JULI PONCE SOLÉ

EL REGLAMENTO DE INTELIGENCIA ARTIFICIAL DE LA UNIÓN EUROPEA DE 2024, EL DERECHO A UNA BUENA ADMINISTRACIÓN DIGITAL Y SU CONTROL JUDICIAL EN ESPAÑA

Prólogo de

Helena Matute

Marcial Pons

MADRID | BARCELONA | BUENOS AIRES | SÃO PAULO

2024

ÍNDICE

Pág.

Pág.

PRÓLOGO

Es un placer prologar este libro del profesor Ponce, un libro que, como el propio autor nos indica, parte de la necesidad de interdisciplinaridad, y de comunicarnos entre personas de muy diferentes ámbitos para abordar los enormes retos que nos plantea el momento actual.

Aborda el Dr. Ponce, desde la perspectiva del derecho, asuntos que nos tienen preocupados también en el ámbito de la psicología. Son temáticas en las que venimos muchos años investigando, como por ejemplo los sesgos cognitivos, y en concreto, el problema de cómo nos relacionamos las personas con la inteligencia artificial (IA) —y ella con nosotros—. Son problemas que desde la psicología llevamos también años insistiendo en que deben ser abordados por las leyes, deben ser reglamentados, y discutidos desde el punto de vista jurídico. Es un placer, por tanto, encontrar este libro en donde se discuten desde el punto de vista ético y legal las implicaciones de todas estas cuestiones que nos afectan tanto a todos.

Entre las muchas propuestas interesantes que realiza el profesor Ponce, yo destacaría por ejemplo su propuesta de educación y títulos conjuntos de derecho, psicología, e IA. Es algo que ya se empieza a ver tímidamente en algunas universidades, y que sin duda habrá de ir a más. Esto es importante si queremos que los problemas que detectamos en la relación humanos-IA, se aborden desde una legislación que ayude a convertir la IA en una herramienta beneficiosa para la humanidad, y que permita el florecimiento que se nos ha prometido (y que por desgracia no siempre se está consiguiendo). Estoy convencida de que solo reglamentando el desarrollo y la buena utilización de la IA seremos capaces de llegar a esa tierra prometida en la que la especie humana sale beneficiada, en su conjunto, de su asociación con la IA.

Por todo ello me congratula saber que ya contamos por fin con un reglamento europeo, y me encanta ver a expertos en leyes, como el Dr. Ponce, comentar el reglamento y hacerlo un poco más accesible para toda la sociedad, desglosando aquellas partes del reglamento con las que está de acuerdo, explicando por qué lo está, así como aquellas otras que cree mejorables, aportando también argumentos convincentes. Me ha resultado muy clarificador, y creo que de esta forma podremos mejorar, y podremos hacer de la IA una herramienta beneficiosa para la humanidad, algo que es en realidad lo único que da sentido al desarrollo de esta tecnología tan poderosa.

Decía el profesor Kahneman, psicólogo y premio nobel de economía por sus investigaciones sobre la irracionalidad humana y su influencia en la economía, que la IA solucionará el problema de los sesgos cognitivos humanos. Muchos autores han secundado estas ideas, y han alabado la introducción de la IA como medio para reducir los sesgos humanos que afectan negativamente a muchas de nuestras decisiones. Pero a pesar de mi enorme admiración por el trabajo del Dr. Kahneman, debo discrepar en este punto, creo que en este detalle se equivocó el maestro. La IA no solo no ha solucionado el problema de los sesgos humanos, sino que los está amplificando. Esto tiene lugar principalmente por dos vías:

(a) Incorporación de sesgos humanos en los modelos de IA, en todas las fases de su desarrollo (programación, entrenamiento con bases de datos sesgadas, contacto con usuarios sesgados en el mundo real, etc.). Estos sesgos, presentes primero en los humanos y sus bases de datos, y luego en los modelos de IA, se transmiten a muchas más personas ahora que cuando el sesgo en una decisión lo tenía una sola persona (por ejemplo, un juez podía tener sus sesgos, pero quizá, o probablemente, fueran diferentes sesgos a los que tenía otro juez, mientras que un mismo modelo de IA será un amplificador de determinados sesgos en todos aquellos contextos de decisión en los que sea utilizado el modelo).

(b) las personas, además, incorporamos y aprendemos sesgos de la IA. Esto es algo que estamos investigando actualmente en nuestro laboratorio de psicología experimental, y estamos comprobando cómo las personas suelen desarrollar a menudo una confianza ciega hacia las IAs con las que trabajan. Si una persona trabaja con una IA sesgada, aprenderá de ella, y es probable, por tanto, que acabe adquiriendo no solo conocimientos y destrezas beneficiosos, sino también los sesgos y errores de su maestro artificial, reproduciendo posteriormente los sesgos del modelo incluso cuando la IA ya no esté presente[1]. En otras

[1] Vicente, L., y Matute, H. (2023), «Humans inherit artificial intelligence biases», *Scientific Reports, 13*, 15737.

palabras, la IA aprende sesgos a partir de los humanos, los amplifica, y los transmite a la siguiente generación de humanos, de manera que podríamos entrar en un bucle del que puede resultar difícil salir.

Los sistemas actuales de inteligencia artificial no son ya como los ordenadores clásicos a los que estábamos acostumbrados hasta hace bien poco, de los de 2+2 = 4. Ahora mismo lo que pedimos a una IA es que tenga una inteligencia flexible y adaptativa, similar a la inteligencia humana y animal. Que sea capaz de relacionar, por ejemplo, una prueba + otra prueba, y decir, «la probabilidad de reincidencia en este caso es elevada». Pero solo probablemente. Y las personas que han de trabajar con esa IA deben saber cuál es el porcentaje de error en ese «probablemente», pues hay situaciones en que el error puede ser irrelevante, pero otras en las que puede ser catastrófico, y las personas deberán decidir teniendo esto en consideración. Los algoritmos de IA nunca podrán dar una respuesta cien por cien segura, pues el mundo en que vivimos es muy incierto y las decisiones que tomamos también lo son. Y los humanos que tomen decisiones asistidos por esa IA, deberán ser muy conscientes de que la IA cometerá errores. Todas estas cuestiones, sus pros y sus contras, los discute magistralmente el profesor Ponce en este libro, que recomiendo efusivamente.

Como bien dice el profesor Ponce,

> *Los jueces y las juezas deben estar muy atentos a los posibles impactos del uso de sistemas de inteligencia artificial en los derechos de los ciudadanos, puesto que, si bien estos sistemas y la IA pueden redundar en una mejora de la gestión pública, también pueden ser sistemas opresivos y con profundos impactos negativos, lo que exige un escrutinio riguroso de la actuación administrativa.*

De ahí también la importancia del Defensor del Pueblo que destaca el profesor Ponce en su libro. El Defensor del Pueblo deberá tener cada vez más peso y más capacidad de investigación si queremos empezar a adelantarnos a los posibles problemas que puedan ir surgiendo, en vez de detectarlos siempre a posteriori.

Solemos decir los profanos en leyes que la ley va siempre por detrás de los problemas. Hace años que muchos de nosotros estamos pidiendo reglamentación para estos temas, y hace muchos años que escucho cosas como que las leyes van muy lentas, siempre por detrás. Ignoro cómo ha sido el proceso en otros campos, pero puedo decir que, en este, que conozco bien, podemos alegrarnos de la rapidez y el interés que ha suscitado la IA entre los juristas. Hace ya mucho tiempo que venimos observando debates legales, intervenciones, propuestas, que entiendo que ahora cristalizan en este gran reglamento europeo al que el Dr. Ponce dedica su libro. El problema no es que las leyes vayan lentas, sino que las grandes empresas tecnológicas van muy rápido, tienen además

mucho poder, muchos recursos, están en continuo cambio, y no es fácil anticiparse a sus decisiones, sus nuevos productos, y todo lo que ello implica. Pero la comprensión del problema es absolutamente necesaria para poder adelantarse, dentro de lo posible, a lo que está por llegar, y este libro nos acerca, sin ninguna duda, a esa comprensión de los muchos matices, que nos puede permitir anticipar al menos una parte de los problemas a los que nos enfrentaremos en los próximos años, y lo que es más importante, a algunas de sus posibles soluciones.

Otro aspecto importante que debemos comprender y con el que debemos tener cuidado, tal y como indica el profesor Ponce, es el sesgo de automatización, que consiste en esa confianza ciega que las personas depositamos a veces en la IA y que nos lleva a menudo a cometer errores. Entraña especial riesgo este sesgo cuando hablamos de procesos *human-in-the-loop* y/o *human oversight* (el requerimiento europeo de contar siempre con un humano que vigile las decisiones del algoritmo en determinados contextos de riesgo). A pesar de las garantías que esta medida pretende introducir, y que introduce, el problema es que lo que consigue a menudo es simplemente trasladar la responsabilidad de los fallos del algoritmo, desde los diseñadores, las empresas que construyeron el algoritmo, y los gerentes que deciden adoptar el algoritmo a pesar de los fallos conocidos, hasta la persona que trabaja con el algoritmo, el *human-in-the-loop* (la persona en el bucle), a la que cargaremos con la responsabilidad de detectar fallos del algoritmo, sin que hayamos tenido tiempo de formarla, y menos aún de investigar lo suficiente sobre la capacidad de respuesta humana en esa situación concreta, y ante ese algoritmo concreto, o sobre qué medidas debemos introducir para mejorar la eficacia de la decisión conjunta cuando la máquina tiene errores.

A menudo la persona se encuentra ante errores algorítmicos que no puede detectar o de los que no puede informar fácilmente, quizá porque no existe aún el protocolo adecuado para ello, o simplemente porque se encuentra trabajando en contextos que no ayudan a oponerse a la decisión de una máquina a la que, de manera general, consideramos de altísima eficacia y neutralidad (sin ser ello necesariamente cierto). Este tipo de normas, por tanto, deben ir acompañadas de la experiencia, formación, motivación, y protocolos (que a veces serán complejos), que faciliten la labor de esa persona que trabaja con la máquina; de lo contrario solo conseguiremos culpar a personas individuales de los fallos del algoritmo. Resulta especialmente interesante a este respecto la discusión sobre los conceptos de discrecionalidad humana, y la reserva de humanidad, a los que dedica también el profesor Ponce una buena parte de su texto. Desde el punto de vista de la psicología resulta algo absolutamente fundamental a tener en cuenta.

El algoritmo, estamos viendo en numerosos experimentos de psicología, hay que probarlo, hay que auditarlo, no solo de forma aislada, sino también en combinación con los humanos con los que interactúa. Y si no funciona bien, o sus respuestas (las de los humanos que trabajan con la IA) no son lo correctas que deben ser, o se observa que el algoritmo los lleva a cometer errores y sesgos, entonces no debemos adoptar el algoritmo. O quizá, en según qué casos, sea suficiente con ajustar los parámetros, o los protocolos, para conseguir que ese algoritmo, en su interacción con la persona, suponga una mejora para la sociedad y no lo contrario. Es mucha la investigación que queda aún por hacer para poder sacar el máximo rendimiento social. Son muchas las preguntas que surgen cada vez que hacemos un nuevo experimento, y debemos ir con sumo cuidado. Como indica también el Dr. Ponce, «un elemental empleo del principio de precaución social exige la regulación del uso de estos sistemas de IA».

En conclusión, es un placer encontrar escritos como el del profesor Ponce, que abordan el derecho de la IA teniendo en cuenta no solo los aspectos legales y técnicos, sino también los aspectos psicológicos que surgen de la interacción de estas máquinas con las personas, pues solo teniendo como objetivo último el beneficio de la humanidad, y solo con un esfuerzo conjunto de toda la sociedad, y de saberes muy diversos, conseguiremos minimizar los riesgos de una IA mal utilizada, y fomentar en cambio los buenos usos y aplicaciones que pueden tener un impacto altamente beneficioso para nuestra especie.

Helena Matute
Catedrática de Psicología experimental
Universidad de Deusto, Bilbao

CAPÍTULO 1

INTRODUCCIÓN

1. MÁS DE SEIS DÉCADAS DE LUCHA CONTRA LAS INMUNIDADES DEL PODER, MÁS DE UNA DÉCADA DE LA DESAPARICIÓN DEL PROFESOR GARCÍA DE ENTERRÍA Y LA REVOLUCIÓN EN MARCHA CON EL REGLAMENTO DE INTELIGENCIA ARTIFICIAL EN LA UNIÓN EUROPEA

En este estudio se aborda la cuestión del uso de sistemas de inteligencia artificial (en adelante, SIA) por el poder público, concretamente por el poder ejecutivo. Consideraremos, pues, el uso de inteligencia artificial (en adelante, IA) para la adopción de decisiones administrativas y su control por el orden judicial contencioso-administrativo.

Varios son los motivos que han impulsado este análisis.

En marzo de 1962, el profesor García de Enterría dictó una importante conferencia en la Facultad de Derecho de la Universidad de Barcelona, *La lucha contra las inmunidades del Poder*, objeto luego de publicación el mismo año en la *Revista de Administración Pública* y posteriormente como libro (García de Enterría, 1962 y 1974).

Esa reflexión, de la que se han cumplido ya más de sesenta años, ha constituido una obra de gran éxito e influencia, que la doctrina de habla hispana ha ensalzado y reconocido como una de las obras más influyentes en la historia del Derecho público (Muñoz Machado, 2009, la considera «un ensayo muy influyente», mientras que de Carreras, 2013, se refiere a su «célebre conferencia en la Facultad de Derecho de

la Universidad de Barcelona» a la que asistió y de la que «comprendí muy poco» pero que estaba «socavando las bases de la dictadura y sentando los principios de un Estado de Derecho»).

Así, por decirlo en la terminología de Kuhn en relación con las ciencias (Kuhn, 1981), la *lucha contra las inmunidades del poder* marca en el siglo xx los *paradigmas* del Derecho administrativo español, con gran influencia en Hispanoamérica y en la línea de los desarrollos en otros países europeos.

Como argumentaremos en este libro, la *Lucha contra las inmunidades del Poder* ha tenido una importancia capital en el Derecho público español, ha marcado el paradigma de control de las Administraciones durante más de medio siglo y es una obra maestra. Ahora bien, su mismo impacto, debido a su enfoque y a los cambios en marcha en el siglo xxi, está suponiendo una inercia en el pensamiento jurídico español que debe ser alterada a la vista de la revolución que las nuevas tecnologías están suponiendo, como intentaremos explicar.

Además, debe señalarse que, en 2013, es decir, hace ahora más de una década, el profesor García de Enterría falleció. Es un buen momento, pues, para reflexionar sobre su fecundo legado y para intentar adaptarlo a la nueva situación generada por los SIA, de los cuales él, así como otros grandes juristas, no pudieron llegar a ocuparse, debido a la distinta situación tecnológica del momento que les tocó vivir[1].

Finalmente, otro motivo para este análisis es la *revolución* generada por la IA, que genera unos *retos* que inciden en el desarrollo de las *políticas públicas*, en el contexto de la transformación digital en curso que en Europa ha cristalizado en el Reglamento de IA publicado en el *DOUE* de 12 de julio de 2024.

Como ha sido destacado (por ejemplo, Ponce, 2019a), es ya casi un lugar común referirse a que está en marcha desde hace años una cuarta revolución industrial, que, sin duda, está teniendo, y va a tener, profundos impactos en las personas, la sociedad y el Derecho, en la línea de los que generaron las revoluciones científicas y humanas de los siglos xviii y xix que dieron lugar al nacimiento del Derecho administrativo, de cuyos paradigmas básicos aún seguimos viviendo en gran medida. Conviene por ello volver la vista a los pensadores y juristas ilustrados del siglo xix, los cuales, salvando las distancias, se enfrentaron también a enormes cambios tecnológicos y sociales en su época[2].

---

[1] Aunque cabe señalar, por cierto, que la relación entre los ordenadores y el Derecho administrativo no es extraña a nuestra doctrina clásica, entre la que podemos destacar un curioso y breve trabajo del profesor Tomás-Ramón Fernández en 1971 cuya lectura actual es deliciosa (Fernández, T.R., 1971).

[2] Así, en ese sentido, se puede citar, por ejemplo, al jurista francés Josserand, quien, al referirse al desarrollo de la industria y sus consecuencias dañosas, señalaba hace más de un siglo

Sin duda, la revolución tecnológica en marcha en este siglo XXI va a hacer que, en las palabras de Josserand, *las cosas inanimadas se tornen más numerosas*, y está por ver si *mucho más terribles* y *obscuras*.

En todo caso, junto a beneficios para la sociedad de la IA, no cabe duda de que costes y riesgos también existen.

Mientras para algunos la IA será un bálsamo de Fierabrás que curará todos los males de la humanidad, lo que se expone en ocasiones con tintes que rozan lo grotesco[3], otros consideran que puede llegar a destruir a la humanidad[4] o, sin llegar a estos extremos, que tendrá impactos muy negativos respecto a la destrucción de empleos[5] o que terminará con la democracia, debido a las manipulaciones a las que la generación de perfiles mediante el acceso a enormes cantidades de datos de todos nosotros (como muestra el caso de la empresa privada *Cambridge Analytica* y su incidencia en la campaña electoral de Trump para las elecciones estadounidenses de 216 y en el *Brexit,* VERCELLI, 2018), el uso de ingentes cantidades de *bots* maliciosos (más del 70 por ciento del total de tráfico en internet en 2023), con el objetivo de difundir bulos y desinformar o incluso al posible uso de la AI generativa con ultrafalsificaciones (*deepfakes*) que pueden afectar gravemente a las campañas electorales[6].

Este trabajo no asume ni un tecno-optimismo ni un tecno-pesimismo, sino que intenta contribuir a un *tecno-realismo*. El uso de IA presenta, efectivamente, riesgos de destrucción de empleos, especialmente aquellos consistentes en funciones rutinarias, tanto en el sector público, como veremos, como en el privado, pero también puede generar nuevos empleos asociados a su uso; puede causar graves problemas a la humanidad, pero también puede contribuir a luchar contra algunos de

---

que (las cursivas son nuestras):

«Porque la industria se mejora se transforma, es que los accidentes ocasionados *o el hecho de las cosas inanimadas se tornan más numerosas, mucho más terribles y también mucho más obscuras*. La injusticia del sistema tradicional de responsabilidad se ve claramente. *En la inmensa mayoría de los casos, las víctimas se encuentran con la imposibilidad de reconstruir la génesis del accidente, de descubrir por qué y de demostrar la culpa del patrón de la gran industria, del conductor, etc*», JOSSERAND, L. (1910).

[3] ANDERSSEN (2023), añadiendo, además, una perspectiva ideológica del mercado como mecanismo infalible.

[4] Recuérdese la carta abierta publicada en 2023 en la que más de 1 000 expertos (entre ellos… Elon Musk) pedían frenar la inteligencia artificial por ser una «amenaza para la humanidad».

[5] Puede consultarse al respecto el informe especial de *The Economist* sobre inteligencia artificial publicado el 25 de junio de 2016 y titulado «The return of the machinery question», aludiendo a los debates sobre destrucción de empleo que ya suscitó la primera revolución industrial en el siglo XIX.

[6] Sobre la cantidad de *bots* operando en internet, puede consultarse el informe de Arkose Labs de finales de 2023 titulado «Breaking (Bad) Bots: Bot Abuse Analysis and Other Fraud Benchmarks». Sobre los impactos de la IA en la democracia, véase, por ejemplo, DE TENA PIERA (2024).