidós.
- Jiménez, Ángel. (28 de julio de 2017). [Facebook apaga una inteligencia artificial que había inventado su propio idioma](#). *El Mundo*. Consultado el 1 de mayo de 2019.
- Lessig, Lawrence. (2001). *El código y otras leyes del ciberespacio*. Madrid: Taurus.

- Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y Garantía de los Derechos Digitales. BOE núm. 294 (2018), 8-68.
- López de Mántaras, Ramón y Messeguer, Pedro. (2017). *Inteligencia artificial*. Madrid: Catarata-CSIC.
- López de Mántaras, Ramón. (2017). [Ética en la inteligencia artificial](#). *Investigación y Ciencia*, (491).
- Mallard, Stéphane. (2018). *Disruption. Préparez-vous à changer du monde*. París: Dunod.
- Marcet, Xavier. (2018). *Esquivar la mediocridad. Notas sobre management: complejidad, estrategia e innovación*. Barcelona: Plataforma Editorial.
- Martínez, Ricard. (2004). *Una aproximación crítica a la autodeterminación informativa*. Madrid: Civitas.
- Martínez, Ricard. (2016). *The challenge of enforcement in the proposal for a General Data Protection Regulation* [Apunte de blog]. *Phaedra Project Blog*. Consultado el 1 de mayo de 2019 desde <http://bit.ly/2UOKJm8>
- Martínez, Ricard. (2019a). Inteligencia artificial, Derecho y derechos fundamentales. En: Tomás De la Quadra Salcedo y José Luis Piñar (dir.), *Sociedad digital y Derecho* (p. 259-278). Madrid: Boletín Oficial del Estado, Ministerio de Industria, Comercio y Turismo y RED.ES.
- Martínez, Ricard. (2019b). [Partidos, algoritmos y campañas electorales](#). *Telos*, (110), 130-135.
- McCarthy, John; Minsky, Marvin L.; Rochester, Nathaniel y Shannon, Claude E. (1955). *A proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. Consultado el 1 de mayo de 2019 desde <https://stanford.io/2JFCgQV>
- McKinsey Global Institute. (2017). *What's now and next in analytics, AI, and automation*. Consultado el 1 de mayo de 2019 desde <https://mck.co/2vthXNJ>
- McKinsey Global Institute. (2018). *How artificial intelligence and data add value to businesses*. Consultado el 1 de mayo de 2019 desde <https://mck.co/2J8uXjI>
- Mercader, Jesús. (2018). El futuro del trabajo y del empleo en la era de la digitalización y la robótica. En: Tomás de la Quadra Salcedo y José Luis Piñar (dir.), *Sociedad digital y Derecho* (p. 611-632). Madrid: Boletín Oficial del Estado, Ministerio de Industria, Comercio y Turismo y RED.ES.
- Moore, Gordon. (1965). [Cramming more components onto integrated circuits](#). *Electronics*, 38(8), 114–117.
- Nadella, Satya. (2017). *Pulsa actualizar*. Madrid: Harper Collins.
- O'Neil, Cathy. (2016). *Weapons of math destruction*. Nueva York: Crown.
- Pariser, Eli. (2017). *El filtro burbuja. Cómo la red decide lo que leemos y lo que pensamos*. Barcelona: Taurus.
- Parlamento Europeo. (2017). [Informe con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica \(2015/2103\[INL\]\)](#).
- Ramió, Carles. (2019). *Inteligencia artificial y Administración pública. Robots y humanos compartiendo el servicio público*. Madrid: Catarata.
- Schwab, Klaus. (2016). *La cuarta revolución industrial*. Barcelona: Debate.
- Stone, Peter; Brooks, Rodney; Brynjolfsson, Erik; Calo, Ryan; Etzioni, Oren; Hager, Greg... y Teller, Astro. (2016). *Artificial intelligence and life in 2030. One hundred year study on artificial intelligence. Report of the 2015-2016 Study Panel*. Stanford: Stanford University, Consultado el 1 de mayo de 2019 desde https://ai100.stanford.edu/sites/g/files/sbiybj9861/f/ai_100_report_0831fnl.pdf

- Supervisor Europeo de Protección de Datos. (2015). [*Dictamen 4/2015. Hacia una nueva ética digital. Datos, dignidad y tecnología.*](#)
- Tegmark, Mark. (2018). *Vida 3.0: Ser humano en la era de la Inteligencia Artificial*. Barcelona: Taurus.
- Trías, Jorge. (1992). Informática y privacidad. ¿Se pueden poner puertas al campo? *Cuenta y Razón*, (63), 98-101.
- Varona, Daniel. (2018). La responsabilidad ética del diseñador de sistemas en inteligencia artificial. *Revista de Occidente*, (446-447), 104-114.