

# IMPLICACIONES JURÍDICAS Y BIOÉTICAS DE LA INTELIGENCIA ARTIFICIAL (IA). ESPECIAL CONSIDERACIÓN AL MARCO NORMATIVO INTERNACIONAL

## LEGAL AND BIOETHICAL IMPLICATIONS OF ARTIFICIAL INTELLIGENCE (AI). SPECIAL CONSIDERATION OF THE INTERNATIONAL REGULATORY FRAMEWORK

DANIEL GARCÍA SAN JOSÉ

*Catedrático de Derecho Internacional Público  
Universidad de Sevilla*

Recibido: 11.12.2020 / Aceptado: 11.01.2021

DOI: <https://doi.org/10.20318/cdt.2021.5959>

**Resumen:** Al igual que otros avances científicos y tecnológicos, la Inteligencia Artificial (AI) plantea tanto beneficios como desafíos y, por algunas de sus características (opacidad, naturaleza indeterminada e imprevisibilidad en cuanto a los riesgos) demanda una regulación internacional para asegurar que no pondrá en peligro los derechos humanos y las libertades fundamentales. En este estudio se plantea varias opciones (tratado internacional, códigos de conducta y activismo judicial por parte de órganos de control internacionales) que podrían servir a tal fin, decantándonos por la que nos parece la mejor opción: la vía de estandarización de principios informadores como la dignidad humana en conexión con el principio de necesidad, en tanto que catalizador de una respuesta de la comunidad mundial sobre una base consensual colectiva.

**Palabras clave:** regulación internacional de la IA, comunidad mundial, Derecho Global, Principios informadores del ordenamiento jurídico internacional, Principios de dignidad humana y de necesidad.

**Abstract:** Like other scientific and technological advances, Artificial Intelligence (AI) poses both benefits and challenges and, due to some of its characteristics (opacity, indeterminate nature, and unpredictability in terms of risks) it demands international regulation to ensure that it will not compromise human rights and fundamental freedoms. In this study, several options are proposed (international treaty, codes of conduct and judicial activism by international control bodies) that could serve this purpose, opting for what seems to us the best option: the standardizing path of informing principles such as human dignity in connection with the principle of necessity, as a catalyst for a response by the world community on a collective consensual basis.

**Keywords:** International Regulation of Artificial Intelligence , World Community , Global Law, Informing Principles of International Legal Order, Principles of Human Dignity and Necessity.

**Sumario:** I. Introducción. La inteligencia artificial (IA): constantes y cambios frente a otros desafíos tecnológicos para el Derecho Internacional. 1. La experiencia aprendida del pasado frente a similares retos. 2. La especificidad de la IA que introduce cambios no contemplados en el pasado. A)

---

\* Este estudio se ha realizado en el marco del Proyecto de Investigación del MICINN Biomedicina, Inteligencia Artificial, Robótica y Derecho: los Retos del Jurista en la Era Digital. (PID2019-108155RB-I00).

La indeterminación conceptual de la IA. B) La opacidad de la IA. C) La imprevisibilidad en cuanto a los riesgos y desafíos que plantea su utilización. II. Opciones sobre la mesa: la vía convencional uniformadora del hard law (tratado), la vía armonizadora del soft law (códigos de conducta) y/o la vía intermedia (interpretación jurisprudencial teleológica de los tratados en vigor). 1. Una solución de hard law: ¿un tratado específico sobre IA? Luces y sombras. 2. Una solución de soft law: ¿códigos de conducta jurídicamente no vinculantes? 3. Una solución intermedia: ¿el activismo judicial de tribunales internacionales previstos en tratados internacionales sin necesidad de modificar éstos? III. La vía de estandarización del Global Law (principios informadores) como catalizador de una respuesta de la comunidad mundial sobre una base consensual colectiva. 1. Los principios de precaución y de gestión responsable del riesgo en el ciclo de vida de los sistemas de IA. 2. El principio de necesidad vinculado al principio de dignidad humana como un emergente principio general del Derecho Global frente a los riesgos de una expansión imparparable y sin control de la IA. IV. Conclusiones.

## I. Introducción. La inteligencia artificial (IA): constantes y cambios frente a otros desafíos tecnológicos para el Derecho Internacional

1. En 1989, Isaac ASIMOV señaló las que se conocen como las tres leyes de la robótica. Primera Ley: “un robot no hará daño a un ser humano ni, por inacción, permitirá que un ser humano sufra daño”. Segunda Ley: “un robot debe cumplir las órdenes dadas por los seres humanos, a excepción de aquellas que entren en conflicto con la primera ley”. Tercera Ley: “un robot debe proteger su propia existencia en la medida en que esta protección no entre en conflicto con la primera o con la segunda ley.”<sup>1</sup> Es relevante, más allá de la simplicidad de estas reglas, la idea subyacente de que la tecnología –representada aquí en forma de robots- debía estar regulada; esto es, sometida a algún tipo de código de conducta de carácter ético, moral e incluso jurídico. Desde la proclamación de las leyes de la robótica hace treinta años, nadie ha cuestionado su validez y este consenso general sirve de punto de partida para el presente estudio. La pregunta desde la que partimos es si esas tres leyes, u otras más sofisticadas, son válidas y/o necesarias para la inteligencia artificial (en adelante “IA”).

2. La IA es una disciplina científica que incluye varios enfoques y técnicas, como el aprendizaje automático (del que el aprendizaje profundo y el aprendizaje por refuerzo constituyen algunos ejemplos), el razonamiento automático (que incluye la planificación, programación, representación y razonamiento de conocimientos, búsqueda y optimización) y la robótica (que incluye el control, la percepción, sensores y accionadores, así como la integración de todas las demás técnicas en sistemas ciberfísicos)<sup>2</sup>.

3. Antes de responder a la pregunta anterior, debe volver a recordarse que la finalidad última de dichas leyes de la robótica era preservar al ser humano<sup>3</sup> y la pertinencia de esta llamada a la memoria se explica por el hecho de que si los avances científicos y tecnológicos de los que somos testigos en los últimos años conllevan tanto beneficios como desafíos para los individuos y las sociedades, tratándose de la IA, los riesgos para los derechos humanos y libertades fundamentales, incluyendo el derecho a la vida, es una realidad no cuestionada en ningún foro académico o institucional. Véanse las distintas Recomendaciones y Resoluciones aprobadas en 2019 y 2020 y que analizamos en este estudio: a nivel de la Unión Europea: *Libro Blanco de la Comisión Europea sobre la Inteligencia Artificial: un enfoque europeo orientado hacia la excelencia y la confianza*, COM (2020) 65 final, de 19 de febrero de 2020<sup>4</sup>, que sigue las *Directrices para una Inteligencia Artificial fiable* del Grupo de Expertos de Alto Nivel, creado por la Comisión Europea

<sup>1</sup> I. ASIMOV, *Círculo vicioso. Los robots*. trad. Domingo Santos. Barcelona, Martínez Roca, 1989.

<sup>2</sup> Así aparece definida en las *Directrices para una Inteligencia Artificial fiable del Grupo de Expertos de Alto Nivel* creado por la Comisión Europea y publicadas en 2019 (disponible en internet en <https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines/#/Top>) Parágrafo 144, en p. 48.

<sup>3</sup> Eran leyes elaboradas por individuos –no por los propios robots- para protegerse de dichas máquinas.

<sup>4</sup> Disponible en internet en [https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020\\_es.pdf](https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_es.pdf) (visitado el 1 de diciembre de 2020)

y publicadas un año antes<sup>5</sup>. El Consejo de Europa, por su parte, ha adoptado la *Recomendación CM/Rec (2020) 1 del Comité de Ministros a los Estados Miembros sobre los impactos de los sistemas algorítmicos en los derechos humanos*, el 8 de abril de 2020<sup>6</sup>. A nivel universal, es posible señalar que la UNESCO sigue a la Unión Europea en su Anteproyecto de Recomendación sobre la Ética de la IA<sup>7</sup>. De igual manera, es posible citar los *Principios Éticos de la OCDE en materia de Inteligencia Artificial* adoptados en mayo de 2019<sup>8</sup> que han sido posteriormente incluidos como anexo a la Declaración Ministerial sobre Comercio y Economía Digital, del G20 en su cumbre de Tsukuba, Japón, en junio de ese mismo año.<sup>9</sup> No apreciamos tal alineamiento en los distintos documentos que hemos analizado en el ámbito de Naciones Unidas, como comentamos críticamente más adelante: la estrategia del Secretario General de esta Organización en materia de nuevas tecnologías, recogida en la Resolución de la Asamblea General A/RES/73/17, de 3 de diciembre de 2018<sup>10</sup> ni tampoco en el documento *United Nations Activities on Artificial Intelligence (AI) 2019*, elaborado por la Unión Internacional de Telecomunicaciones<sup>11</sup> que recoge anualmente los progresos de la alianza *AI for Good* liderada por la Unión Internacional de Telecomunicaciones<sup>12</sup>.

4. A nivel doctrinal la IA aparece definida como “la ciencia y la ingeniería de crear máquinas inteligentes, especialmente programadas por computación inteligente”<sup>13</sup>. En esto ofrecen una clara conexión con los robots del Profesor ASIMOV: son máquinas capaces de pensar por sí mismas pero no por sí solas, al menos, no en un primer momento, en la medida en que requieren de un aprendizaje de la mano de programadores de sistemas. A medida que progresan en su aprendizaje, éste se va haciendo cada vez más autónomo e independiente gracias a las interacciones que realizan con seres humanos y otras máquinas y a su capacidad de procesar datos a gran escala (Big Data). Es un hecho temido como inevitable, el que no será posible controlar de manera indefinida una tecnología como la IA que tiene el potencial de superarnos en inteligencia a los seres humanos. Es lo que se denomina “la singularidad de las máquinas”<sup>14</sup> y, se crea o no, ese momento está más cerca de la realidad que de la ciencia ficción<sup>15</sup>.

5. Hasta que ese día llegue –si finalmente ocurre– podemos pensar a más corto plazo y enfocar nuestra pregunta inicial “¿qué normas debían regular la IA en el presente?” en la medida en que hoy día la IA ya está presente en numerosos ámbitos de nuestra vida. Así se destaca en publicaciones de divulgación científica como *Revista de Robots*<sup>16</sup>: en la movilidad y el transporte (vehículos privados y de servicio público conducidos de manera autónoma por sistemas de IA); proyectos de ciudades inteligentes con tecnología 5G y miles de sensores y de cámaras recogiendo datos que analizan en tiempo real; chatbots que interactúan con usuarios de servicios públicos (Administración) y privados (bancos y otras empresas);

<sup>5</sup> Disponible en internet en <https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#/Top> (visitado el 1 de diciembre de 2020)

<sup>6</sup> Disponible en <https://rm.coe.int/09000016809e1154> (visitado el 1 de diciembre de 2020)

<sup>7</sup> SHS/BIO/AHEG-AI/2020/4 REV 2, Doc. Final, Paris, 7 de septiembre de 2020, disponible en <https://unesdoc.unesco.org/ark:/48223/pf0000373434/PDF/373434eng.pdf.multi> Visitado el 1 de diciembre de 2020)

<sup>8</sup> Disponibles en <https://www.oecd.org/going-digital/ai/principles/> (visitado el 1 de diciembre de 2020)

<sup>9</sup> Disponible en [https://trade.ec.europa.eu/doclib/docs/2019/june/tradoc\\_157920.pdf](https://trade.ec.europa.eu/doclib/docs/2019/june/tradoc_157920.pdf) (visitado el 1 de diciembre de 2020)

<sup>10</sup> Disponible en internet en <https://undocs.org/pdf?symbol=es/A/RES/73/17> (visitado el 1 de diciembre de 2020)

<sup>11</sup> Disponible en <https://www.un.org/en/newtechnologies/> (visitado el 1 de diciembre de 2020)

<sup>12</sup> Disponible en <https://www.aiforgood.itu.int> (visitada el 1 de diciembre de 2020).

<sup>13</sup> Definición atribuida al Profesor John MacCarthy quien la utilizó por primera vez en una conferencia impartida en Dartmouth (USA) en 1956. *Vid.*: “¿Qué es la Inteligencia Artificial y para qué sirve la IA?”, Editorial de *Revista de Robots*, septiembre 1, 2020. Disponible en internet: <https://revistaderobots.com/inteligencia-artificial/que-es-la-inteligencia-artificial/> Visitada el 01/12/2020

<sup>14</sup> Elon MUSK, fundador de empresas que utilizan la inteligencia artificial como Tesla, SpaceX y Paypal, afirma incluso que las IA es mucho más peligrosa que las armas nucleares por su capacidad de mejora exponencial, más allá de lo que podamos imaginar. *Confer*: “¿Qué es la Inteligencia Artificial y para qué sirve la IA?”, en *Revista de Robots*, septiembre 1, 2020, *op. cit.*

<sup>15</sup> Como se señala en “¿Qué es la Inteligencia Artificial y para qué sirve la IA?”, en *Revista de Robots*, septiembre 1, 2020, *op. cit.*: “un suceso que no debe pasar desapercibido sucedió en 2017, cuando los ingenieros del Departamento de Desarrollo de IA de Facebook tuvieron que desconectar dos máquinas urgentemente al comprobar que habían perdido el control sobre ellas (...) En pocas horas habían sido capaces de crear un lenguaje nuevo con el que se comunicaban entre sí, además de haber desarrollado un sistema que impedía que pudiesen ser desconectadas.”

<sup>16</sup> Ver *supra* cita en nota a pie núm. 14.

algoritmos y sistemas de IA capaces de realizar diagnósticos médicos y de predicción de sufrir enfermedades con un índice de fiabilidad del 99 por ciento, a partir de herramientas como el Big data, etc.

6. Como veremos en páginas siguientes, si a nivel doctrinal existe una cierta coincidencia a la hora de definir la IA, no ocurre lo mismo en cuanto a su regulación internacional pudiendo apreciarse diversidad en su conceptualización: sistemas de IA<sup>17</sup>, IA<sup>18</sup>, sistemas algorítmicos<sup>19</sup>. Así, por ejemplo, para la Comisión Europea el término IA se aplica a los sistemas que manifiestan un comportamiento inteligente, pues son capaces de analizar el entorno y pasar a la acción –con cierto grado de autonomía– con el fin de alcanzar objetivos específicos. Los sistemas basados en la IA pueden consistir simplemente en un programa informático (ej. Asistentes de voz, programas de análisis de imágenes, motores de búsqueda, sistemas de reconocimiento facial y de voz), pero la IA también puede estar incorporada en dispositivos de hardware (ej. Robots avanzados, automóviles autónomos, drones o aplicaciones del internet de las cosas).<sup>20</sup> Para el Consejo de Europa por *sistemas algorítmicos* se entienden “las aplicaciones que, a menudo, mediante técnicas de optimización matemática, realizan una o más tareas como recopilar, combinar, limpiar, ordenar, clasificar e inferir datos, así como la selección, priorización, elaboración de recomendaciones y la toma de decisiones. Basándose en uno o más algoritmos para cumplir con sus requisitos en los entornos en los que se aplican, los sistemas algorítmicos automatizan las actividades de una manera que permite la creación de servicios adaptables a escala y en tiempo real.”<sup>21</sup>

7. En el ámbito de Naciones Unidas, la iniciativa de cumbres globales a celebrar con carácter anual bajo el lema *AI for Good*<sup>22</sup> es el buque insignia de la toma en consideración de la IA por parte de Naciones Unidas<sup>23</sup>. Liderada por la Unión Internacional de Telecomunicaciones, mediante esta plataforma en la que participan más de 37 agencias y organismos de la ONU<sup>24</sup>, se promueve un diálogo global e inclusivo sobre AI, en especial, en conexión con los objetivos del desarrollo sostenible previstos en la Agenda 2030. La última de estas reuniones tuvo lugar el 29 de mayo 2019 y de la misma surgió un documento<sup>25</sup> en el que se da detalles sobre la experiencia de las agencias de la ONU con IA para mejorar su respuesta a los desafíos globales sin contener una lectura crítica de la misma, siguiendo así la Resolución

<sup>17</sup> *Libro Blanco de la Comisión Europea sobre la Inteligencia Artificial: un enfoque europeo orientado hacia la excelencia y la confianza*, COM (2020) 65 final, de 19 de febrero de 2020 que sigue las *Directrices para una Inteligencia Artificial fiable* del Grupo de Expertos de Alto Nivel, creado por la Comisión Europea y publicadas un año antes, *op. cit.* La UNESCO sigue a la Unión Europea al referirse a sistemas de inteligencia artificial en su Anteproyecto de Recomendación sobre la Ética de la IA, SHS/BIO/AHEG-AI/2020/4 REV 2, Doc. Final, Paris, 7.09.2020, *op. cit.* En la actualidad, dicho documento está en fase de consulta de los 193 Estados miembros de la Organización con vista a su posible aprobación en la Conferencia General de la UNESCO prevista en noviembre de 2021. *Vid.* <https://es.unesco.org/news/unesco-da-gran-paso-primer-instrumento-normativo-etica-ia> (consultado el 1 de diciembre de 2020).

<sup>18</sup> Principios Éticos de la OCDE en materia de Inteligencia Artificial, *op. cit.*, adoptados en mayo de 2019 y posteriormente incluidos como anexo a la Declaración Ministerial sobre Comercio y Economía Digital, del G20 en su cumbre de Tsukuba, Japón, en junio de ese mismo año. Esta es también la terminología preferida en el ámbito de Naciones Unidas. *Vid. United Nations Activities on Artificial Intelligence (AI) 2019*, documento elaborado por la Unión Internacional de Telecomunicaciones y disponible en internet en [https://www.itu.int/dms\\_pub/itu-s/opb/gen/S-GEN-UNACT-2019-1-PDF-E.pdf](https://www.itu.int/dms_pub/itu-s/opb/gen/S-GEN-UNACT-2019-1-PDF-E.pdf) (consultados el 1 de diciembre de 2020).

<sup>19</sup> Es el caso del Consejo de Europa: *Recomendación CM/Rec (2020) 1 del Comité de Ministros a los Estados Miembros sobre los impactos de los sistemas algorítmicos en los derechos humanos*, *op. cit.*

<sup>20</sup> COM (2018) 237, final, p.1.

<sup>21</sup> Parágrafo 2 de las Directrices para abordar los impactos de los sistemas algorítmicos en los derechos humanos, incluida como apéndice a la *Recomendación CM/Rec (2020) 1 del Comité de Ministros a los Estados Miembros sobre los impactos de los sistemas algorítmicos en los derechos humanos*, *op. cit.*

<sup>22</sup> Liderada por la Unión Internacional de Telecomunicaciones <https://www.aiforgood.itu.int>

<sup>23</sup> <https://www.un.org/en/newtechnologies/#>

<sup>24</sup> A saber, CTBTO, FAO, OACI, OIT, OMI, OIM, ONUSIDA, UNCTAD, UNDESA, UNDP (ex UNDP), UNECE, PNUMA, UNESCO, UNFCCC, UNFPA, UNGP, ONU-Hábitat, ACNUR, UNICEF, UNICRI, UNIDIR, ONUDI, UNDRR (antes UNISDR), UNITAR, UNODA, UNODC, UNOOSA, UNOPS, UNRISD, UNU, ONU Mujeres, OMT, PMA, OMS, OMPI, OMM y Grupo del Banco Mundial

<sup>25</sup> *United Nations Activities on Artificial Intelligence (AI) 2019*, documento elaborado por la Unión Internacional de Telecomunicaciones y disponible en internet en [https://www.itu.int/dms\\_pub/itu-s/opb/gen/S-GEN-UNACT-2019-1-PDF-E.pdf](https://www.itu.int/dms_pub/itu-s/opb/gen/S-GEN-UNACT-2019-1-PDF-E.pdf) (consultado el 1 de diciembre de 2020)

de la Asamblea General de Naciones Unidas sobre el impacto del cambio tecnológico rápido en la consecución de los Objetivos de Desarrollo Sostenible y sus metas<sup>26</sup> y la estrategia del Secretario General de las Naciones Unidas en materia de nuevas tecnologías, de septiembre de 2018<sup>27</sup>.

**8.** Así, pese a que en dicha estrategia el Secretario General de Naciones Unidas recuerda la necesidad de colaborar estrechamente para superar los desafíos y conciliar los intereses, especialmente en las esferas de la privacidad y los derechos humanos, la ética, la igualdad y la equidad, la soberanía y la responsabilidad, la transparencia y la rendición de cuentas, sin embargo, los cinco principios sobre los que articula la relación entre la principal organización internacional del mundo y la IA no parecen ser tan audaces como los enunciados en otros ámbitos institucionalizados: la protección y promoción de los valores globales establecidos en la Carta de las Naciones Unidas y la Declaración Universal de Derechos Humanos; el fomento de la inclusión y la transparencia entre gobiernos, empresas y sociedad civil para que se tomen decisiones colectivas con respecto a las nuevas tecnologías; el trabajo en colaboración estableciendo alianzas como el proyecto *breakthrough* dentro del Pacto Mundial con el sector empresarial<sup>28</sup>; el aprovechamiento de las capacidades y los mandatos existentes; finalmente, la humildad y el aprendizaje permanente.

**9.** Puede sorprender esta actitud confiada sobre los parabienes de la IA viniendo desde la Secretaría General de Naciones Unidas cuando desde el Fondo Monetario Internacional se avanza en un informe técnico cómo la inteligencia artificial podría ampliar la brecha entre naciones ricas y pobre<sup>29</sup> y en esa misma línea, el anterior relator para la pobreza de la Naciones Unidas, Profesor Philip ALSTON, presentó en 2019 su informe sobre extrema pobreza y derechos humanos<sup>30</sup> ante la Asamblea General antes de ser sustituido en 2020 por el profesor Olivier DE SCHUTTER, alertando sobre el uso de la tecnología y la IA para vigilar y castigar a las personas más pobres. En su opinión, las grandes empresas tecnológicas están desregularizándose, sin someterse a las leyes en lo que denomina zonas francas de derechos humanos. Incluso, el Profesor ALSTON se permite dejar un mensaje a su sucesor cuando señala en el párrafo 73 de su último informe como relator para la extrema pobreza y los derechos humanos: “Se objetará razonablemente que este informe es desequilibrado, o unilateral, porque se centra en los riesgos más que en las muchas ventajas que potencialmente se derivan del estado de bienestar digital. La justificación es sencilla. Hay un gran número de animadoras (cheerleaders) que ensalzan los beneficios, pero muy pocas aconsejan una reflexión sobria sobre las desventajas.”

**10.** Para la UNESCO, que también utiliza el término sistemas de IA, se entiende por tales aquellos sistemas tecnológicos capaces de procesar información de manera que se asemeje a un comportamiento inteligente<sup>31</sup>; y la OCDE omite incluir cualquier definición de IA, siguiendo una aproximación pragmática a la cuestión.

<sup>26</sup> Resolución A/RES/73/17 de 3 de diciembre de 2018. Disponible en internet en <https://undocs.org/pdf?symbol=es/A/RES/73/17> (consultado el 1 de diciembre de 2020)

<sup>27</sup> Disponible en <https://www.un.org/en/newtechnologies/>

<sup>28</sup> [https://breakthrough.unglobalcompact.org/site/assets/files/1454/hhw-16-0017-d\\_c\\_artificial\\_intelligence.pdf](https://breakthrough.unglobalcompact.org/site/assets/files/1454/hhw-16-0017-d_c_artificial_intelligence.pdf) (consultado el 1 de diciembre de 2020).

<sup>29</sup> IMF Working Paper: “Will the AI Revolution Cause a Great Divergence?” CH. ALONSO, A. BERG, S. KOTHARI, C. PAPAGEORGIOU y S. REHMAN, WP/20/184, Disponible en <https://www.imf.org/en/Publications/WP/Issues/2020/09/11/Will-the-AI-Revolution-Cause-a-Great-Divergence-49734> (consultado el 1 de diciembre de 2020)

<sup>30</sup> Report of the Special rapporteur on extreme poverty and human rights, Doc. A/74/48037, de 11 de Octubre de 2019. Disponible en internet en [https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKewje-sDS\\_sPtAhWEiVwKHaZiDNIQFjABegQIAhAC&url=https%3A%2F%2Fwww.ohchr.org%2FDocuments%2FIssues%2FPoverty%2FA\\_74\\_48037\\_AdvanceUneditedVersion.docx&usq=AOvVaw1jRH0rb0geojazfZCC8f3V](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKewje-sDS_sPtAhWEiVwKHaZiDNIQFjABegQIAhAC&url=https%3A%2F%2Fwww.ohchr.org%2FDocuments%2FIssues%2FPoverty%2FA_74_48037_AdvanceUneditedVersion.docx&usq=AOvVaw1jRH0rb0geojazfZCC8f3V) (consultado el 1 de diciembre de 2020)

<sup>31</sup> La definición que se da en el párrafo 2.a) de su Anteproyecto de Recomendación sobre la Ética de la IA, SHS/BIO/AHEG-AI/2020/4 REV 2, Doc. Final es la siguiente: “los sistemas de IA son tecnologías de procesamiento de la información que incorporan modelos y algoritmos que producen una capacidad para aprender y realizar tareas cognitivas, dando lugar a resultados como la predicción y la adopción de decisiones en entornos reales y virtuales. Los sistemas de IA están diseñados para funcionar con una cierta autonomía, mediante la modelización y representación del conocimiento y la explotación de datos y el cálculo de correlaciones.”

**11.** Lo que sí parece ser elemento común a los distintos foros internacionales institucionalizados es la necesidad de que la IA esté controlada por el Derecho con un enfoque centrado en la persona<sup>32</sup>. En el ámbito europeo, por ejemplo, una IA con un enfoque centrado en la persona se esfuerza por asegurar que los valores humanos ocupen un lugar central en el desarrollo, despliegue, utilización y supervisión de los sistemas de IA, garantizando el respeto de los derechos fundamentales, incluidos los recogidos en los Tratados de la Unión Europea y en la Carta de los Derechos Fundamentales de la Unión Europea; todos ellos constituyen una referencia unitaria a un fundamento común arraigado en el respeto de la dignidad humana, en el que el ser humano disfruta de una condición moral única e inalienable. Esto requiere asimismo tener en cuenta el entorno natural y el resto de seres vivos que forman parte del ecosistema humano, así como un enfoque sostenible que permita la prosperidad de las generaciones futuras.<sup>33</sup>

**12.** Así se recoge, por ejemplo, en el *Libro Blanco de la Comisión Europea sobre la IA*: “El futuro marco normativo para la IA en Europa... Debe velar por el cumplimiento de las normas de la UE, especialmente la normas de protección de los derechos fundamentales y los derechos de los consumidores, y en concreto, con relación a los sistemas de IA que operan en la UE y presentan un riesgo elevado.”<sup>34</sup> Idea que es reiterada después, para el ámbito de la Acción Exterior de esta organización: “La Comisión está convencida de que la cooperación internacional sobre cuestiones relativas a la IA debe basarse en un enfoque que promueva el respeto de los derechos fundamentales, especialmente la dignidad humana, el pluralismo, la inclusión, la ausencia de discriminación y la protección de la privacidad y de los datos personales y se esforzará por exportar esos valores al resto del mundo.”<sup>35</sup>

**13.** En otros foros, esta visión antropocéntrica de la regulación de la IA está igualmente presente. Así, para la UNESCO, la dignidad humana ha de ser central en cualquier instrumento normativo internacional de la IA: “(la UNESCO) ... Convencida de que el instrumento normativo que aquí se presenta, basado en el Derecho Internacional y en un enfoque normativo mundial y centrado en la dignidad humana y de los derechos humanos, así como en la igualdad de género, la justicia social y económica, el bienestar físico y mental, la diversidad, la interconexión, la inclusión y la protección del medio ambiente y de los ecosistemas, puede dar una orientación responsable a las tecnologías de la IA.”<sup>36</sup> Más explícitamente, se señala en la parte dispositiva de dicha recomendación que: “(Los Estados deben) asegurarse de que los derechos humanos y el Estado de Derecho son respetados en el mundo digital al igual que en el mundo físico.”<sup>37</sup> Incluso, en el ámbito de la OCDE se asume esta visión antropocéntrica cuando se señala que: “Los actores de la IA deben respetar el estado de derecho, los derechos humanos y los valores democráticos durante todo el ciclo de vida del sistema de IA. Estos incluyen libertad, dignidad y autonomía, privacidad y protección de datos, no discriminación e igualdad, diversidad, equidad, justicia social y los derechos laborales reconocidos internacionalmente.”<sup>38</sup>

**14.** El lenguaje utilizado en la Recomendación de la UNESCO –un foro institucionalizado de ámbito más amplio que el europeo e incluso, de la OCDE- explica el diferente alcance de las obligaciones asumidas por los Estados con relación a dicha recomendación: “Los Estados y todas las demás

<sup>32</sup> Un principio recogido en la Recomendación de la UNESCO y no reflejado en otras Recomendaciones del Consejo de Europa, de la UE o de la OCDE, es el de la soberanía de los datos (párrafo 46 de la Recomendación), en virtud del cual, los Estados, en cumplimiento del Derecho Internacional, regulan los datos generados dentro de su territorio o que pasen por ellos y adoptan medidas para la regulación efectiva de los datos sobre la base del respeto del derecho a la privacidad y otros derechos humanos.

<sup>33</sup> *Directrices para una Inteligencia Artificial fiable del Grupo de Expertos de Alto Nivel*, *op. cit.*, Párrafo 153 en p. 49.

<sup>34</sup> *Libro Blanco de la Comisión Europea sobre la Inteligencia Artificial: un enfoque europea orientado hacia la excelencia y la confianza*, COM (2020) 65 final, *op. cit.*, p. 3.

<sup>35</sup> *Ibidem*, p. 11.

<sup>36</sup> Anteproyecto de Recomendación sobre la Ética de la IA, SHS/BIO/AHEG-AI/2020/4 REV 2, Doc. Final, *op. cit.*, párrafo 4 del Preámbulo.

<sup>37</sup> Anteproyecto de Recomendación sobre la Ética de la IA, *op. cit.*, par. 53.

<sup>38</sup> Principio 1.2. a) Valores y equidad centrados en el ser humano de los Principios Éticos de la OCDE en materia de Inteligencia Artificial, *op. cit.*, adoptados en mayo de 2019 y posteriormente incluidos como anexo a la Declaración Ministerial sobre Comercio y Economía Digital, del G20 en su cumbre de Tsukuba, Japón, en junio de ese mismo año.

partes interesadas que se indican en la presente Recomendación *deben respetar, promover y proteger* los principios y normas éticas relativos a la IA que se establecen en la Recomendación y *deberían adoptar todas las medidas posibles* para dar efecto a sus recomendaciones de actuación.” (La cursiva es añadida)

## 1. La experiencia aprendida del pasado frente a similares retos.

15. Una mirada atrás en la Historia muestra que frente a los nuevos descubrimientos científicos sólo caben cuatro aproximaciones posibles: el incentivo de la actividad mediante la financiación gubernamental y la protección de patentes; la abstención reguladora, consecuente con el clima de *laissez faire* imperante en las sociedades europeas y occidentales de la revolución industrial del siglo XIX; la opción de la regulación legal, controlando y desincentivando una actividad o parte de la misma, como sucedió a raíz de la Gran Depresión de 1929 que siguió a la expansión económica sin precedentes del siglo XIX; finalmente, la cuarta opción posible sería la que más frecuentemente se habría repetido en el pasado: la prohibición de la idea o descubrimiento en cuestión<sup>39</sup>. De las cuatro, la prohibición y la abstención reguladora no parecen ser opciones particularmente aconsejables por los riesgos de la aparición de zonas francas tecnológicas en las que la investigación se lleva a cabo sin límites normativos y pueda extrapolar riesgos o eventuales daños para terceros en nuestro presente globalizado sin restricciones a los flujos de mercancías y servicios. La respuesta legal de incentivos mediante apoyo gubernamental y la protección de patentes no es descartable pero parece claramente insuficiente si no va acompañada de la que nos parece la aproximación legal más conveniente para la IA: su regulación normativa<sup>40</sup>. Así pues, llegados a este punto, ¿qué normas deben regular la IA desde un enfoque antropocéntrico? sigue siendo la pregunta relevante. De nuevo un ejercicio de reflexión es necesario.

16. La regulación normativa de la IA no es tarea exenta de dificultades, como no lo ha sido –y sigue siendo– la regulación de cuestiones técnicas como la nanotecnología o de temas con implicaciones morales como la investigación embrionaria humana o la gestación por sustitución<sup>41</sup>. De entrada, la Ciencia siempre avanza más deprisa que el Derecho y por ello, éste sólo puede dar respuestas tardías, imperfectas y a menudo provisionales a los numerosos desafíos que día a día plantea el incesante avance en los descubrimientos científicos<sup>42</sup>. De otro lado, resulta necesario aplicar la *programación neurolingüística* (PNL) al ámbito del Derecho y asumir que la aproximación como juristas a la IA debe hacerse trascendiendo la partición de esta Ciencia en distintas parcelas del conocimiento (Derecho civil, constitucional, administrativo, penal, etc.) para asumir una nueva perspectiva jurídica caracterizada por cuatro notas: la *transversalidad* –tomando como referencia en este ámbito las normas nacionales, internacionales y supranacionales–, la *multidisciplinariedad* –reconociendo que son varias las áreas jurídicas tanto de derecho público como privado concernidas– y aceptando el *carácter integrador* y no reduccionista de dicha aproximación, esto es, no descartando *a priori* las consideraciones éticas y morales pero tampoco condicionando nuestro análisis por las mismas. A estas notas habría que añadir aún una cuarta característica: la *flexibilidad*; esto es, asumiendo la *superación de la distinción entre Derecho duro vinculante y Derecho blando no vinculante* (hard y soft law) como la mejor opción para una eventual regulación internacional de IA.

<sup>39</sup> S. MARTYN, “Human Cloning: The Role of Law”, University of Toledo Law Review, vol. 32, 2001, p. 375.

<sup>40</sup> Así se viene defendiendo por los autores considerando otros desarrollos científicos y tecnológicos de fuerte impacto ético y jurídico como son la investigación embrionaria humana y la gestación por sustitución. Vid. D. GARCÍA SAN JOSÉ, *European Normative Framework for Biomedical research in Human Embryos*, Cizur menor, Thomson Reuters Aranzadi, 2013. Del mismo autor “Epigenética y gestación por sustitución: más razones a favor de una regulación internacional para un negocio global”. Anuario Mexicano de Derecho Internacional, XVII, 2017, pp. 329-368.

<sup>41</sup> Para formarse una idea de tales dificultades a nivel universal basta considerarlas a nivel europeo, como han sido puestas de manifiesto en relación con la Carta de Derechos Fundamentales de la Unión Europea en D. GARCÍA SAN JOSÉ, *La Europa de los derechos ante los avances científicos y tecnológicos*. Valencia, Tirant lo blanch, 2018.

<sup>42</sup> D. GARCÍA SAN JOSÉ, “De la Bioética al Bioderecho: retos para el jurista español”, Revista General de Legislación y Jurisprudencia, nº 1, 2012, p. 8.

## 2. La especificidad de la IA que introduce cambios no contemplados en el pasado.

17. A las dificultades para una regulación internacional de la IA que se acaban de apuntar, hay que añadir la especificidad de la IA caracterizada por rasgos que dificultan su aprehensión por el Derecho: su indeterminación conceptual –de la que es ejemplo la diversidad de términos utilizados en los diferentes instrumentos no vinculantes apuntados al comienzo de este estudio–, su opacidad –motivada por la multiplicidad de actores que intervienen en el proceso de vida de la IA<sup>43</sup> y las consideraciones económicas en presencia- y su imprevisibilidad en cuanto a los riesgos y desafíos que plantea su utilización.

### A) La indeterminación conceptual de la IA

18. Algunos autores, como RUSELL y NORVIG, sugieren que dentro de la IA es posible distinguir cuatro subcategorías (las diseñadas para pensar como los seres humanos, para pensar de modo racional, para actuar como seres humanos y para actuar racionalmente)<sup>44</sup>. Otros autores, como TZIMAS sugieren una clasificación diferente entre una IA débil y otra fuerte (weak and strong AI), de modo que la IA fuerte se definiría sobre la base de su autonomía intelectual y de su capacidad para aprender por sí misma de manera autónoma a la presencia de seres humanos<sup>45</sup>.

19. Asumiendo la pertinencia de distinguir entre una IA débil de otra fuerte; la primera sería “buena” y un claro ejemplo de la misma es el descubrimiento gracias al sistema de IA de Google *Deepmind* del enigma del plegado de proteínas mediante el análisis de las cadenas de aminoácidos que integran los trescientos millones de genes que conforman el genoma humano.<sup>46</sup> Incluso, tratándose de una IA beneficiosa, se asume que puede causar riesgos de ahí que se postule una IA fiable<sup>47</sup>. De hecho, la preocupación por los derechos humanos es manifiesta en el ámbito del Consejo de Europa. Una novedad de la Recomendación del Comité de Ministros del Consejo de Europa a diferencia de las otras recomendaciones analizadas en el contexto de la UE, la UNESCO y la OCDE es que concreta las obligaciones de los Estados con respecto a la protección y promoción de los derechos humanos y las libertades fundamentales en el contexto de los sistemas algorítmicos como principios de aplicación general: en cuanto a la legislación<sup>48</sup>, a la

<sup>43</sup> El ciclo de vida de un sistema de IA abarca las fases de desarrollo (incluidas las tareas de investigación, diseño, provisión de datos y realización de ensayos limitados), despliegue (incluida la aplicación) y utilización de dicho sistema. *Directrices para una Inteligencia Artificial fiable del Grupo de Expertos de Alto Nivel*, op. cit., parágrafo 147, en p. 48.

<sup>44</sup> S. RUSSELL, Y P. NORVIG, *Artificial Intelligence a Modern Approach*, New Jersey, Ed. A. Simon & Schuster, 1995, pp. 4-5.

<sup>45</sup> T. TZIMAS, “Artificial Intelligence as Global Commons and The International Law Supremacy Principle”, *Advances in Social Science, Education and Humanities Research*, vol. 211, 2018, p. 83.

<sup>46</sup> <https://elpais.com/ciencia/2020-12-02/la-inteligencia-artificial-arrasa-en-uno-de-los-problemas-mas-importantes-de-la-biologia.html> (visitado el 1 de diciembre de 2020)

<sup>47</sup> Ese es el eje del *Libro Blanco de la Comisión Europea sobre la Inteligencia Artificial: un enfoque europea orientado hacia la excelencia y la confianza* siguiendo las *Directrices del Grupo de Expertos de Alto Nivel*: 1) debe ser lícita, es decir, cumplir todas las leyes y reglamentos aplicables; 2) ha de ser ética, demostrando el respeto y garantizando el cumplimiento de los principios y valores éticos, y 3) debe ser robusta. La solidez de un sistema de IA abarca tanto su solidez técnica (que resulta adecuada en un contexto determinado, como el ámbito de aplicación o la fase del ciclo de vida) así como desde el punto de vista social (garantizando que el sistema de IA tenga debidamente en cuenta el contexto y el entorno en el que opera). Esto es crucial para asegurar que, incluso si las intenciones son buenas, el sistema no provoque daños involuntarios. *Directrices para una Inteligencia Artificial fiable del Grupo de Expertos de Alto Nivel*, op. cit., párrafos 139, 156 y 160 en pp. 46 y 50.

<sup>48</sup> Principio 1.1. de la Obligaciones de los Estados, en el apéndice a la Recomendación CM /Rec (2020) 1 del Comité de Ministros del Consejo de Europa, op. cit., “El proceso de redacción, promulgación y evaluación de políticas y legislación o regulación aplicable al diseño, desarrollo y despliegue continuo de sistemas algorítmicos debe ser transparente, responsable e inclusivo. Los Estados deberían consultar periódicamente con todas las partes interesadas pertinentes y las partes afectadas, incluso a nivel sectorial cuando sea apropiado. Los Estados deben garantizar la aplicabilidad y el cumplimiento de las leyes, incluso exigiendo que los actores relevantes produzcan la documentación adecuada para verificar el cumplimiento legal. Cuando los actores del sector público y privado no cumplan con sus obligaciones legales, deben ser considerados responsables.”



revisión continua<sup>49</sup>, a la participación democrática y concienciación<sup>50</sup> y en lo que a los marcos institucionales se refiere<sup>51</sup>.

## B) La opacidad de la IA

20. Siguiendo con el ejemplo señalado en el apartado anterior, la importancia de este descubrimiento radica en que gracias a un sistema de algoritmos que aprende por sí mismo —y cuyo software Google se ha negado a revelar a la comunidad científica— será posible fabricar medicamentos a la carta que neutralicen aquellas proteínas que son dañinas por estar “mal plegadas o construidas”, y que son la causa de enfermedades tan temidas como el parkinson, el alzheimer o recientemente, los efectos asociados a la covid 19. No debe sorprender, por consiguiente, la actitud de Google reacia a poner a disposición de la comunidad científica las claves de su descubrimiento para aprovechamiento general y gratuito. La IA es un gran negocio en todo el mundo. Estudios recientes señalan que a la cabeza en inversión en este ámbito está Estados Unidos con cerca del 3.3% de su PIB, seguido de China, con un 2.1% y de Europa con un 1,7%<sup>52</sup>. No debe extrañar, tampoco, que la Comisión Europea estableciera en abril de 2018 una estrategia europea para la IA<sup>53</sup> con la intención de recuperar el terreno perdido, y que en su Libro Blanco sobre la IA haya querido movilizar al sector privado a la par que la acción concertada de los gobiernos europeos, recordando que en el ciclo de vida de un sistema de IA existen numerosas partes involucradas: el desarrollador, el implementador (la persona que utiliza un producto o servicio provisto de IA) y otras partes potenciales (productor, distribuidor o importador, proveedor de servicios, usuario profesional o particular)<sup>54</sup>. Además, en cuanto a la cuestión del alcance geográfico de la intervención legislativa, para la Comisión Europea resulta esencial que todos los agentes que ofrezcan productos o servicios provistos de IA en la Unión Europea, independientemente de que se hayan o no establecidos en la Unión, estén sujetos a requisitos comunes. De lo contrario, los objetivos de la regulación normativa no podrán alcanzarse plenamente.<sup>55</sup> Frente a la opacidad se ha defendido en la recomendación del Comité de Ministros del Consejo Europeo la exigencia de niveles de transparencia<sup>56</sup> y la necesaria identificación

---

<sup>49</sup> Principio 1.2. de la Obligaciones de los Estados, en el apéndice a la Recomendación CM /Rec (2020) 1, *op. cit.*, “A lo largo de todo el ciclo de vida de un sistema algorítmico, desde la etapa de propuesta hasta la evaluación de los efectos, los impactos sobre los derechos humanos de los sistemas individuales y su interacción con otras tecnologías deben evaluarse periódicamente. Esto es necesario debido a la velocidad y escala a la que funcionan estos sistemas y el entorno tecnológico en rápida evolución en el que operan. Esto debe hacerse sobre la base de consultas amplias y efectivas con los afectados o que puedan verse afectados.”

<sup>50</sup> Principio 1.1. de la Obligaciones de los Estados, en el apéndice a la Recomendación CM /Rec (2020) 1, *op. cit.*, “A fin de garantizar el pleno ejercicio de los derechos humanos y las libertades democráticas, los Estados deben fomentar la conciencia pública en general sobre la capacidad, el poder y los impactos consiguientes de los sistemas algorítmicos, incluido su uso potencial para manipular, explotar, engañar o distribuir recursos, con miras a Permitir que todas las personas y grupos conozcan sus derechos y sepan cómo ponerlos en práctica y cómo utilizar las tecnologías digitales para su propio beneficio. Además, todos los actores relevantes, incluidos los de los sectores público, privado y de la sociedad civil en los que se contemplan o se utilizan sistemas algorítmicos, deben promover, alentar y apoyar de manera personalizada e inclusiva un nivel de medios (teniendo en cuenta la diversidad con respecto a, por ejemplo, edad, género, raza, etnia, antecedentes culturales o socioeconómicos).

<sup>51</sup> Principio 1.4. de la Obligaciones de los Estados, en el apéndice a la Recomendación CM /Rec (2020) 1, *op. cit.*, “Los Estados deben identificar y/o desarrollar marcos y estándares institucionales y regulatorios apropiados que establezcan puntos de referencia y salvaguardas generales o sectoriales para asegurar la compatibilidad del diseño, desarrollo y despliegue continuo de sistemas algorítmicos con los derechos humanos.”

<sup>52</sup> “¿Qué es la Inteligencia Artificial y para qué sirve la IA?”, Editorial, en *Revista de Robots*, septiembre 1, 2020, *op. cit.*

<sup>53</sup> COM (2018) 237.

<sup>54</sup> *Libro Blanco de la Comisión Europea sobre la Inteligencia Artificial: un enfoque europeo orientado hacia la excelencia y la confianza*, COM (2020) 65 final, de 19 de febrero de 2020, *op. cit.*, p. 27.

<sup>55</sup> *Ibidem*.

<sup>56</sup> “Los Estados deben establecer niveles apropiados de transparencia con respecto a la contratación pública, el uso, el diseño y los criterios y métodos básicos de procesamiento de los sistemas algorítmicos implementados por y para ellos, o por actores del sector privado. Los marcos legislativos para la propiedad intelectual o los secretos comerciales no deben impedir tal transparencia, ni los Estados o las partes privadas deben tratar de explotarlos para este propósito. Los niveles de transparencia deben ser lo más altos posible y proporcionales a la gravedad de los impactos adversos sobre los derechos humanos,

de la toma de decisiones algorítmicas<sup>57</sup>. En otros foros institucionalizados, UNESCO, OCDE, se han postulado medidas similares.

### C) La imprevisibilidad en cuanto a los riesgos y desafíos que plantea su utilización

21. Precisamente, un sistema de IA dotado de la capacidad de aprendizaje autónomo y la consiguiente toma de decisiones y su ejecución sin intervención humana es lo que, en opinión de TZIMAS, dotaría de imprevisibilidad en cuanto a los riesgos y desafíos que plantea su utilización<sup>58</sup>. Un sistema de IA que es consciente de que posee conciencia propia que puede fortalecerse sobre la base de la reflexión crítica, dispone de un elevado nivel de autonomía. De ahí la necesidad de aceptar la imprevisibilidad de los riesgos que pueden derivarse de su funcionamiento<sup>59</sup> y, de modo especial, allí donde ya no llega la presencia humana<sup>60</sup>. En todo caso, se ha insistido a modo paliativo en reafirmar la rendición de cuentas<sup>61</sup> y ofrecer a los individuos que puedan ser afectados negativamente por tales decisiones de recursos efectivos a través de los cuáles puedan hacer valer sus derechos<sup>62</sup>.

---

incluidas etiquetas o sellos éticos para sistemas algorítmicos que permitan a los usuarios navegar entre sistemas. El uso de sistemas algorítmicos en los procesos de toma de decisiones que conllevan altos riesgos para los derechos humanos debe estar sujeto a estándares particularmente altos en lo que respecta a la explicación de los procesos y resultados.” Punto 4.1. de las directrices para abordar los impactos de los sistemas algorítmicos en los derechos humanos, anexo a la Recomendación CM/ Rec (2020) 1, *op. cit.*

<sup>57</sup> “Los Estados deben garantizar que todos los procesos de selección o decisiones tomadas o asistidas por sistemas algorítmicos que puedan tener un impacto significativo en el ejercicio de los derechos humanos, ya sea en el ámbito público o privado, sean identificables y trazables como tales en la interacción inicial, de forma clara y accesible”. Punto 4.2. de las directrices para abordar los impactos de los sistemas algorítmicos en los derechos humanos, anexo a la Recomendación CM/ Rec (2020) 1, *op. cit.*

<sup>58</sup> T. TZIMAS, “Artificial Intelligence as Global Commons and The International Law Supremacy Principle”, *op. cit.*, p. 84.

<sup>59</sup> *Imprevisibilidad*. Parágrafo 5 de los *Principios Éticos de la OCDE en materia de Inteligencia Artificial*, *op. cit.*, incluidos como anexo a la Declaración Ministerial sobre Comercio y Economía Digital, del G20 en su cumbre de Tsukuba, Japón, en junio de ese mismo año: “(...) es importante tener en cuenta que la mayoría de los sistemas algorítmicos se basan en modelos estadísticos en los que los errores forman una parte inevitable, a veces con ciclos de retroalimentación que mantienen, replican y refuerzan los datos previos. -sesgos, errores y suposiciones existentes. Aunque pueda parecer que los conjuntos de datos más grandes brindan mejores posibilidades de encontrar patrones y correlaciones recurrentes, las tasas de precisión no aumentan automáticamente con el tamaño del conjunto de datos. Como resultado del gran número de personas afectadas por los sistemas algorítmicos, el número de errores en forma de falsos positivos y falsos negativos, y de personas que se ven afectadas por estos errores y sesgos incorporados, también se expandirá, lo que provocará interferencias adicionales con el ejercicio de los derechos humanos de múltiples formas.”

<sup>60</sup> T. TZIMAS, “Artificial Intelligence as Global Commons and The International Law Supremacy Principle”, *op. cit.*, p. 85.

<sup>61</sup> Principio 4.3 de las obligaciones de los Estados con respecto a la protección y promoción de los derechos humanos y las libertades fundamentales en el contexto de los sistemas algorítmicos, dentro del apéndice a la Recomendación CM/ Rec (2020) 1 del Comité de Ministros del Consejo de Europa sobre los impactos de los sistemas algorítmicos en los derechos humanos: “Los individuos y grupos afectados deben disponer de medios eficaces para impugnar las determinaciones y decisiones pertinentes. Como condición previa necesaria, la existencia, el proceso, la justificación, el razonamiento y el posible resultado de los sistemas algorítmicos a nivel individual y colectivo deben explicarse y aclararse de manera oportuna, imparcial, fácilmente legible y accesible a las personas cuyos derechos o intereses legítimos puedan ser afectados, así como a las autoridades públicas pertinentes. La impugnación debe incluir la oportunidad de ser escuchado, una revisión completa de la decisión y la posibilidad de obtener una decisión no automatizada. Este derecho no puede ser renunciado y debe ser asequible y fácil de hacer cumplir antes, durante y después del despliegue, incluso mediante la provisión de puntos de contacto y líneas directas de fácil acceso.”

<sup>62</sup> Principio 4.4 de las obligaciones de los Estados con respecto a la protección y promoción de los derechos humanos y las libertades fundamentales en el contexto de los sistemas algorítmicos, dentro del apéndice a la Recomendación CM/ Rec (2020) 1 del Comité de Ministros del Consejo de Europa: “Los Estados deben garantizar procedimientos judiciales y extrajudiciales iguales, accesibles, asequibles, independientes y efectivos que garanticen una revisión imparcial, de conformidad con los artículos 6, 13 y 14 de la Convención, de todas las denuncias de violaciones de los derechos de la Convención mediante el uso de sistemas algorítmicos, provenientes de actores del sector público o privado. A través de sus marcos legislativos, los Estados deben garantizar que las personas y los grupos tengan acceso a recursos efectivos, rápidos, transparentes, funcionales y efectivos con respecto a sus quejas. La reparación judicial debe permanecer disponible y accesible cuando los mecanismos internos y alternativos de solución de controversias resulten insuficientes o cuando alguna de las partes afectadas opte por la revisión judicial o la apelación.”

## II. Opciones sobre la mesa: la vía convencional uniformadora del hard law (tratado), la vía armonizadora del soft law (códigos de conducta) y/o la vía intermedia (interpretación jurisprudencial teleológica de los tratados en vigor)

22. El Derecho Internacional ha afrontado riesgos clásicos como emergencias y catástrofes (naturales y derivadas de la acción humana, ya sean éstas accidentales o no)<sup>63</sup> y parece estar listo para encarar nuevos riesgos como la IA a través de diversos instrumentos con los que cuenta: o bien se recurre a la aproximación clásica ante cuestiones de interés general de establecer conferencias multilaterales de las que surjan convenciones de alcance universal, con mayor o menor éxito, en razón no tanto de los Estados firmantes sino de los ausentes en las mismas en virtud del principio de libre consentimiento de los Estados a la hora de asumir compromisos internacionales (sólo basta pensar en las notables ausencias en el Protocolo de Kyoto); o bien se acude, como segunda vía, a la fuerza imperativa del Consejo de Seguridad de Naciones Unidas actuando en el marco del Capítulo VII de la Carta, para lo cual, debe calificar una situación o emergencia como una amenaza para la paz y seguridad internacionales<sup>64</sup>; o finalmente –y es por esta última solución por la que nos decantamos<sup>65</sup>– por la propia naturaleza del Derecho internacional público es posible invocar una nueva base legitimadora de obligaciones internacionales, complementaria al consentimiento estatal de sumisión a dichas normas<sup>66</sup> a partir de los principios generales de dicho ordenamiento (“prima principia”, esto es, principios inferidos o de carácter estructural del mismo).

### 1. Una solución de hard law: ¿un tratado específico sobre IA? Luces y sombras

23. Al igual que se aprecia en relación con otras cuestiones suscitadas por los avances y prácticas científicas en el campo de la biomedicina y sus implicaciones para el ser humano<sup>67</sup>, la gestación por subrogación se caracteriza por una regulación internacional inexistente y las aproximaciones nacionales que se evidencian desde una perspectiva de Derecho comparado –particularmente evidente a nivel europeo– pueden ser calificadas de “normatividad de geometría variable”. Geometría, en primer lugar, por cuanto las posibles aproximaciones normativas a esta cuestión son, respectivamente, la permisividad, la restricción, la prohibición y la no regulación<sup>68</sup>. Variable, además, por cuanto las combinaciones de las aproximaciones normativas a nivel nacional se multiplican en función de que se distinga en nuestra comparativa la posición de los Estados en relación con la gestación por subrogación o respecto de las consecuencias jurídicas para las partes implicadas en este tipo de prácticas<sup>69</sup>.

<sup>63</sup> M. I. TORRES CAZORLA, “Las emergencias y catástrofes como riesgo para la seguridad: una visión desde la perspectiva del Derecho Internacional Público a la luz de la Estrategia de Seguridad Nacional de mayo de 2013”, ICADE. Revista cuatrimestral de las Facultades de Derecho y Ciencias Económicas y Empresariales, nº 92, 2014, p. 82. *Vid.* En relación con esta cuestión el interesante estudio de J. M. SÁNCHEZ PATRÓN, “Las actividades peligrosas para el medio ambiente y la salud humana en la jurisprudencia del Tribunal Europeo de Derechos Humanos”, en *Bioderecho Internacional y Europeo. Desafíos actuales*, J.M. SÁNCHEZ PATRÓN (coord.), Valencia, Tirant lo Blanch, 2014, pp. 131-148.

<sup>64</sup> No es frecuente pero tampoco imposible. Nótese que en la Resolución 2177 (2014), adoptada por el Consejo de Seguridad de Naciones Unidas el 18 de septiembre de 2015, se calificaba un brote del virus ébola como una amenaza para la paz y seguridad internacionales.

<sup>65</sup> Descartando por el momento la vía del Capítulo VII de la Carta de Naciones Unidas prevista para supuestos en los que por ahora y afortunadamente no ha entrado la IA, procederemos a examinar la pertinencia de las otras posibles soluciones.

<sup>66</sup> A. TRUYOL Y SERRA, “Théorie du droit international public”, *Recueil des Cours de l’Académie de Droit International de La Haye*, 1981-IV, tomo 183, pp. 142-143. J. A., CARRILLO SALCEDO, “Influencia de los derechos humanos en la consolidación de los principios generales del Derecho Internacional”, en *El Derecho Internacional: Normas, Hechos y Valores. Liber Amicorum José Antonio Pastor Ridruejo*, Madrid, Universidad Complutense, 2005, p. 82. D. GARCÍA SAN JOSÉ, *El Derecho Internacional Postcontemporáneo*, op. cit., pp. 226 y ss.

<sup>67</sup> Es el caso, por ejemplo, de la investigación con células embrionarias de origen humano. D. GARCÍA SAN JOSÉ, “Investigación embrionaria humana y patentes biológicas: la necesidad de contar con un corpus iuris gentium europaeum”. *Marco Jurídico Europeo relativo a la transferencia y reprogramación celular*, AA.VV., Cizur menor, Thomson Aranzadi, 2012, pp. 285 - 301.

<sup>68</sup> Confer a nivel internacional *The Parentage/Surrogacy Project: An Updating Note*. Preliminary Document No. 3ª, Febrero 2015, preparado en el seno de la Conferencia de Derecho Internacional Privado de La Haya (disponible en [www.hcch.net](http://www.hcch.net)). El último informe, de octubre de 2020 puede consultarse en <https://assets.hcch.net/docs/a6aa2fd2-5aef-44fa-8088-514e93ae251d.pdf>

<sup>69</sup> Así, por ejemplo, un Estado puede no permitir la gestación por subrogación en su territorio al tiempo que permite

**24.** La Conferencia de La Haya sobre Derecho internacional Privado desde 2012 lleva ocupándose del tema, interesándose en especial por los problemas relativos a la filiación de los hijos tenidos mediante esta práctica. En uno de sus primeros informes publicado en 2016, concluía que la dificultad de la cuestión y la variedad de perspectivas por parte de los Estados hacen inviable cualquier avance en esta materia y, a medio plazo, no parece previsible cambio alguno en este *impasse*<sup>70</sup>. Incluso si se lograra concluir un tratado internacional en una materia como la IA, se corre el riesgo de que su eficacia se viera altamente condicionada bien por su escaso número de ratificaciones, bien por el carácter inconcreto del alcance de sus obligaciones debido al abuso de las reservas y declaraciones interpretativas en caso de que se permitiera.

**25.** El Convenio del Consejo de Europa nº 211 relativo a la falsificación de productos médicos y delitos similares que supongan una amenaza para la salud pública (Convenio MEDICRIME), en vigor desde el 1 de enero de 2016<sup>71</sup> es un buen ejemplo de esta posibilidad. El Convenio MEDICRIME es un “non self-executing treaty on criminal-law”; esto es, un tratado internacional de carácter no auto ejecutorio que, por su contenido material, requiere de los Estados que sean partes en el mismo, la adopción de medidas en su Derecho interno tendentes a considerar como delitos penales la falsificación de medicamentos y de productos similares, sin perjuicio de incluir otras disposiciones de carácter procesal, por ejemplo, en relación con las víctimas de estas conductas criminales. En este sentido, es un ejemplo clásico de instrumento convencional regido con carácter supletorio a lo convenido expresamente por los Estados parte en el mismo por lo dispuesto en la Convención de Viena de 23 de mayo de 1969 sobre Derecho de tratados<sup>72</sup>.

**26.** Adoptado en el seno de una Organización Internacional regional, el Consejo de Europa, está abierto a todos los Estados de la comunidad internacional, por lo tanto, puede considerarse como un tratado multilateral sobre una cuestión de interés general para todos los Estados. Cuenta con un mecanismo de control del cumplimiento de obligaciones asumidas por los Estados parte muy blando<sup>73</sup>: un órgano intergubernamental, el *Comité de las Partes* que está previsto en el artículo 23 y siguientes –por su composición, en principio más politizado que un órgano integrado de carácter judicial o cuasi judicial– es el encargado de velar por la correcta ejecución de los compromisos asumidos por los Estados parte en este Convenio, no concretándose cómo desarrollará esta función de control, por ejemplo, exigiendo memorias o informes anuales a los Estados parte en el mismo.

**27.** Al adoptar un mecanismo de supervisión blando, el Convenio MEDICRIME sigue la línea tradicional de los tratados multilaterales de carácter universal en los que la aspiración de vincular en los mismos al mayor número posible de países conlleva el precio de rebajar el contenido de los compromisos asumidos y el control de su cumplimiento<sup>74</sup>. El carácter blando de este mecanismo se reafirma en el artículo 31 que prevé un mecanismo no vinculante de solución de controversias entre los Estados parte en el Convenio MEDICRIME en relación con su aplicación<sup>75</sup>, por ejemplo, para el supuesto de que uno de los Estados pretenda su terminación o la suspensión de su aplicación como consecuencia del incum-

---

regularizar a sus nacionales las relaciones parentales-filiales derivadas de una gestación por subrogación realizada en el extranjero.

<sup>70</sup> Confer The Hague Conference on Private International Law, Report of the February 2016 Meeting Experts’ Group on Parentage / Surrogacy, <https://assets.hcch.net/docs/t92c95b5-4364-4461-bb04-2382e3c0d50d.pdf> Visitado el 1 de junio de 2018. En la actualidad se han presentado anualmente informes respecto de esta cuestión siendo el más reciente el de octubre de 2020.

<sup>71</sup> Convenio MEDICRIME, CETS No. 211. BOE núm. 286, de 30 de noviembre de 2015.

<sup>72</sup> BOE núm. 142, de 13 de junio de 1980.

<sup>73</sup> Capítulo VIII, artículos 23 a 25 del Convenio MEDICRIME.

<sup>74</sup> El Acuerdo de París de 12 de diciembre de 2015 sobre el Cambio climático es un ejemplo de esta práctica internacional. Vid. Instrumento de ratificación por el Reino de España de este instrumento publicado en BOE núm. 28 de 2 de febrero de 2017.

<sup>75</sup> Artículo 31 del Convenio MEDICRIME. *Solución amistosa*. El Comité de las Partes, en estrecha colaboración con el Comité Europeo para Problemas Penales (CEPC) y los demás comités intergubernamentales o científicos competentes del Consejo de Europa, supervisará la aplicación del presente Convenio y facilitará en su caso la solución amistosa de toda dificultad que pudiera surgir en su aplicación.

plimiento de las obligaciones asumidas por otro Estado parte<sup>76</sup>. Pese a ello, el Convenio MEDICRIME es un fracaso pues si problemas globales requieren soluciones globales<sup>77</sup>, un tratado que a fecha de 1 de diciembre de 2020 compromete sólo a dieciocho Estados<sup>78</sup> de los cerca de doscientos Estados que integran la comunidad internacional, no admite otro calificativo en mi opinión

**28.** La experiencia demuestra que los tratados internacionales adoptados en el seno del Consejo de Europa no tienen una masiva adhesión de Estados fuera del contexto de la propia organización regional. No alcanzan el grado de tratados multilaterales de alcance global, ni directa ni indirectamente; esto es, por el número de Estados que son parte en los mismos o en la medida en que sirven de inspiración para otros tratados de ámbito universal. Todavía se sigue viendo con recelo en diversas partes del mundo las propuestas normativas que llegan de Europa como reflejo de aspiraciones neocoloniales, irrespetuosas con la diversidad cultural de la sociedad mundial. Este sentimiento de rechazo en parte se explica por el hecho de que algunos términos utilizados en este Convenio no habrían sido consensuados con actores no participantes en el mismo.<sup>79</sup>

## 2. Una solución de soft law: ¿códigos de conducta jurídicamente no vinculantes?

**29.** Los códigos de conducta, como fenómeno reciente asociado a la globalización económica, son una política formulada por escrito o una declaración de principios destinada a servir de base al compromiso de conducta de una empresa particular<sup>80</sup>. En virtud de la misma, una empresa, con frecuencia se trata de una multinacional, se compromete a respetar unas determinadas prácticas en materia de derechos humanos, de estándares laborales, de protección del medio ambiente y de no fomentar la corrupción de las autoridades públicas en la cadena de producción de un producto que comercialice, lo cual implica que los efectos de dicho compromiso se extienden a los centros filiales que participan en la cadena de producción, generalmente, ubicados en países en desarrollo. Al enunciar estándares mínimos, los códigos de conducta han contribuido a consolidar un núcleo de derechos humanos y de estándares laborales y medioambientales en el elevado número de países donde no existen medios de aplicación junto con un reforzamiento de las buenas prácticas de los gobiernos, y son un claro ejemplo de la incidencia de la globalización en la interacción de los actores y sujetos internacionales, Estados y Organizaciones Internacionales universales (OIT, ONU), a través del conocido como Pacto Mundial (Global Compact en inglés).

**30.** Respondiendo a una iniciativa personal del Secretario General de la Organización de las Naciones Unidas lanzada en Davos, Suiza, el 31 de enero de 1999, el Pacto Mundial se articuló a nivel institucional en la Oficina del Pacto Mundial<sup>81</sup>, la Oficina del Alto Comisionado para los Derechos Humanos, la Organización Internacional del Trabajo, el Programa de las Naciones Unidas para el Desarrollo y el Programa de las Naciones Unidas para el Medio Ambiente<sup>82</sup>. A nivel de contenidos, el

<sup>76</sup> De conformidad con el artículo 60 de la Convención de Viena de 23 de mayo 1969 sobre Derecho de Tratados. BOE núm. 142, de 13 de junio de 1980.

<sup>77</sup> Así se entiende desde el Derecho internacional público en relación con cuestiones biojurídicas: *vid.* D. GARCÍA SAN JOSÉ, “Epigenética y gestación por sustitución: más razones a favor de una regulación internacional para un negocio global”, *op. cit.*, y, en el ámbito específico de la falsificación de medicamentos: *vid.* Ó. ALARCÓN-JIMÉNEZ, “The MEDICRIME Convention. Fighting against counterfeit medicine”, *Eurohealth*, vol. 21, 4, 2015, p. 26.

<sup>78</sup> [https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/211/signatures?p\\_auth=0c4cwzL9](https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/211/signatures?p_auth=0c4cwzL9)

<sup>79</sup> Así, por ejemplo, el concepto de falsificación recogido en el Convenio MEDICRIME no coincide con el establecido por la Organización Mundial del Comercio ni por la Organización Mundial de la Salud. *Vid.* F. CVETANOVSKI *et al.*, “Counterfeiting of medicines as an infringement of the intellectual property rights”, *Macedonian pharmaceutical bulletin*, vol. 62, 1, 2016, p. 85.

<sup>80</sup> *Examen general de los acontecimientos globales y de las actividades de la Oficina en relación con los códigos de conducta, el etiquetado social y otras iniciativas del sector privado que revisten importancia para las cuestiones laborales*. Documento de la OIT, GB.273/WP/SDL/1 (rev. 1), Consejo de Administración, 273ª reunión, Ginebra, noviembre de 1998, párr. 68.

<sup>81</sup> Cuya página web es <https://www.globalcompact.org>

<sup>82</sup> Asimismo, en el Pacto participan los agentes sociales pertinentes: los gobiernos, que definieron los principios en los que se basa esta iniciativa; las empresas, cuyas acciones se pretende influenciar, los trabajadores, en cuyas manos se concreta el proceso

Pacto Mundial tiene su base en el apoyo de los actores económicos, en particular, el mundo empresarial a un conjunto de valores fundamentales en cuatro ámbitos interconectados en la sociedad económica globalizada: los derechos humanos, el medio ambiente, las normas laborales y, recientemente, las prácticas de buen gobierno a través del principio contra la corrupción<sup>83</sup>. En el contexto del Pacto Mundial hay que señalar el *Project Breakthrough*<sup>84</sup> como una herramienta no institucionalizada para implicar al mundo empresarial en los nuevos objetivos de Desarrollo Sostenible previstos en la Agenda 2030. El potencial de esta alianza a través de la aplicación de la IA y la Oficina del Pacto Mundial es indudable considerando el número de empresas que en ella participan y las sedes nacionales del Pacto Mundial en más de 160 países. Los distintos modos en los que la IA eventualmente contribuirá a hacer realidad dichos objetivos de Desarrollo Sostenible aparecen profusamente descritos en el documento *Artificial Intelligence. A more intelligent future*.<sup>85</sup>

31. El Pacto Mundial no consiste en un instrumento regulador, que vigila, impone, o evalúa la conducta o las acciones de las empresas sino un medio de canalizar la incentiva privada previendo como única sanción para las empresas participantes en caso de incumplimiento de lo acordado el salir de la lista publicada en el sitio web del pacto mundial<sup>86</sup>. Sin embargo, la no exigencia jurídica del cumplimiento de los compromisos voluntariamente asumidos por las empresas multinacionales en el Pacto Mundial resulta difícil de aceptar y casa mal con las exigencias que desde los distintos foros institucionalizados se viene haciendo para que la IA sea fiable, robusta, ética y respetuosa de los derechos humanos<sup>87</sup>.

### 3. Una solución intermedia: ¿el activismo judicial de tribunales internacionales previstos en tratados internacionales sin necesidad de modificar éstos?

32. Esta solución toma como referente la aproximación seguida por el Tribunal Europeo de Derechos Humanos a los desafíos que la asimétrica regulación de la gestación por sustitución a nivel europeo planteaba para los derechos y libertades fundamentales recogidos en el Convenio de Derechos Humanos de 1950: sentencias del Tribunal Europeo de Derechos Humanos de 26 de junio de 2014, *casos Menesson y Labasse c. Francia*, y de 27 de enero de 2015 (Sala) y de 24 de enero de 2017 (Gran Sala), ambas en el *caso Paradiso y Campanelli c. Italia*, al asumir el concepto de “familia de facto” y defender el interés superior del menor. En dicha jurisprudencia ha jugado un papel esencial el principio del interés superior del menor pese al hecho de que en el ámbito específico de la gestación subrogada, no resulta fácil determinar dónde se halla el interés superior del menor: ¿en mantenerlo con los padres de intención que lo han cuidado desde sus primeras horas de vida o con su madre biológica que, en algunos casos, no desarrolló lazos afectivos con el ser humano que concebía al considerarlo como el objeto de un negocio jurídico?

33. Al aplicar la *teoría del justo equilibrio de los distintos intereses en juego*, los principios desarrollados por el Tribunal Europeo en esta cuestión constituyen un marco inspirador para el legislador internacional a la hora de elaborar un eventual convenio de protección de derechos humanos *ad hoc* para

---

de mundialización, las organizaciones de la sociedad civil, que representan a la más amplia comunidad de interesados; y las naciones Unidas, que son verdaderamente el único foro político mundial, una voz autorizada, con poder convocador y catalizador.

<sup>83</sup> El principio 10º del Pacto Mundial, incluido el 24 de junio de 2004 establece que “los negocios deberán actuar en contra de la corrupción en todas sus formas, incluyendo la extorsión y el soborno”.

<sup>84</sup> <http://breakthrough.unglobalcompact.org/> (visitado el 1 de diciembre de 2020)

<sup>85</sup> Disponible en internet en [https://breakthrough.unglobalcompact.org/site/assets/files/1454/hhw-16-0017-d\\_c\\_artificial\\_intelligence.pdf](https://breakthrough.unglobalcompact.org/site/assets/files/1454/hhw-16-0017-d_c_artificial_intelligence.pdf) (visitado el 1 de diciembre de 2020)

<sup>86</sup> Se trata de la “movilización de la vergüenza” una sanción moral bien conocida en el ámbito de los mecanismos extraconvencionales de protección de derechos humanos de Naciones Unidas que, sin embargo, la práctica parece confirmar su limitada eficacia.

<sup>87</sup> Pese a sus críticos, el Pacto Mundial ofrece una sistematización necesaria para salvaguardar los contenidos mínimos (el “núcleo duro”) que en materia laboral, medioambiental y de protección de derechos humanos deben ser respetados en cualquier parte del mundo donde quiera que tenga lugar una actividad productiva.

la gestación por sustitución. De hecho, el Comité de Ministros del Consejo de Europa, en su referida Recomendación CM/Rec (2020)1, sobre los impactos de los sistemas algorítmicos en los derechos humanos, recuerda entre sus Principios de aplicación general, la responsabilidad de respetar los derechos humanos en el ámbito de la IA<sup>88</sup>: “Considerando que los Estados miembros del Consejo de Europa se han comprometido a garantizar los derechos y libertades consagrados en el Convenio para la Protección de los Derechos Humanos y las Libertades Fundamentales (ETS No. 5, “el Convenio”) a todas las personas dentro de su jurisdicción y que este compromiso se mantiene a lo largo de los continuos procesos de avance tecnológico y transformación digital que están viviendo las sociedades europeas.”<sup>89</sup>; “Recordando la obligación de los Estados miembros en virtud de la Convención de abstenerse de cometer violaciones de derechos humanos, incluso a través de sistemas algorítmicos, ya sea que se utilicen por sí mismos o como resultado de sus acciones, y su obligación de establecer marcos legislativos, regulatorios y de supervisión efectivos y predecibles que prevenir, detectar, prohibir y reparar las violaciones de derechos humanos, ya sean de origen público o privado y que afecten a las relaciones entre empresas, entre empresas y consumidores o entre empresas y otras personas y grupos afectados.”<sup>90</sup>

**34.** Sin embargo, el velar por posibles violaciones causadas durante el ciclo de vida de la IA sobre los derechos y libertades protegidos en un tratado internacional que, como el Convenio Europeo de Derechos Humanos, cuenta con un mecanismo de supervisión de las obligaciones asumidas por los Estados parte en el mismo en el sentido de respetar y hacer respetar dichos derechos y libertades, puede no ser la solución más idónea para la IA. Como recuerda la Comisión en su Libro Blanco sobre la IA, “los desarrolladores e implementadores de la IA ya están sujetos a la legislación europea en materia de derechos fundamentales (la protección de datos, la privacidad o la no discriminación, entre otras), de protección de los consumidores y normas sobre la seguridad de los productos y la responsabilidad civil. *Sin embargo, algunas características específicas de la IA (como la opacidad) pueden hacer que la aplicación y ejecución de dicha legislación sea más compleja.*”<sup>91</sup> (La cursiva es añadida).

**35.** Esa opacidad no sólo es atribuible al secretismo o falta de transparencia de los fabricantes y diseñadores de IA, sino que -como hemos tenido ocasión de señalar páginas atrás al referirnos a los factores que dificultan su regulación internacional- se debe también a las propias características de la IA, pues como reconoce la Comisión Europea: “un determinado riesgo para los derechos fundamentales puede existir en el momento de la comercialización de un producto que lleva incorporada IA, o surgir después, como resultado de la actualización de los programas informáticos y del aprendizaje automático de dicho producto una vez esté en funcionamiento.”<sup>92</sup> Por ello, concluye la Comisión afirmando que “no sólo es necesaria la adaptación de la legislación vigente relativa a la protección de los derechos fundamentales y los derechos de los consumidores, sino que es necesaria nueva legislación específica sobre IA.”<sup>93</sup>

---

<sup>88</sup> Principio 1.1. de la Recomendación CM/Rec (2020)1 del Comité de Ministros del Consejo de Europa: “Los actores del sector privado involucrados en el diseño, desarrollo, venta, despliegue, implementación y servicio de sistemas algorítmicos, ya sea en la esfera pública o privada, deben ejercer la debida diligencia en el respeto de los derechos humanos. Tienen la responsabilidad de respetar los derechos humanos y las libertades fundamentales reconocidos internacionalmente de sus clientes y de otras partes afectadas por sus actividades. Esta responsabilidad existe independientemente de la capacidad o voluntad de los Estados para cumplir con sus obligaciones en materia de derechos humanos. Como parte del cumplimiento de esta responsabilidad, los actores del sector privado deben tomar medidas continuas, proactivas y reactivas para garantizar que no causen ni contribuyan a violaciones de los derechos humanos y que sus acciones, incluidos sus procesos innovadores, respeten los derechos humanos. También deben ser conscientes de su responsabilidad hacia la sociedad y los valores de la sociedad democrática. Deben documentarse los esfuerzos para garantizar el cumplimiento de los derechos humanos.”

<sup>89</sup> Párrafo 2 del Preámbulo de la Recomendación CM/Rec (2020)1 del Comité de Ministros del Consejo de Europa, sobre los impactos de los sistemas algorítmicos en los derechos humanos.

<sup>90</sup> Párrafo 10 del Preámbulo de la Recomendación CM/Rec (2020)1 del Comité de Ministros del Consejo de Europa, sobre los impactos de los sistemas algorítmicos en los derechos humanos.

<sup>91</sup> *Libro Blanco de la Comisión Europea sobre la Inteligencia Artificial: un enfoque europea orientado hacia la excelencia y la confianza*, COM (2020) 65 final, de 19 de febrero de 2020, *op. cit.*, p. 12.

<sup>92</sup> *Libro Blanco de la Comisión Europea sobre la Inteligencia Artificial*, *op. cit.*, pp. 17-18.

<sup>93</sup> *Libro Blanco de la Comisión Europea sobre la Inteligencia Artificial*, *op. cit.*, p. 20.

### III. La vía de estandarización del Global Law (principios informadores) como catalizador de una respuesta de la comunidad mundial sobre una base consensual colectiva

36. Desde hace años creo que es posible identificar la presencia de una incipiente comunidad mundial que reivindica un paradigma de Derecho internacional post contemporáneo (Global Law) integrado por el conjunto de obligaciones asumidas por los sujetos y actores internacionales en virtud de su consentimiento o al margen del mismo. Un paradigma de Derecho internacional postcontemporáneo que, conforme a sus valores universales expresados en la Carta de las Naciones Unidas, aspira a regular las relaciones de coexistencia, cooperación e interdependencia de los integrantes de esa incipiente comunidad global, junto al bien común de la propia comunidad en su conjunto a través de una gestión responsable, y por ello solidaria, de las competencias reconocidas a cada integrante de la misma<sup>94</sup>. De ahí la soberanía responsable como piedra angular de esta estructura comunitaria y, en consecuencia, que se prefieran los medios institucionalizados, representados en las Organizaciones Internacionales, para la gestión responsable de intereses generales. Se caracterizaría esta comunidad mundial por la aspiración de hacer primar los intereses generales sobre los particulares de cada uno de ellos y comunes a todos o a una gran parte de los mismos. La idea de comunidad mundial toma a los Estados y a las Organizaciones Internacionales como sus pilares centrales, reflejando el adjetivo de “mundial” los vínculos de solidaridad que estos establecen entre sí y que son consecuencia directa de las relaciones de interdependencia que les unen: en cuanto a los problemas y a sus posibles soluciones.

37. Esa Comunidad mundial reclama un concepto de Global Law -coincidente con el defendido por la profesora IGLESIAS VÁZQUEZ<sup>95</sup>- cuya vía de realización idónea es la de los principios generales del Derecho internacional general, entendidos como los *prima principia* o principios constitucionales del orden jurídico internacional. Serían, como dejó escrito mi maestro, el profesor CARRILLO SALCEDO, “los axiomas o postulados éticos que deben inspirar todo el orden jurídico, bien porque el Derecho se remite a ellos como término o canon de aspiración o porque forman parte integrante del ordenamiento.”<sup>96</sup> Los dos principios estructurales en esta comunidad mundial aparecen enunciados en la Carta de las Naciones Unidas: por una parte, está el principio de la igualdad soberana de los Estados y por otra parte, el principio de la dignidad intrínseca del ser humano<sup>97</sup>. Ahora bien, ¿podemos identificar principios informadores de este Global Law aplicables a la IA?

#### 1. Los principios de precaución y de gestión responsable del riesgo en el ciclo de vida de los sistemas de IA

38. Sorprende que ninguno de los instrumentos analizados en el contexto de la Unión Europea se haya invocado el principio de precaución en el ámbito de la IA, pese a ser un principio asumido por esta organización desde hace años para hacer frente a los desafíos tecnológicos<sup>98</sup>, en especial, para afrontar los riesgos de la nanotecnología y los nanomateriales. En efecto, desde los primeros documentos europeos consagrados a las nanotecnologías se ha venido poniendo énfasis en la necesidad de respetar

<sup>94</sup> D. GARCÍA SAN JOSÉ, *El Derecho Internacional post contemporáneo*, op. cit., p. 319.

<sup>95</sup> M. IGLESIAS VÁZQUEZ, “Globalización, Globalización Jurídica, *Global Law* y Derecho Internacional Privado”, Cuadernos de Derecho Transnacional, vol. 9, 1, 2017, p. 234.

<sup>96</sup> J. A. CARRILLO SALCEDO, *Soberanía de los Estados y Derechos Humanos en Derecho internacional contemporáneo*, 2ª edición, Madrid, Tecnos, 2001, p. 147.

<sup>97</sup> Vid. F. MARIÑO MENÉNDEZ, “Los límites de la noción de orden público internacional”, en *Soberanía del Estado y Derecho Internacional. Homenaje al Profesor Juan Antonio Carrillo Salcedo*, Universidades de Córdoba, Sevilla y Málaga, 2005, p. 825.

<sup>98</sup> COM (2008) 366 final, p. 9. Al respecto es de lectura obligada la Comunicación de la Comisión Europea sobre el recurso al principio de precaución, de 2 de febrero de 2000 COM (2000) 1 final. En resumen, las medidas derivadas del recurso al principio de precaución deben respetar los principios generales de la gestión de los riesgos y, por tanto, deben ser medidas proporcionadas, no discriminatorias y coherentes, y han de basarse en el análisis de los posibles beneficios y costes de la acción o de la falta de acción, así como en el estudio del desarrollo científico.



los principios éticos que caracterizan la sociedad europea<sup>99</sup>. De manera expresa, la Comisión Europea en doble cualidad de responsable de la elaboración de las políticas de la Unión y como organismo de financiación de la investigación en innovación, se ha comprometido a garantizar “que la I+D financiada por la Comunidad en el ámbito de las N+N (nanociencias y nanotecnologías) siga desarrollándose de manera responsable recurriendo, por ejemplo, a análisis éticos.”<sup>100</sup>

39. Es en este contexto en el que debe verse la Recomendación de la Comisión Europea de 7 de febrero de 2008 sobre un código de conducta para una investigación responsable en el campo de las nanociencias y las nanotecnologías. Este código de conducta, es de carácter voluntario para las partes interesadas<sup>101</sup> pero hay una expresa llamada a los Estados Miembros de la Unión Europea para que fomenten a todos los niveles de su jurisdicción la adopción voluntaria del mismo<sup>102</sup> y se aspira a utilizarlo en las relaciones comerciales a escala internacional<sup>103</sup>. Por esta razón aparece anexo a una Recomendación de la Comisión Europea, típico acto comunitario sin carácter vinculante y a falta de una legislación en la Unión Europea específica sobre esta materia, *de facto* ejerce un valor normativo –aún de soft law- complementando la normativa existente en la Unión Europea<sup>104</sup>. Responde a la idea eje de que la investigación sobre las nanociencias y las nanotecnologías en toda la Unión Europea se lleve a cabo dentro de un marco seguro, ético y eficaz, pues ha sido deseo expreso de las instituciones europeas integrar desde el principio la dimensión de la seguridad en esta investigación<sup>105</sup>. Se basa el código en siete principios generales: sentido<sup>106</sup>, sostenibilidad<sup>107</sup>, precaución<sup>108</sup>, participación<sup>109</sup>, excelencia<sup>110</sup>, innovación<sup>111</sup> y rendición de cuentas<sup>112</sup>, que sustentan unas directrices sobre las medidas que deben tomar las partes interesadas.

<sup>99</sup> Principios como, por ejemplo, el respeto de la dignidad, el principio de la autonomía del individuo, el principio de justicia y de beneficencia, el principio de libertad de investigación y el principio de proporcionalidad. Comunicación de la Comisión Europea: “Hacia una estrategia europea para las nanotecnologías”, COM (2004) 338 final, de 12.5.2004, p. 22. Se insistía desde la Comisión Europea, además, en que algunas aplicaciones, como por ejemplo los sensores miniaturizados, podían tener implicaciones específicas para la protección de la privacidad y de los datos personales. *Ibidem*.

<sup>100</sup> COM (2005) 243, “Nanociencias y nanotecnologías: un plan de acción para Europa 2005-2009”, p. 9.

<sup>101</sup> Según el artículo 2. c) de la Recomendación de la Comisión C(2008) 424, éstas son los Estados miembros, los empleadores, las entidades financieras de la investigación, los investigadores y, más en general, todas las personas y organizaciones de la sociedad civil que participan, están involucradas o están interesadas en la investigación sobre N+N.

<sup>102</sup> En todos los apartados de la Recomendación de la Comisión y, expresamente, en el cuarto.

<sup>103</sup> *Vid.* Punto 7.1. a) (p. 12) de *supra* citada COM (2005) 243

<sup>104</sup> Nótese que aparece publicado en el Diario Oficial de la Unión Europea con la nomenclatura “L” (legislación) y no “C” (actos preparatorios)

<sup>105</sup> Como se indica en el considerando 11 del preámbulo de la Recomendación de la Comisión que anexa este código de conducta: “(su) objetivo es promover una investigación integrada, segura y responsable en materia de nanociencias y nanotecnologías en Europa en beneficio de la sociedad en su conjunto.”

<sup>106</sup> Artículo 3.1.: “Las actividades de investigación sobre N+N deben ser comprensibles para el público. Deben respetar los derechos fundamentales y llevarse a cabo en interés del bienestar de las personas y de la sociedad en su diseño, ejecución, difusión y uso.”

<sup>107</sup> Artículo 3.2.: “Las actividades de investigación sobre N+N deben ser seguras y éticas, y contribuir al desarrollo sostenible, estando al servicio de los objetivos de sostenibilidad de la Comunidad y contribuyendo a los objetivos de Desarrollo del Milenio de las Naciones Unidas. No deben resultar perjudiciales ni crear una amenaza biológica, física o moral para las personas, los animales, los vegetales o el medio ambiente, ni en el presente ni en el futuro.”

<sup>108</sup> Artículo 3.3.: “La actividades de investigación sobre N+N deben llevarse a cabo de conformidad con el principio de precaución, previendo los posibles impactos de sus resultados sobre el medio ambiente, la salud y la seguridad y adoptando las precauciones debidas en proporción con el nivel de protección, al tiempo que se fomenta el progreso en beneficio de la sociedad y del medio ambiente.”

<sup>109</sup> Artículo 3.4.: “La gobernanza de las actividades de investigación sobre N+N debe guiarse por los principios de apertura a todas las partes interesadas, transparencia y respeto por el derecho legítimo de acceso a la información. Debe permitir que participen en los procesos de adopción de decisiones todas las partes interesadas o preocupadas por las actividades de investigación sobre N+N.”

<sup>110</sup> Artículo 3.5.: “Las actividades de investigación sobre N+N deben ajustarse a los mejores patrones científicos incluidos los que fundamentan la integridad de la investigación y los relacionados con las buenas prácticas de laboratorio.”

<sup>111</sup> Artículo 3.6.: “La gobernanza de las actividades de investigación sobre N+N debe fomentar la máxima creatividad, flexibilidad y capacidad de planificación para la innovación y el crecimiento.”

<sup>112</sup> Artículo 3.7.: “Los investigadores y las organizaciones de investigación deben rendir cuentas por los impactos sociales, ambientales y sobre la salud humana que su investigación sobre N+N pueda tener para las generaciones presentes y futuras.”

40. El Consejo de Europa ha asumido de manera explícita un principio de precaución en el ámbito de la IA en su Recomendación CM/Rec (2020) 1 del Comité de Ministros a los Estados miembros sobre los impactos de los sistemas algorítmicos en los derechos humanos<sup>113</sup>, y desarrollado en su apéndice que contiene las Directrices para abordar dichos impactos: “5. *Medidas de precaución*.” 5.1. Evaluación continua<sup>114</sup>, 5.2. Capacitación del personal<sup>115</sup>, 5.3. Evaluaciones del impacto en los derechos humanos<sup>116</sup>, 5.4. Seguimiento<sup>117</sup>.

41. A diferencia de la aproximación seguida por el Comité de Ministros del Consejo de Europa en su Recomendación sobre los impactos de los sistemas algorítmicos en los derechos humanos, la Comisión Europea ha apostado en su Libro Blanco sobre la IA, por una nueva regulación basada no en el principio de precaución sino en la gestión del riesgo (a mayor riesgo, se requerirá una regulación más estricta)<sup>118</sup>. En el plano teórico, la Comisión considera que una aplicación determinada de IA debe considerarse de riesgo elevado en función de lo que está en juego y considerando si tanto el sector como el uso previsto suponen riesgos significativos, en especial, desde la perspectiva de la protección de la seguridad, los derechos de los consumidores y los derechos fundamentales<sup>119</sup>.

42. Es una aproximación que, en principio, podría parecer similar a la seguida en el Derecho Internacional general respecto de actividades lícitas pero altamente peligrosas, como la energía nuclear, el transporte de hidrocarburos o el lanzamiento de artefactos al espacio. En virtud de esta regulación internacional, cuando estemos en estos sectores de actividades ultra peligrosas, se presume una responsabilidad objetiva o por riesgo siempre que se materialice un daño transfronterizo como resultante de una de dichas actividades. Diversos autores han venido defendiendo la tesis de que la obligación impuesta a los Estados en virtud del principio de prevención sería de comportamiento, la diligencia debida, y no

---

<sup>113</sup> Principio de precaución. Parágrafo 15 de la Recomendación CM/Rec (2020)1 del Comité de Ministros del Consejo de Europa: “(...) Los Estados miembros deben adoptar un enfoque de precaución y exigir el rechazo de determinados sistemas cuando su despliegue conlleve un alto riesgo de daños irreversibles o cuando, debido a su opacidad, el control y la supervisión humanos se vuelvan impracticables.”

<sup>114</sup> “Los actores del sector privado deben desarrollar y documentar procesos internos para asegurar que su diseño, desarrollo y despliegue continuo de sistemas algorítmicos se evalúen y prueben continuamente, no solo para detectar posibles errores técnicos sino también los posibles impactos legales, sociales y éticos que los sistemas puedan generar. Cuando la aplicación de sistemas algorítmicos conlleva altos riesgos para los derechos humanos, incluso a través de procesos de micro-focalización que pueden evitar o mitigar por sí mismos, los actores del sector privado deben tener la posibilidad de notificar y consultar a las autoridades supervisoras en todas las jurisdicciones relevantes para buscar asesoramiento y orientación sobre cómo gestionar estos riesgos, incluso mediante el rediseño de los servicios en cuestión.”

<sup>115</sup> “Todos los miembros relevantes del personal involucrados en las evaluaciones de impacto sobre los derechos humanos y en la revisión de los sistemas algorítmicos deben estar adecuadamente capacitados y ser conscientes de sus responsabilidades con respecto a los derechos humanos, incluidos, entre otros, los estándares de privacidad y protección de datos personales aplicables.”

<sup>116</sup> “Las evaluaciones del impacto en los derechos humanos deben realizarse de la manera más abierta posible y con la participación activa de las personas y grupos afectados. En caso de despliegue de sistemas algorítmicos de alto riesgo, los resultados de las evaluaciones de impacto en los derechos humanos en curso, las técnicas identificadas para la mitigación de riesgos y los procesos de monitoreo y revisión relevantes deben ponerse a disposición del público, sin perjuicio del secreto protegido por la ley. Cuando sea necesario hacer cumplir las reglas del secreto, cualquier información confidencial debe proporcionarse en un apéndice separado del informe de evaluación. Este apéndice debe ser accesible a las autoridades de supervisión pertinentes.”

<sup>117</sup> “Los actores del sector privado deben garantizar un seguimiento adecuado de sus evaluaciones de impacto en los derechos humanos tomando las medidas adecuadas en función de los hallazgos registrados durante todo el ciclo de vida del sistema algorítmico y monitoreando la efectividad de las respuestas, con miras a evitar o mitigar los efectos adversos y riesgos para el ejercicio de los derechos humanos. Las fallas identificadas deben resolverse lo antes posible y las actividades relacionadas deben suspenderse cuando corresponda. Esto requiere controles de control de calidad regulares y continuos y auditorías en tiempo real a lo largo de las etapas de diseño, prueba e implementación. Además, requiere una consulta regular con las personas afectadas para monitorear los sistemas algorítmicos en busca de impactos sobre los derechos humanos en contexto e *in situ* y corregir errores y daños de manera adecuada y oportuna. Esto es particularmente importante dado el riesgo de ciclos de retroalimentación que pueden exacerbar y afianzar los impactos adversos sobre los derechos humanos.”

<sup>118</sup> *Libro Blanco de la Comisión Europea sobre la Inteligencia Artificial, op. cit.*, pág. 21.

<sup>119</sup> *Ibidem*.

solo de resultado<sup>120</sup>. Esto es, la responsabilidad internacional del Estado que planifique o realice actividades no prohibidas por el Derecho Internacional aunque ecológicamente peligrosas, podría originarse con independencia de que se hubiera o no producido un daño ecológico a terceros Estados. En esta línea de pensamiento, otros autores, como Pierre-Marie DUPUY -con quien coincido-, han estimado que el principio de prevención impone dos obligaciones complementarias y diferentes, de comportamiento y de resultado, respectivamente: de un lado, la obligación de gestionar el riesgo de la manera en que lo haría “un buen gobierno” (principio de la debida diligencia); de otro lado, la obligación de no permitir que un Estado sufra en su territorio o bajo su control, un daño derivado de actividades realizadas bajo su soberanía, jurisdicción o control (principio de la no utilización perjudicial del territorio)<sup>121</sup>.

43. La idea de la soberanía responsable parece así estar cada vez más presente en este Derecho Global, aún en gestación, como muestra la labor codificadora de la Comisión de Derecho Internacional en relación con el principio de prevención al imponer al Estado en cuyo territorio, o bajo cuya jurisdicción o control en otros lugares, se planifican o realizan actividades no prohibidas por el Derecho Internacional, que entrañen el riesgo de causar, por sus consecuencias físicas, un daño transfronterizo sensible, dos obligaciones complementarias aunque diferentes: la obligación genérica de la debida diligencia *ex ante* (anterior al daño y que se traduce en la obligación de gestionar el riesgo de una manera apropiada) y *ex post facto* (una vez causado éste, minimizando sus efectos, esto es, la obligación de conducir estas actividades peligrosas de una forma respetuosa con las obligaciones internacionalmente asumidas frente a otros estados, de manera individual o colectiva)<sup>122</sup>.

44. Sin embargo, una lectura atenta al Libro Blanco de la Comisión sobre IA permite concluir la tesis contraria a esta idea que acaba de defenderse, en la medida en que la Comisión acepta como premisa de una eventual regulación europea, el hecho de que no toda utilización de la IA en sectores previamente señalados (sanidad, transporte, energía, etc.) implica necesariamente riesgos significativos. En concreto, para la Comisión, una aplicación de IA debe cumplir dos requisitos para considerarse que conlleva un riesgo elevado: de una parte, se exige que la aplicación se emplee en un sector en el que por las características o actividades que se llevan a cabo normalmente, es previsible que existan riesgos significativos (sanidad, transporte, energía y determinados ámbitos del sector público); de otra parte, se requiere además que la aplicación de IA en el sector en cuestión, se use de manera que pueda hacer surgir riesgos significativos (ej. El uso de aplicaciones de IA con efectos jurídicos o similares en los derechos de un particular o de una empresa; aplicaciones que presenten el riesgo de causar lesiones, muerte o daños materiales o inmateriales significativos)<sup>123</sup>. Para las aplicaciones de IA de alto riesgo la Comisión Europea señala que la nueva legislación que deba adoptarse ha de centrarse en el proceso de entrenamiento de los sistemas inteligente<sup>124</sup>.

<sup>120</sup> J. BARBOZA, “International liability for the injurious consequences of acts not prohibited by International Law and Protection of the Environment”, *Recueil des Cours de l’Académie de Droit International*, 1994-III, tomo 247, p. 336. En los propios términos de este autor: “es una obligación de diligencia debida y en modo alguno una obligación de resultado, y su incumplimiento por el Estado de origen genera responsabilidad por hecho ilícito.” (La cursiva es añadida).

<sup>121</sup> *Proyecto de artículos sobre responsabilidad de los Estados por hechos internacionalmente ilícitos*, adoptado por la Comisión de Derecho Internacional de las Naciones Unidas en su 53º Período de sesiones (2001), párrafo 76, págs. 21 a 38 del Informe de la Comisión de Derecho Internacional en su 53º período de sesiones (de 23 de abril a 1 de junio y de 2 de julio a 10 de agosto de 2001), Asamblea General, Documentos Oficiales, 56º Período de sesiones, Suplemento nº 10 (A/56/19). P. -M. DUPUY, “Où en est le droit international de l’environnement à la fin du siècle?” *Revue Général de Droit International Public*, 1997, 4, p. 881.

<sup>122</sup> *Vid.* D. GARCÍA SAN JOSÉ, “La prevención de daños medioambientales transfronterizos resultantes de actividades peligrosas: recientes desarrollos en Derecho Internacional”, en *Soberanía del Estado y Derecho Internacional. Homenaje al Profesor Juan Antonio Carrillo Salcedo*, *op. cit.*, pp. 661-678.

<sup>123</sup> *Libro Blanco de la Comisión Europea sobre la Inteligencia Artificial*, *op. cit.*, p. 21.

<sup>124</sup> *Libro Blanco de la Comisión Europea sobre la Inteligencia Artificial*, *op. cit.*, p. 23: “deben adoptarse las medidas necesarias para garantizar que, en lo que se refiere a los datos utilizados para entrenar los sistemas de IA, se respeten los valores y normas de la UE, concretamente con relación a la seguridad y la legislación vigente para la protección de los derechos fundamentales.”

45. Recordando la fallida experiencia del Convenio MEDICRIME, el hecho de que incluso a nivel europeo se evidencie una discrepancia en cuanto a los principios informadores eventualmente aplicables a la regulación internacional de la IA, parece presagiar que será difícil consensuar esos mismos principios u otros a escala universal, conforme entren en juego otros sujetos y actores internacionales externos al contexto europeo. Es por esta razón por la que desde estas páginas se aboga por aprovechar la indiscutida validez universal de un principio informador, como es el principio de la dignidad intrínseca del ser humano, para releerlo en el contexto de los desafíos y riesgos que suscita la IA a la luz de otro emergente principio general del Derecho internacional: el principio de necesidad.

## 2. El principio de necesidad vinculado al principio de dignidad humana como un emergente principio general del Derecho Global frente a los riesgos de una expansión imparable y sin control de la IA

46. La legitimidad de este nuevo orden normativo global, aún en formación, se basa en la percepción de algunas amenazas como cuestiones de interés general de la Comunidad mundial (la Comunidad Internacional de Estados en su Conjunto y el resto de actores con una presencia significativa en relación con dichas cuestiones), sobre una base consensual colectiva que prevalecería sobre la base consensual individual y en la superación de la distinción entre normas duras y blandas como asunción de obligaciones de comportamiento y de resultado. En este sentido, sería posible una regulación internacional en el ámbito de la IA considerando los principios generales del Derecho Internacional tales como el principio de necesidad conectado con la dignidad humana.

47. Desde hace años se viene defendiendo la tesis de que cuando frente a amenazas globales para la seguridad humana sólo puedan adoptarse aproximaciones multilaterales, entonces, la posición unilateral de un simple Estado o de un reducido grupo de Estados no puede ser un obstáculo para la solución<sup>125</sup>. Este dato puede llevar a justificar la necesidad de defender el principio de necesidad como preferible al principio de precaución por varias razones: en primer lugar, el principio de necesidad vinculado a la dignidad humana reclamaría una aproximación multilateral –como sucede con los principales tratados internacionales de protección de derechos humanos ampliamente ratificados o las normas esenciales del Derecho internacional humanitario que gozan, igualmente, de una aceptación universal- mientras que el principio precautorio normalmente es invocado desde una posición unilateral; en segundo lugar, el principio precautorio es utilizado como una reacción preventiva de un daño cuando una amenaza es detectada y no existe una suficiente base científica, mientras que el principio de necesidad, estando estrechamente ligado a consideraciones elementales de humanidad, sugeriría una aproximación más amplia, esto es, no sólo como reacción sino como una estrategia vista como un *continuum*: antes, durante y después de que la concreta amenaza es percibida. Finalmente, el principio de necesidad conectado con la dignidad y la protección de los derechos fundamentales debería establecerse sobre la base de una nueva lectura de la soberanía en un sentido funcional, en virtud del cual, se toma en consideración los derechos pero especialmente los deberes de los Estados frente a otros Estados, sus propios ciudadanos y la Comunidad mundial<sup>126</sup>.

48. En consecuencia, el principio de necesidad como nuevo y emergente principio general del Derecho Global, vinculado al principio informador de la dignidad intrínseca del ser humano, de indis-

<sup>125</sup> En otras palabras, su posición unilateral no puede seguir siendo relevante en esta cuestión. D. GARCÍA SAN JOSÉ, “El principio de necesidad ligado a la seguridad medioambiental como instrumento idóneo para reforzar la acción internacional, normativa e institucional, en materia de cambio climático”, en R. GILES CARNERO, (Coord.): *Cambio Climático, Energía y Derecho Internacional: Perspectivas de Futuro*, Cizur Menor, Thomson Reuter Aranzadi, 2012, p. 81.

<sup>126</sup> En este sentido, mientras que el principio precautorio podría ser visto como una barrera de un Estado frente al resto del mundo, el principio de necesidad-seguridad medioambiental debería ser considerado, más bien, como el puente que vincula a ese Estado con la Comunidad Internacional en la que se integra. D. GARCÍA SAN JOSÉ, “El principio de necesidad ligado a la seguridad medioambiental como instrumento idóneo para reforzar la acción internacional, normativa e institucional, en materia de cambio climático”, *op. cit.*, p. 82.

tida validez en el Derecho Internacional contemporáneo desde la Carta de San Francisco de 1945, podría y, de hecho debería, ser considerado como la piedra angular de un *corpus iuris* internacional sobre IA del que se derivasen obligaciones jurídicas vinculantes para todos los actores que participan en el ciclo de vida de los sistemas de IA con independencia de su naturaleza pública o privada, o de su dimensión internacional, supranacional o nacional.

49. Es común a las distintas resoluciones y recomendaciones aprobadas en los últimos años sobre la conveniencia de una regulación internacional de la IA, la preferencia por la vía de principios generales dirigidos no sólo a los Estados sino también a otros actores de carácter público y privado relacionados con la IA. Esos principios tienen en común los valores subyacentes a los mismos, siendo el principal de ellos como ya se ha señalado al comienzo de este estudio, el *enfoque antropocéntrico de cualquier posible regulación de la IA*. Recuérdese que la Comisión europea publicó en 2019 la Comunicación COM (2019) 168 en virtud de la cual acogía favorablemente los siete requisitos esenciales contemplados en las directrices del Grupo de Expertos de Alto Nivel: acción y supervisión humanas; solidez técnica y seguridad; gestión de la privacidad y de los datos; transparencia; diversidad, no discriminación y equidad; bienestar social y medioambiental; y la rendición de cuentas.

50. Asimismo, nótese como la Recomendación de la UNESCO distingue entre valores y principios, precisando que: "...Mientras que el conjunto de valores que se enuncian a continuación inspira, por tanto, un comportamiento deseable y representa los cimientos de los principios, éstos, por su parte, revelan los valores subyacentes de manera más concreta, de modo que los valores pueden aplicarse más fácilmente en las declaraciones y acciones políticas."<sup>127</sup> La distinción entre valores y principios puede ser relevante, por ejemplo, en otros contextos como la Carta de Derechos Fundamentales de la Unión Europea. En las *Explicaciones* formuladas por el *Praesidium* de la Convención de 2000, y actualizadas por el *Praesidium* de la Convención Europea en 2007, se aclara la distinción entre derechos y principios, indicándose que los derechos subjetivos deberán respetarse, mientras que los principios deben observarse.<sup>128</sup> En todo caso, de conformidad con el párrafo 140 de dicha Recomendación del Comité de Ministros de Consejo de Europa, ésta ha de entenderse como un todo, los valores<sup>129</sup> y principios fundamentales<sup>130</sup> en ella recogidos "deben considerarse complementarios y relacionados entre sí".

#### IV. Conclusiones

51. Al finalizar estas páginas, algunas ideas pueden presentarse a modo de resumen de las principales conclusiones alcanzadas. En primer lugar, en los desafíos y riesgos que conlleva el empleo de la IA en cualquiera de sus posibles manifestaciones, es posible apreciar unas constantes y unos cambios con respecto a situaciones similares afrontadas en el pasado como consecuencia de los incesantes avances científicos y tecnológicos. La indeterminación de lo que debe entenderse por IA, que se evidencia a

<sup>127</sup> Anteproyecto de Recomendación sobre la Ética de la IA, SHS/BIO/AHEG-AI/2020/4 REV 2, Doc. Final, París, 7 de septiembre de 2020., párrafo 10.

<sup>128</sup> Publicadas junto al texto de la Carta en su versión de 2007 en el Diario Oficial de la Unión Europea (DO C 303, de 14 de diciembre de 2007, pp. 17 y ss.). Y se añade: "Los principios pueden aplicarse mediante actos legislativos y ejecutivos adoptados por la Unión en función de sus competencias y por los Estados miembros solamente en aplicación del Derecho de la Unión); por consiguiente, son importantes para los tribunales sólo cuando se trata de la interpretación o revisión de dichos actos pero no dan lugar a derechos inmediatos de acciones positivas de las Instituciones de la Unión o de las autoridades de los Estados miembros."

<sup>129</sup> Valores: 1) Respeto, protección y promoción de la dignidad humana, los derechos humanos y las libertades fundamentales (parás. 13-16); 2) Prosperidad del Medioambiente y los ecosistemas (parás. 17-18); 3) Respeto, protección y promoción de la diversidad y la inclusión (parás. 19-21); 4) Vivir en armonía y paz (parás. 22-24).

<sup>130</sup> Principios: 1) Proporcionalidad e inocuidad (parás. 25-26); 2) Seguridad y protección (pará. 27); 3) Equidad y no discriminación (parás. 28-30); 4) Sostenibilidad (pará. 31); 5) Privacidad (parás. 32-34); 6) Supervisión y decisión humanas (parás. 35-36); 7) Transparencia y explicación (parás. 37-41); 8) Responsabilidad y rendición de cuentas (parás. 42-45); 10) Gobernanza y colaboración adaptativas y de múltiples partes interesadas (parás. 46-47).

nivel institucional, junto con la opacidad que rodea este sector y la consiguiente imprevisibilidad de sus riesgos, dificultan pero no imposibilitan una eventual regulación internacional de la IA.

**52.** En segundo lugar, si algo podemos aprender del pasado reciente en el que se ha tenido que dar respuesta a similares desafíos para los individuos y las sociedades de todo el mundo, traídos de la mano de la investigación embrionaria humana, de los avances en nanomateriales, o de la gestación por sustitución, es que el Derecho internacional cuenta con herramientas válidas para hacer frente a los riesgos y desafíos que comporta el desarrollo exponencial de la IA y su imparable presencia en nuestra vida diaria. Examinadas las posibles opciones: la vía convencional uniformadora del *hard law* (tratado), la vía armonizadora del *soft law* (códigos de conducta) y/o la vía intermedia (interpretación jurisprudencial teleológica de los tratados en vigor), nos hemos decantado por otra distinta consistente en la vía de estandarización del *Global Law* (principios informadores) como catalizador de una respuesta de la comunidad mundial sobre una base consensual colectiva.

**53.** De este modo, y como tercera idea concluyente, aun considerando que pueden ser relevantes los principios de precaución y de gestión responsable del riesgo en el ciclo de vida de los sistemas de IA, el hecho de que no sean principios recogidos de manera idéntica en los distintos documentos analizados, nos lleva a pensar que puedan tener alguna dificultad para convertirse en el necesario catalizador de dicha respuesta de la comunidad mundial. Preferimos la solución que puede aportar el principio de necesidad vinculado al principio de dignidad humana como un emergente principio general del Derecho Global frente a los riesgos de una expansión imparable y sin control de la IA.

**54.** Es evidente, como avanzábamos al principio de estas páginas, que para que esta propuesta sea operativa, hace falta asumir que la aproximación como juristas a la IA debe hacerse desde las premisas de transversalidad, multidisciplinariedad, carácter integrador y no reduccionista, así como desde la superación de la distinción entre Derecho duro vinculante y Derecho blando no vinculante (*hard* y *soft law*). Si asumimos estas premisas ya se habrá logrado la mitad del empeño (saber lo que hay que hacer). Sólo restará la otra mitad (hacer lo que hay que hacer), siempre que al Derecho Global le acompañe la voluntad política de sus protagonistas.