



Conferencia dictada en el Colegio Notarial de Madrid, Salón Académico, el 12 de diciembre de 2019

La inteligencia artificial. Una mirada desde el Derecho

PABLO GARCÍA MEXÍA, PH.D.

Jurista digital,
Letrado de las Cortes



¿A qué nos referimos al hablar de “inteligencia artificial” (IA)? Aquí lo haremos a sistemas que justamente son “inteligentes” por aprender (Alan Turing), hasta el punto de que ni siquiera sus creadores son conscientes del modo como han alcanzado sus resultados, al tratarse de sistemas que incluso “aprenden a aprender” (José Ignacio Latorre).

Es ya muy común distinguir entre dos tipos de IA en este sentido, la que se suele llamar “fuerte” y la IA “débil”. La IA *fuerte* es la que, en su momento, si se llega a él, permitirá hablar de sistemas “conscientes de su propia existencia” (Daniel Tennett). La IA fuerte tardará aún bastante tiempo

en llegar. Los riesgos de tipo apocalíptico que suelen predicarse de la IA se refieren a este tipo de IA, la fuerte. De entre muchos, me quedo con quienes los comparan con los de la energía nuclear empleada con fines bélicos. En segundo lugar, tenemos la IA *débil*, es decir, la actual, fundamentalmente compuesta por el análisis de macrodatos (o Big data Analytics), que justamente emplea algoritmos en una combinación de sistemas inteligentes (*machine learning*) y Big data.

No pensemos sin embargo que los riesgos de estos algoritmos son desdeñables. El más evidente radica en una característica consustancial al proceso algorítmico, cual es la opacidad con la que funciona, como si

fuese una auténtica *Black Box*, o “caja negra”, pues “podemos acceder a los *inputs* o insumos, [que son los datos sobre personas], y a los *outputs* o resultados [que son las clasificaciones], pero no terminamos de entender lo que sucede por el camino, ni cómo se obtienen algunos resultados, incluyendo entre ellos decisiones y acciones (Comisión Europea)”.

Las Ciencias sociales, por su naturaleza práctica, desde la Economía, a la Sociología o al propio Derecho, están saliendo al paso de los retos de la IA en sus respectivos campos. Aunque es cierto que la Humanidad no puede acudir más que a la Filosofía para encontrar respuestas a problemas con dimensión ética, que este tipo de sistemas está generando ya hoy: por solo apelar al ejemplo más frecuente, el del dilema del vehículo autónomo a la hora de atropellar un carrito de bebé o chocar de frente contra otro vehículo. Por eso estamos ante una verdadera eclosión de declaraciones, manifiestos, guías o códigos éticos orientados a ahorrar la IA, cuyo número se acerca ya a las 90. Al menos a estos efectos, corren pues *buenos tiempos para la Filosofía*.

También desde el Derecho comienzan a afrontarse los riesgos de la IA, que no en vano afecta a derechos y bienes jurídicos esenciales. Un primer caso es la dignidad de las personas, y de su mano, su misma integridad física. De hecho, uno de los más tempranos textos sobre los principios éticos aplicables a la IA, las “architectadas” tres leyes de la robótica de Isaac Asimov, publicadas en 1942, aluden de lleno a ella. Basta mencionar una parte de la primera: “Un robot no hará daño a un ser humano”. Asimismo, también es hoy real el empuje de la IA sobre un componente consustancial a nuestra dignidad, como es la idea de identidad personal. La identidad online, de hecho, puede llegar a definirse, “no desde la autonomía de la persona, sino heterónomamente, a través del poder de los algoritmos, merced a flujos de información sobre nosotros mismos que podemos no llegar a controlar en absoluto” (José Luis Piñar).

El de los riesgos para la equidad y la igualdad es muy probablemente el tema sobre el que más se ha escrito hasta ahora, de cuantos atañen al cruce IA-Derecho. Es ya axiomático afirmar que el funcionamiento algorítmico genera sesgos. La principal razón es que los propios humanos, como seres consustancialmente sociales y provistos de determinados valores, actuamos siempre con prejuicios, a su vez causantes de sesgos. Los sesgos se deben normalmente a errores técnicos, más o menos reprochables, de los diseñadores del algoritmo en cuestión. Y pueden provocar resultados como el del algoritmo de la tarjeta bancaria Apple Card, que en un concreto caso establecía para una mujer un límite de crédito 20 veces inferior al que ofrecía a su esposo, en condiciones financieras de total equiparación. Lo relevante aquí es que el sesgo algorítmico genera discriminación, es decir, un trato injustificadamente más favorable para determinados grupos o personas.

La privacidad es otro de los derechos frontalmente amenazados por los avances en IA, en cuanto la opacidad algorítmica compromete principios como la finalidad del tratamiento, su transparencia o la minimización del dato, a resultados de la voracidad del algoritmo a la hora de recoger información. ¿Cómo justificar, por ejemplo, que un ente público recopile datos para otorgar una ayuda social y que, tras el proceso algorítmico, el resultado sea denegarla? No son tampoco pequeños los riesgos del trazado de perfiles, que la IA puede llevar a cabo gracias a su potencial para generar correlaciones entre las miríadas de piezas de información que maneja. Riesgos que nos atañen muy singularmente en nuestra condición de consumidores y usuarios, frecuentes víctimas de asimetrías de poder (a la vista del plus de información que una plataforma puede llegar a adquirir gracias al algoritmo inteligente), ajustes algorítmicos de precios o discriminaciones por precio, gracias a la posibilidad de rastrear precios en línea y de cruzarlos con los datos de

Estamos ante una verdadera eclosión de declaraciones, manifiestos, guías o códigos éticos orientados a ahorrar la inteligencia artificial, cuyo número se acerca ya a las 90. Al menos a estos efectos, corren pues buenos tiempos para la Filosofía



navegación del consumidor (Organización europea de consumidores).

El último de los retos de la IA que comentaremos es el referido a la responsabilidad por los daños que un sistema pueda causar. No revisten especial dificultad los daños de naturaleza criminal. Hoy por hoy, y hasta un futuro aún lejano, ningún sistema inteligente puede equipararse a una *mens criminis* humana, es decir, a un cerebro que, como el humano, acumule el conocimiento y voluntad propios del dolo penal. De ahí que resulte claro que quien en estas situaciones deba responder sea “la mente de atrás”, es decir, el fabricante o comercializador de estos sistemas.

En el caso de la responsabilidad civil, en cambio, el impacto es mucho más complejo. Primero, porque la posible responsabilidad objetiva de fabricantes o comercializadores solo operará en caso de defectos. Segundo, porque la responsabilidad derivada de riesgos resultará inútil para usos en principio inocuos de un sistema inteligente, como puede ser el de, pongamos por caso, recomendar el consumo de una determinada marca de agua mineral. Si en estos últimos supuestos, y

por la razón que fuera, llegara a generarse algún daño, porque, siguiendo con el ejemplo, el manantial del agua en cuestión resultase estar contaminado, no estaría sino acudir al criterio de la culpa o negligencia. Aunque también aquí sería necesario que el daño hubiera sido previsible. Y la previsibilidad puede fácilmente topar con la oscura impredecibilidad del algoritmo...

Es claro pues que el Derecho está ya sometido a un considerable zarandeo por parte de la IA. E igualmente lo es que se debe reaccionar al respecto. Como ha ocurrido ya en otros ámbitos tecnológico-digitales, la regulación jurídica tendrá que entrar a ordenar también determinados aspectos de la IA, máxime cuando los principios y derechos en juego (dignidad, igualdad, privacidad...) son de la mayor importancia.

Para ello, dos principios jurídicos reinan sobre todos los demás: *centralidad de la persona humana*, en forma de *control* sobre los sistemas inteligentes; y *responsabilidad proactiva* (o *accountability*) por parte de dichos sistemas, en cuanto “obligación de justificar sus decisiones y posibilidad de

enfrentarse a sanciones si tal justificación resultase inadecuada (Parlamento Europeo).” El principio de responsabilidad proactiva ha logrado la primacía sobre otros tres principios adyacentes, que resultan absolutamente tópicos en la ya profusa literatura sobre este tema: *lealtad* (“fairness”) algorítmica, que las mejores definiciones equiparan a “ausencia de sesgos indeseados”; *transparencia*, que de nuevo el Parlamento Europeo considera como “disponibilidad del código del sistema y de la documentación correspondiente” y *explicabilidad*, o “interpretabilidad o comprensibilidad de un sistema inteligente”.

Todos estos principios están llamados a ir plasmándose en normas jurídicas. Hasta ahora, eso sí, esto no ha sucedido más que en un concreto ámbito, el de la privacidad, concretamente, en el RGPD de la UE, primera norma jurídica del mundo en regular el impacto de la IA sobre la privacidad, en cuanto concede a los ciudadanos el “derecho a una mínima intervención humana” en decisiones (inteligentes y automatizadas) con impacto en la privacidad. Y en cuanto prevé un derecho de información a los interesados acerca de la



lógica seguida por sistemas inteligentes que les afecten, así como de las consecuencias que todo ello les pueda generar.

Más allá de la privacidad, apuntaré dos ideas principales. La primera es la de que, merced a su severo impacto sobre tres áreas legales clásicas como son el consumo, la competencia y la responsabilidad civil, a su vez intrínsecamente imbricadas entre sí, el algoritmo inteligente está operando como vector de convergencia entre todas ellas. Baste citar que autoridades de competencia como la alemana, francesa o portuguesa vienen demostrando que algunas de las empresas Big Tech pueden llegar a utilizar su hegemonía sobre el dato y su automatización inteligente para menoscabar la competencia, dañando así también a los consumidores.

La segunda se refiere a los obstáculos de la IA para la responsabilidad civil, respecto de los que descuellan la propuesta de 2017 del Parlamento Europeo de atribuir una "personalidad electrónica" a determinados sistemas de IA, que permitiera por ejemplo indemnizaciones pecuniarias. La idea tiene el fuste necesario para ser to-

mada en serio: lo demuestra por ejemplo la perfecta viabilidad de otras ficciones en el Derecho, como la misma noción de persona jurídica. Con todo, esta "personalidad electrónica" no constituye un instrumento adecuado para resolver problemas como los expuestos, porque este mecanismo extendería a sistemas o ingenios artificialmente inteligentes un atributo, la propia personalidad, que es patrimonio exclusivo de la esencia humana. Con ello se diluiría esa exclusividad que los humanos tenemos sobre esa nuestra nota singular. Al tiempo que se difuminaría el atributo capital de cuantos han de caracterizar un uso confiable de la IA, como es la centralidad del ser humano, de manera que sea la IA la que ponga en el centro a la persona, y no la persona quien deba vivir a expensas de la IA.

Inteligencia artificial, al fin y al cabo, es sin duda "conocimiento artificial", pero nunca será "sabiduría artificial". La sabiduría es una virtud, por eso es y será siempre un atributo exclusivamente humano. La máquina conoce, no sabe. La máquina desconoce la medida de su ignorancia. ●

Como ha ocurrido ya en otros ámbitos tecnológico-digitales, la regulación jurídica tendrá que entrar a ordenar también determinados aspectos de la inteligencia artificial



Dos principios jurídicos reinan sobre todos los demás: *centralidad de la persona humana*, en forma de *control* sobre los sistemas inteligentes; y *responsabilidad proactiva* (o *accountability*) por parte de dichos sistemas