

La Unión Europea y el futuro de la Inteligencia Artificial

The European Union and the Future of AI

Por MIGLE LAUKYTE
Universidad Pompeu Fabra

RESUMEN

El presente escrito trata la relación entre el sector público y el sector privado a través del uso de la IA que se da en el sector público: utilizar la IA para los servicios esenciales y funciones fundamentales del Estado no es tanto una cuestión de factibilidad tecnológica, como una cuestión de valores que el sector público y el privado no comparten. Cómo esta diferencia de valores afecta el interés común de la sociedad, y cuáles son sus aspectos más problemáticos es el centro de este escrito que también brevemente intenta abordar el posible futuro de la IA y pregunta si hay maneras de gobernarlo.

Palabras clave: inteligencia artificial, sector público, sector privado, trasgresiones sectoriales, La Ley de la IA, derechos fundamentales.

ABSTRACT

This work addresses the relationship between public and private sectors through the use of AI by the former: to use the AI for the essential services and fundamental functions of the State is not a question of technological feasibility, but rather a question of values that public and private sectors do not share. How this difference in values affect the common interest of

society, what its most relevant aspects are is the focus of this article, which also presents a few ideas on the future of AI and asks whether it is possible to govern it.

Key words: artificial intelligence, public sector, private sector, sector transgressions, AI Act, fundamental rights.

SUMARIO: 1. INTRODUCCIÓN.—2. EJEMPLOS DEL USO DE LA INTELIGENCIA ARTIFICIAL (IA). 2.1 *Videovigilancia inteligente en el metro a Londres*; 2.2 *Analítica de los datos de los tribunales*.—3. TRASGRESIONES SECTORIALES O EL AVANCE DE LA IA.—4. LA RESPUESTA DE LA UNIÓN EUROPEA (UE) Y EL FUTURO: LA LEY DE LA IA.—5. CONCLUSIONES Y REFLEXIONES FINALES.

SUMMARY: 1. INTRODUCTION. —2. EXAMPLES OF THE USE OF ARTIFICIAL INTELLIGENCE (AI). 2.1 *SMART SURVEILLANCE IN METRO STATION OF LONDON*; 2.2 *DATA ANALYTIC OF COURTS*. —3. SECTORIAL TRANSGRESSIONS OR THE ADVANCEMENT OF THE AI. —4. THE ANSWER OF THE EUROPEAN UNION (EU) AND THE FUTURE: THE AI ACT. —5. CONCLUSIONS AND FINAL REMARKS.

1. INTRODUCCIÓN

El tema tan amplio –Unión Europea (UE) y la Inteligencia Artificial (IA)– obviamente queda tratado de manera muy limitada sea cual sea la extensión del escrito. Aún más limitado resulta el enfoque de su futuro, que es siempre incierto y cualquier previsión sobre el mismo al fin y al cabo se basa en (y acaba siendo) una especulación. En este contexto de limitaciones manifiestas, el presente escrito aborda algunos aspectos del tema sobre el futuro de IA en la UE para contribuir al (y fomentar el) debate entre los filósofos políticos y filósofos del Derecho en España y en los otros países de habla española.

También seguramente cualquier escrito sobre la IA está destinado a ser obsoleto: el avance tecnológico es tan rápido que cada día descubrimos nuevas aplicaciones, usos, logros, riesgos y problemas de la IA. Tratar cada uno de sus avances de manera puntual no es factible ni deseable. Por lo tanto, en este escrito se proporcionan algunas ideas que emergieron no en base a los avances de IA concretos –entre otros muchos, chatGPT y otros sistemas de IA de propósito general como las llama la Ley de la IA¹– sino en base a las obser-

¹ Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial

vaciones de patrones, tendencias, escritos de muchos estudiosos europeos y de otros continentes que desde hace tiempo están observando las dinámicas tanto en el sector público como en el sector privado, y cuestionando las estrategias de la carrera por el santo grial de la IA.

Entre muchas visiones que se pueden adoptar al analizar el fenómeno de la IA, el presente escrito presenta una visión más bien crítica, refiriéndose al uso de la IA muchas veces sin tener en cuenta sus riesgos ni entender que el sector público –servicios públicos en particular– es muy distinto del sector privado y está sujeto (o, mejor dicho, tendría que estar sujeto) a dinámicas bien distintas: mientras el sector privado privilegia la eficiencia, eficacia, y los beneficios económicos, el sector público tiene que priorizar el interés de la sociedad, tener en cuenta las generaciones futuras, rendir las cuentas, garantizar el apoyo del Estado a quienes lo necesitan, defender los derechos, sobre todo de los más vulnerables, y, en general, pensar al bien común en vez de pensar en el bien individual².

Este escrito está dedicado a argumentar que la IA no siempre ha sido utilizada para defender y promover los valores del sector público y que las más de las veces la IA ha sido el caballo de Troya para que el sector privado pueda conquistar más y crear nuevos espacios para su beneficio comercial. Es más, últimamente podemos ver como en distintos países no (tan) democráticos, la tecnología y la IA en particular están ampliamente desplegados por el sector público, pero ya no como una herramienta del sector privado (con objetivo de incrementar los beneficios económicos), sino como arma de control y vigilancia de masas³.

Para tratar el tema de manera comprensible y coherente, el presente escrito está organizado de la siguiente manera: en primer lugar, enfoco dos ejemplos reales del uso de la IA en dos áreas del interés público: movilidad (2.1 con el ejemplo del metro y videovigilancia inteligente) y justicia (2.2 con el ejemplo de los tribunales y la analítica de los datos que los mismos generan). Luego, en el tercer apartado,

y por el que se modifican los Reglamentos (CE) núm. 300/2008, (UE) núm. 167/2013, (UE) núm. 168/2013, (UE) 2018/858, (UE) 2018/1139 y (UE) 2019/2144 y las Directivas 2014/90/UE, (UE) 2016/797 y (UE) 2020/1828 (Reglamento de Inteligencia Artificial) (Texto pertinente a efectos del EEE). Disponible en: <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689>. Consultado a 19 de diciembre de 2024.

² Para una lista de valores y su posicionamiento en cada uno de sectores, véase VAN DER WAL, Zeger, DE GRAAF, Gjalt., LASTHUIZEN, Karin. «What's valued most? Similarities and differences between the organizational values of the public and private sector», *Public Administration*, vol. 86, núm. 2, 2008, pp. 465-482, en particular, p. 473.

³ Por ejemplo, Mark COELCKERBERG ve la participación de la IA en la erosión democrática y en el aumento de regímenes autoritarios y totalitarios en el mundo y la conecta con «la manipulación, polarización, discriminación, vigilancia, poder y represión». COECKELBERGH, Mark, *Por qué la IA debilita la democracia y que hacer al respecto*. Trad. L. Álvarez Canga, Madrid, Cátedra, 2024, p. 12.

sitúo estos ejemplos en el cuadro más general de las así llamadas trasgresiones sectoriales que el sector público está viviendo desde hace tiempo y que explica por qué el avance tecnológico no es solo un avance de medios y de herramientas, sino y sobretudo un avance del sector privado a los espacios que hasta hace poco eran solo y exclusivamente públicos: de esta manera trazo una línea lógica entre los ejemplos prácticos y la tendencia más global que describen estas trasgresiones sectoriales. En la cuarta parte entro ya en el análisis del futuro de la IA en la Unión Europea (UE) y lo enfoco a través de la ley (la nueva Ley de la IA)⁴ y también hago referencia a la protección y garantías de los derechos fundamentales⁵ sin perder de vista el problema de las trasgresiones sectoriales. Finalizo el escrito con las conclusiones y algunas reflexiones.

Antes de empezar, una premisa: el presente escrito en ningún caso es un escrito en contra de la IA en sí, pero es un escrito en contra de la IA carente de valores públicos que ignora el interés y el bien común.

2. EJEMPLOS DEL USO DE LA INTELIGENCIA ARTIFICIAL (IA)

La IA hoy en día se usa en muchos sectores y para una gran variedad de tareas: en este apartado se enfocan dos ejemplos concretos. El primero, el caso de una estación de metro en Londres, es un ejemplo del despliegue de la IA a nivel local; el segundo, relacionado con el uso de la analítica de los datos generados por los tribunales, trata el uso de la IA a nivel institucional. Luego, en la cuarta parte enfocaremos el patrón que desde hace tiempo se creó en relación con el uso de la IA en el sector público. Esta manera de enfocar el uso de la IA a través de estos ejemplos nos ofrece una matriz de la visión multidimensional del uso de la IA en el sector público y nos ayuda a entender mejor la situación actual.

⁴ En la ponencia original se trató también el tema de la ética de la IA que por motivos de espacio no se tratará en este escrito.

⁵ En el primer artículo de la Ley, leemos que «El objetivo del presente Reglamento es mejorar el funcionamiento del mercado interior y promover la adopción de una inteligencia artificial (IA) centrada en el ser humano y fiable, garantizando al mismo tiempo un elevado nivel de protección de la salud, la seguridad y los derechos fundamentales consagrados en la Carta, incluidos la democracia, el Estado de Derecho y la protección del medio ambiente, frente a los efectos perjudiciales de los sistemas de IA [...] en la Unión así como prestar apoyo a la innovación». Esta referencia, entre muchas otras, explica la importancia de los derechos fundamentales para la visión normativa de la IA en EU.

2.1 Videovigilancia inteligente en el metro a Londres

Se trata de un experimento en una pequeña estación de metro en Londres (Inglaterra)⁶: En la estación de metro *Willesden Green* 23 se instalaron las videocámaras con la IA. La idea era que gracias a estas videocámaras se identificara a las personas que entran sin billete y luchar contra la praxis de viajar sin pagar.

El experimento duró poco más de un año y se instalaron no solo las videocámaras con IA donde se controla acceso a la estación (donde se timbra el billete), sino también se actualizaron los dispositivos en el espacio de control, donde se visualizan las grabaciones de las videocámaras. Los periodistas descubrieron que no se trataba de un par de videocámaras en el control de acceso (como se comunicó al público), sino de muchas videocámaras en todo el edificio de la estación. Por lo tanto, el acceso a los trenes era solo la parte muy pequeña donde videocámaras dotadas de IA habían sido instaladas.

De esta manera y gracias al mucho más amplio despliegue de las videocámaras inteligentes, se descubrieron muchas más cosas que los viajes sin pagar. Obviamente no hay que olvidar que así se trataron algunos incidentes en las vías, como, por ejemplo, si alguien se caía en las vías, automáticamente se enviaba la notificación al personal de la estación garantizando así su rápida intervención. AI también detectaba a quiénes estaban en la línea amarilla, y por tanto, demasiado cerca del tren que entraba en la estación: en estos casos y teniendo en cuenta que a veces esta posición viene adoptada por las personas que consideran el suicidio, también se activaba la alarma para el personal de estación. En todos estos casos el personal tenía una traza de estos eventos en sus dispositivos y podían ver si un caso concreto había sido solucionado o bien si sus colegas aún estaban en ello y necesitaban ayuda.

Es decir, se identificaron numerosos comportamientos más allá del acceso ilegal (objetivo primario para el cual estas videocámaras estaban desplegadas): desde actividades criminales⁷ hasta comportamientos antisociales, molestos o agresivos; se detectaron cosas perdidas y quien las perdió, pasajeros en dificultades, desorientados, per-

⁶ O'MALLEY, James, «TfL's AI Tube Station experiment is amazing and slightly terrifying», del 12 del febrero del 2024. Disponible en: <https://takes.jamesomalley.co.uk/p/tfls-ai-tube-station-experiment-is>. Consultado a 19 de diciembre de 2024. También véase BURGESS, Matt, «London Underground is testing real-time AI surveillance tools to spot crime», *Wired* del 8 de febrero de 2024 y LYDALL, Ross «London Underground: Artificial Intelligence software is being used to detect fare dodgers», *The Standard* del 4 de diciembre de 2023. Disponible en: <https://www.standard.co.uk/news/london/tfl-ai-fare-dodgers-london-underground-artificial-intelligence-b1124563.html>. Consultado a 19 de diciembre de 2024.

⁷ Se quiso también enseñar al sistema a detectar las armas: para lograrlo, la policía ferroviaria paseaba con las armas blancas y de fuego para que las videocámaras inteligentes aprendieran a reconocerlas.

didados, y muchas cosas más. También las videocámaras detectaban las personas en sillas de ruedas y avisaba al personal de la estación: la estación no tenía elevadores y estas personas necesitaban asistencia para bajar a las vías.

Pero el potencial de este sistema de videovigilancia va más allá: cualquier comportamiento fuera de lo que los algoritmos de la IA «consideraban» normal, quedaba señalado como «no normal» y por lo tanto negativo, en el sentido de peligroso e inaceptable⁸. Por ejemplo, estar sentado en la estación y dejar pasar los trenes sin subir a ninguno durante una hora ¿es algo normal o no⁹? Ello puede responder a varias razones (la persona espera a un amigo o bien quiere retrasar la vuelta a casa a causa de una situación familiar) y hoy en día nadie haría caso, pero con la IA esto ya no entra dentro del cuadro de la normalidad algorítmica y quedaría señalado (por ejemplo, enviando un aviso a los dispositivos del personal de la estación)¹⁰.

Sin duda, este sistema de videocámaras, aparte de garantizar más seguridad y mejor asistencia a los pasajeros, también es útil para gestionar los flujos de ciudadanos durante las horas de más afluencia y calcular la disponibilidad de los trenes en distintos horarios del día y del periodo del año.

Al principio ningún ciudadano era identificado a través del reconocimiento facial, pero después se decidió identificar solo a los reincidentes que no pagaban el billete más de X veces. Esto, antes de la IA, se hacía a mano, ya que el personal de estación conoce a los que viajan sin pagar, pues normalmente registraban estos hechos y una vez acumulado un número dado, paraban a la persona y la informaban sobre este dossier acumulado.

Y todo esto es solo lo que podemos hacer hoy: el día de mañana la tecnología será aun más sofisticada, sabrá reconocer más comportamientos y actitudes y podemos imaginar millones de otras oportunidades que esta información y estos datos podrían realizar: desde organizar la limpieza, mejorar la iluminación, hasta eliminar los billetes

⁸ Los estudios sobre el comportamiento de los consumidores desde hace tiempo han comprobado que la IA puede restringir el comportamiento de las personas, entre otras cosas, proporcionándoles un abanico limitado de opciones que impide a las personas explorar alternativas (porque ni saben que dichas alternativas existen ya que la IA no las ofrece, porque para la IA dichas opciones no son opciones para ofrecer, es decir, no son opciones adecuadas, no son aceptables o, en términos del presente escrito, no son normales). VALENZUELA, Ana *et al.*, «How Artificial Intelligence Constrains the Human Experience», *Journal of the Association for Consumer Research*, vol. 9, núm. 3, 2024, pp. 241-256.

⁹ Obviamente en algunos casos extremos –personas que se sienten mal y se desmayan– clasificar este hecho como «no normal» y avisar al personal de estación sería deseable.

¹⁰ Es curioso también especular como la IA modifique la perspectiva del personal de la estación de metro: ¿la realidad es la que pasa en la estación de metro o es la que señala IA y, es lo que de consecuencia aparece en sus dispositivos? Es decir, si algo no viene capturado por las videocámaras, ¿este algo pasó o no?

(digitales o en papel) y pasar directamente al reconocimiento facial como manera de pago entrelazando la cara de una persona con su tarjeta de crédito¹¹.

Si vamos más allá de las estaciones podemos pensar que este uso de videocámaras inteligentes podría ser desplegado también en las calles, plazas, supermercados, parques, autovías, museos, auditorios, iglesias, oficinas, universidades, palacios, correos, hospitales. Seguramente en cada uno de estos sitios desplegando la videovigilancia con videocámaras inteligentes podremos obtener una serie de ventajas para todos en términos de seguridad, eficiencia, organización, ahorro de recursos, entre otros.

Lo preocupante es que para una parte de la sociedad esta realidad podría ser más que aceptable, porque en todos estos lugares aumentaría la seguridad y accesibilidad, y llegaría antes la ayuda que necesitamos¹². En este sentido hay que tener en cuenta que la Ley de la IA –que trataremos en las siguientes secciones de este escrito– en el artículo 5 dedicado a las prácticas de la IA prohibidas, incluye solo «la utilización de sistemas de identificación biométrica a distancia en tiempo real en espacios de acceso público con fines policiales»¹³. Es decir, la prohibición se aplica solo a unos usos muy limitados, dejando abiertas posibilidades para un uso distinto, bajo el apartado de los sistemas de riesgo alto. En este caso concreto, sabemos que la clase de sistemas de alto riesgo incluyen los sistemas biométricos de identificación remota, de categorización biométrica y de reconocimiento de emociones que

¹¹ Sobre el tema del uso de los datos biométricos como sistema de pago, véase entre muchos, desde el punto de vista tecnológico, el reciente trabajo de NOSRATI, L., BIGDOLI, Amir Massoud, SEYYED JAVADI, Hamis Haj, «Identifying People's Faces in Smart Banking Systems Using Artificial Neural Networks», *International Journal of Computational Intelligence Systems*, vol. 17, 2024, 1-21. Para una visión comparativa entre el uso de estos sistemas en EE. UU. y China desde el punto de vista de aceptación y confianza de los usuarios, véase LEE, Jung-Chieh; BI, Lei, LIU, Haotian, «User stickiness to facial recognition payment technology: insights from Sako's trust typology, privacy concerns, and a cross-cultural context», *Humanities and Social Sciences Communications*, vol. 11, 2024. Disponible en <https://www.nature.com/articles/s41599-024-03569-2>. Consultado el 31 de diciembre de 2024.

¹² Más sobre la aceptación y/o necesidad de la video vigilancia, véase FONTES, Catarina *et al.*, «AI-powered public surveillance systems: why we (might) need them and how we want them», *Technology in Society*, vol. 71, 2022, pp. 1-12. También es interesante en este sentido el trabajo de Sarah Kasthuri Francis sobre el uso de IA en videovigilancia a nivel global, preparado para el fórum dedicado a la ciencia, tecnología y la innovación para los Objetivos del Desarrollo Sostenible en mayo de 2024: KASTHURI FRANCIS, Sarah Kasthuri, «Navigating the Intersection of AI, Surveillance, and Privacy: A Global Perspective», *Science-Policy Brief for the Multistakeholder Forum on Science, Technology and Innovation for the SDGs, May 2024*. Disponible en https://sdgs.un.org/sites/default/files/2024-05/Francis_Navigating%20the%20Intersection%20of%20AI%2C%20Surveillance%2C%20and%20Privacy.pdf. Consultado el 20 de diciembre de 2024.

¹³ Y esta prohibición está sujeta a numerosas excepciones.

no sean desplegados en ambientes de trabajo y de educación¹⁴. Por lo tanto, el reconocimiento biométrico en la estación de metro, al menos en teoría no sería contrario a la normativa de la Ley de la IA (no representa el caso del sistema prohibido), si bien obviamente la cuestión de la protección de datos sería una cuestión aparte¹⁵.

La pregunta es: ¿que nos enseña el caso del metro de Londres? El objetivo principal –lucha contra los viajeros que no pagan el viaje– se cumple, pero no gracias a las videocámaras, sino gracias al reconocimiento facial. Pues el medio –videocámaras con IA– no ayudan a reducir infracciones y no cumplen sus objetivos pero ofrecen otras posibilidades que –y es una sospecha sin confirmar– no hubiesen sido aprobadas por las autoridades si se hubieran declarado explícitamente, públicamente y desde el primer momento. Explorar estas posibilidades que responden a las necesidades de eficiencia y seguridad son las que llaman la atención, si bien no son compatibles con los derechos fundamentales: por ejemplo, hemos visto antes que la entrada en una estación de una persona en silla de ruedas queda señalada porque al no existir elevador, los empleados tienen que ayudar a esta persona a bajar las escaleras. ¿Podría parecer una buena acción? Sí y no: sí porque así las personas en sillas de ruedas pueden coger transporte público, y no, porque así se señala a las personas con discapacidad, sus movimientos quedan registrados y grabados, su derecho al anonimato violado y su privacidad quebrantada.

Después de este ejemplo, el siguiente ejemplo enfoca el uso de la IA para analizar la actividad de los tribunales a través de los datos que dichos tribunales generan. Estos datos, entre otras cosas, también ayudan a establecer los patrones y generar los perfiles de los jueces: es un ejemplo del uso de IA que ya no es a nivel local y puntual, como hemos visto en el ejemplo del metro de Londres, sino un ejemplo que plantea preguntas sobre algunos aspectos basilares del siste-

¹⁴ Véase Anexo III «Sistemas de IA de alto riesgo a que se refiere el artículo 6, apartado 2» de la Ley de la IA (*op. cit.*).

¹⁵ No tenemos espacio tratarlo en este escrito: de todas formas, lo que si que nos tendría que hacer pensar es como fácilmente una notablemente amplia parte de la sociedad renuncia a la misma o, mejor dicho, la vende, como hemos visto en el caso del escaneo de iris dentro de un marco de proyecto *Worldcoin*. Sobre este caso, la Agencia Española de Protección de Datos (AEPD) en el marzo del 2024 ha impuesto la medida cautelar para que la empresa responsable por la recopilación de datos biométricos no continúe su actividad en España y que los datos que ya se recopilaban, se bloqueen: en seguida en diciembre 2024 y gracias a las autoridades de protección de datos alemanas –el país donde la empresa en cuestión tiene su establecimiento principal en Europa– se obligó a la empresa a eliminar todos los datos recopilados ilegalmente. Más sobre esta cuestión, véase <https://www.aepd.es/prensa-y-comunicacion/notas-de-prensa/worldcoin-tendra-que-eliminar-todos-los-codigos-de-iris>. Consultado a 19 de diciembre de 2024. También SAN PEDRO ARAÚJO, Carlos, «Protección de datos y escaneo de iris: el “caso Worldcoin” en España», *Actualidad Jurídica Aranzadi*, vol. 1005, 2024.

ma democrático, como, entre muchos, la independencia de los jueces y su autonomía.

2.2 Analítica de los datos de los tribunales

El avance de la IA ha llegado también al sistema de la justicia: en este ejemplo miramos los sistemas que ayudan a entender la jurisprudencia y elaboran los patrones del comportamiento de los jueces. A través de estos sistemas¹⁶, sobre todo desarrollados y ampliamente desplegados en los EE. UU., podemos descubrir las tendencias de juicio (a quién da razón normalmente un juez X en caso de divorcio, si al marido o a la mujer; cómo un concreto juez trata los casos de litigio por los derechos de autor, etc.), la rapidez en solucionar el caso, cuantas sentencias de un mismo juez han sido confirmadas o no en apelación, sus estadísticas comparadas con los demás jueces de un mismo distrito, etc.

Hablando de este tipo de sistemas, podemos hablar de dos tipos de herramientas: las que enfoca la jurisprudencia para establecer los patrones de la casuística (en el caso de divorcio y en presencia de un menor y teniendo acceso a todos los juicios de divorcio de toda la Costa Oeste, ¿cuáles son las probabilidades que el padre se quede con la custodia del hijo?) o perfilar el juez mismo, analizando sus juicios y estableciendo los patrones de su manera de trabajar. Dicha perfilación obviamente lleva no solo a poder entender la manera de tratar distintas cuestiones de un juez concreto, sino también elaborar en base a esto conocimiento estrategias de litigación y predicción del resultado del juicio¹⁷.

La pregunta que plantean este tipo de sistemas es su impacto: si por un lado hacen a los tribunales más transparentes y accesibles, porque les podemos entender mejor, tener los datos para analizar su funcionamiento; por otro lado, puede también negativamente afectar los jueces. Por ejemplo, en base a los variables, se puede extraer datos

¹⁶ Un ejemplo de este tipo de programas está disponible en <https://legal.thomson-reuters.com/en/products/westlaw-edge/litigation-analytics>. Consultado el 31 de diciembre 2024. Otra empresa, entre muchas, que ofrece este tipo de programas es Pre/Dicta, más sobre la misma <https://www.pre-dicta.com/>. Consultado el 31 de diciembre del 2024.

¹⁷ Por cierto, es algo que los abogados siempre han hecho: con los años de experiencia, los abogados ya conocen como mínimo los jueces en cada uno de los tribunales de su ciudad, que tipo de personalidad tienen, y mucha información más. Lo que sí es distinto hoy en día es el poder que este tipo de conocimiento –mucho más preciso y basado no en experiencia individual, percepciones, impresiones individuales, sino en los datos de los juicios– da tanto a las partes del juicio, como a las empresas que desarrollan estos sistemas y podrían utilizarlas también para motivos ilegales, como por ejemplo, para desacreditar un juez concreto manipulando sus datos o representando solo algunos datos parciales, así creando una imagen negativa de la persona.

que podrían manipular la opinión pública o utilizar estos datos para producir daño (por ejemplo, reputacional) a los jueces¹⁸.

En la UE el uso de estos sistemas ha encontrado la oposición explícita en Francia, donde un investigador, gracias a los algoritmos de Machine Learning (ML) ha descubierto que los jueces del mismo tribunal aplican la normativa de inmigración a los refugiados (en las situaciones idénticas) en modo distinto. Por lo tanto, el permiso de asilo no es una certeza que se basa en cumplimiento de requisitos establecidos en la ley, sino una cuestión aleatoria que depende del juez que trata el caso¹⁹.

La reacción del legislador francés ha sido contundente: el artículo 33 de la Ley 2019-222 establece que «los datos relativos a la identidad de los jueces [...] no pueden ser reutilizados para evaluar, analizar, comparar o predecir sus actividades profesionales reales o supuestas»²⁰, pero si la norma protege a los jueces, ¿quién protege a los solicitantes de asilo?

En este contexto de analíticas empoderadas por la IA surgen además algunas preguntas sobre la naturaleza de nuestras instituciones, sobre la comercialización de la justicia, la legitimización de perfilación de algunas profesiones para el bien común, como, por ejemplo, la lucha contra las «cajas negras» de sesgos y discriminación invisible en la justicia. También surgen dudas del verdadero estado de la salud del sistema de justicia y se plantean las preguntas sobre cómo entendemos este estado: ¿aquel generado por los datos lo es? ¿A lo mejor lo que nos ofrecen

¹⁸ Si lo pensamos, el uso de estos sistemas para evaluar y perfil cualquier profesión –médico, profesor universitario, jugador de básquet– podría llevar a un perfil que será definido por los variables introducidos que definirá el resultado del análisis. Este resultado no necesariamente refleja la verdad: aquí tenemos que acordarnos de que ningún sistema refleja la verdad, sino el modelo de la realidad que sus programadores eligieron. Más sobre este tema, O'NEIL, Cathy, *Armas de destrucción matemática: Cómo el Big Data aumenta la desigualdad y amenaza la democracia*, Madrid, Capitan Swing, 2018.

¹⁹ La versión de hechos por el mismo Michaël Benesty está disponible en BENESTY, Michael, «The Judge Statistical Data Ban –My Story– Michaël Benesty», *Artificial Lawyer* del 7 del junio de 2019. Disponible en <https://www.artificiallawyer.com/2019/06/07/the-judge-statistical-data-ban-my-story-michael-benesty/>. Consultado el 21 de diciembre de 2024.

²⁰ La versión original: [...] «Les données d'identité des magistrats et des membres du greffe ne peuvent faire l'objet d'une réutilisation ayant pour objet ou pour effet d'évaluer, d'analyser, de comparer ou de prédire leurs pratiques professionnelles réelles ou supposées», art. 33 de La Ley 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice (1), JORF n. 0071 del 24 marzo 2019. La dicha ley se ampara a la normativa europea sobre la privacidad y protección de datos que permite al sujeto oponerse a la perfilación cuando la misma «produzca efectos jurídicos en él o le afecte significativamente de modo similar», si bien existen excepciones a esta norma general, como establece el art. 22 del 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 (Reglamento general de protección de datos) Disponible en <https://www.boe.es/doue/2016/119/L00001-00088.pdf>. Consultado a 19 de diciembre de 2024.

estos datos no es otra cosa que una forma de vigilancia de masas que disrumpe las principales instituciones del Estado de Derecho y hace tremar el equilibrio de los poderes? O, al contrario, ¿hablamos de nuevos niveles de transparencia en un sector público tan importante?

¿Qué tienen en común y que nos cuentan estos dos ejemplos? En ambos casos podemos observar el uso de la IA para vigilar directa o indirectamente las personas, su trabajo, su modo de ser, sus patrones de comportamiento, razonamiento, interacción con los demás: esta vigilancia origina muchas posibilidades para violar los derechos fundamentales, como la privacidad, libertad, autonomía, anonimato, entre otros. Pero también esta vigilancia nos revela aspectos del comportamiento humano –sobre todo, por parte de las personas que toman decisiones, sean estas personas encargadas de la seguridad de la estación o los jueces– que nos deja preocupados por los sesgos y actitudes discriminatorias, fallos y faltas, que sin IA hubiese quedado escondidos y nunca hubieran visto la luz del sol. Y si bien estas revelaciones nos generan la sensación de inseguridad y arbitrariedad y nos hacen dudar de todo y de todos, seguimos sin entender si lo que dichos sistemas revelan son la verdad o la verdad algorítmica, que es fruto de la visión selectiva y limitada de la realidad.

Y es en esta óptica que entramos en la tercera parte del presente escrito donde extraeremos alguna reflexión más de estos ejemplos y los situaremos dentro de un nuevo patrón de la interacción entre el Estado y las grandes multinacionales –entre el sector público y el sector privado– que señala el cambio del paradigma en la interacción entre ambos.

3. TRASGRESIONES SECTORIALES O EL AVANCE DE IA

Las trasgresiones sectoriales es un término que ha sido utilizado por varios investigadores para explicar la expansión del sector privado (tecnológico sobre todo, pero en particular se tiene en cuenta las grandes multinacionales tecnológicas como *Microsoft*, *Apple*, *Meta*) en los ámbitos que tradicionalmente y por el tipo de sociedad en la que vivimos eran siempre ámbitos gestionados por el sector público: en particular, la educación, la comunicación, la agricultura, la sanidad, la movilidad, la pública administración (el funcionamiento del Estado) y la política en general. Bajo el lema de la eficiencia y ahorro de los costes y otros valores e intereses (más bien) corporativos, las empresas penetran en muchos sectores con el objetivo de expandir sus zonas de control y poder²¹.

²¹ Entre muchos, SHARON, Tamar «Blind-sided by privacy? Digital contact tracing, the Apple/Google API and big tech's newfound role as global health policy makers», *Ethics and Information Technology*, vol. 23, 2021, pp. 45-57; TAYLOR, Linnet. «Public

Es más: en estas dinámicas de trasgresión, el conocimiento y habilidades sectoriales –médicas, administrativas, políticas, que han representado la fuente de legitimidad de las decisiones– son sustituidas por el conocimiento y habilidades digitales. En otras palabras, tener los datos convierte a cualquiera en experto en la materia, sea cual sea ésta. Pero este desplazamiento de conocimientos sectoriales también significa eliminar los valores y las normas que definen estos sectores automatizándolos de manera reduccionista y re-orientándolos a los valores del sector privado.

Lo que pasó es que no solo se traspasaron los valores del sector privado, sino también las consecuencias negativas: entre muchas, la falta de rendición de cuentas se hizo endémica en el funcionamiento del Estado, cuando cualquier solicitud del acceso a la información pública y de la transparencia en las actuaciones se encuentra con el muro de la protección de propiedad intelectual, en particular, protección del secreto comercial. En este sentido el caso del Bosco es bien conocido por todos²²: no se puede conocer el código fuente del sistema el Bosco –si bien los algoritmos utilizados por públicas administraciones están sujetos al derecho del acceso a la información pública– porque, como señala Cotino Hueso, el Gobierno ha utilizado la propiedad intelectual para proteger la seguridad informática del dicho sistema. Según las autoridades responsables, abrir el código fuente a la asociación de la sociedad civil significaría exponer este sistema (y otras con ella conectadas) a los ataques informáticos²³.

El caso prueba otro aspecto: las trasgresiones sectoriales también aportan justificaciones y herramientas sectoriales. En este caso específico esta herramienta es el régimen de la propiedad intelectual: ¿cómo es posible que un algoritmo desarrollado con el dinero público sea protegido por este régimen que protege el retorno de la inversión privada y garantiza la ventaja competitiva en el mercado y, es más, al menos según algunos autores, « [...] transforma la creatividad, la información, la ciencia y tecnología de bienes comunes a bienes privados»²⁴?

Otro problema es la falta del equilibrio entre las inversiones públicas, intereses de los ciudadanos y los beneficios que las grandes empresas tecnológicas obtienen del uso de datos que son de toda la ciudadanía: por ejemplo, el desarrollo de los algoritmos para las apli-

actors without public values: Legitimacy, domination and the regulation of the technology sector», *Philosophy & Technology*, vol. 34, núm. 4, 2021, pp. 897–922.

²² Sobre el caso, mirase la página web del Civio, disponible en <https://civio.es/tag/bono-social/>. Consultado el 20 de diciembre del 2024.

²³ COTINO HUESO, LORENZO «Caso Bosco, a la tercera tampoco va la vencida. Mal camino en el acceso los algoritmos públicos», *Diario La Ley*, núm. 84, del 17 de mayo de 2024. Ahora se espera la sentencia del Tribunal Supremo que ha admitido a trámite el recurso del Civio en diciembre de 2024.

²⁴ Mi traducción de SCHAUER, Lea «The right to science and culture», *Wisconsin Law Review*, vol. 1, 2010, p. 124.

caciones médicas que son desarrolladas en base de los datos de muchos ciudadanos, al final y al cabo son algoritmos privados, desarrollados por empresas privadas que no comparten los mismos valores de reciprocidad, altruismo y solidaridad que definen y son banderas del sector público²⁵.

Así llegamos al problema que ya ha sido identificado en la literatura²⁶: el problema de los datos no es el control sobre los mismos, sino su uso y en este particular contexto, su uso para entrenar la IA. Nos hemos olvidado de formular una de las preguntas más importantes: ¿el uso concreto de la IA beneficia a la sociedad²⁷? ¿Cómo? ¿Hay perjuicios para algún grupo social? En particular, ¿este uso afecta a grupos sociales vulnerables, como menores de edad, personas mayores, personas con discapacidad? ¿El uso podría discriminar injustamente? Si es así y dicha discriminación no puede ser eliminada, no podría existir ningún apoyo público a dicho uso: en este sentido podemos preguntarnos ¿a quién beneficia la videovigilancia inteligente en el metro de Londres? El uso de este sistema era presentado al público como beneficioso en el sentido de parar la gente que no paga por sus viajes, pero al final hemos visto que el resultado no fue este. Es más, también hemos visto que el uso era discriminatorio para las personas en silla de ruedas porque su llegada a la estación era anunciada a los empleados que les tenían que ayudar a bajar en la estación; por lo tanto, su libertad de movimiento y su derecho al anonimato han sido quebrantadas. ¿Tenemos que apoyar este uso de IA? Obviamente no.

La misma pregunta podemos hacernos en el caso de la analítica del trabajo de jueces: ¿quién se beneficia? Dicho en otras palabras, ¿cómo se beneficia la sociedad de estos sistemas? En este caso no lo tenemos tan claro: la sociedad se queda en manos de los desarrolladores de

²⁵ Sobre el historial de la interacción entre las empresas tecnológicas y *National Healthcare Services (NHS)* de Inglaterra, véase HORN, Ruth; KERASIDOU, Angeliki, «Sharing whilst caring: solidarity and public trust in a data-driven healthcare system», *BMC Medical Ethics*, vol. 21, 2020. Disponible en <https://bmcomedethics.biomedcentral.com/articles/10.1186/s12910-020-00553-8>. Consultado a 19 de diciembre de 2024.

²⁶ PRAINSACK, Barbara *et al.*, «Data solidarity: a blueprint for governing health futures», *The Lancet Digital Health*, vol. 4, núm. 11, 2022, pp. 773-774; PRAINSACK, Barbara; EL-SAYED, Seliem, «Beyond individual rights: How data solidarity gives people meaningful control over data», *The American Journal of Bioethics*, vol. 23, núm. 11, 2023, pp. 36-39; PRAINSACK, Barbara; KICKBUSCH, Ilona, «A new public health approach to data: why we need data solidarity?», *British Medical Journal*, vol. 386, 2024. Disponible en www.bmj.com/content/bmj/386/bmj.q2076.full.pdf. Consultado el 20 de diciembre del 2024.

²⁷ ACEMOGLU y JOHNSON formulan preguntas muy parecidas y explícitamente se refieren a la «utilidad de las máquinas» que debería llevar al empoderamiento humano y soporte a sus habilidades. Es más, estos autores preguntan por qué enfocamos tanto la inteligencia de las máquinas, cuando lo que nos tendría que interesar es su utilidad. ACEMOGLU, Daron; JOHNSON, Simon, *Power and progress: our thousand-year struggle over technology and prosperity*, Londres, Basic Books, 2023, p. 299; pp. 304-305.

estos sistemas que en base a distintos variables –variables a los que por cierto nunca tendremos acceso gracias a los secretos comerciales– pueden manipular resultados del análisis de trabajo de los jueces en base a sus intereses o en base a los intereses del poder ejecutivo y/o poder legislativo (es decir, en base a los intereses de quien paga por estos sistemas). Por lo tanto, la «verdad» generada por este tipo de sistemas opacos no hace ningún favor a la democracia, transparencia del sistema judicial, separación de los poderes y en general a la confianza social en las instituciones.

Este uso social de la tecnología se alinea con el razonamiento de Adela Cortina que en su último trabajo observa: « [...] en un mundo en que es una realidad sangrante del sufrimiento causado por las guerras, la pobreza, la aporofobia y la injusticia, [...] ¿No es una exigencia ética palmaria utilizar los grandes beneficios de la IA para resolver estos problemas acuciantes?»²⁸. Y es verdad que no oímos mucho hablar de los sistemas de la IA que ayudan a buscar la paz entre Israel y Gaza o acabar con la hambruna en Somalia, Burkina Faso o Mali.

La influencia del sector privado sobre el sector público lleva también a preguntarnos como el poder privado impacta en la distribución de las inversiones públicas y la priorización de sectores de la inversión: dichas decisiones tendrían que ser fruto de la deliberación pública que incluye y refleja la pluralidad de intereses, que obviamente no existe en la toma de decisiones delineada por el sector privado, quien además ya ha envuelto el sector público en sus infraestructuras informáticas de las que no hay manera de librarse: el *Drive*, *Cloud*, *Gmail*, *Windows*, todos sistemas que utiliza el sector público desde las administraciones locales hasta las universidades públicas nacionales.

En este sentido, cuanto más el sector privado toma funciones públicas, tanto más los espacios públicos (virtuales o físicos) se vuelven opacos y orientados a las dinámicas comerciales y, como cualquier espacio privado, pueden en cualquier momento ser vendidos, cerrados, modificados: pensemos en cualquier espacio virtual abierto al público que no sea una red social o una plataforma digital. En ambos casos son espacios privados. La recreación de espacios públicos en un medio privado es una de las tantas maneras en las que la relación entre los ciudadanos y el Estado se quiebra. Estamos observando el cambio de paradigma del concepto de lo que es público, porque lo público ya no es público, en el sentido de ser de todos y representar los intereses de todos, sino lo público es del Estado y el Estado es un poder, una infraestructura, un sistema que representa sus intereses de sustentarse y mantenerse alineándose a las dinámicas del sector privado, pero no para defender y representar las preocupaciones ciudadanas. Y es por este motivo que vivimos un evidente distanciamiento y desconfianza del Estado y, en consecuencia, en la cosa pública y estamos más abiertos a los movimientos populistas y antidemocráticos.

²⁸ CORTINA, Adela *¿Ética o ideología de la inteligencia artificial? El eclipse de la razón comunicativa en una sociedad tecnologizada*, Barcelona, Paidós, 2024, p. 38.

Lo podemos observar en el auge de movimientos ciudadanos, sociedad civil y otras entidades del tercer sector que crecen cada año²⁹ y también nuevas fórmulas en entender las colaboraciones entre el público y el privado, donde siempre más frecuentemente encontramos ya no solo la colaboración público-privada sino también las colaboraciones del público-privado-de personas (*public-private-people*), donde el papel de la sociedad civil toma protagonismo que nunca tuvo antes³⁰.

En este apartado hemos retratado brevemente la situación actual donde el sector privado, liderado por grandes empresas multinacionales tecnológicas está utilizando la IA para acabar conquistando y adueñando el sector público. Hemos visto también los distintos desafíos y líneas del pensamiento que se entrelazan y se entretejen en tratar, observar, recapacitar sobre esta realidad y ofrecer alguna manera de darle sentido. El objetivo de estas críticas, al fin y al cabo, es dirigir esta coyuntura en la dirección más favorable al interés y el bien común, que en este momento el Estado falla en defender y solo la ciudadanía a través de la sociedad civil mantiene de pie. En el siguiente apartado abocetaremos la visión de la IA por parte de la UE, que si bien «presenta muchos problemas de conexión y de entendimiento», y «el pensamiento comunitario [...] no termina de funcionar a pleno rendimiento»³¹, sigue siendo la fuente principal de las políticas comunes –no solo, pero también sobre los temas de la IA– que adoptan todos los países miembros.

²⁹ Véase, por ejemplo, el resumen ejecutivo del «Barómetro del tercer sector de acción social en España 2022», preparado por Plataforma ONG de Acción Social, Ministerio de Derechos Sociales y Agenda 2030 de Gobierno de España en colaboración con otros, publicado en el 2023 y accesible en https://www.plataformaong.org/ARCHIVO/documentos/biblioteca/1676295113_resumenejecutivo-barometro-3-sector.pdf. Consultado a 19 de diciembre de 2024; SIEBER, René et al., «Who are the publics engaging in AI?», *Public Understanding of Science*, vol. 33, núm. 5, 2024, pp. 634-653.

³⁰ KOBASHI, Yurie et al., «Maturing of public-private-people partnership (4P): Lessons from 4P for triple disaster and subsequently COVID-19 pandemic in Fukushima», *Journal of Global Health*, vol. 12. Disponible en <https://pmc.ncbi.nlm.nih.gov/articles/PMC9172965/pdf/jogh-12-03028.pdf>. Consultado el 21 de diciembre del 2024; Perjo, Liisa «Public-Private-People partnerships – a new concept to bring public and private actors and citizens together», *Nordregio News*, núm. 4, 2016. Disponible en <https://norden.diva-portal.org/smash/record.jsf?pid=diva2%3A1128939&dswid=1797>. Consultado el día 20 de diciembre del 2024.

³¹ OTERO PARGA, Milagros. «¿Existe en realidad una Unión Europea?», *Anuario de Filosofía Del Derecho*, vol. 40, 2024, p. 141.

4. LA RESPUESTA DE LA UNIÓN EUROPEA (UE) Y EL FUTURO: LA LEY DE LA IA

En el presente apartado analizo la visión europea del futuro de la IA a través de la Ley de la IA. La Ley de la IA –que mira a la IA a través del prisma de los riesgos– es una iniciativa que en sí merece más espacio del que aquí le podemos dedicar. Me limito al enfoque de las trasgresiones sectoriales: miraré una serie de prospectivas que ayudan a cristalizar y entender mejor a donde vamos, visto lo que hemos descrito en la sección anterior.

En primer lugar, la ley no se aplica a la investigación de la IA: por este motivo se crea más espacio para las instituciones públicas (universidades, centros de la investigación, ...) para buscar la manera de desarrollar IA útil para el bien común. Seguramente la investigación es libre de límites normativos también para el sector privado, pero es el sector público el que tiene más limitaciones para asumir el cumplimiento de la normativa y es por eso por lo que librar de este cumplimiento al sector más débil desde el punto de vista económico, es beneficioso para la sociedad y el interés común.

En segundo lugar, la ley reconoce (¡modestamente!) el papel de la sociedad civil en el desarrollo y el despliegue de la IA: por ejemplo, reconoce el papel que la sociedad civil podría jugar en la evaluación del impacto de los sistemas de IA –sobre todo destinada al sector público– abre las puertas para las colaboraciones público-privadas de personas (*public-private-people partnerships*) a las que nos hemos referido antes, y abre la puerta a la función e importancia de la sociedad civil en el juego de equilibrios de poderes³².

La Ley de la IA también proporciona otra herramienta que, si implementada correctamente, empoderaría la sociedad y no permitiría quitarle la autonomía decisional y libertad de pensamiento: hablamos de la alfabetización en la IA que está construida por distintos elementos como las habilidades de reconocer la IA, enten-

³² Por ejemplo, el considerando 96 incluye lo siguiente «Cuando proceda, para recopilar la información pertinente necesaria para llevar a cabo la evaluación de impacto, los responsables del despliegue de un sistema de IA de alto riesgo, en particular cuando el sistema de IA se utilice en el sector público, pueden contar con la participación de las partes interesadas pertinentes, como, por ejemplo, los representantes de colectivos de personas que probablemente se vean afectados por el sistema de IA, expertos independientes u organizaciones de la sociedad civil, en la realización de dichas evaluaciones de impacto y en el diseño de las medidas que deben adoptarse en caso de materialización de los riesgos» o en el Considerando 116 subraya que «Para garantizar que los códigos de buenas prácticas reflejen el estado actual de la técnica y tengan debidamente en cuenta perspectivas distintas, la Oficina de IA debe colaborar con las autoridades nacionales competentes pertinentes y, cuando proceda, podrá consultar a organizaciones de la sociedad civil y a otras partes interesadas y expertos pertinentes, incluido el Grupo de Expertos Científicos, por lo que respecta a la elaboración de dichos códigos». También prevé la participación de la sociedad civil en el foro consultivo del Consejo de la IA y de la Comisión Europea (art. 67).

derla, saber utilizarla correctamente, evaluarla y también poder crearla y poder entender sus aspectos éticos³³. El considerando 20 lo adelanta con las siguientes palabras: «Con el fin de obtener los mayores beneficios de los sistemas de IA, protegiendo al mismo tiempo los derechos fundamentales, la salud y la seguridad, y de posibilitar el control democrático, la alfabetización en materia de IA debe dotar a los proveedores, responsables del despliegue y personas afectadas de los conceptos necesarios para tomar decisiones con conocimiento de causa en relación con los sistemas de IA»³⁴. Dicho objetivo toma su forma en el artículo 4, donde la alfabetización sobre la IA se relaciona con los «[...] conocimientos técnicos, su experiencia, su educación y su formación, así como el contexto previsto de uso de los sistemas de IA y las personas o los colectivos de personas en que se van a utilizar dichos sistemas»³⁵. A este punto son muy pertinentes las palabras de Adela Cortina quien apunta que para la ética de la IA es indispensable «el derecho de cada persona a comprender aquello que le afecta, para poder asumir su respuesta de forma autónoma, y el derecho a que su respuesta sea tomada realmente en serio por personas, [...], que se hacen responsables de ello»³⁶.

Con estas consideraciones, la Ley avanza la idea que la ignorancia sobre la IA ya no tiene justificación: es más, los conocimientos sobre la IA también aseguran que no caigamos en la trampa de marketing de que la IA (sobre todo, la generativa) sea la respuesta a todo. Más bien al contrario: conocer la IA también significa ser conscientes de sus limitaciones, sesgos, fallos, y vulnerabilidades en términos de seguridad y adoptar una visión crítica en relación con sus resultados³⁷.

³³ ALMATRAFI, Omayma; JOHRI, Aditya; LEE, Hyuna, «A systematic review of AI literacy conceptualization, constructs, and implementation and assessment efforts (2019–2023)», *Computers and Education Open*, vol. 6, junio 2024. Disponible en <https://www.sciencedirect.com/science/article/pii/S2666557324000144>. Consultado el 28 de diciembre del 2024.

³⁴ Y sigue: «[...] En el contexto de la aplicación del presente Reglamento, la alfabetización en materia de IA debe proporcionar a todos los agentes pertinentes de la cadena de valor de la IA los conocimientos necesarios para garantizar el cumplimiento adecuado y la correcta ejecución. Además, la puesta en práctica general de medidas de alfabetización en materia de IA y la introducción de acciones de seguimiento adecuadas podrían contribuir a mejorar las condiciones de trabajo y, en última instancia, sostener la consolidación y la senda de innovación de una IA fiable en la Unión. El Consejo Europeo de Inteligencia Artificial (en lo sucesivo, «Consejo de IA») debe apoyar a la Comisión para promover las herramientas de alfabetización en materia de IA, la sensibilización pública y la comprensión de los beneficios, los riesgos, las salvaguardias, los derechos y las obligaciones en relación con el uso de sistemas de IA [...]».

³⁵ La Ley de la IA, *op. cit.*

³⁶ CORTINA, Adela, *op. cit.*, p. 197.

³⁷ HEIKKILÄ, Melissa, «Nobody knows how AI works», *MIT Technology Review* de 5 de marzo del 2024. Disponible en <https://www.technologyreview>.

Otro aspecto relevante de esta Ley es la atención dedicada a la IA con la licencia libre y del código abierto³⁸: tanto la IA con la licencia libre, que se basa en las libertades de usuarios, sea la IA del código abierto, que se basa en la colaboración abierta entre los programadores, no sería regulada por la Ley, a menos que se comercialice como sistemas de IA de alto riesgo, o represente sistemas prohibidos y/o sistemas especiales –por ejemplo, sistemas de IA que interactúan con las personas, o sistemas de IA que pueden generar imágenes, videos, textos, etc.– a las que se aplican requisitos de transparencia.

El código abierto es un tema complejo, porque ser código abierto no es una característica monolítica: por ejemplo, se distinguen 14 dimensiones de la apertura que un sistema podría haber y por lo tanto tenemos que entender dicha apertura del código no solo como apertura compuesta de muchos ingredientes, sino también como la apertura gradiente, es decir, que tiene distintos niveles³⁹. Por lo tanto, clasificar los sistemas de IA en abiertos y no es una visión muy limitada porque hay muchas maneras de ser abiertos y el sector utiliza el término «código abierto» aun cuando dicho código no es abierto ni mucho menos⁴⁰. Por lo tanto la Ley de la IA seguramente necesitará tratar este tema en el futuro para evitar que la IA que no sea de código abierto pase por IA de código abierto y así evita cumplimiento de la normativa.

La Ley de la IA también enfatiza la transparencia de la IA, que entendemos como un término paraguas que englobe otros términos cuales la explicabilidad, la interpretabilidad, la auditabilidad y demás⁴¹. Sin duda, la transparencia viene ligada a otro requisito que

com/2024/03/05/1089449/nobody-knows-how-ai-works/. Consultado el 28 de diciembre del 2024. Sobre los mitos, ideas falsas y preconcepciones sobre la IA, véase BEWERSDORFF, Arne *et al.*, «Myths, mis- and preconceptions of artificial intelligence: A review of the literature», *Computers and Education: Artificial Intelligence*, vol. 4, 2023. Disponible en <https://www.sciencedirect.com/science/article/pii/S2666920X2300022X>. Consultado el 29 de diciembre del 2024.

³⁸ La definición de la IA del código abierto ha sido presentada el noviembre pasado por parte de *Open Source Initiative* quienes la describen como un sistema de IA mesa a disposición bajo las condiciones y en la manera que permita las libertades de utilizar el sistema para cualquier objetivo y sin tener que pedir permisos; estudiar como el sistema funcione y inspeccionar sus componentes; modificar el sistema por sea cual sea motivo, incluido también el cambio de los objetivos del sistema mismo; compartir el sistema con los demás para su utilizo con o sin modificaciones por cualquier motivo (mi traducción). El original en inglés disponible en <https://opensource.org/ai/open-source-ai-definition>. Consultado el 28 de diciembre del 2024.

³⁹ LIESENFELD, Andreas; DINGEMANSE, Mark, «Rethinking open source generative AI: open-washing and the EU AI Act», en *The 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24)*, 3-6 junio, 2024, Rio de Janeiro, Brazil. ACM. Disponible en https://pure.mpg.de/pubman/item/item_3588217_2/component/file_3588218/liesenfeld_dingemanse_2024_FAccT_generative_AI_open-washing_EU_AI_Act.pdf. Consultado el 28 de diciembre del 2024.

⁴⁰ *Ibid.*

⁴¹ Más sobre este tema, véase COTINO HUESO, LORENZO; CASTELLANOS CLARAMUNT, Jorge (eds.), *Transparencia y explicabilidad de la Inteligencia Artificial*.

siempre se enfatiza más: la supervisión humana de los sistemas de la IA. En la supervisión humana confluyen muchos de los requisitos antes vistos: por ejemplo, la alfabetización sobre la IA y los sistemas a código abierto (en otras palabras, supervisadas por una pluralidad de actores que comparten visión de *software* como iniciativa común y no como producto comercial) son maneras de permitir a las personas ser más activas e involucradas en supervisión de los sistemas de IA.

Sin ninguna duda, la supervisión humana de los sistemas de la IA dará mucho que hablar por sus limitaciones, fallos y otras imperfecciones, muchas veces relacionadas con la naturaleza humana y con el hecho de que los seres humanos no somos unas máquinas: nos cansamos, tenemos nuestros sesgos y prejuicios, no prestamos la debida atención, elegimos cerrar los ojos y mucho más, sin hablar de los obstáculos objetivos relacionados con la complejidad algorítmica, rapidez de los sistemas, o su autonomía⁴². Queda mucho para crear una colaboración ser humano-sistema IA que sea por un lado efectiva y por otro lado adaptada al ser humano, donde el ser humano sea plenamente consciente y capacitado para desplegar esta función en la mejor manera posible.

La visión de la IA del legislador europeo se orienta más a ver la IA como producto y no como una herramienta que va más allá de ser un producto o un servicio, pero como hemos visto en los ejemplos analizados al inicio de este trabajo, el problema no es solo (ni puede ser reducida a solo) cuestión de la seguridad de producto o servicio digital como podría sugerir la Ley de la IA. Por ejemplo, si miramos el ejemplo de la analítica de los datos de las sentencias en los tribunales (y, en consecuencia, si bien a lo mejor sin explicitarlo, de los jueces mismos), tenemos el problema de la igualdad de acceso a estos sistemas, la manipulación de las pruebas con IA generativa

Valencia, Tirant lo Blanch, 2022; VELASCO RICO, Clara I., «Marco regulatorio de los sistemas algorítmicos y de Inteligencia Artificial: el papel de la administración», INAP, 2024. Disponible en <https://www.aepda.es/VerArchivo.aspx?ID=4185>. Consultado el 14 de diciembre del 2024.

⁴² HOLZINGER, Andreas; ZATLOUKAL, Kurt; MÜLLER, Heimo, «Is human oversight to AI systems still possible?», *New Biotechnology*, vol. 85, 2024, pp. 59-62; VAN VOORST, Roanne, «Challenges and Limitations of Human Oversight in Ethical Artificial Intelligence Implementation in Health Care: Balancing Digital Literacy and Professional Strain», *Mayo Clinic Proceedings: Digital Health*, vol. 2, núm. 4, 2024, pp. 559-563. También PONCE SOLE, Juli, «Reserva de humanidad y supervisión humana de la Inteligencia Artificial», *El Cronista del Estado Social y Democrático de Derecho*, vol. 100, 2022, pp. 58-67; OBREGÓN FERNANDEZ, Aritz; LAZCOZ MORATINOS, Guillermo, «La supervisión humana de los sistemas de inteligencia artificial de alto riesgo. Aportaciones desde el Derecho Internacional Humanitario y el Derecho de la Unión Europea», *Revista electrónica de estudios internacionales (REEI)*, vol. 42, 2021. Disponible en <https://dialnet.unirioja.es/servlet/articulo?codigo=8202329>. Consultado el 18 de diciembre de 2024.

(que reverbera sobre la admisibilidad e interpretación de las pruebas digitales en general), la generación y falsificación documental (de solicitudes, dictámenes, expedientes), necesidad de la normativa sobre el uso permitido de los sistemas IA en los procesos judiciales y sobre todo la necesidad de reformar el trayecto de estudios y preparación de los jueces, incluida la educación continua en la alfabetización sobre IA.

Es decir, lo que podría ser uno de los problemas fundamentales de la Ley de la IA es no entender que la IA es una forma de invasión sin control de muchas de nuestras instituciones democráticas, tanto a nivel local como a nivel nacional y en muchos contextos. Y es por eso que, por lo menos al principio (y ojalá sea solo al principio), su manejo y control por parte de la ciudadanía será problemático⁴³.

Una breve mención aparte es para los derechos fundamentales, que son también el punto de referencia en la Ley de la IA y representa una dificultad para una completa invasión de la IA. La centralidad de los derechos fundamentales en el debate sobre la gobernanza y el futuro de la IA ha sido uno de los pocos aspectos de consenso en la comunidad académica⁴⁴ y también entre distintas instituciones europeas, estadounidenses y también internacionales⁴⁵.

El primero de los derechos que dificulta el camino es el derecho a la privacidad y la protección de datos. Asistimos al cambio del paradigma en este sentido: estamos atravesando el cambio de ver el derecho a la privacidad como derecho negativo a tener los demás fuera de nuestro espacio más íntimo al derecho positivo de exigir la libertad a

⁴³ Por motivos de espacio no se trata otros problemas de esta ley, como, por ejemplo, la gestión de riesgos de la IA es en las manos de los proveedores mismos quienes por cierto son primeros interesados en no reconocer los riesgos de sus sistemas o los estándares para la IA. Hoy en día para acceder a dichos estándares, hay que pagar y por lo tanto el público tiene poco acceso, poco conocimiento y, aún menos, impacto sobre su contenido.

⁴⁴ Véase en este sentido los excelentes trabajos de, entre muchos, PRESNO LINERA, Miguel Ángel, *Derechos fundamentales e inteligencia artificial*, Madrid, Marcial Pons, 2022; VILLEGAS DELGADO, César; MARTÍN-RÍOS, Pilar, *El derecho en la encrucijada tecnológica: Estudios sobre derechos fundamentales, nuevas tecnologías e inteligencia artificial*, Valencia, Tirant lo Blanch, 2022; CASTELLANOS CLARAMUNT, Jorge, «Una reflexión acerca de la influencia de la inteligencia artificial en los derechos fundamentales», en RAMÓN FERNÁNDEZ, Francisca (ed.), *Ciencia de datos y perspectivas de la inteligencia artificial*, Valencia, Tirant lo Blanch, 2024, pp. 271-300; QUINTAVALLA, Alberto; TEMPERMAN, Jeroen, *Artificial Intelligence and Human Rights*, Oxford, Oxford University Press, 2023, sin olvidar uno de los primeros trabajos que tocó el tema en profundidad en España de DE ASÍS, Rafael, *Una mirada a la robótica desde los derechos humanos*, Madrid, Dykinson, 2015.

⁴⁵ Véase, por ejemplo, informe para el Parlamento Europeo, comisionado por subcomité de Derechos humanos «Artificial intelligence (AI) and human rights: Using AI as a weapon of repression and its impact on human rights», del mayo de 2024. Disponible en [https://www.europarl.europa.eu/RegData/etudes/IDAN/2024/754450/EXPO_IDA\(2024\)754450_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/IDAN/2024/754450/EXPO_IDA(2024)754450_EN.pdf). Consultado el 29 de diciembre de 2024.

desarrollarse como una persona sin ser vigilada, observada y «datificada». Esta libertad también incluye el derecho a poder elegir si y con quien compartir los datos y, en el caso de no compartir, el derecho de no sufrir en cambio repercusiones en el acceso a los servicios esenciales.

También tenemos que volver a la alfabetización de la IA como una parte del derecho a la educación: conocer y entender el funcionamiento de IA empodera como empodera cualquier conocimiento y representa la principal herramienta para entender los riesgos y daños a nivel individual y garantizar que los demás les entienden. En este sentido podemos defender el derecho a la educación en la IA como manera de abatir la ventaja del sector privado que por mucho tiempo se ha beneficiado de la falta de la comprensión y entendimiento de la ciudadanía de la tecnología en general y de la IA en particular: la narrativa sobre la complejidad de IA, basada por un lado en la asimetría de los conocimientos técnico-informáticos y por otra en «la terminología desconocida, los temas extraordinariamente míticos, empaquetamiento novedoso y el drama»⁴⁶, es una verdad parcial y con el tiempo se vuelve aún menos verdad por el interés de la ciudadanía y divulgación técnico-científica, accesibilidad de información e inclusión de ciencias informáticas en los currículos académicos⁴⁷.

Tenemos que pensar también en nuevos derechos como por ejemplo el derecho a la participación en la toma de decisiones sobre la IA, construyendo y apoyándonos en el pensamiento de Sheila Jasanoff y sus ideas sobre las tecnologías de humildad, que nos recuerda que la cuestión no es solo controlar las tecnologías que generan riesgos para la sociedad, sino también hacer a la sociedad participe en la co-construcción de la epistemología civil, a través del pluralismo y deliberación participativa⁴⁸.

Por lo tanto volvemos a la importancia no solo de mirar derechos fundamentales a través de (las posibilidades, riesgos, desarrollos de) la IA, sino también ver a la IA a través del contenido de los derechos fundamentales (existentes o hipotéticos), porque no es la IA la que define los derechos, establece sus contornos, sus límites y sus alcances, sino los derechos tendrían que ser quienes definen, moldean y dan forma a la IA.

⁴⁶ STANLEY, Dan, «What we don't talk about when we talk about AI», *Joseph Raintree Foundation*, 2024. Disponible en <https://www.jrf.org.uk/ai-for-public-good/what-we-dont-talk-about-when-we-talk-about-ai>. Consultado el 30 de diciembre del 2024.

⁴⁷ Sobre la educación y la IA, véase CORTINA, Adela, *op. cit.*, p. 203 ss.

⁴⁸ JASANOFF, Sheila, «Technologies of humility: citizens participation in governing science», *Minerva*, vol. 41, núm. 3, 2003, pp. 223-244.

5. CONCLUSIONES Y REFLEXIONES FINALES

En este escrito hemos analizado brevemente algunos de los problemas que el avance de IA genera para nuestra sociedad, sobre todo identificando aquellos que no necesariamente son visibles pero que afectan a todos, en particular a los que son más vulnerables. A través de los ejemplos del metro de Londres y de los sistemas analíticos de los tribunales hemos evidenciado la manera en la que la IA invade e impregna los servicios y tareas del sector público, propagando la narrativa de la eficiencia, operatividad, ahorro de tiempos y recursos, mejoría y utilidad como únicos valores sociales. Hemos argumentado que dichos valores son solo una parte y no representan los valores de base del sector público: los verdaderos valores públicos son el bien y el interés común, rendición de cuentas, el apoyo a los más vulnerables, defensa de los derechos fundamentales, transparencia, justicia, entre otros. Por este motivo, la IA también ha sido llamada «la herramienta para rehacer la sociedad»⁴⁹ y por esta diferencia entre valores, las trasgresiones sectoriales –efectuadas por parte de las grandes multinacionales tecnológicas– son tan problemáticas para el bien común y el interés social.

Hablando del futuro de la IA a la luz de estas preocupaciones, hemos visto que la Ley de la IA no es perfecta pero de todas formas es indispensable. Lo que sí que es innegociable es la necesidad del control social, y cuanto más variado mejor: de los ciudadanos, del sector público, de las organizaciones de la sociedad civil y en este sentido los derechos fundamentales, entendidos en el sentido amplio, han sido siempre y siguen siendo la fuente principal de inspiración. La Ley abre la puerta para que todos los sectores, tanto privado como el público, involucre más a la sociedad civil en la gobernanza y uso de la IA. Si esta invitación es acogida y cómo, lo veremos en los próximos años.

⁴⁹ ACEMOGLU, Daron; JOHNSON, Simon., *op. cit.*, p. 296.