



Revista Administración & Cidadanía, EGAP

Vol. 18_núm. 2_2023 | pp. 161-182

Santiago de Compostela, 2023

<https://doi.org/10.36402/ac.v18i2.5173>

© Lorenzo Cotino Hueso

ISSN-L: 1887-0279 | ISSN: 1887-5270

Recibido: 29/01/2024 | Aceptado: 15/02/2024

Editado bajo licencia Creative Commons Atribución 4.0 International License

O Convenio sobre intelixencia artificial, dereitos humanos, democracia e Estado de Dereito do Consello de Europa

El Convenio sobre inteligencia artificial, derechos humanos, democracia y Estado de derecho del Consejo de Europa

The Council of Europe's Convention on artificial intelligence, human rights, democracy and the rule of law



LORENZO COTINO HUESO¹

Catedrático de Derecho Constitucional
Universidad de Valencia. Valgrai

<https://orcid.org/0000-0003-2661-0010>

cotino@uv.es

Resumo: Analízase o Convenio do Consello de Europa sobre intelixencia artificial. Primeiroponse no contexto da regulación europea e mundial da IA, especialmente acelerada en 2023 con desenvolvementos como ChatGPT e a culminación do Regulamento da UE. Examínase o potencial xurídico e normativo do Convenio, destácase que non está limitado aos sistemas de alto risco como o Regulamento de IA. Tamén, a importancia da proclamación de “principios” normativos, os dereitos e obrigas específicas que establece, así como a relevancia especial do enfoque baseado en riscos e as súas garantías. A pesar de quedar á sombra do Regulamento de IA e o dereito da UE, sublíñanse os aspectos que o convenio complementa e afírmase a súa certa superioridade polo valor case constitucional e tamén simbólico dun Convenio do Consello de Europa.

Palabras clave: Intelixencia artificial, algoritmos, dereitos fundamentais, tratado de dereitos humanos, Consello de Europa, tratado internacional, Regulamento IA Unión Europea.

Resumen: Se analiza el Convenio del Consejo de Europa sobre inteligencia artificial. Primero se pone en el contexto de la regulación europea y mundial de la IA, especialmente acelerada

en 2023 con desarrollos como ChatGPT y la culminación del Reglamento de la UE. Se examina el potencial jurídico y normativo del Convenio, se destaca que no está limitado a los sistemas de alto riesgo como el Reglamento de IA. También, la importancia de la proclamación de “principios” normativos, los derechos y obligaciones específicas que establece, así como la relevancia especial del enfoque basado en riesgos y sus garantías. Pese a quedar a la sombra del Reglamento de IA y el derecho de la UE, se subrayan los aspectos que el convenio complementa y se afirma su cierta superioridad por el valor casi constitucional y también simbólico de un Convenio del Consejo de Europa.

Palabras clave: Inteligencia artificial, algoritmos, derechos fundamentales, tratado de derechos humanos, Consejo de Europa, tratado internacional, Reglamento IA Unión Europea.

Abstract: The Council of Europe Convention on Artificial Intelligence is analyzed. It is put in the context of European and global regulation of AI, especially accelerated in 2023 with developments such as ChatGPT and the completion of the EU Regulation. It is examined the legal and regulatory potential of the Convention, stressing that it is not limited to high-risk systems such as the IA Regulation. Also, the importance of the proclamation of regulatory “principles”; the specific rights and guarantees that are recognized, as well as the particular relevance of the risk-based approach and its concrete safeguards. Despite being in the shadow of the IA Regulation and EU law, the aspects that the Convention complements are underlined and its certain superiority is affirmed caused by the quasi-constitutional and also symbolic value of a Council of Europe Convention.

Key words: Artificial intelligence, algorithms, fundamental rights, human rights treaty, Council of Europe, convention, AI Act.

SUMARIO: 1 El Consejo de Europa se suma a la regulación de la inteligencia artificial con el “Convenio IA”. 2 El valor normativo, interpretativo y simbólico de un Convenio IA. 3 Las disposiciones generales: objeto, ámbito de aplicación y las exclusiones en investigación, defensa y seguridad nacional. 4 La importante regulación de “principios” y el enfoque de los “sujetos afectados” y los grupos y colectivos. 5 Los derechos a la documentación, registros o notificaciones y la garantía de una autoridad independiente. 6 Las relevantes obligaciones de evaluación y mitigación de riesgos e impactos. 7 Algunas cláusulas habituales de los convenios de derechos y el “Mecanismo de seguimiento y cooperación”. 8 La interrelación del Convenio con el derecho de la Unión Europea y, en especial, con el Reglamento de inteligencia artificial. 9 Para concluir: el Convenio incorpora la “lírica” a la “prosa” que supone el Reglamento de la Unión Europea. 10 Bibliografía empleada.

1 EL CONSEJO DE EUROPA SE SUMA A LA REGULACIÓN DE LA INTELIGENCIA ARTIFICIAL CON EL “CONVENIO IA”

Resulta de interés ubicar el “Convenio marco sobre inteligencia artificial, derechos humanos, democracia y Estado de derecho” del Consejo de Europa (en adelante, Convenio IA) en el contexto de aceleración de la regulación de la inteligencia artificial (IA). Así, la regulación de la IA ha figurado en la agenda de la UE desde 2016, pero particularmente con el *Libro blanco sobre la IA* en 2020², seguido por la propuesta del Reglamento de IA (RIA), en 2021³, que culminó políticamente con un acuerdo el 9 de diciembre de 2023 y la aprobación final a inicios de 2024⁴. Este proceso regulatorio de la IA de la UE inició el interés regulatorio de la IA en otras regiones del mundo. Así, el 16 de junio de 2022 en Canadá se presentó la Ley de implementación de la carta digital 2022 (*Digital Charter Implementation Act*)⁵ en el marco de la Ley de datos e inteligencia artificial (*Artificial Intelligence and Data Act*, “AIDA”)⁶. En diciembre de 2022, Brasil inició la tramitación del proyecto de ley n. 2338 en el Senado, orientado a la regulación de la IA⁷. También

y entre otras, cabe tener en cuenta las perspectivas de regulación débil desde el Reino Unido tras su salida de la UE⁸. Estas iniciativas parece que estaban a la espera de la adopción del RIA de la UE.

Especialmente con la aparición del ChatGPT desde finales de 2022 el proceso se ha acelerado. Así, caben destacar en China las “Medidas provisionales para la gestión de servicios de inteligencia artificial generativa”, adoptadas el 13 de julio de 2023, en vigor desde 24 de agosto⁹. En Canadá en septiembre de 2023 se anunció y se abrió a la adopción voluntaria el “Código de conducta voluntario para el desarrollo y gestión responsable de sistemas generativos avanzados de IA”¹⁰. En el contexto del G7, el llamado “Proceso de Hiroshima” que se sustancia en el acuerdo en el “Código internacional de conducta para organizaciones que desarrollan sistemas avanzados de IA”, de 30 de octubre de 2023, un instrumento voluntario para los desarrolladores de IA¹¹. El mismo día, en Estados Unidos se adoptó la Orden ejecutiva sobre el desarrollo y la utilización seguros y fiables de la inteligencia artificial, de 30 de octubre de 2023¹². Con carácter más parcial, el 3 de enero de 2024 California ha regulado su *Artificial Intelligence Accountability Act*¹³ o Illinois cuenta con su *Artificial Intelligence Video Interview Act (820 ILCS 42/)*¹⁴. En Iberoamérica existen múltiples propuestas legislativas¹⁵.

En este contexto y en especial bajo la *larga sombra* del RIA de la UE, cabe ubicar la adopción del Convenio IA que se adopta en 2024, si bien con precedentes desde 2019. Así, el Consejo de Europa, a través de su Comité de Ministros, instauró el Comité Ad Hoc sobre Inteligencia Artificial (CAHAI)¹⁶ el 11 de septiembre de 2019. Posteriormente, se otorgó un mandato al Comité sobre Inteligencia Artificial (CAI) desde enero de 2022 hasta diciembre de 2024¹⁷, con los “Términos de Referencia del CAI”¹⁸. El CAI fue el encargado de crear un “instrumento jurídico vinculante de carácter transversal”¹⁹ que promueva la innovación y establezca principios “sólidos y claros” para el diseño, desarrollo y aplicación de sistemas de IA, reconociendo que “actualmente no puede regular todos los aspectos del desarrollo y uso de los sistemas de IA”.

Este mandato tenía una clara intención de trascender fronteras, buscando crear un “instrumento atractivo no sólo para los Estados de Europa, sino para el mayor número posible de Estados de todas las regiones del mundo”, involucrando a “Observadores”²⁰ como Israel, Canadá²¹, Estados Unidos, Japón, la Asociación Mundial sobre Inteligencia Artificial (GPAI), empresas de Internet y organizaciones de la sociedad civil. Participan en la redacción y podrían adherirse al Convenio IA una vez que este entre en vigor, conforme lo estipula su artículo 31.

El CAI se construyó sobre las bases establecidas por CAHAI entre 2019 y 2021²². El considerando 9 subraya la urgencia de “un marco jurídico aplicable a escala mundial”. Así, se definieron unos contenidos básicos²³ que se han ido articulando en los sucesivos documentos: objeto y ámbito de aplicación del Convenio IA; definiciones de sistema de IA, ciclo de vida, proveedor, usuario y “sujeto de IA”; principios fundamentales, incluidas las garantías procedimentales y los derechos de los sujetos de IA; medidas adicionales para el sector público, así como los sistemas de IA que planteen niveles de riesgo

“inaceptables” y “significativos”; un mecanismo de seguimiento y cooperación entre las partes y disposiciones finales.

El 6 de enero 2023 (4.ª sesión plenaria) se publicó un “borrador cero” revisado²⁴. El 7 de julio de 2023 (6.ª reunión) se dio a conocer el borrador de trabajo consolidado²⁵, base para futuras negociaciones. El 18 de diciembre de 2023 (8.ª sesión), se ha hecho público el proyecto de convenio marco²⁶. Este proyecto se espera que sea la base para la última lectura prevista para 2024. El texto al que se hace mención en el presente estudio es relativo a este texto. Se trata de un convenio extenso que incluye diecisiete considerandos y treinta y seis artículos repartidos en ocho capítulos²⁷.

2 EL VALOR NORMATIVO, INTERPRETATIVO Y SIMBÓLICO DE UN CONVENIO IA

Se trata de un convenio de normas mínimas, con escasas obligaciones y derechos. El CAHAI señaló el objetivo de crear un “marco jurídico común que contenga ciertas normas mínimas para el desarrollo, diseño y aplicación de la IA en relación con los derechos humanos, la democracia y el Estado de derecho”²⁸. Se habla de “principios y normas generales comunes” (considerando 9) y del “carácter marco del convenio, que podrá completarse con otros instrumentos” (considerando 11).

En cuanto a la concreción de las obligaciones en el texto del Convenio IA, el artículo 4 en general afirma que “cada parte adoptará o mantendrá medidas para garantizar que las actividades dentro del ciclo de vida de los sistemas de inteligencia artificial sean compatibles con las obligaciones de protección de los derechos humanos, consagradas en el derecho internacional aplicable, y en su derecho interno”. El artículo 5 presenta de forma genérica la obligación de adoptar medidas respecto de la “Integridad de los procesos democráticos y respeto del Estado de derecho” así como “para respetar la dignidad humana y la autonomía individual” (art. 6). Hasta un total de en 31 ocasiones se afirma que las partes tomarán o adoptarán “medidas” (“*each party shall take*”) “necesarias”, “apropiadas”, “políticas” o “legislativas o de otro tipo”; estas son expresiones, habituales en el contexto internacional. Como se expondrá, todas las medidas “se graduarán y diferenciarán” según los impactos y riesgos que generen los sistemas de IA bajo un enfoque de riesgos (art. 1.2.º).

Aunque en general el Convenio IA no se caracteriza por establecer nítidas obligaciones y derechos específicos, hay varios motivos para tenerlo normativamente en cuenta. Como luego se explicará, los “principios” generales que se regulan en el capítulo III tienen un alto potencial en manos de los operadores jurídicos. Asimismo, sí que se regulan algunos derechos y garantías en el capítulo IV, y el artículo 16 (capítulo V), sobre evaluación y mitigación de riesgos e impactos, impone las obligaciones más relevantes del Convenio IA. Además, no hay que perder de vista el efecto directo de los tratados, su valor casi constitucional y su potencial efecto interpretativo y su integración en el ordenamiento estatal. Así, cabe recordar que, tras su ratificación y publicación, los

tratados internacionales se incorporan al ordenamiento interno (art. 96 CE), permitiendo su invocación directa ante autoridades y tribunales²⁹. Por ello, aunque no sea sencillo derivar derechos y obligaciones del Convenio IA sin legislación intermedia, hay que reconocer el potencial que tiene su aplicación directa y su preeminencia sobre normas internas (arts. 29 y 30 de la Ley 25/2014). Además y especialmente, hay que destacar el valor casi constitucional de los tratados relacionados con derechos fundamentales conforme al artículo 10.2 CE, pues pasan a ser elementos obligatorios para interpretar la Constitución y sus derechos. De esta manera, el tratado “se convierte en cierto modo en el contenido constitucionalmente declarado de los derechos y libertades” (STC 36/1991, FJ 4) y determina “los perfiles exactos de su contenido” (STC 28/1991, FJ 5). Así, un convenio, una vez ratificado, posee un notable valor normativo e interpretativo.

Incluso yendo más allá de su valor propiamente jurídico normativo, no hay que olvidar el valor simbólico y metajurídico de los Tratados de derechos del Consejo de Europa, pues representan un compromiso con valores compartidos y establecen ideales y directrices políticas. Su valor simbólico y su capacidad para guiar la interpretación de derechos fundamentales y fomentar políticas y legislación son cruciales. A diferencia de normativas con disposiciones más específicas como el RIA o el RGPD, este Convenio IA cuenta también con estas características de tanta potencialidad.

3 LAS DISPOSICIONES GENERALES: OBJETO, ÁMBITO DE APLICACIÓN Y LAS EXCLUSIONES EN INVESTIGACIÓN, DEFENSA Y SEGURIDAD NACIONAL

El capítulo I, con sus disposiciones generales, resalta el objetivo general de que “las actividades dentro del ciclo de vida de los sistemas de inteligencia artificial sean plenamente coherentes con los derechos humanos, la democracia y el Estado de derecho” y menciona la graduación de obligaciones según la gravedad de “impactos adversos” (art. 1.1.º y 2.º). El artículo 16 desarrolla este enfoque de riesgo.

El ámbito de aplicación del Convenio IA se asocia al concepto de “sistema de inteligencia artificial”: “es un sistema basado en una máquina que, para objetivos explícitos o implícitos, infiere, a partir de los datos que recibe, cómo generar resultados tales como predicciones, contenidos, recomendaciones o decisiones que puedan influir en entornos físicos o virtuales”. “Los distintos sistemas de inteligencia artificial varían en sus niveles de autonomía y adaptabilidad tras su despliegue” (art. 2). Se observa una convergencia internacional en la definición de IA, reflejada en la evolución del Reglamento de la UE y los ajustes de la OCDE en noviembre de 2023³⁰, por lo que no se prevén divergencias. El Convenio IA se aplica a sistemas IA en cuanto “puedan interferir con los derechos humanos, la democracia y el Estado de derecho” (art. 3.1.º). El Convenio IA puede implicar obligaciones para sistemas que no son de “alto riesgo”, puesto que no se maneja esta noción que, como es sabido, es esencial para que se apliquen las obligaciones del RIA.

En lo que respecta a la aplicación del Convenio IA en el sector público y privado, la versión de 7 de julio de 2023 no lo especificaba. La versión del 18 de diciembre plantea opciones. La opción A no diferencia sectores, implicando posibles obligaciones en ambos. La opción B hace referencia a “actividades que [...] lleven a cabo las autoridades públicas o entidades que actúen en su nombre” si bien se afirma expresamente la obligación de las partes de hacer cumplir respecto de las entidades privadas. La opción C no distingue, aunque refuerza para el público la necesidad de medidas adecuadas y afirma la introducción de adoptar medidas “progresivamente” para las “entidades privadas”. El Supervisor Europeo de Protección de Datos (SEPD) considera positivamente y da por hecho que el Convenio IA alcanza a “los proveedores como a los usuarios públicos y privados [...], independientemente de si los proveedores y usuarios de sistemas de IA son entidades públicas o privadas” (n. 23 y 24). Asimismo, valoraba también positivamente que se darían “medidas adicionales para el sector público”, que alcanzarían también a las entidades privadas cuando presten servicios públicos (n. 25)³¹.

Las exclusiones de aplicación a la investigación y respecto de la seguridad nacional no son totales. Cabe señalar que la exclusión de la investigación es un tema que ha variado en el RIA. En la propuesta inicial de la Comisión no se mencionaba –y por tanto no excluía– la investigación. La exclusión de sistemas IA con “el único fin de la investigación y el desarrollo científicos” se introdujo en la versión del Consejo de la UE de diciembre de 2022 (art. 2) y se especificó y restringió en la versión del Parlamento de junio de 2023, incluyendo, por ejemplo, la posibilidad de actos delegados para precisar el alcance de la exclusión, que en ningún caso abarcaría las pruebas en condiciones reales³². La exclusión en el RIA tiene mayor lógica, puesto que es un reglamento para la puesta en el mercado de sistemas IA, y en la investigación la finalidad no es poner el producto en el mercado. Cabe recordar que en el ámbito de la protección de datos hay modulaciones y flexibilizaciones pero en ningún caso una exclusión³³.

Para el futuro Convenio IA respecto de la investigación se barajan varias opciones: exclusión “a menos que los sistemas se prueben o se utilicen de otro modo” (opción A 3.2); “cuando el diseño, el desarrollo, el uso y el desmantelamiento de los sistemas de IA impliquen investigación, dicha investigación se incluirá en el ámbito de aplicación del presente Convenio en el artículo 2 (en su opción B 3.3)³⁴. La opción C 3.2 delega la cuestión en la legislación interna. Cabe esperar una solución normativa que delimite diversos criterios, como la necesaria utilidad pública, que no haya una puesta en el mercado o que no se dé el uso del sistema de IA que pueda impactar en entornos generales o no controlados en la investigación. En todo caso y como es obvio, la exclusión de la investigación no puede suponer ni la vigencia de los derechos fundamentales reconocidos, ni de otra normativa que sí que sea aplicable.

En cuanto a la seguridad y defensa nacional, se trata de un ámbito en el que se concede una muy fuerte discrecionalidad a los Estados³⁵. Ahora bien, estas limitaciones o discrecionalidad no implican por sí la exclusión de la aplicación del derecho o los derechos y, en su caso, es necesario establecerla expresamente. En la UE, el RIA “no

se aplicará a los sistemas de IA desarrollados o utilizados exclusivamente con fines militares” (art. 2.3). También está excluida la aplicación en estas áreas de la normativa de protección de datos, que muchas veces se aplica al ámbito de la IA (art. 2.2.º RGPD). Sin embargo, en el caso del Convenio IA no hay una exclusión total. El borrador del Convenio de 7 julio de 2023 hacía referencia a las “restricciones, derogaciones o excepciones” (capítulo III)³⁶. En la versión del 18 de diciembre se establece una exclusión general para la defensa, pero que no es total respecto de la seguridad nacional. A partir de ahí hay varias alternativas con matices y más o menos poder de las partes respecto de esta exclusión³⁷. Por su parte, el Consejo de la UE en su Decisión sobre el Convenio recuerda la responsabilidad exclusiva de cada Estado en la materia³⁸. Considero positivo que, al menos potencialmente, se reconozca la vigencia del Convenio IA en materia de seguridad nacional, aunque sea de mínimos.

4 LA IMPORTANTE REGULACIÓN DE “PRINCIPIOS” Y EL ENFOQUE DE LOS “SUJETOS AFECTADOS” Y LOS GRUPOS Y COLECTIVOS

Considero relevante la regulación en el Convenio IA de “principios” generales aplicables a todos los sistemas de IA. A este respecto, cabe recordar que desde hace años, entre decenas de declaraciones y documentos, se han ido visibilizando y destilando unos principios éticos esenciales de la IA³⁹. Desde Harvard⁴⁰ se analizaron más de treinta de las principales declaraciones internacionales y corporativas de ética de la IA y se sintetizaron en privacidad, rendición de cuentas, seguridad, transparencia y explicabilidad, equidad y no discriminación, control humano, responsabilidad profesional, valores humanos y sostenibilidad. El futuro convenio es positivo por cuanto va más allá de las declaraciones en el ámbito del *soft law* y regula estos principios; si se me permite, pasa de las *musas* de la ética al *teatro* del derecho.

En el Convenio IA los principios se integran bajo la fórmula de que cada “Parte adoptará o mantendrá medidas para garantizar” (*each party shall adopt or maintain measures to ensure*) estos principios. Así, el capítulo III “establece” “los principios generales comunes que cada parte deberá aplicar [...] de manera adecuada a su ordenamiento jurídico interno” (como “nota explicativa”). Ocho artículos (arts. 6 a 13) expresan y afirman tales “principios”: dignidad humana y autonomía individual (art. 6), transparencia y supervisión (art. 7), rendición de cuentas y responsabilidad (art. 8), igualdad y no discriminación (art. 9), privacidad y protección de datos personales (art. 10), preservación de la salud [y del medio ambiente] (art. 11), fiabilidad y confianza (art. 12), innovación segura (art. 13). Son pocos los elementos normativos más específicos, como la obligación de la “detección y transparencia del contenido generado por sistemas de inteligencia artificial” (art. 7); la mención de “estándares y marcos nacionales e internacionales aplicables en materia de protección de datos personales y [gobernanza de datos]” (art. 10)⁴¹ o de “la seguridad, la protección, la exactitud, el rendimiento, la calidad [...] requisitos de integridad, seguridad de los datos, gobernanza, ciberseguridad y solidez” (art. 12). Destaca que se “anime”

(*encouraged*) a las partes a “fomentar la innovación” y, para ello, “el establecimiento de [un] entorno[s] [reglamentario] controlado[s] para [el desarrollo y la experimentación] / [pruebas]” (art. 13).

Pese a que es una regulación flexible y abierta, caracterizada por mandatos amplios o remisiones a la normativa existente, su potencialidad normativa es importante. Los principios desempeñan un papel esencial en la configuración y estructuración del ordenamiento jurídico, informan el ordenamiento y son herramientas clave para la interpretación y aplicación de las normas, así como fuente de inspiración para resolver conflictos y derivar nuevas interpretaciones. Y ello tiene particular relevancia en sectores digitales y disruptivos, donde se da una considerable incertidumbre, como es el caso de la IA.

Además, no olvidemos que en un ámbito tan relacionado como el de la protección de datos, durante más de treinta años, los “principios” (art. 5 RGPD) han desempeñado y continúan desempeñando estas funciones generales, siendo los pilares fundamentales. Es más, los “principios” de la protección de datos han constituido reglas concretas aplicables a los tratamientos. Tanto es así que su mero incumplimiento directamente implica la comisión de infracciones.

Llama la atención que la regulación de unos principios en el RIA de la UE ha ido a remolque del Convenio del Consejo de Europa. Ni la propuesta de la Comisión de 2021 ni el texto del Consejo de diciembre de 2022 incluía una proclamación de principios. Ha sido el Parlamento de la UE en junio de 2023 quien propuso incluir una proclamación estructural de los “Principios generales aplicables a todos los sistemas de IA” en los primeros artículos del RIA (art. 4 bis, nuevo, enmienda 213). Se pretende seguir el esquema del RGPD (art. 5) y con la finalidad de que estos principios sean aplicables tanto a sistemas IA –de alto riesgo o no– como también a los modelos fundacionales. Se proclaman los principios de “intervención y vigilancia humanas” (a), “solidez y seguridad técnicas” (b), “privacidad y gobernanza de datos” (c), “transparencia” (d), “diversidad, no discriminación y equidad” (e) y “bienestar social y medioambiental” (f), y cabe decir, que se regulan con bastante más detalle que el futuro convenio. Sin embargo, en contraste con el convenio, se modula y restringe expresamente el alcance de estos principios, pues se proclaman “sin crear nuevas obligaciones en virtud del presente reglamento” (art. 4 bis. 2.º). Ello no obstante, sí que se afirma que deben inspirar los procesos de normalización y las orientaciones técnicas (art. 4 bis. 2.º).

En cuanto al reconocimiento de los “sujetos afectados”, en el “*Zero draft*” de 6 de enero de 2023, artículo 2 e), apareció la definición de “sujeto de inteligencia artificial”⁴² y el SEPDP la acogió “con satisfacción”⁴³. Este enfoque contrastaba notoriamente con el RIA de la UE, el cual ha sido criticado por su total ignorancia de los afectados por un sistema de IA. Se llegó a afirmar que “Mientras que el AIA se centra en el mercado único digital y no crea nuevos derechos para las personas, el convenio podría llenar estos vacíos [...] Mientras la Comisión Europea hace hincapié en la economía y en la integración de los mercados, el Consejo de Europa se centra en los derechos humanos, la democracia y

el Estado de derecho⁴⁴. Como un efecto reflejo, tras este borrador cero del Convenio de enero de 2023, en el RIA de la UE en junio se incluyó entre las definiciones la de “persona afectada”, junto con diversos derechos en las enmiendas del Parlamento (enmienda 174, art. 3. 1.º, 8 bis). No deja de ser irónico que en el futuro Convenio ha desaparecido el concepto de “sujeto de IA”, tanto en la versión de 7 de julio como en la de 18 de diciembre de 2023.

Un tema más relevante es el reconocimiento y la protección de colectivos y grupos afectados por la IA. Como he tenido ocasión de insistir hace años, hay que superar un enfoque limitado a la afectación directa del sistema automatizado o de IA en el individuo⁴⁵. Hay que tener en cuenta el impacto estructural y masivo del uso de los sistemas IA, que en muchos casos se utilizan para apoyar la toma de decisiones generales ya sea en el sector público o privado. Respecto de los grupos, es esencial estimular y, cuando sea necesario, garantizar una transparencia específica, así como la participación de la sociedad civil y de colectivos afectados en los diversos usos de sistemas de IA. Así lo propuse para la Carta de Derechos Digitales en un apartado sobre “Garantías de impacto social”⁴⁶. Sobre el tema resulta especialmente inspirador Mantelero⁴⁷. También la Recomendación sobre IA de la Unesco insiste en la necesidad de un enfoque colectivo y sobre todo de la inclusión de mecanismos de participación social respecto del uso de sistemas IA⁴⁸.

Desde Canadá se ha resaltado una carencia en el Convenio IA, subrayando la importancia de “pasar de la privacidad individual a la privacidad colectiva”⁴⁹. Especialmente el SEPD en sus conclusiones, insta a incorporar en el Convenio IA “la especificación de que los riesgos sociales o de grupo que plantean los sistemas de IA deben también ser evaluados y mitigados” (n. 6)⁵⁰.

Lo cierto es que sí que se aprecia esta sensibilidad con los grupos y colectivos afectados en el futuro Convenio, aunque las garantías no se delinean con extrema precisión. En todo caso, supone un avance significativo respecto a lo actual, abriendo camino y motivando a los Estados a desarrollar y refinar estos mecanismos de protección colectiva. Así, el artículo 15 regula la necesidad de identificar y evaluar riesgos e “integrar la perspectiva de todas las partes interesadas pertinentes, incluida cualquier persona cuyos derechos puedan verse potencialmente afectados” (art. 15. 2.º c). La participación de la sociedad civil se menciona en el artículo 5, sobre “Integridad de los procesos democráticos y respeto del Estado de derecho”. En su apartado 2.º, “en el contexto de las actividades dentro del ciclo de vida de los sistemas de inteligencia artificial”, obliga a adoptar medidas, en su caso, para “proteger la participación [de las personas] en los procesos democráticos, el acceso justo al debate público [y la capacidad de las personas para tomar decisiones libres de influencias o manipulaciones externas indebidas/ [perjudiciales y malintencionadas]”. Es más destacable, si cabe, que el artículo 20 sobre “consulta pública” disponga que “cada parte se esforzará por garantizar que las cuestiones importantes [...] sean, según proceda, objeto de debate público y de consulta a las múltiples partes interesadas a la luz, en particular, de las implicaciones sociales,

económicas, jurídicas, éticas y medioambientales pertinentes”. Incluso hay algunas referencias más indirectas de interés.

5 LOS DERECHOS A LA DOCUMENTACIÓN, REGISTROS O NOTIFICACIONES Y LA GARANTÍA DE UNA AUTORIDAD INDEPENDIENTE

Aunque se afirma que uno de los “aspectos clave” del Convenio IA es su enfoque en los derechos, en términos generales no establece nuevos derechos subjetivos. Hay más de cuarenta remisiones o menciones a los “derechos humanos” o “derechos” preexistentes ya reconocidos. También hay referencias particulares a la “igualdad y no discriminación” (arts. 9 y 17), a la “privacidad y protección de datos” (art. 10) o a los “derechos de las personas con discapacidad y de los niños” (art. 18), pero no se regula ningún contenido sustantivo y en modo alguno un nuevo derecho.

Como excepción, pueden considerarse como derechos en el Convenio IA algunas obligaciones de documentación, registros y otras medidas para garantizar los “recursos” efectivos del artículo 14, así como las garantías de supervisión y notificación del artículo 15.

Ya en 2014 en el contexto norteamericano, Crawford y Shultz acuñaron el *data due process* para redefinir las garantías del debido proceso en el contexto de la toma de decisiones algorítmicas. Este nuevo derecho incluía la garantía de una información mínima a los afectados sobre qué predijo el algoritmo, a partir de qué datos y la metodología empleada. Siguiendo esta línea, el artículo 14 sobre “Recursos” (“*Remedies*”) obliga a los Estados a “adoptar medidas” que garanticen recursos efectivos. Se garantiza para ello la generación de documentación, registros y evidencias y que los afectados puedan acceder a ellas. Así, se afirma la necesidad de que “los sistemas de inteligencia artificial [...] estén debidamente documentados y que se facilite información adecuada sobre el uso [...] operaciones realizadas [...] y esta información] se registre, se facilite a los organismos autorizados de conformidad con su derecho interno para acceder a dicha información y, cuando proceda y sea aplicable, se ponga a disposición o se comunique a las personas afectadas” (art. 14 a). Además, esta información ha de ser “suficiente y proporcionada para que las personas afectadas impugnen” (art. 14 b). Resulta sin duda del todo positiva la regulación expresa de estas garantías, que habrán de ser reguladas por las partes del convenio. En todo caso, considero que estas obligaciones no se pueden generalizar a todo sistema de IA. Para ello cabe apostar por una interpretación restrictiva de la referencia del artículo 14 a sistemas que puedan “afectar significativamente” (“*significantly affecting*”) a los derechos humanos.

Entre los “recursos” (capítulo IV), el artículo 15 regula “garantías procesales” (“*procedural safeguards*”). Su apartado primero afirma “la importancia de la revisión humana” y la necesidad de garantías efectivas en los casos en los que un sistema de IA “informe o adopte decisiones o actos que afecten sustancialmente a los derechos humanos” (art.

15.1.º). Pese a que no se exprese con la concreción con que se debería, el artículo 15 guarda claras conexiones con las garantías que el derecho europeo reconoce respecto de las decisiones automatizadas: artículo 22 RGPD, el artículo 11 de la Directiva (UE) 2016/680 o el nuevo artículo 9.1.º Convenio 108 del Consejo de Europa que reconoció en 2018 también este “derecho”. Ahora bien, las garantías del artículo 15 del Convenio IA no estarían limitadas a una decisión “basada únicamente en el tratamiento automatizado”, sino que extiende explícitamente las garantías no solo a las decisiones tomadas directamente por sistemas de IA, sino también a situaciones donde la IA “influye de manera significativa” en la toma de decisiones humanas. Ello va en la línea correcta que el TJUE adopta recientemente en diciembre de 2023. Asimismo, el Convenio IA sigue la corriente del RIA en su delimitación de los sistemas de “alto riesgo” que lo son, “salvo que la información de salida del sistema sea meramente accesoria respecto de la acción o decisión pertinente que deba adoptarse” (art. 6.3.º RIA, versión del Consejo de diciembre de 2022).

El segundo apartado del artículo 15 añade la obligación, y potencialmente el derecho, a la “notificación”: “se notifica[rá] a las personas que interactúan con un sistema de inteligencia artificial y no con un ser humano”. En la UE desde 2018 el HLEG incluyó en la transparencia el derecho a la “comunicación”, esencialmente como “derecho a saber que [las personas] están interactuando con un sistema de IA” (n. 78). El artículo 52 del RIA recoge esta notificación entre las diversas “obligaciones de transparencia para determinados sistemas de IA”. El artículo 15 parece seguir al *soft law* internacional, en el que la comunicación se bifurca en un derecho u obligación a la “notificación cuando se interactúa con un sistema de IA” o “principio de interacción” y en las garantías frente a las decisiones automatizadas del artículo 15.1.º.

El reconocimiento de los derechos y garantías en el artículo 15 es, indudablemente, un paso positivo. No obstante, sería deseable introducir mayor concreción de sus requisitos, contenidos y facultades mínimas.

También hay que destacar la importante garantía de una autoridad independiente ante la que poder presentar recursos efectivos. En este sentido, cabe partir de que el artículo 13 del Convenio europeo de derechos humanos (CEDH) reconoce el derecho a un recurso interno efectivo ante una instancia nacional para garantizar los derechos del propio CEDH (y no otros). Este recurso debe posibilitar que una autoridad independiente, no necesariamente judicial, examine el fondo de la petición y, si procede, conceder la reparación adecuada. En esta línea, el artículo 15 del Convenio IA sobre “Recursos” en su apartado 3.º prevé que se regule “una posibilidad efectiva para las personas afectadas de presentar una reclamación [ante las autoridades públicas, incluido, en su caso, ante el mecanismo de supervisión a que se refiere el artículo 26, de conformidad con su derecho interno]”. Por su parte, el artículo 26 sobre “Mecanismos de supervisión eficaces” es bien relevante: “Cada parte establecerá o designará uno o más mecanismos eficaces para supervisar el cumplimiento de las obligaciones contenidas en el Convenio IA” (1.º). Y “garantizará que dichos mecanismos ejerzan sus funciones de manera independiente e

imparcial y que tengan los poderes, la experiencia y los recursos necesarios para cumplir eficazmente sus tareas de supervisar el cumplimiento de las obligaciones del convenio” (2.º). Si hay más de un mecanismo, habrán de cooperar entre ellos (3.º).

El RAI no exigía la independencia e imparcialidad de la “autoridad de vigilancia del mercado”. Pero en la enmienda 123 del Parlamento al Considerando 77 y la 558 con el nuevo artículo 59 –apartado 4 pasa a exigir la “total independencia”–, se afirma su carácter “independiente, imparcial y objetiva”. Llama la atención que la primera autoridad creada expresamente para estas funciones en la UE, la Agencia Española de Supervisión de Inteligencia Artificial (AESIA), ha sido regulada por la Ley 28/2022, y en especial el artículo 8 del Real decreto 729/2023, de 22 de agosto no cumple con los criterios de independencia establecidos tanto en el RAI como en el Convenio IA.

Se trata de una regulación valiente con suficiente concreción dada la naturaleza del convenio. El SEPD subraya en especial que el convenio expresamente ha de “conceder a las autoridades de control competentes poderes adecuados de investigación y ejecución.” (n. 47 y 49.º y conclusión 10.ª) Y tiene en cuenta que es muy fácil que se dé una convergencia de autoridades sectoriales con competencias en materia de uso de IA, por ello será necesaria la cooperación (n. 46), así como la necesidad de “cooperación transfronteriza entre autoridades competentes que serán designadas por las partes en el acuerdo” (conclusión 11).

6 LAS RELEVANTES OBLIGACIONES DE EVALUACIÓN Y MITIGACIÓN DE RIESGOS E IMPACTOS

La regulación de la evaluación de los riesgos e impactos es posiblemente lo más relevante y destacable del Convenio IA. La UE marchó por la senda de un enfoque basado en el riesgo desde 2018, bajo la marca de IA Europa del “*Ethics & rule of law by design (X-by-design)*”. También el alto grupo de expertos de la UE con su lista exhaustiva de evaluación. El *Libro blanco de la IA* de 2020 hizo explícito el modelo de riesgos, que se concretaría un año más tarde en el RIA. Este enfoque implica la mayor imposición de obligaciones y garantías cuanto mayor impacto o riesgo implique el sistema de IA. Bajo la filosofía del *compliance* o cumplimiento normativo, se identifican posibles riesgos y daños y se han de disponer los mecanismos preventivos y proactivos adecuados para evitar que ocurran. Cabe destacar especialmente las aportaciones de Mantelero en el marco de la protección de datos y luego para los sistemas IA.

El enfoque basado en riesgos se integró entre 2019 y 2021 por parte del Comité Ad Hoc sobre Inteligencia Artificial (CAHAI), entre otros, a través de la participación de Mantelero. En diciembre de 2021 el CAHAI definió los “Elementos básicos” del futuro convenio y dejó claro que iba a “centrarse en la prevención y/o mitigación de los riesgos [...] Los requisitos legales para el diseño, desarrollo y uso de sistemas de IA deben ser proporcionales a la naturaleza del riesgo que plantean para los derechos humanos, la democracia y el Estado de derecho” (parte I, n. 5). Además dedicó diversos pasajes a la

cuestión (V 18-21 y XII 45-53). Se introdujo la posibilidad de incluir una “Evaluación de impacto sobre los derechos humanos, la democracia y el Estado de derecho” (HUDERIA, *Human Rights, Democracy and Rule of Law Impact Assessment*), si bien se afirmaba que “no tiene por qué formar parte de un posible instrumento jurídicamente vinculante” (n. 19) y se baraja “complementar con un instrumento transversal no vinculante jurídicamente para evaluar el impacto de los sistemas de IA” (n. 45) con “legislación nacional o internacional, con otros mecanismos de cumplimiento, como la certificación y el etiquetado de calidad, las auditorías, los espacios aislados de regulación y la supervisión periódica” (n. 47). En el marco del Convenio IA el estudio de impacto sólo se daría “si existen indicios claros y objetivos de riesgos relevantes” a partir de una revisión inicial de todos los sistemas de IA (n. 48) y habría de actualizarse “de forma sistemática y periódica” (n. 49). En el convenio se detallaban los pasos principales a realizar: identificación de riesgos; evaluación de impacto; evaluación de la gobernanza; y mitigación y evaluación (n. 50), y en concreto los elementos mínimos de la evaluación de impacto (n. 51). Asimismo, adelantaba la importancia de involucrar a la sociedad civil en la materia y garantizar fórmulas de participación (n. 53). Sobre estas bases, el “borrador cero de 6 de enero de 2023 reafirmó el “enfoque basado en el riesgo” (n. 5), que ha sido bien recibido por la doctrina y en la UE en su Decisión de la UE sobre el convenio, por cuanto sirve para “minimizar los riesgos [...] evitando al mismo tiempo cargas o restricciones innecesarias y desproporcionadas” (anexo, n. 4). Por su parte, el SEPD afirma que “la Comisión debería aspirar a incluir en el convenio una metodología para evaluar los riesgos que plantean los sistemas de IA en aspectos fundamentales” (n. 19).

Procede, pues, advertir la regulación de la “evaluación y mitigación de riesgos e impactos adversos” en el convenio. Pues bien, en la misma puerta de entrada se establece la obligación general de adoptar “medidas legislativas, administrativas o de otro tipo”, que “se graduarán y diferenciarán según sea necesario en vista de la gravedad y la probabilidad de que se produzcan impactos adversos sobre los derechos humanos, la democracia y el Estado de derecho” (art. 1). Y se regula en lo que posiblemente es lo más relevante y destacable del Convenio IA, el artículo 16 que integra el Capítulo V impone a la obligación general de adoptar medidas “para la identificación, evaluación, prevención y mitigación de los riesgos e impactos” (art. 16.1.º). “Dichas medidas tendrán en cuenta el enfoque basado en el riesgo a que se refiere el artículo 1” (art. 16.2.º). Entre las obligaciones concretas, hay que tener “debidamente en cuenta el contexto y el uso previsto de los sistemas” (a); “la gravedad, duración y reversibilidad de los posibles riesgos” (b); e “integrar la perspectiva de todas las partes interesadas relevantes, incluida cualquier persona cuyos derechos puedan verse potencialmente afectados” (c). También se incluye “el registro, seguimiento y la debida consideración de los impactos” (d); “que de manera continua se lleven a cabo procesos de gestión de riesgos e impactos” (e); y que estos análisis de riesgos se plasmen en “la documentación adecuada” (f). Incluso se afirma “en su caso, la publicación de información sobre los esfuerzos adoptados” (g).

Como consecuencia lógica de la identificación y evaluación de riesgos, el Convenio IA también exige su mitigación, es decir, “la aplicación de medidas preventivas y paliativas suficientes para hacer frente a los riesgos e impactos adversos detectados, incluida, si procede, la exigencia de pruebas previas del sistema antes de que esté disponible para su primer uso” (h). Asimismo, si los riesgos se consideran “incompatibles”, se impone “establecer mecanismos de moratoria o prohibición u otras medidas apropiadas” (art. 16.3.º). Llama la atención que –a diferencia por ejemplo del artículo 5 RIA de la UE– no hay en el Convenio IA ninguna prohibición de concretos sistemas IA. Ello es así a pesar de que inicialmente la CAHAI, en diciembre de 2021, sugería prohibir total o parcialmente ciertas aplicaciones de IA.

7 ALGUNAS CLÁUSULAS HABITUALES DE LOS CONVENIOS DE DERECHOS Y EL “MECANISMO DE SEGUIMIENTO Y COOPERACIÓN”

El Convenio IA incluye habituales cláusulas de estándar mínimo de un convenio de derechos: no se puede interpretar para “limitar, derogar o afectar” derechos ya garantizados en la legislación interna de una Parte o en otros tratados (art. 22), ni para impedir dar una “protección más amplia” (art. 23). Con especial importancia para la Unión Europea y sus Estados miembros, se da la prevención de la aplicación preferente de los acuerdos o tratados que existan entre las partes sobre la materia (art. 27).

No hay un mecanismo duro de cumplimiento del Convenio IA, sino un “Mecanismo de seguimiento y cooperación” (art. 24) y se hace referencia a una “Conferencia de las partes” (1.º y 3.º a 8.º) para consultas periódicas para identificación de problemas, complementos y enmiendas, recomendaciones interpretativas, intercambio de información, solución de controversias, cooperación e incluso audiencias públicas (2.º). También se prevé la cooperación internacional (art. 25) y se invita a sumarse al Convenio IA, estando prevista la adhesión por Estados no miembros del Consejo de Europa (art. 31). Se regula la posibilidad de “enmiendas” (art. 28), “solución de controversias” (art. 29, no para los Estados de la UE respecto de su regulación).

En lo que respecta a la “Firma y entrada en vigor”, es significativo que se requiera únicamente la ratificación de cinco Estados, de los cuales al menos tres deben ser miembros del Consejo de Europa, evidenciando la intención de activar el Convenio IA lo antes posible (como, por ejemplo, el Convenio del Cibercrimen, n. 185).

8 LA INTERRELACIÓN DEL CONVENIO CON EL DERECHO DE LA UNIÓN EUROPEA Y, EN ESPECIAL, CON EL REGLAMENTO DE INTELIGENCIA ARTIFICIAL

En principio, todo lo regulado por el Convenio IA ya está protegido y con mayores garantías por el derecho de la UE, en especial en lo concerniente con el RIA y el RGPD. Resulta crucial considerar la superposición e interacción entre el Convenio IA y el derecho de la UE y por eso hay que tener en cuenta las acciones de la UE y las previsiones del convenio para garantizar la coherencia y evitar problemas respecto de la interacción del Convenio IA con el derecho de la Unión. En todo caso, no hay que obviar el valor casi constitucional del Convenio IA y su potencial de desplegar interpretaciones o estimular regulaciones y, también, que hay aspectos concretos del Convenio IA que pueden ir algo más allá de la regulación de la UE.

Respecto de las medidas para garantizar la coherencia entre el derecho de la UE y el convenio, en las negociaciones la UE busca: “1. Que el convenio sea compatible con el derecho del mercado único de la UE y otros ámbitos del derecho de la UE” y “2. Que el convenio sea compatible con la propuesta de Ley de inteligencia artificial (Ley de IA)”. La Comisión Europea llamó la atención de que “existe un solapamiento muy significativo entre el borrador preliminar del convenio y el Reglamento de IA”. Para el SEPD esto era una “importante oportunidad de complementar la propuesta de ley de IA fortaleciendo la protección de derechos fundamentales”. Por ello, aboga por “la inclusión en el convenio de disposiciones destinadas a reforzar los derechos de las personas afectadas”.

Desde la UE se propone que “las negociaciones se lleven a cabo en nombre de la Unión” para garantizar la “coherencia y la uniformidad”. El consejo enfatiza a los Estados que “en las negociaciones del convenio deben [...] apoyar, con pleno respeto mutuo, al negociador de la Unión” y “cooperarán estrechamente a lo largo del proceso de negociación, con el fin de garantizar la unidad de la representación exterior de la Unión” (art. 2. Decisión (UE) 2022/2349 del Consejo, de 21 de noviembre de 2022). En todo caso, para evitar una *mala combustión*, se consideró “necesario” incluir “una cláusula de desconexión que permita a los Estados miembros de la UE que se conviertan en partes en el Convenio IA regular las relaciones entre ellos con arreglo al derecho de la UE: “Entre los Estados miembros de la UE debe prevalecer la Ley de IA propuesta”. Así, en el Convenio IA específicamente se dispone que “Las Partes que sean miembros de la Unión Europea aplicarán, en sus relaciones mutuas, las normas de la Unión Europea que regulen las materias comprendidas en el ámbito de aplicación del presente Convenio” (cláusulas finales, capítulo VIII, art. 27 “Efectos del convenio”). Ahora bien, el ya mencionado artículo 22, obliga a aplicar el derecho más favorable y considero que en algunos ámbitos el Convenio IA podría subir el estándar y garantías del derecho de la Unión. En estos casos, habría que aplicar el Convenio IA, salvo en lo relativo a las “relaciones mutuas” (art. 27).

Finalmente, se regula la posible adhesión de la Unión Europea en sí al Convenio IA y se prevé que “dentro de los ámbitos de su competencia” la Unión Europea participará en la Conferencia de las Partes con los votos de los Estados miembros en (art. 26.4.º).

En cuanto a la superposición e interacción del convenio con el derecho de la Unión Europea, se adoptan medidas, por lo que los conflictos si los hay, no serán significativos. Como punto de partida, el RIA o el RGPD de la UE regulan y con mayores garantías lo que regula el Convenio IA. Sin embargo, hay algunos elementos concretos del convenio que pueden ir algo más allá del RAI y el derecho de la UE.

- Respecto del ámbito de aplicación, hay obligaciones en el Convenio IA que no están limitadas a sistemas IA de “alto riesgo”, como sucede con casi todas las obligaciones del RIA. En particular, las garantías de evaluación y mitigación de riesgos e impactos (artículo 16) podrían exigirse a sistemas que no son de alto riesgo.
- También, mientras que las normas del RIA o el RGPD están excluidas del ámbito de la seguridad nacional, el Convenio IA no excluye *totalmente* su aplicación.
- En el Convenio IA la proclamación y regulación de los “principios” tiene un notable potencial interpretativo, mientras que en el RIA tienen limitada esta potencialidad de desplegar efectos.
- Con relación a los derechos, podrían adquirir relevancia las garantías de acceso a la documentación, registros y evidencias para poder recurrir del artículo 14.
- También el convenio amplía el alcance de las garantías del artículo 15 cuando un sistema IA “informa sustancialmente” una decisión humana.

Considero que en estos casos habría de aplicarse el propio Convenio como regulación más garantista (art. 22), a salvo de lo previsto en el artículo 27. En todo caso, parece improbable que se dé –si se me permite– algún *cortocircuito* o mala *combustión* del Convenio IA con los Estados de la UE o la misma UE como miembro.

9 PARA CONCLUIR: EL CONVENIO INCORPORA LA “LÍRICA” A LA “PROSA” QUE SUPONE EL REGLAMENTO DE LA UNIÓN EUROPEA

Nos situamos en un momento crucial para la regulación de la inteligencia artificial. La acción desde hace años de la UE y en especial su RIA, el desarrollo de esta tecnología y el impacto de tecnologías como ChatGPT han acelerado la normatividad a nivel global. En 2024 culmina el proceso de regulación de un Convenio IA desde el Consejo de Europa. Se trata de un esfuerzo significativo por armonizar y establecer un marco común y mínimo para Europa, pero con una ambición mundial centrada en los derechos humanos, la democracia y el Estado de derecho y no tanto en la economía y el mercado. Europa se quiere posicionar como un *faro luminoso*, una piedra angular en el paisaje normativo de la IA no sólo para todo el continente, sino a nivel global. Para los Estados que no son miembros de la UE, el Convenio IA obviamente puede desplegar un importante

papel normativo y jurídico. Y, para la UE y sus Estados miembros que lo ratifiquen, el Convenio tiene el potencial de complementar e incluso elevar y reforzar la protección de los derechos. Se ha destacado en este sentido que el Convenio IA no se limita a los sistemas de IA de alto riesgo. Además, integra en el ámbito jurídico y normativo como “principios” lo que hasta ahora no pasaban de ser consensuados principios de la ética de la IA. Estos principios jurídicos tienen gran potencial en manos de los operadores jurídicos para que de ellos se destilen normas y obligaciones concretas como sucede desde hace décadas con los principios de la protección de datos. Asimismo, el Convenio IA reconoce a los sujetos afectados por la IA algunos derechos y especialmente garantías y mecanismos para la defensa de sus derechos ante autoridades independientes. Especialmente relevante es el enfoque basado en el riesgo y, en particular, su artículo 16, que exige una evaluación y mitigación continuas de los riesgos e impactos adversos de cualquier tipo de sistema de IA.

Pero, más allá de estas concretas aportaciones que pueden complementar al RAI, el valor del convenio va más allá. Si se me permite, *el convenio pone la lírica a la prosa que supone el RIA*. El RIA establece las bases y estructuras de un ecosistema de IA seguro y confiable, e el Convenio centra en su impacto en las personas y la sociedad democrática. El RIA es metódico, detallado y preciso, trazando un camino claro a través de la complejidad técnica y jurídica, estableciendo estándares firmes y obligaciones concretas para los proveedores y usuarios o implantadores de sistemas de IA. En contraste, en cuanto a la *lírica*, el convenio se eleva para integrar normativamente los valores fundamentales, los principios éticos y los derechos humanos que deben guiar la evolución de la IA. El convenio no sólo tiene un valor simbólico y metajurídico, sino que es un instrumento normativo, con capacidad de integración casi constitucional en los ordenamientos jurídicos de los Estados parte y cuenta con gran potencial interpretativo. Por ello, el Convenio IA supera a decenas de instrumentos declarativos y de *soft law* que ya resultaban superfluos, inocuos e incluso tediosos.

No es posible ahora prever cuánto y cuándo podrá desplegar el convenio el potencial que tiene. No obstante, sólo con su adopción Europa lidera la regulación de la inteligencia artificial, ahora desde los derechos y la democracia, e incorpora la “lírica” a la “prosa” que supone el Reglamento de la Unión Europea.

10 REFERENCIAS BIBLIOGRÁFICAS

- Brandusescu, A. y Sieber, R. 2023. *Comments on Preliminary Discussions with the Government of Canada on Council of Europe Treaty Negotiations on Artificial Intelligence*, 31/08/2023. Disponible en: <https://ssrn.com/abstract=4559139>.
- CAHAI. 2020a. *First progress report*, 1384th meeting, 23/09/2020. Disponible en: https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=09000016809ed062.

- CAHAI. 2020b. *Feasibility study on a legal framework on AI design, development and application based on CoE standards*, 17/12/2020. Disponible en: <https://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da>.
- CAHAI. 2020c. «Towards regulation of AI systems, Global perspectives on the development of a legal framework on Artificial Intelligence systems based on the Council of Europe's standards on human rights, democracy and the rule of law», *Compilation of contributions DGI (2020)16*, CAHAI Secretariat, 12/2020. Disponible en: <https://rm.coe.int/prems-107320-gbr-2018-compli-cahai-couv-texte-a4-bat-web/1680a0c17a>.
- CAHAI. 2021a. *Possible elements of a legal framework on artificial intelligence, based on the Council of Europe's standards on human rights, democracy and the rule of law*, 3/12/2021.
- CAI. 2023a. *Revised zero draft [framework] convention on artificial intelligence, human rights, democracy and the rule of law*, Estrasburgo, 6/01/2023. Disponible en: <https://rm.coe.int/cai-2023-01-revised-zero-draft-framework-convention-public/1680aa193f>.
- CAI. 2023b. *Consolidated Working Draft of the Framework Convention*, Estrasburgo, 7/07/2023. Disponible en: <https://rm.coe.int/cai-2023-18-consolidated-working-draft-framework-convention/1680abde66>.
- CAI. 2023c. *Consolidated working draft*, 18/12/2023. Disponible en: <https://rm.coe.int/cai-2023-28-draft-framework-convention/1680ade043>.
- Canadian Bar Association. 2023a. *Council of Europe Treaty Negotiations on Artificial Intelligence*, 10/2023. Disponible en: [https://www.cba.org/Our-Work/Submissions-\(1\)/Submissions/2023/October/Council-of-Europe-Treaty-Negotiations-on-Artificia](https://www.cba.org/Our-Work/Submissions-(1)/Submissions/2023/October/Council-of-Europe-Treaty-Negotiations-on-Artificia).
- Canadian Bar Association. 2023b. *Privacy And Access Law Section, Immigration Law Section, Ethics And Professional Responsibility Subcommittee*, 10/2023. Disponible en: [https://www.cba.org/Our-Work/Submissions-\(1\)/Submissions/2023/October/Council-of-Europe-Treaty-Negotiations-on-Artificia](https://www.cba.org/Our-Work/Submissions-(1)/Submissions/2023/October/Council-of-Europe-Treaty-Negotiations-on-Artificia).
- Castillo Parrilla, J.A. 2023. «Privacidad de grupo: un reto para el derecho a la protección de datos a la luz de la evolución de la inteligencia artificial», en *Derecho Privado y Constitución*, 43: 53-88. Disponible en: <https://doi.org/10.18042/cepc/dpc.43.02>
- Comisión Europea. 2019. *Directrices éticas para una IA fiable* (redactado por el Grupo independiente de expertos de alto nivel sobre inteligencia artificial).
- Comisión Europea. 2020. *Libro Blanco. Sobre la Inteligencia Artificial - Un enfoque europeo para la excelencia y la confianza*, COM(2020) 65 final, Bruselas, 19/02/2020. Disponible en: <https://op.europa.eu/es/publication-detail/-/publication/ace9398-594d-11ea-8b81-01aa75ed71a1>.
- Comisión Europea. 2022a. *Recomendación de Decisión del Consejo por la que se autoriza la apertura de negociaciones en nombre de la Unión Europea con vistas a un Convenio del Consejo de Europa sobre inteligencia artificial, derechos humanos, democracia y Estado de Derecho*, COM/2022/414 final, Bruselas, 18/08/2022: 5. Disponible en: <https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52022PC0414>
- Consejo de Europa. 2019. *Comité de Ministros Decisión CM/Del/Dec(2019)1353/1.5, 11 de septiembre de 2019. 353.ª reunión.*

- Consejo de Europa. 2021. *Comité de Ministros Mandato de Comité Ad Hoc (CM(2021)131-addfinal) Terms of reference*. Disponible en: <https://rm.coe.int/terms-of-reference-of-the-committee-on-artificial-intelligence-for-202/1680a74d2f>.
- Consejo de Europa. 2022. *Guide on Article 13 of the European Convention on Human Rights Right to an effective remedy*, actualizado 31/08/2022. Disponible en: <https://ks.echr.coe.int/web/echr-ks/article-13>
- Consejo de la UE. 2022. *Decisión (UE) 2022/2349 del Consejo, de 21 de noviembre de 2022, por la que se autoriza la apertura de negociaciones en nombre de la Unión Europea con vistas a un convenio del Consejo de Europa sobre inteligencia artificial, derechos humanos, democracia y Estado de Derecho*. Disponible en: <https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:32022D2349>
- Cotino Hueso, L. 2019. «Derechos y garantías ante el uso público y privado de inteligencia artificial, robótica y big data», en M. Bauzá (dir.), *El Derecho de las TIC en Iberoamérica*. Montevideo: La Ley-Thompson-Reuters. Disponible en: <http://links.uv.es/BmO8AU7>.
- Cotino Hueso, L. 2019. «Ética en el diseño para el desarrollo de una inteligencia artificial, robótica y big data confiables y su utilidad desde el derecho», en *Revista Catalana de Derecho Público*, 58. Disponible en: <http://dx.doi.org/10.2436/rcdp.i58.2019.3303>
- Cotino Hueso, L. 2021. *Guía de Protección de Datos en IA y Espacios de Datos*. Valencia: ITI. Disponible en: <https://www.iti.es/downloads/GUIA-DE-PROTECCION-DE-DATOS-EN-IA-Y-ESPACIOS-DE-DATOS.pdf>.
- Cotino Hueso, L. 2022. «Nuevo paradigma en las garantías de los derechos fundamentales y una nueva protección de datos frente al impacto social y colectivo de la inteligencia artificial», en L. Cotino Hueso (ed.), *Derechos y garantías ante la inteligencia artificial y las decisiones automatizadas*. Cizur Menor: Thompson-Reuters Aranzadi, FIADI (Federación Iberoamericana de Asociaciones de Derecho e Informática).
- Cotino Hueso, L. 2023. «Los tratados internacionales y el Derecho de la Unión Europea. Integración y relaciones con el ordenamiento español», en J.M. Castellá Andreu (ed.), *Derecho Constitucional Básico*, 7.^a ed. Barcelona: Huygens (Col. Lex Academica).
- Cotino Hueso, L. 2024. «La primera sentencia del Tribunal de Justicia de la Unión Europea sobre decisiones automatizadas y sus implicaciones para la protección de datos y el Reglamento de inteligencia artificial», en *Diario La Ley*, 01/2024.
- Cotino Hueso, L. y Gómez de Ágreda, Á. 2024. «Criterios éticos y de Derecho Internacional Humanitario en el uso de sistemas militares dotados de inteligencia artificial», en *Novum Jus*, 18(1). Disponible en: <https://novumjus.ucatolica.edu.co/issue/archive>
- Council of Europe. 2022. *Guide on Article 13 of the European Convention on Human Rights*, 31/08/2022. Disponible en: https://www.echr.coe.int/documents/d/echr/guide_art_13_eng
- Crawford, K. y Schultz, J. 2014. «Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms», en *Boston College Law Review*, 55: 93-128.

- Fjeld, J., et al. 2020. «Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI», en *Berkman Klein Center for Internet & Society Research at Harvard University*, enero. Disponible en: <https://dash.harvard.edu/handle/1/42160420>
- Gobierno de Reino Unido. 2023. *A pro-innovation approach to AI regulation*. Londres: Department for Science, Innovation and Technology and Office for Artificial Intelligence. 29/03/2023, actualización 3/08/2023. Disponible en: <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper>
- Grupo del Artículo 29. 2017. *Directrices sobre decisiones individuales automatizadas y elaboración de perfiles a los efectos del Reglamento 2016/679*. 3/10/2017, revisadas el 6/02/2018.
- Kolfschooten, H. y Shachar, C. 2023. «The Council of Europe's AI Convention (2023-2024): Promises and pitfalls for health protection», en *Health Policy*, 138. Disponible en: <https://doi.org/10.1016/j.healthpol.2023.104935>.
- Leslie, D., Burr, C., Aitken, M., Cows, J., Katell, M. y Briggs, M. 2021. *Artificial Intelligence, Human Rights, Democracy and the Rule of Law: a Primer*. Disponible en: <https://rm.coe.int/primer-en-new-cover-pages-coe-english-compressed-2754-7186-0228-v-1/1680a2fd4a>.
- Mantelero, A. 2014. «Toward a New Approach to Data Protection in the Big Data Era», en U. Gasser, et al. (dir.), *Internet Monitor 2014: Reflections on the Digital World*. Cambridge (MA): Berkman Center for Internet and Society at Harvard University.
- Mantelero, A. 2017. «From group privacy to collective privacy: towards a new dimension of privacy and data protection in the big data era», en L. Taylor, B. Van Der Sloot y L. Floridi, L. (eds.), *Group Privacy. New Challenges of Data Technologies*. Berlín: Springer Verlag.
- Mantelero, A. 2022. *Beyond Data. Human Rights, Ethical and Social Impact Assessment*. Torino: Springer (Col. Information Technology and Law Series, 36). Disponible en: <https://link.springer.com/book/10.1007/978-94-6265-531-7>.
- Martínez Martínez, R. 2019. «Inteligencia artificial desde el diseño. Retos y estrategias para el cumplimiento normativo», en *Revista catalana de dret públic*, 58: 64-81.
- OCDE. 2023. *AI terms & concepts*, 10/2023. Disponible en: <https://oecd.ai/en/ai-principles>.
- Palma Ortigosa, A. 2022. *Decisiones automatizadas y protección de datos personales. Especial atención a los sistemas de inteligencia artificial*, Dykinson, 2022.
- RED DAIA. 2020. *Aportación DAIA a consulta pública Carta de derechos digitales*, 12/2020. Disponible en: <https://bit.ly/3paFOH0>.
- Roig I Batalla, A. 2021. *Las garantías frente a las decisiones automatizadas del Reglamento general de Protección de Datos a la gobernanza algorítmica*. Barcelona: J.M. Bosch.
- Sánchez-Molina, P. 2018. «El origen de la cláusula de la mayor protección de los derechos humanos», en *Estudios de Deusto*, 66(1): 375-391. Disponible en: [http://dx.doi.org/10.18543/ed-66\(1\)-2018pp375-391](http://dx.doi.org/10.18543/ed-66(1)-2018pp375-391).

- 29 Sobre el tema, puede seguirse Cotino Hueso, 2023: 241-258.
- 30 Al respecto, cabe tener en cuenta la "nueva" definición en OCDE, 2023.
- 31 SEPD, 2022: 8.
- 32 Considerando 2 septies (nuevo), enmienda 11 y en la regulación del ámbito en el artículo 2, apartado 5 quinquies (nuevo).
- 33 El tema lo analizo en profundidad en Cotino Hueso, 2021.
- 34 "3. El presente convenio no se aplicará a las actividades de investigación [y desarrollo] relativas a los sistemas de inteligencia artificial, a menos que los sistemas se diseñen, desarrollen, utilicen o desmantelen [se prueben o se utilicen de otro modo] de manera que puedan interferir con los derechos humanos, la democracia y el Estado de derecho".
- 35 Sobre el tema, puede seguirse Cotino Hueso y Gómez de Ágreda, 2024.
- 36 <https://rm.coe.int/cai-2023-18-consolidated-working-draft-framework-convention/1680abde66>.
- 37 Así, se flexibiliza la aplicación del Convenio para "proteger intereses esenciales de seguridad nacional [incluyendo en su caso], incluso mediante actividades relacionadas con la inteligencia y la contrainteligencia extranjeras" (opción A); o simplemente por "intereses de seguridad nacional (opción B). En la opción C, la parte "podrá restringir" la aplicación del convenio "para proteger intereses esenciales de seguridad nacional".
- 38 Considerando 10, con mención del artículo 216.2.º, del TFUE. *Decisión (UE) 2022/2349 del Consejo, de 21 de noviembre de 2022, por la que se autoriza la apertura de negociaciones en nombre de la Unión Europea con vistas a un convenio del Consejo de Europa sobre inteligencia artificial, derechos humanos, democracia y Estado de Derecho*, <https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:32022D2349>
- 39 La Declaración del Parlamento de la UE de 2017 sobre robótica es una clara manifestación de estos principios. Destacan sin duda las Directrices éticas para una IA fiable desde la UE, la Recomendación del Consejo sobre inteligencia artificial de la OCDE del 22 de mayo de 2019 y la Recomendación sobre la ética de la inteligencia artificial de la UNESCO de noviembre de 2021. Ya en 2018, el proyecto *AI4People* contabilizó 47 principios éticos proclamados internacionalmente y los destiló en cinco puntos: beneficencia ("hacer el bien"), no maleficencia ("no hacer daño"), autonomía o acción humana ("human agency") ("respeto por la autodeterminación y elección de los individuos") y justicia ("Trato justo y equitativo para todos"). Todo ello puede seguirse Cotino Hueso, 2019.
- 40 Fjeld *et al.*, 2020.
- 41 El SEPD, 2022, recomienda una referencia explícita al cumplimiento de la normativa UE (3.3) y que haya un enfoque de protección de datos desde el diseño y por defecto (apartado 6, p. 13 y ss. n. 36 y ss.).
- 42 Así, en el artículo 2e. "sujeto de inteligencia artificial": toda persona física o jurídica cuyos derechos humanos y libertades fundamentales o derechos jurídicos conexos garantizados por el Derecho nacional o internacional aplicable se vean afectados por la aplicación de un sistema de inteligencia artificial, incluidas las decisiones adoptadas o sustancialmente informadas por la aplicación de dicho sistema.
- 43 SEPD, 2022: 10 (n. 28).
- 44 Valcke y Hendrickx, 2023.
- 45 Ver "La creación dinámica de grupos algorítmicos, la privacidad colectiva y de grupo", en Cotino Hueso, 2019. Sobre el tema recientemente, Castillo Parrilla, 2023: 53-88.
- 46 Ver mi aportación en RED DAIA, 2020.
- 47 Además de la obra reseñada, por ejemplo, Mantelero, 2017: cap. 8.
- 48 En especial apartados 28, 30 (igualdad) 46 y 47 (Gobernanza), 50, 52, 53, 60 y 64 (impacto ético) o en el n. 91 (género).
- 49 Brandusescu y Sieber, 2023: 6.
- 50 SEPD, 2022: n. 6, apartado 8 de "Conclusiones", p. 15.