



---

# Universidad de Valladolid

## Facultad de Derecho

### Grado en Derecho y Administración y Dirección de Empresas

## Consideraciones éticas y jurídicas en torno a la Inteligencia Artificial

Presentado por:

***ADRIAN FERRERAS FERNÁNDEZ***

Tutelado por:

***ENRIQUE MARCANO BUENAGA***

*Valladolid, 14 de Julio de 2023*

## **RESUMEN DEL TRABAJO**

En el presente trabajo se busca llevar a cabo una aproximación ética y jurídica de la inteligencia artificial (IA) en la actualidad. Y para ello se examina en primer lugar los conceptos básicos de la (IA) en el sector jurídico, distinguiendo diferentes aspectos relacionados con su definición y tipos, seguido de un análisis del marco regulatorio donde se estudia a nivel internacional, europeo y nacional, destacando recomendaciones éticas y directrices. Finalmente se abordan los principales principios éticos que rigen la IA profundizando en la equidad, diversidad, responsabilidad y privacidad.

## **PALABRAS CLAVE**

Inteligencia artificial, sistema experto, derechos fundamentales, ética, Unión Europea, principios éticos, algoritmo, machine learning

## **ABSTRACT**

This paper seeks to carry out an ethical and legal approach to artificial intelligence (AI) today. To this purpose, we analyze the key concepts of AI in the legal sector, distinguishing different aspects related to its definition and types, followed by a study of the regulatory framework where we examine it at international, European and national levels, highlighting ethical recommendations and guidelines. Finally, the main ethical principles governing AI are addressed, delving into equity, diversity, responsibility and privacy.

## **KEY WORDS**

Artificial intelligence, expert system, fundamental rights, ethics, European Union, ethical principles, algorithm, machine learning

## INDICE

1. INTRODUCCION.....	4
2. CONCEPTOS BASICOS DE LA INTELIGENCIA ARTIFICIAL EN EL SECTOR JURIDICO.....	5
2.1. Definición la inteligencia artificial.....	5
2.2. Tipos de inteligencia artificial.....	8
2.3. Los sistemas inteligentes y los sistemas expertos.....	10
3. MARCO REGULATORIO DE LA INTELIGENCIA ARTIFICIAL.....	12
3.1. La IA y su enfoque jurídico en el mundo.....	12
3.1.1. Unesco: Recomendación sobre la ética de la inteligencia artificial, 21 de noviembre de 2021.....	13
3.1.2. Visión general de los Principios de IA de la OCDE.....	14
3.2. La IA y su enfoque europeo.....	18
3.2.1. El grupo de expertos y las directrices éticas.....	18
3.2.2. Marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas.....	23
3.2.3. El desafío regulatorio.....	25
3.3. La estrategia nacional de inteligencia artificial.....	29
4. CONTENIDO Y PRINCIPIOS ETICOS.....	32
4.1. Principios éticos básicos.....	33
4.2. Dilemas éticos asociados a la IA.....	36
5. DIVERSIDAD, NO DISCRIMINACION Y EQUIDAD.....	38
5.1. Algoritmos, sesgos y discriminación.....	38
5.1.1. ¿Por qué discriminan los algoritmos? El diseño de los algoritmos y elaboración de perfiles.....	41
5.1.2. Probar la discriminación algorítmica.....	42
5.1.3. Detección de la discriminación por algoritmos.....	43
6. LA RESPONSABILIDAD Y LA IA.....	43
6.1. Concepto de responsabilidad.....	43
6.2. La responsabilidad y su relación con la IA.....	44
6.3. La gobernanza de la IA.....	46
7. PRIVACIDAD.....	48
7.1. Concepto de privacidad.....	48

7.2. Gestión de la privacidad y de los datos.....	49
8. CONCLUSIONES.....	51
9. BIBLIOGRAFIA.....	53

# 1. INTRODUCCION

La Inteligencia Artificial (IA) ha emergido como una poderosa fuerza impulsora de la innovación y el progreso en nuestra sociedad. A medida que las capacidades de la IA continúan expandiéndose, se plantea una pregunta fundamental: ¿cómo podemos garantizar que su desarrollo y aplicación estén en línea con los principios éticos y jurídicos que rigen nuestras vidas?

En este trabajo, exploraremos el enfoque normativo y ético que se le ha dado a la Inteligencia Artificial desde el ámbito internacional, europeo y nacional. La relevancia del tema radica en el hecho de que la IA se ha vuelto omnipresente en nuestra vida cotidiana, desde los sistemas de recomendación en línea hasta los asistentes virtuales y los vehículos autónomos. Estas tecnologías impactan cada vez más en nuestras decisiones, interacciones sociales y estructuras legales.

Sin embargo, a medida que la IA se convierte en una parte integral de nuestras vidas, surgen desafíos significativos que deben abordarse. Uno de los desafíos radica en la necesidad de establecer un marco ético sólido que guíe el desarrollo y la implementación de la IA. Esto implica considerar cuestiones como la transparencia algorítmica, la privacidad, la equidad y la responsabilidad en el uso de datos, así como los impactos sociales y económicos de estas tecnologías.

En este contexto, la realización de este trabajo se justifica por la necesidad imperante de explorar y comprender los aspectos éticos y jurídicos de la Inteligencia Artificial.

En las siguientes secciones, exploraremos los principios éticos que deben guiar el desarrollo de la IA, así como los desafíos jurídicos que se presentan en su implementación. Al hacerlo, esperamos fomentar una discusión informada y multidisciplinaria que permita aprovechar plenamente el potencial de la IA, al tiempo que mitigamos sus riesgos y aseguramos un futuro más ético y equitativo.

## 2. CONCEPTOS BASICOS DE LA INTELIGENCIA ARTIFICIAL EN EL SECTOR JURIDICO

### 2.1. Definición la inteligencia artificial

Llegar a establecer una definición precisa de lo que es una inteligencia artificial, de ahora en adelante IA, supone un trabajo arduo el cual no es el fin de este trabajo. Esta misma conclusión fue a la que llegó el Comité Económico y Social Europeo en su *“Dictamen del Comité Económico y Social Europeo sobre la «Inteligencia artificial: las consecuencias de la inteligencia artificial para el mercado único (digital), la producción, el consumo, el empleo y la sociedad”*<sup>1</sup> donde en su apartado 2.1 establece lo siguiente: *“No existe una definición establecida y unánimemente aceptada de la IA. La IA es un concepto que engloba muchas otras (sub)áreas como la informática cognitiva [...], el aprendizaje automático [...], la inteligencia aumentada [...], o la robótica con IA.”*

Por otro lado, pese a la dificultad que supone dicho trabajo, no han sido pocos los científicos y expertos que se han aventurado a establecer una definición de lo que es la IA.

Será el científico americano John McCarthy quien en uno de sus escritos realizados para «Falmouth Summer Research Project on AI» en 1955 acuñó la expresión de la IA, así como su definición, siendo esta “la ciencia e ingeniería de hacer maquinas que se comporten de una forma que llamaríamos inteligente si el humano tuviese ese comportamiento”<sup>2</sup>

Asi mismo, su compañero y cofundador del Laboratorio de Inteligencia Artificial del MIT (Instituto Tecnológico de Massachusetts), Marvin Minsky, definió

---

<sup>1</sup> Comité Económico y Social Europeo: *Inteligencia artificial: las consecuencias de la inteligencia artificial para el mercado único (digital), la producción, el consumo, el empleo y la sociedad*, Diario Oficial de la Unión Europea C 288/1, 31 de agosto de 2017, p. 3.

<sup>2</sup> M. J. DÍAZ, F. HERRERA TRIGUERO, O. CORDÓN GARCÍA: *Una visión actual de la inteligencia artificial: recorrido histórico, datos y aprendizaje confiabilidad y datos*, en Alfonso Peralta, Leopoldo Salvador Torres y Francisco Herrera (ed.): *El derecho y la Inteligencia Artificial*, Título II Aplicaciones técnicas de la inteligencia artificial, Universidad de Granada, Granada, 2022, p. 52.

la IA como “la ciencia de hacer que las máquinas hagan cosas que requerirían inteligencia si las hicieran las personas”.<sup>3</sup>

Para un mayor consenso se puede tener en cuenta la definición que señaló el propio Parlamento Europeo en uno de sus artículos publicados en su boletín de noticias definió a la IA de la siguiente forma: “La inteligencia artificial es la habilidad de una máquina de presentar las mismas capacidades que los seres humanos, como el razonamiento, el aprendizaje, la creatividad y la capacidad de planear”.<sup>4</sup>

En definitiva, el concepto de IA se puede resumir, según señala África María Morales Moreno<sup>5</sup>, en función de los objetivos que tenga esta:

- En primer lugar, uno donde la inteligencia artificial (IA) se enfoca en el estudio de los procesos cognitivos con el fin de adquirir un desarrollo teórico y sistematizado de las diversas actividades del intelecto.
- Y un segundo, donde el objetivo de la IA es el desarrollo de sistemas automáticos como dejaron constancia en sus definiciones John McCarthy y Marvin Minsky.

Centrándonos en el sector jurídico, se podría emplear la definición que establece la Real Academia Española (RAE): “Disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico.”<sup>6</sup>

O la definición que María Hernández Giménez empleó en su artículo “Inteligencia artificial y derecho penal”. Siendo esta definición aportada por la consultora Gartner, que en su informe “Cool Vendors en AI for Legal Affairs, 2017” ofrece una definición legal de la IA: “conjunto de todos aquellos sistemas

---

<sup>3</sup> M. DELGADO: *La inteligencia artificial. Realidad de un mito moderno*, Granada, Servicio de Publicaciones de la Universidad de Granada, 1996, p. 18.

<sup>4</sup> Parlamento Europeo (Bruselas). *¿Qué es la inteligencia artificial y como se usa?* [en línea]: en *Portada / Sociedad*: Noticias Parlamento Europeo. Creado el 8 de septiembre de 2020, actualizado el 23 de marzo de 2021, [Consulta: 28 mar. 2023]. Disponible en: <https://www.europarl.europa.eu/news/es/headlines/society/20200827STO85804/que-es-la-inteligencia-artificial-y-como-se-usa>, [consulta:28/03/2023]

<sup>5</sup> Á. M. MORALES MORENO: “Inteligencia artificial y derecho penal: primeras aproximaciones”, *Revista jurídica de Castilla y León* (53), p. 182

<sup>6</sup> REAL ACADEMIA ESPAÑOLA: *Diccionario de la lengua española*, 23.ª ed., [versión 23.6 en línea]. <<https://dle.rae.es>> [28/03/2023].

*expertos tales como el procesamiento de lenguaje natural, las redes neuronales profundas o el aprendizaje automático aplicados en tareas legales como el análisis de un contrato, la preparación de una estrategia procesal o la investigación legal, con la finalidad de automatizar la tarea y mejorar la toma de decisiones*<sup>7</sup>

De esta definición es preciso destacar el concepto de “sistemas expertos”, un término que se desarrollara en el punto siguiente.

Hoy en día estamos viviendo la conocida como “Cuarta Revolución Industrial” y al igual que las anteriores revoluciones, está encaminada a la automatización de ciertas actividades o tareas, afectando más precisamente a las actividades de carácter cognitivo o intelectual.

Con todo ello y de lo que se puede extraer de estas definiciones, la relación que va a existir entre la IA y el Derecho no va a ser unidireccional.

Resultado de esto, José Ignacio Solar Cayón en su artículo “La inteligencia artificial jurídica: nuevas herramientas y perspectivas metodológicas para el jurista” ya habla de una “inteligencia artificial en el Derecho” define este campo de interés de la siguiente manera: *“conjunto de herramientas de inteligencia artificial diseñadas y/o utilizadas para la automatización de las diversas tareas jurídicas-, de los desafíos e implicaciones epistemológicas que comporta el desarrollo o la adaptación de dichas herramientas a las peculiaridades del razonamiento jurídico, y de los efectos que su utilización puede tener tanto sobre los esquemas mentales y los métodos de trabajo tradicionales del jurista como sobre la práctica profesional del Derecho”*.<sup>8</sup>

---

<sup>7</sup> M. HERNÁNDEZ GIMÉNEZ: “Inteligencia artificial y derecho penal”, *Actualidad Jurídica Iberoamericana* N° 10 bis, junio 2019, p. 797.

<sup>8</sup> J. I. SOLAR CAYÓN (2020): “La inteligencia artificial jurídica: nuevas herramientas y perspectivas metodológicas para el jurista”, *Revus. Journal for Constitutional Theory and Philosophy of Law/Revija za ustavno teorijo in filozofijo prava*, (41). Disponible en: <https://journals.openedition.org/revus/6547>, [consulta 28/03/2023]

## 2.2. TIPOS DE INTELIGENCIA ARTIFICIAL.

Desde que se empezó a hablar del término “IA” en 1956 con John McCarthy hasta nuestros días, el desarrollo de la IA ha pasado por múltiples etapas y resultando de éstas en diferentes tipos de IA. Por este motivo en el presente punto se estudiarán las IA más importantes actualmente y cuál puede ser proyectada a ser empleada en el sector jurídico.

Inicialmente se pueden encontrar distintos tipos de IA si se estudian según su capacidad y su funcionalidad.

En función de su capacidad las IA se clasifican en:

- **Inteligencia artificial estrecha (ANI-*Artificial Narrow Intelligence*):** este tipo de IA también recibe el nombre de “IA débil”. La IA débil se refiere a sistemas que están diseñados para realizar tareas específicas y limitadas, pero no tienen la capacidad de razonar o comprender el mundo de la manera en que lo hacen los seres humanos. Es decir, están enfocados en solucionar problemas muy específicos para los cuales han sido diseñados.
- **Inteligencia artificial general (AGI-*Artificial General Intelligence*):** al contrario que la anterior, esta emplea el nombre de “IA fuerte”, aunque también se la puede llamar “IA verdadera”. es un tipo de inteligencia artificial que tendría una capacidad de razonamiento y comprensión similar a la humana. La IA fuerte no solo estaría diseñada para realizar tareas específicas, sino que tendría una amplia gama de habilidades cognitivas, como la comprensión del lenguaje natural, la capacidad de razonar y resolver problemas abstractos y la habilidad de aprender de la experiencia.

En segundo lugar, la clasificación de la IA según su funcionalidad fue establecida por Arend Hintze<sup>9</sup> dividiéndose esta en 4 tipos:

---

<sup>9</sup> A. HINTZE (2016): “Understanding the four types of AI, from reactive robots to self-aware beings”. *The Conversation*, 14. Disponible en: <https://theconversation.com/understanding-the-four-types-of-ai-from-reactive-robots-to-self-aware-beings-67616>, [consulta: 30/03/2023]

- **Máquinas reactivas:** se trata de sistema de inteligencia artificial que a partir de lo que perciba de su entorno desarrolla una toma de decisiones en tiempo real en función de esa percepción. Esto la lleva a ser una de las IAs más básica dado que solo actúa en función de su entorno y a partir de un rango de respuestas para unas peticiones concretas. A diferencia de los humanos esta IA carece de memoria por lo que no podrá emplear situaciones anteriores para la toma de una decisión actual. La IA más conocida de este tipo es Deep Blue creada por IBM en la década de los noventa.
- **Máquinas de memoria limitada:** otro tipo de IA que, a diferencia del anterior, tiene la capacidad de almacenar datos de experiencias pasadas y procesarlos, pero de una forma limitada y para un periodo de tiempo corto. Este tipo de IA están especializadas en resolver tareas específicas en las que se necesita una respuesta rápida y en tiempo real, como el control de procesos en tiempo real en una planta de producción.
- **Máquinas con una teoría de la mente:** para entender este modelo de IA es preciso definir lo que es la teoría de la mente, siendo esta la capacidad del ser humano de atribuir estados mentales a otros seres y con esa información poder comprender su comportamiento. Y este mismo proceso es el que seguiría este tipo de IA comprendiendo y modelando los estados mentales de los diferentes agentes y con dicha información poder predecir su comportamiento. Pueden ser útiles en aplicaciones en las que se requiere comprender el comportamiento humano en situaciones sociales y de interacción, como en la robótica social, la interacción hombre-máquina, el análisis de sentimientos y la toma de decisiones en entornos complejos. También pueden ser útiles en aplicaciones de seguridad y defensa, donde se necesita predecir el comportamiento de otros agentes para evitar amenazas o actos hostiles.
- **Máquinas con autoconciencia:** con los sistemas de computación con los que se dispone actualmente resultaría imposible poder desarrollar una programación la cual cuente con la capacidad de ser consciente de sí misma, así como comprender la existencia de individuos con emociones y pensamientos propios. Pese a todo no deja de ser un camino que muchos científicos se han dispuesto seguir.

Es importante señalar que dentro del sector jurídico no se busca un sistema de IA que se base en una inteligencia general, sino más bien una inteligencia específica. Este requerimiento se logra encontrar en los sistemas de IA débil que se han definido previamente en este apartado.

Relacionado con los sistemas de IA débil están los sistemas inteligentes. Un sistema inteligente es aquel capaz de emplear técnicas de inteligencia artificial para llevar a cabo tareas específicas. Puede incluir tanto sistemas de inteligencia artificial débil como sistemas de inteligencia artificial más avanzados, como redes neuronales y sistemas expertos. La principal diferencia es que un sistema inteligente puede tener una gama de habilidades más amplia.

Cabe recordar que previamente en la definición aportada por la consultora Gartner y mencionada en su artículo por María Hernández Giménez, ya se emplea este concepto de “sistema experto”. Concepto que se desarrollará en el próximo punto.

### **2.3. LOS SISTEMAS INTELIGENTES Y LOS SISTEMAS EXPERTOS**

Aunque el campo de desarrollo y estudio de las IA es demasiado extenso como para abarcarlo en un único punto, sin ánimos de ser exhaustivos, en este último punto se estudiará de forma pormenorizada los sistemas inteligentes y su subclase los sistemas expertos.

Los sistemas expertos suponen un área de investigación muy importante, no solo actualmente dentro del ámbito jurídico, sino que, además, la mayor parte de las restantes áreas de investigación tienen un cierto componente de sistemas expertos.

Partimos de la idea que con la que se concluyó en el apartado anterior, donde se explicó que los sistemas inteligentes y los sistemas expertos son dos tipos de sistemas informáticos que empleando la inteligencia artificial son capaces de desarrollar tareas específicas, es decir, se tratan de IA débiles.

Un sistema inteligente es un sistema informático capaz de realizar tareas como el reconocimiento de voz, la toma de decisiones, el procesamiento del lenguaje natural, el razonamiento y la resolución de problemas. Los sistemas inteligentes son capaces de aprender de forma autónoma y así como adaptar su capacidad de realizar tareas específicas a medida que se les proporciona más información.

Por otro lado, un sistema experto es una subclase de sistema inteligente con la que se busca simular el razonamiento humano en un campo de conocimiento especializado. Emplean una base de datos extensa y una serie de reglas de inferencia para el desempeño tareas específicas dentro del campo de conocimiento para el que han sido requeridas. Los sistemas expertos pueden ser utilizados para resolver problemas en campos como la medicina, la ingeniería, las finanzas y la logística.

De esta breve explicación se obtienen dos elementos clave dentro del sistema experto: la base de datos y las reglas de inferencia. En las siguientes líneas se explicará su importancia

Lo más habitual es que los datos y el lenguaje con el que trabaja y funciona el sistema experto haya sido el resultado de la colaboración de uno o varios expertos humanos en el tema que será objeto de análisis para el sistema experto. En un principio el componente humano será primordial para el correcto desempeño del sistema experto, ya que estos serán los que suministren a los ingenieros el conocimiento necesario sobre la materia para establecer una definida base de datos.

Por otro lado, el sistema experto puede llegar a desarrollar experiencia siguiendo los conocimientos aportados en la base de datos. Esto se consigue a través del motor de inferencia. Este componente busca obtener conclusiones a partir de los conocimientos que posee aplicándolos a los datos que se introducen.

Pero, qué pasaría si el conocimiento que se dispone es insuficiente para alcanzar una conclusión por parte del motor de inferencia. En esta situación entraría a funcionar el subsistema de adquisición de conocimiento que buscará

obtener dicho conocimiento y así poder llegar a una conclusión óptima por parte del motor de inferencia. Se puede dar la situación que sean los propios expertos humanos los que aporten directamente el conocimiento faltante.

Como se puede deducir de estos sistemas expertos es que todo el conocimiento y datos de los que dispondrán serán de tipo específico, de una materia o de un campo de conocimiento concreto sin poder llegar a alcanzar un conocimiento de tipo general.

En conclusión, según lo definen Enrique Castillo, José Manuel Gutiérrez, y Ali S. Hadi, *“Un sistema experto puede definirse como un sistema informático (hardware y software) que simula a los expertos humanos en un área de especialización dada”*<sup>10</sup>

### **3. MARCO REGULATORIO DE LA INTELIGENCIA ARTIFICIAL.**

#### **3.1. LA IA Y SU ENFOQUE JURIDICO EN EL MUNDO**

A nivel global es complicado encontrar una regulación que desarrolle de forma precisa el tratamiento de la IA. Hasta el momento han sido la mayoría de las empresas desarrolladoras de IA las que han provisto de ciertos límites al desarrollo y empleo de la IA, dictando normas sectoriales, directrices éticas, etc.

Pese a todo, se pueden encontrar dos actores importantes en el ámbito internacional, los cuales han publicado en sus respectivas plataformas guías y recomendaciones con las que pretenden establecer una base para el futuro desarrollo de los principios que regirán el desarrollo de la IA, al igual que transmitir una visión general de la misma. Estos dos actores serán por un lado la UNESCO y por el otro la OCDE.

---

<sup>10</sup> Enrique CASTILLO, José Manuel GUTIÉRREZ y Ali S HADI: Sistemas expertos y modelos de redes probabilísticas, Academia de Ingeniería, 1997, pp. 2-9.

### **3.1.1. Unesco: Recomendación sobre la ética de la inteligencia artificial, 21 de noviembre de 2021.**

Será la UNESCO la que el 21 de noviembre de 2021 emita la “Recomendación sobre la Ética de la Inteligencia Artificial”. El objetivo que persigue dicha recomendación se puede resumir en “hacer realidad las ventajas que la IA aporta a la sociedad y reducir los riesgos que conlleva”.<sup>11</sup>

De la misma forma esta recomendación busca estar presente a nivel global estableciendo un marco ético, así como señalar diferentes acciones políticas. A modo de resumen, se pueden establecer cuatro puntos principales que se pueden extraer de esta Recomendación:

- Por un lado, tenemos la protección de los datos: se trata de establecer medidas cuyo objetivo sea garantizar una adecuada protección de los datos personales. Entre las medidas recomendadas se encuentran la transparencia en el manejo de datos personales, el derecho de las personas a acceder y borrar sus propios registros de datos personales, la mejora de la protección de los datos y la defensa del derecho del individuo a controlar sus propios datos. Asimismo, se promueve el fortalecimiento de la capacidad de los organismos reguladores a nivel global para garantizar el cumplimiento de estas medidas.
- La prohibición de los marcadores sociales y la vigilancia masiva: ante la preocupación del uso malicioso que se le puede dar a la IA en el ámbito social, la Recomendación trata de mostrar la importancia de que los Estados sean los que establezcan las regulaciones necesarias ante este tipo tecnologías intrusivas y capaces de vulnerar los derechos y libertades fundamentales de las personas. La conclusión más importante a la que llega la Recomendación en este punto es la necesidad de que la responsabilidad final debe recaer en los seres humanos y que las

---

<sup>11</sup> UNESCO. (2021): *Ética de la inteligencia artificial*. Oficina de la UNESCO en Montevideo. Disponible en:

<https://es.unesco.org/fieldoffice/montevideo/EticaInteligenciaArtificial#:~:text=La%20Recomendaci%C3%B3n%20proh%C3%ADbe%20expl%C3%ADcitamente%20el,se%20utilizan%20de%20forma%20generalizada>, [consulta: 30/03/2023]

tecnologías de inteligencia artificial no deben ser consideradas como entidades jurídicas autónomas.

- Ayudar a supervisar y evaluar: como se mencionó anteriormente, el objetivo que se busca con esta Recomendación es reducir los riesgos que puede conllevar la IA. Por ello la UNESCO propone una herramienta que pueden desarrollar los Estados y las empresas privadas con el fin de evaluar el impacto de los sistemas de IA en las personas, la sociedad y el medio ambiente. Esta herramienta está recogida en el “ámbito de actuación 1: Evaluación del impacto ético”. Asimismo, la Recomendación establece la opción designar a un oficial de ética de IA independiente u otro mecanismo que permita supervisar los esfuerzos de monitoreo y auditoría en curso.
- Como último punto aparece la protección del medio ambiente: para la UNESCO el desarrollo de la IA también puede suponer un impacto para el medio ambiente, por lo que instan a los gobiernos y a las empresas privadas a analicen desde el consumo de energía hasta el impacto que supone la extracción de las materias necesarias para el desarrollo de la IA.

### **3.1.2. Visión general de los Principios de IA de la OCDE**

Unos años antes que la UNESCO la OCDE también lanzó una guía titulada “Visión general de los Principios de IA de la OCDE”<sup>12</sup>. En esta se reconoce el punto de inflexión que puede suponer la IA para la innovación, para la economía o para el bienestar de las personas. Pero en contrapartida el empleo de estas herramientas puede poner en riesgo derechos humanos como la privacidad o la protección de datos, así como desembocar en cambios económicos y sociales tan trascendentales que sean de difícil encaje dentro de la sociedad.

---

<sup>12</sup> OCDE (s. f): *Artificial Intelligence (AI) Principles*. OECD.AI. Disponible en: <https://oecd.ai/en/ai-principles>, [consulta: 30/03/2023]

En esta guía se encuentran un término que más adelante se desarrollará pero que conviene introducir. Se trata de la “IA confiable”, aprovechar las oportunidades ofrecidas y abordar los desafíos planteados, mediante aplicaciones de IA, y empoderar a las partes interesadas para que participen es esencial para fomentar la adopción de una IA confiable en la sociedad y convertir la IA confiable en un parámetro competitivo en el mercado global”.<sup>13</sup>

En la recomendación se pueden distinguir dos secciones:

- Una primera sección dedicada a los principios para la admisión responsable de la IA confiable. En esta sección se describen 5 valores y principios que son de recomendable aplicación por parte de las administraciones estatales y los desarrolladores de la IA para alcanzar una administración responsable de una IA confiable. Estos 5 principios son:
  - Crecimiento inclusivo, desarrollo sostenible y bienestar: en este punto se destaca el papel que puede jugar el empleo de herramientas IA confiable para promover los Objetivos de Desarrollo Sostenible en las diferentes áreas sociales como la salud, medioambiente, etc. Pero para ello es necesario que las partes destacadas busquen fomentar el desarrollo de estas herramientas con el fin de obtener los resultados requeridos. Con este principio lo que busca la OCDE es que exista una constante colaboración entre las distintas partes, un dialogo con el fin de promover la confianza pública y la comprensión de la IA.
  - Valores centrados en el ser humano y equidad: unido a uno de los riesgos que reconoce la OCDE en su introducción, “estas transformaciones pueden tener efectos dispares. dentro de las sociedades y economías y entre ellas [...], desigualdades e implicaciones por la democracia y los derechos humanos, la privacidad y la protección de datos, y la tecnología digital

---

<sup>13</sup>OCDE (s. f): *Recommendation of the Council on Artificial Intelligence*, OECD/LEGAL/0449. Disponible en: <https://oecd.ai/en/assets/files/OECD-LEGAL-0449-en.pdf>, [consulta: 30/03/2023]

seguridad”<sup>14</sup>, este principio trata de solucionarlo. Con éste se busca que sean los propios actores los que incluyan los medios necesarios para respetar los derechos humanos, el Estado de Derecho, los valores democráticos, así como salvaguardar la privacidad, igualdad, la protección de los datos, etcétera.

- Transparencia y explicabilidad: de este principio hay que destacar la necesidad de que los actores puedan garantizar una transparencia y divulgación responsable de los sistemas IA, garantizando de esta forma que los usuarios puedan entenderla. Ello busca que los afectados por sistemas IA puedan cuestionar los resultados.
- Robustez, seguridad y protección: se ha mencionado a lo largo de este apartado las posibles contingencias que se pueden dar por el empleo de las herramientas IA. Por ello, en este punto la OCDE se enfoca en la importancia de la gestión de los riesgos, de tratar los posibles desafíos de seguridad y protección que se pueden dar en estas herramientas por parte de sus desarrolladores, así como de las administraciones. La OCDE recomienda en este punto dos formas de conseguir una IA robusta y segura: 1) trazar y analizar los posibles riesgos. 2) establecer un correcto enfoque de gestión de riesgos.
- Responsabilidad: como último principio, hace responsables del respeto de estos principios a los propios actores de la IA.
- La segunda sección propone 5 pautas sobre cómo enfocar las políticas nacionales con el fin de alcanzar una IA fiable:
  - Invertir en investigación y desarrollo de la IA: se motiva a los gobiernos a destinar parte de la inversión pública a proyectos de investigación y desarrollo de esta IA confiable, sobre todo en cuestiones sociales, legales y éticas. Al igual que se gusta esta inversión pública, se insta a los gobiernos a alentar la inversión privada.

---

<sup>14</sup> *idídem*

- Fomentar un ecosistema digital para la IA: se debe entender como “ecosistema digital” un entorno donde los gobiernos hayan procurado el desarrollo de una infraestructura adecuada que propicie los mecanismos para compartir los conocimientos. Estos mecanismos son tales como la confianza, el intercambio justo, legal y ético de los datos.
- Propiciar un entorno político para la IA: la OCDE considera necesario que para que se dé un entorno viable para el desarrollo y experimentación de la IA, primero se debe establecer un entorno político flexible que lo permita. Para ello, en la guía se menciona la experimentación como el mejor medio para conseguir un entorno propicio. Éste sería un entorno controlado donde los sistemas IA podrían mejorar la adaptabilidad, la reactividad, la versatilidad, así como la correcta aplicación los instrumentos políticos. Pero este fin no sería posible si no se lleva a cabo una correcta supervisión del entorno. En este punto los gobiernos deben de ser los encargados de supervisar que sus marcos normativos se están aplicando a los sistemas IA.
- Desarrollo de la capacidad humana y preparación para la transición al mercado laboral: ya se está haciendo notorio en la actualidad que el desarrollo de la IA puede suponer un gran impacto en el mundo laboral, se reemplazarán oficios por herramientas IA, otros podrán beneficiarse del empleo de estas y surgirán nuevos empleos. Pero lo que está claro es que, si esta transición se gestiona de forma apresurada y mal, podrá suponer un elevado costo social y económico. En este punto, la guía insta a los responsables políticos a promover políticas de aprendizaje, desarrollar las competencias y fomentar la formación de los trabajadores en estos campos. No se busca que el sistema IA solape al ser humano, añadidas a estas políticas, en la guía se manifiesta la necesidad de que se tomen en cuenta políticas de desarrollo de habilidades humanas (el juicio, el pensamiento creativo y crítico y la comunicación interpersonal), imprescindibles para completar los sistemas IA.

- Cooperación internacional para una IA fiable: este último punto actúa a modo de llamamiento a todos los gobiernos para compartir conocimientos sobre la IA y de este modo reducir el riesgo que supone para la seguridad, la privacidad y los derechos humanos el desarrollo de una IA sin control.

A modo de conclusión, podemos considerar esta Guía como un marco importante para el desarrollo y uso responsable de la IA. Como se ha podido comprobar durante la explicación esta recomendación se basa en principios fundamentales como la igualdad de oportunidades, el bienestar humano, la transparencia, la robustez y la gobernanza responsable. Su objetivo es orientar a los países en la adopción ética y segura de esta tecnología. Al fomentar la inclusión, el respeto de los derechos humanos y la rendición de cuentas, esta recomendación busca generar confianza pública y asegurar que la IA sea una herramienta beneficiosa para la sociedad en general. De esta forma, es esencial que los países actualicen sus políticas y marcos regulatorios en línea con estas recomendaciones para garantizar un desarrollo responsable y ético de la Inteligencia Artificial.

### **3.2. LA IA Y SU ENFOQUE EUROPEO.**

La Comisión Europea desde marzo del 2018 trabaja para dotar a la Unión de un marco jurídico que garantice un entorno seguro donde el desarrollo, la innovación y la investigación de los sistemas IA sean objetivos a al mismo tiempo que se establecen unos estándares de seguridad y de protección de los derechos fundamentales.

#### **3.2.1. El grupo de expertos y las directrices éticas.**

Con el rápido desarrollo de la IA en los últimos años del siglo XXI, en la Unión Europea empiezan a surgir cuestiones enfocadas al futuro de estos sistemas y sus implicaciones, sobre todo en si la legislación existente será capaz de adaptarse a los nuevos cambios que tarde o temprano llegarán.

Es por ello por lo que en marzo de 2018 se decide desde la Comisión Europea la creación de un grupo de expertos que trabaje en la creación de una propuesta de directrices éticas para la IA, tomando como referencia la declaración del Grupo Europeo sobre Ética en la Ciencia y las Nuevas Tecnologías.

De las palabras pronunciadas por la comisaria de Economía y Sociedad Digitales Mariya Gabriel, se puede desprender el espíritu con el que nace este grupo de expertos: *“Para aprovechar todos los beneficios de la inteligencia artificial, la tecnología siempre debe utilizarse en interés de los ciudadanos y respetar los más altos estándares éticos, promover los valores europeos y defender los derechos fundamentales. Por eso estamos en constante diálogo con las principales partes interesadas, incluidos investigadores, proveedores, implementadores y usuarios de esta tecnología. Nuestro trabajo para construir un mercado único digital es esencial para fomentar el desarrollo y la adopción de nuevas tecnologías”*.<sup>15</sup>

De estas palabras se puede extraer lo que pretende lograr el grupo de expertos, que es una IA en la que se pueda confiar, una IA que pueda ser implantada en la sociedad y que reporte un beneficio a la misma.

Pero ¿cómo se puede llegar a establecer una IA que cumpla estos requisitos?, ¿cuáles son los fundamentos necesarios para llegar a ella?

La respuesta a estas preguntas está en la denominada “IA fiable o IA confiable”. Para ello el grupo de expertos en su comunicado establece tres pilares necesarios, una serie de directrices que ayudan a enfocar una IA fiable fundamentadas en los derechos humanos:

- Una IA lícita, una IA conforme a la ley: no es ni más ni menos que una IA que garantice el respeto y el cumplimiento de los preceptos legales y normativos.

---

<sup>15</sup> Comisión Europea (2018): *Artificial intelligence: Commission kicks off work on marrying cutting-edge technology and ethical standards*. Comunicado de prensa. Disponible en: [https://ec.europa.eu/commission/presscorner/detail/es/IP\\_18\\_1381](https://ec.europa.eu/commission/presscorner/detail/es/IP_18_1381), [consulta: 02/04/2023]

- Una IA ética: siendo ya casi un dogma, respetar la ley no significa un respeto integro por los principios éticos. Es necesario por ello que la IA sea capaz de garantizar los principios y valores éticos. Para ello, la Comisión Europea fija que la IA tiene que:<sup>16</sup>
  - Proteger la igualdad, la no discriminación y la solidaridad entre los ciudadanos. Este principio está ligado con el tratamiento de los datos personales.
  - El respeto por la democracia, la justicia y el estado de Derecho. Se observa que este es un principio muy relacionado con el primero de los pilares mencionado, una IA lícita.
  - Respeto a la dignidad humana. Considerar a los humanos como personas con valores y no como serie de datos a analizar.
  - Respeto por la libertad individual. La IA debe evitar la intromisión o influencia en la toma de decisiones de cada persona. Eso implica que debe evitar emplear coerciones ilegítimas, amenazas, etc.
  - Por último, está el respeto a los derechos de los ciudadanos,
- Como último pilar esta la IA robusta: habiendo garantizado un fin ético y legal, es necesario garantizar que los sistemas IA actuarán de forma segura y no provocarán daños involuntarios. Desde el grupo de expertos diferencian dentro de esta característica dos ámbitos. Por un lado, tenemos el ámbito técnico, donde debe existir cierta garantía técnica ya sea su ámbito de aplicación o su ciclo de vida. Y también un ámbito social, donde se debe tener en cuenta el contexto y el entorno social en el que operará el sistema IA.

Como se extrae de la recomendación, los expertos no buscan con estas directrices establecer valoraciones o recomendaciones que busquen reemplazar

---

<sup>16</sup> M. J. DÍAZ, F. HERRERA TRIGUERO, O. CORDÓN GARCÍA: *Una visión actual de la inteligencia artificial:...* op. cit. p.69

las políticas o reglamentos actuales. Sino que buscan centrarse en los principios y requisitos éticos que la IA debe respetar.

Estos requisitos se encuentran definidos en el capítulo II de la recomendación. Son siete puntos que los expertos consideran esenciales para que los sistemas IA puedan lograr la fiabilidad y robustez necesarias. La mayoría de estos requisitos ya se han visto anteriormente en otras recomendaciones internacionales siendo estas influidas, al igual que esta recomendación, por la normativa internacional sobre derechos humanos. Estos siete requisitos son los siguientes:<sup>17</sup>

- Acción y supervisión humana: con este requisito los expertos expresan la necesidad de establecer medidas de control para los sistemas IA. Ya sea con una participación humana directa en el control en el sistema IA, con una supervisión humana o que sea el humano quien directamente comande la IA.
- Solidez y seguridad técnica: un requisito íntimamente relacionado con la característica de robustez. En este punto los expertos exponen una realidad cada vez más frecuente, los ataques a los sistemas informáticos, y por consiguiente también a los sistemas IA. Ante esta situación, consideran pertinente que los sistemas IA sean capaces de garantizar una robustez ante estos ataques o comportamientos anómalos.
- Privacidad y gestión de datos: un requisito ligado al anterior y por cercanía a la característica de robustez.
- Transparencia: una IA que no puede ser monitorizada perdería todo el sentido ya que no se podría controlar su flujo de datos ni el uso que hace de ellos. Por ello es de imperiosa necesidad que los sistemas IA sean explicables, que se pueda registrar de donde obtiene los datos, de esta forma que expertos que la supervisen, controlen o comanden puedan entender las decisiones que toma el sistema.

---

<sup>17</sup> M. J. DÍAZ, F. HERRERA TRIGUERO, O. CORDÓN GARCÍA: *Una visión actual de la inteligencia artificial:...* op. cit. p.70

- Diversidad, discriminación y equidad: requisito ligado a la característica de ética. Como su enunciado indica los sistemas IA no deberán actuar de forma discriminante o abusiva a la hora de obtener los datos con los que trabajarán.
- Bienestar social y medioambiental: al igual que cualquier otro sistema informático, los sistemas IA no están excluidos de su huella de carbono y de su impacto ambiental. De esta forma, durante el proceso de desarrollo y de explotación de la IA, es necesario respetar este requisito.
- Rendición de cuentas: los sistemas IA carecen de personalidad, siendo esto un punto importante a la hora de dirimir responsabilidades. Por ello es necesario que los sistemas IA estén auditados. De esta forma se garantiza una responsabilidad y una reparación en el caso de que la IA provoque daños o perjuicios.

Dentro de la recomendación los expertos introducen un apartado donde incluyen métodos técnicos y no técnicos para hacer realidad la IA fiable. Pero este punto no es objeto de estudio para este trabajo.

A modo de conclusión, decir que el conjunto de imperativos éticos que exponen estos expertos en la recomendación son principalmente un compendio de los ya expuestos en otras instituciones a nivel internacional. Resumiendo, estas siete directrices o requisitos que se han explicado anteriormente, en cuatro grandes imperativos éticos: <sup>18</sup>

- El respeto a la autonomía humana.
- La prevención del daño.
- La equidad.
- La explicabilidad.

---

<sup>18</sup> M. PEGUERA POCH: *En búsqueda de un marco normativo para la inteligencia artificial*, en Agustín Cerrillo y Martínez y Miguel Peguera Poch (eds.): *Retos jurídicos de la inteligencia artificial*, Título I Impacto, ética y marco legal, Aranzadi, Pamplona, 2020, pp. 48-51

### **3.2.2. Marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas.**

La resolución del Parlamento Europeo con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas fue aprobada el 20 de octubre de 2020<sup>19</sup> con el fin de presentar una propuesta legislativa para la tramitación de un Reglamento sobre los principios éticos para el desarrollo, el despliegue y el uso de dichas tecnologías.

La mayor parte del contenido de esta resolución encuentra su base en lo ya propuesto en el Libro Blanco sobre Inteligencia Artificial y en las Directrices éticas para una IA fiable elaboradas por el Grupo independiente de expertos de alto nivel, expuesto en el punto anterior.

De esta forma podemos dividir el contenido de la resolución tres puntos básicos:

1. Aplicaciones de alto riesgo.
2. Supervisión humana integral.
3. Sesgo y discriminación.

En cuanto a las aplicaciones de alto riesgo, la resolución en su Consideración número 55, concluye la necesidad de que el marco regulatorio debe ajustarse al nivel de riesgo que puedan representar los diferentes sistemas IA y sus aplicaciones. Esta misma visión la expone la Comisión en su Libro Blanco sobre la IA<sup>20</sup>. Para ello, es necesario poder establecer unos criterios claros para poder medir el nivel de riesgo de estas tecnologías.

---

<sup>19</sup> Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas 2020/2012INL. Disponible en:

[https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275\\_ES.html](https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_ES.html),  
[consulta:05/04/2023]

<sup>20</sup> Comisión Europea (2020): *Libro Blanco sobre la inteligencia artificial - un enfoque europeo orientado a la excelencia y la confianza*, p. 21. Disponible en:

<https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52020DC0065>,  
[consulta:05/04/2023]

Esto se resuelve en el Consideración número 11 de la evaluación de riesgos de la misma resolución, donde se dice que una IA de alto riesgo es la que contiene *“un riesgo significativo de causar lesiones o daños a particulares o a la sociedad, vulnerando los derechos fundamentales y las normas de seguridad establecidas en el Derecho de la Unión”*.<sup>21</sup>

En el mismo punto se establecen diferentes criterios que se pueden seguir a efectos de evaluar si las tecnologías de inteligencia artificial entrañan un riesgo de ese tipo:

- el sector al que pertenece.
- el uso o finalidad específica al que se dedica.
- la gravedad del daño que cabe esperar que produzca.

En el siguiente apartado hablaremos de los diferentes niveles de riesgo.

En segundo lugar, tenemos la supervisión humana integral. Este es uno de los siete requisitos que, como se vio en el punto anterior, el grupo de expertos considera necesarios para el desarrollo de una IA fiable.

Acudiendo a la Consideración 10 de la Resolución se observa cómo se establecen cuatro formas de supervisión humana *“Las decisiones adoptadas por la inteligencia artificial, la robótica y las tecnologías conexas o basadas en ellas deben seguir siendo objeto de revisión, evaluación, intervención y control humanos significativos”*. Es decir, la revisión, la evaluación, intervención y el control.

De esta forma en el artículo 7 de la Resolución se concluye que *“las tecnologías de inteligencia artificial de alto riesgo, incluidos los programas informáticos, los algoritmos y los datos utilizados o producidos por dichas tecnologías, se desarrollarán, desplegarán y utilizarán de forma que se garantice*

---

<sup>21</sup> Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas (2020/2012(INL)). Disponible en: [https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275\\_ES.html](https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_ES.html), [consulta:05/04/2023]

*en todo momento una supervisión humana integral*". Simplemente establece que dicho control humano debe darse en todo el ciclo de vida de la IA.

En tercer lugar, sesgo y discriminación. Estos vienen definidos en el artículo 4 apartados L y M de la Resolución. Este artículo define "sesgo" como *"toda percepción personal o social prejuiciosa de una persona o de un grupo de personas sobre la base de sus características personales"* y define "discriminación" como *"todo trato diferenciado de una persona o de un grupo de personas basado en un motivo que no tiene justificación objetiva o razonable alguna y que, por tanto, está prohibido por el Derecho de la Unión"*.

En esta Resolución se muestra la preocupación que muestra el Parlamento con esta cuestión al determinar la ausencia de estos dos preceptos como un principio ético de obligado cumplimiento para los sistemas IA, además de establecer en su artículo 9 apartado primero una enumeración bastante detallada *"Los programas informáticos, los algoritmos o los datos utilizados o producidos por la inteligencia artificial, la robótica y las tecnologías conexas desarrolladas, desplegadas o utilizadas en la Unión serán imparciales [...] no discriminarán por motivos de raza, sexo, orientación sexual, embarazo, discapacidad, características físicas o genéticas, edad, minoría nacional, origen étnico o social, lengua, religión o creencias, opiniones políticas o participación cívica, nacionalidad, estado civil o económico, educación o antecedentes penales."*

### **3.2.3. El desafío regulatorio.**

Pese a que actualmente la legislación permita disponer de una protección frente a los sistemas IA, esta es insuficiente para abordar los retos que pueden llegar a surgir en el futuro.

Sin ir más lejos hoy día resulta complicado entender el porqué un sistema IA llega a una toma de decisión en un supuesto particular. Ya sea por ejemplo ante una contratación o ante una resolución judicial, pudiendo de esto desprenderse un perjuicio para el particular.

Es por ello por lo que con este marco jurídico se pretende que los europeos puedan confiar en los sistemas IA. Entre las normas propuestas encontramos<sup>22</sup>:

- abordar los riesgos creados específicamente por las aplicaciones de IA.
- proponer una lista de aplicaciones de alto riesgo.
- establecer requisitos claros para los sistemas de IA para aplicaciones de alto riesgo.
- definir obligaciones específicas para los usuarios de IA y los proveedores de aplicaciones de alto riesgo.
- proponer una evaluación de la conformidad antes de que el sistema de IA se ponga en servicio o se comercialice.
- proponer la aplicación después de la introducción en el mercado de un sistema de IA de este tipo.
- proponer una estructura de gobernanza a nivel europeo y nacional.

El resultado obtenido es que la Unión Europea va a optar por un marco jurídico horizontal basado en el riesgo. De esta forma se esquivo la necesidad de establecer un enfoque sectorial, un sistema de etiquetado voluntario o establecer requisitos para todos los sistemas IA<sup>23</sup>. La Unión Europea ha definido cuatro niveles de riesgo en la IA.

1. Riesgo inaceptable.
2. Alto riesgo.
3. Riesgo limitado.
4. Riesgo mínimo o nulo.

---

<sup>22</sup> Comisión Europea. (s. f): *Regulatory Framework for Artificial Intelligence*. Disponible en:<https://digital-strategy.ec.europa.eu/es/policies/regulatory-framework-ai>, [consulta:05/04/2023]

<sup>23</sup> Mar MORENO REBATO: *Principios éticos y derecho. Autorregulación y/o regulación: Inteligencia artificial (umbrales éticos, derecho y administraciones públicas)*, Capítulo I Principios éticos y derecho, Aranzadi, Pamplona, 2020, p 34.

En primer lugar, tenemos el riesgo inaceptable. En este punto, cualquier sistema IA que sea considerado una amenaza para la seguridad, los medios de vida y los derechos de las personas, serán prohibidos.

En segundo lugar, está el riesgo alto. Dentro de este riesgo la Unión europea incluye los sistemas IA que se emplean en<sup>24</sup>:

- infraestructuras críticas (transporte), que podrían poner en peligro la vida y la salud de los ciudadanos.
- formación educativa o profesional, que puede determinar el acceso a la educación y al curso profesional de la vida de una persona (la puntuación de los exámenes).
- componentes de seguridad de los productos (aplicación de IA en cirugía asistida por robot).
- empleo, gestión de los trabajadores y acceso al trabajo por cuenta propia (*software* de clasificación de currículos para los procedimientos de contratación).
- servicios públicos y privados esenciales (la calificación crediticia que deniega a los ciudadanos la oportunidad de obtener un préstamo).
- aplicación de la ley que pueda interferir con los derechos fundamentales de las personas (evaluación de la fiabilidad de las pruebas).
- gestión de la migración, el asilo y el control de fronteras (verificación de la autenticidad de los documentos de viaje).
- administración de justicia y procesos democráticos (aplicación de la ley a un conjunto concreto de hechos).

La Unión europea no solo expone los ámbitos donde los sistemas IA pueden considerarse de alto riesgo, sino que además imponen obligaciones que deben cumplir antes de que estos sistemas IA puedan salir al mercado.

---

<sup>24</sup> Comisión Europea. (s. f). *Regulatory Framework for Artificial Intelligence*. Disponible en: <https://digital-strategy.ec.europa.eu/es/policies/regulatory-framework-ai>, [consulta:05/04/2023]

Algunas de estas obligaciones son por ejemplo que los sistemas IA cuenten con un alto nivel de robustez, seguridad y precisión, que la información sea clara para el usuario, así como que cuenten con medidas de supervisión humana para mitigar los posibles riesgos.

En tercer lugar, el riesgo limitado. Aquí la Unión Europea imponen ciertas obligaciones de transparencia a los sistemas IA. Se refieren por ejemplo a los *chatbots*. Los usuarios deberán ser conscientes en todo momento que está hablando e interactuando con una IA.

En cuarto lugar y último, están los sistemas IA considerados de riesgo mínimo o nulo. Se trataría de videojuegos o filtros de spam.

Con el fin de conseguir una IA de confianza, la Comisión decidió emprender tres iniciativas jurídicas relacionadas entre sí con el fin de crear una IA fiable. Estas son las tres iniciativas:

1. Un marco jurídico europeo a fin de abordar los riesgos para los derechos fundamentales y la seguridad específicos de los sistemas de IA.
2. Una Directiva sobre responsabilidad por IA: adaptación de las normas de responsabilidad a la era digital y a la IA.
3. Una propuesta de Directiva sobre responsabilidad por productos defectuosos.

A modo de conclusión de este apartado, cabe mencionar el Considerando 5 de la propuesta de Reglamento del Parlamento Europeo y del Consejo estableciendo normas armonizadas sobre IA:

*“Se necesita un marco jurídico de la Unión que defina unas normas armonizadas en materia de inteligencia artificial orientadas a impulsar el desarrollo, la utilización y la adopción en el mercado interior de la inteligencia artificial y que, al mismo tiempo, ofrezca un nivel elevado de protección de los intereses públicos, como la salud y la seguridad, y de los derechos fundamentales reconocidos y protegidos por el Derecho de la Unión. Para alcanzar dicho objetivo, conviene establecer normas que regulen la introducción*

*en el mercado y la puesta en servicio de determinados sistemas de IA, lo que garantizará el buen funcionamiento del mercado interior y permitirá que dichos sistemas se beneficien del principio de la libre circulación de bienes y servicios. Al establecer tales normas, el presente Reglamento respalda el objetivo de la Unión de ser un líder mundial en el desarrollo de inteligencia artificial segura, digna de confianza y ética, como indicó el Consejo Europeo, y garantiza la protección de los principios éticos, como solicitó específicamente el Parlamento Europeo”.*<sup>25</sup>

### **3.3. LA ESTRATEGIA NACIONAL DE INTELIGENCIA ARTIFICIAL**

En España no existe un marco normativo definido que regule la inteligencia artificial. Actualmente está en vigor la Estrategia Nacional de Inteligencia Artificial (ENIA) que tiene como objetivo proporcionar un marco de referencia para el desarrollo de una IA inclusiva, sostenible y centrada en la ciudadanía<sup>26</sup>.

Sin ánimo de ser exhaustivos en este análisis, ya que aborda cuestiones alejadas del ámbito de estudio, este plan estratégico se puede resumir en los objetivos que busca alcanzar y en las medidas que se proponen para alcanzarlos.

En un principio este plan estratégico busca los siguientes objetivos<sup>27</sup>:

- Situar a España en la excelencia científica e innovación en Inteligencia Artificial.

---

<sup>25</sup> Propuesta de Reglamento del Parlamento Europeo y del Consejo estableciendo normas armonizadas sobre inteligencia artificial (Ley de Inteligencia Artificial) y modificando ciertos actos legislativos de la Unión. Bruselas, 21-4-2021. Disponible en: <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=celex%3A52021PC0206>, [consulta:05/04/2023]

<sup>26</sup> Ministerio de asuntos económicos y transformación digital: Inteligencia artificial. Disponible en: <https://portal.mineco.gob.es/es-es/ministerio/areas-prioritarias/Paginas/inteligencia-artificial.aspx>, [consulta:05/04/2023]

<sup>27</sup> Ministerio de asuntos económicos y transformación digital: *Plan Estratégico de Inteligencia Artificial*, Madrid, 2020, p 15.

- Proyección de la lengua española. Con este objetivo se trata de posicionar a la lengua española como un referente en el ámbito de aplicación de la IA.
- Creación de empleo cualificado.
- Transformación del tejido productivo.
- Entorno de confianza en relación con la Inteligencia Artificial.
- Valores humanistas en la Inteligencia Artificial.
- Inteligencia Artificial inclusiva y sostenible.

La mayor parte de estos objetivos se encuentran íntimamente relacionado con el Eje Estratégico 6 del plan de medidas que propone el plan estratégico, que es *“Establecer un marco ético y normativo que refuerce la protección de los derechos individuales y colectivos, a efectos de garantizar la inclusión y el bienestar social”*.

Para la consecución de este objetivo, el plan propone una serie de medidas<sup>28</sup>:

- Desarrollo de un sello nacional de calidad IA.
- Poner en marcha observatorios para evaluar el impacto social de los algoritmos.
- Desarrollar la Carta de Derechos Digitales.
- Puesta en marcha de un modelo de gobernanza nacional de la ética en la IA (Consejo Asesor IA).
- Promoción de foros de diálogo, sensibilización y participación nacionales e internacionales en relación con la IA.

Algunas de las medidas destacadas son por ejemplo la creación del Consejo Asesor de la IA, cuyo objetivo principal es estudiar el alcance de la aplicación de la ENIA, así como la gobernanza nacional de la ética en la IA y la creación del sello nacional de calidad de IA.

---

<sup>28</sup> Ministerio de asuntos económicos y transformación digital: *Plan Estratégico de Inteligencia Artificial*, Madrid, 2020, *op. cit.*, p 73.

Relevante es la publicación de la Carta de derechos Digitales<sup>29</sup>. Publicada en julio de 2021 esta carta busca afianzar los derechos individuales y colectivos de los escenarios digitales. Se actualizan los derechos ya existentes en la Declaración de derechos Humanos o de la misma Constitución Española adaptándolos a la actualidad de la era digital. Respecto a la IA, se da un enfoque más centrado en las personas, en el desarrollo de sistemas para garantizar el derecho de no discriminación.

Por último, habría que hacer referencia a cómo enfoca los principios de la IA una de las empresas más importantes en España, Telefónica. Ésta incluye principios éticos<sup>30</sup> que ya se han ido viendo a lo largo de este punto, en el documento aprobado por el Comité Ejecutivo de Telefónica en el año 2018. Estos principios son los siguientes:

- Equidad: basada en aplicaciones de la IA justas y ni discriminación por razón de la raza, el origen étnico, la religión, el sexo, la orientación sexual, la discapacidad o cualquier otra condición de las personas.
- Transparencia y explicabilidad: en este tipo de tecnologías las decisiones que tomen los sistemas IA podrían llegar a afectar significativamente a la vida de las personas, por lo que es necesario entender la lógica que hay detrás de las conclusiones. Esto se consigue realizando una revisión de los datos con el fin de entender para que se van a emplear.
- IA centrada en las personas, al servicio de la sociedad, evitando provocar un impacto negativo en los derechos humanos o en el logro de los Objetivos de Desarrollo Sostenible de la ONU.
- Privacidad y seguridad de la persona: búsqueda de políticas de privacidad y seguridad en todos los procesos de diseño.

---

<sup>29</sup> Ministerio de Economía y Transformación Digital. (s. f.): *Carta de los Derechos Digitales*. Disponible en: [https://portal.mineco.gob.es/RecursosArticulo/mineco/ministerio/participacion\\_publica/audiencia\\_ficheros/SEDIACartaDerechosDigitales.pdf](https://portal.mineco.gob.es/RecursosArticulo/mineco/ministerio/participacion_publica/audiencia_ficheros/SEDIACartaDerechosDigitales.pdf), [consulta:10/04/2023]

<sup>30</sup> Telefónica (2018): *Principios para la Inteligencia Artificial*. Disponible en: <https://www.telefonica.com/es/wp-content/uploads/sites/4/2021/08/principios-ai-esp-2018.pdf>, [consulta:10/04/2023]

## 4. CONTENIDO Y PRINCIPIOS ÉTICOS

Un punto de partida para las premisas éticas de la IA la encontramos en las numerosas declaraciones que se han ido exponiendo a lo largo del trabajo, centradas sobre todo en la dignidad y los derechos fundamentales.

La marca de la IA confiable promovida por la Unión Europea parte de esas mismas dos ideas. Pensamiento que ha sido matizado por el propio Supervisor Europeo de Protección de Datos cuando expone que, *“un mayor respeto de la dignidad humana y una mayor salvaguardia de esta podrían servir de contrapeso a la vigilancia generalizada y la asimetría de poder a la que se enfrentan las personas”*.<sup>31</sup>

Por otro lado, en la Declaración de robótica del Parlamento de la UE de 2017 exponen en su principio ético número 13 que *“este marco de orientaciones éticas debe basarse en los principios de beneficencia, no maleficencia, autonomía y justicia, así como en los principios consagrados en la Carta de los Derechos Fundamentales de la Unión Europea, como la dignidad humana, la igualdad, la justicia y la equidad, la no discriminación, el consentimiento informado, la vida privada y familiar y la protección de datos, así como en otros principios y valores inherentes al Derecho de la Unión, como la no estigmatización, la transparencia, la autonomía, la responsabilidad individual, y la responsabilidad social, sin olvidar las actuales prácticas y códigos éticos”*<sup>32</sup>

Esto significa, que los derechos fundamentales serán los que nutran los principios de esta ética basada en la dignidad.

---

<sup>31</sup> Supervisor Europeo de Protección de Datos: *Dictamen 4/2015. Hacia una nueva ética digital. Datos, dignidad y tecnología*, 2015, p 14.

<sup>32</sup> Parlamento Europeo: *Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL))*, 2017, p 8.

#### 4.1. PRINCIPIOS ÉTICOS BÁSICOS.

Se entiende que la Ética es una disciplina filosófica con un carácter normativo y práctico. El individuo interviene en el entorno social siguiendo unos principios éticos con el fin de conciliarlos con de las personas con las que interactúa.

En la perspectiva consecuencialista, una actitud ética es aquella que mide los efectos que pueden tener sus acciones y que elige aquellos que tienen resultados menos lesivos.

Según la ética deontológica, una actitud ética es aquella que se limita a conformarse con algún principio ético, prescindiendo de los resultados. Las personas que tienen una ética deontológica, también conocida como ética de la obligación o del deber, tan solo buscan la conformidad con dichos principios.

En el campo tecnológico y de los sistemas IA se aúnan estas dos actitudes éticas, por un lado, está el respeto a los principios éticos, pero en contrapartida la implementación de estos principios no debe limitar injustificadamente el desarrollo tecnológico

Es por ello por lo que los principios son un eje de referencia y siguiendo con la ya mencionada declaración de robótica del Parlamento de la UE de 2017 y lo expuesto en su punto número 13 de principios éticos, se pueden apreciar ya resumidos en buena medida los cuatro grandes principios éticos que van a regir a la IA: principios de beneficencia, no maleficencia, autonomía y justicia.

Se debe añadir un quinto principio, el principio de transparencia, recogido en el punto número 12 de la declaración: *“Pone de relieve el principio de transparencia, que consiste en que siempre ha de ser posible justificar cualquier decisión que se haya adoptado con ayuda de la inteligencia artificial y que pueda tener un impacto significativo sobre la vida de una o varias personas; considera que siempre debe ser posible reducir los cálculos del sistema de inteligencia artificial a una forma comprensible para los humanos; estima que los robots avanzados deberían estar equipados con una «caja negra» que registre los datos*

*de todas las operaciones efectuadas por la máquina, incluidos, en su caso, los pasos lógicos que han conducido a la formulación de sus decisiones”.*<sup>33</sup>

El principio de beneficencia establece que la IA debe desarrollarse con el objetivo de promover el bienestar general y el progreso humano, es decir, mejorar el bienestar tanto individual como colectivo, al mismo tiempo que se promueve la prosperidad y se crea valor de manera sostenible. Es imperioso que la IA se esfuerce por construir una sociedad justa, inclusiva y pacífica aumentando la autonomía de los ciudadanos y asegurando la distribución equitativa de las oportunidades económicas, sociales y políticas.

La IA debe tener como objetivos principales la protección del proceso democrático y el Estado de Derecho. Trabajar en la provisión de bienes y servicios comunes a bajo costo manteniendo altos estándares de calidad.

También es fundamental promover la alfabetización y representación de datos para garantizar la coherencia y evitar fallas en su funcionamiento. Además, la IA debe dedicarse a la prevención de daños potenciales y la mejora de la confianza que los usuarios depositan en ella.

En resumen, la IA debe desarrollarse con responsabilidad y ética, velando siempre por el bien común y el beneficio de la Humanidad. Al mismo tiempo, debe trabajar por la creación de una sociedad justa, inclusiva y pacífica, que promueva la autonomía ciudadana y asegure la distribución equitativa de las oportunidades económicas, sociales y políticas, así como protegiendo la democracia, el Estado de Derecho y la confianza de los usuarios.

Por otro lado, el principio de no maleficencia tiene un componente preventivo, se busca que la IA evite provocar daños físicos, psicológicos, financieros o sociales.

Dentro de este principio cabe anotar varios bienes en riesgo: todo derecho fundamental, la “privacidad, seguridad y precaución de capacidad”.

---

<sup>33</sup> Parlamento Europeo: *Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL))*, 2017, p 8

Dentro de los bienes mencionados anteriormente, también debe protegerse el proceso democrático, prevenir la fragmentación ideológica y el determinismo algorítmico. Respecto al uso de datos, se debe tener cautela con el empleo que se dan a los datos sobre minorías. De forma, para asegurar la inclusión y la diversidad.

Otro daño que se busca evitar con este principio es sobre el medio ambiente. Los sistemas IA están muy lejos de considerarse unos sistemas “verdes y ecológicos”, sobre todo si tenemos en cuenta el consumo que requieren los cómputos masivos, el almacenamiento de datos, así como los minerales para las baterías necesarias.

El tercer principio ético básico es el de autonomía o acción humana. El Grupo Europeo de Ética en Ciencia y Nuevas Tecnologías recuerda: “*que la autonomía como derecho a ser libre solo puede atribuirse a los seres humanos, por lo que no cabe atribuir autonomía a los sistemas por inteligentes que sean*”<sup>34</sup>. Así como que “*la autonomía humana queda bajo el principio de la libertad del ser humano que se traduce en responsabilidad humana y, por lo tanto, control y conocimiento sobre los sistemas autónomos y su mayor transparencia y previsibilidad*”.<sup>35</sup>

Con lo expuesto, se comprueba que se tiende al llamado principio de supervisión humana de la IA. Es decir, más control humano a mayor autonomía e impacto social de la IA, así como más registros auditables y una gobernanza más estricta. La capacidad de desviarse de la decisión elegida por la máquina tiene que ser algo propio del ser humano.

El principio de justicia está muy ligado con el principio de beneficencia. Eliminar la discriminación mediante el uso de la IA para distribuir y acceder a los recursos de manera equitativa. Incluye la eliminación de las discriminaciones que han existido en el pasado, la prevención de prejuicios, la estigmatización y la prevención de nuevas discriminaciones sociales. Además, se trata de generar

---

<sup>34</sup> Grupo Europeo de Ética en Ciencia y Nuevas Tecnologías: *Statement on Artificial Intelligence, Robotics and “Autonomous” Systems*. Comisión Europea, Dirección General de Investigación e Innovación, 2018, pp 11-12.

<sup>35</sup> *Ibidem*, p 16.

beneficios y prosperidad compartidos. El principio de justicia también implica altos estándares de responsabilidad y reparación o remedio efectivo en caso de daño. También se menciona la evaluación del desempeño de la IA para garantizar la disponibilidad de tratamientos nuevos y la atención médica.

Por último, está el principio de transparencia. Este es el principio más básico de los cinco, ya que complementa a los cuatro principios anteriores. Con este principio se debe dar respuesta al ¿Cómo funciona la IA? ¿Quién es el responsable de la forma en que funciona? En definitiva, se trata de entender y dar cuenta de la responsabilidad, ya que, sin transparencia, no se podría llegar a entender el alcance del daño de la actuación de la IA.

#### **4.2. DILEMAS ÉTICOS ASOCIADOS A LA IA.**

Kvalnes define el dilema moral como “*situaciones en las que el tomador de decisiones debe considerar dos o más valores o deberes morales, pero solo puede elegir uno de ellos; por lo tanto, el individuo violará al menos un deber moral importante, independientemente de la decisión.*”<sup>36</sup>

Como resultado, se produce un dilema ético cuando los valores morales y las responsabilidades se encuentran en conflicto y cualquier respuesta posible conlleva una lesión moral. Esto contribuye a cambiar las normas y el sistema de valores sociales. Estos cambian y se convierten en un problema social cuando afectan a una parte importante de la sociedad.

Resulta importante delimitar cuestiones que se han ido viendo a lo largo del trabajo:

- La primera es que la IA está enfocada a la resolución de problemas, pero siempre siguiendo el comportamiento humano desde un punto de vista cognitivo y no emocional.

---

<sup>36</sup> Ø KVALNES, : *Moral Reasoning at Work. Rethinking Ethics in Organizations*. (Second ed.). Capítulo 2 Dilemas Morales. Department of Leadership and Organizational Development, Oslo, Noruega, p 11 Disponible en: <https://doi.org/10.1007/978-3-030-15191-1> , [consulta:13/04/2023]

- La segunda cuestión es que los sistemas IA no son autónomos, sino que son guiados, por lo que es importante tener en cuenta la naturaleza humana en este punto. El uso malicioso que se le pueda dar a las IA siempre centra de conductas inadecuadas de quienes desarrollan las IA o las controlan.
- Ligado al anterior punto está el considerar a las IA como sistemas o tecnologías amorales. Éstas no cuentan con la capacidad de diferenciar el bien del mal ya que toman como suyas la naturaleza que sus creadores hayan programado
- Por último, de la misma forma que las IA tomaran la naturaleza programada por sus creadores, hace considerar que los dilemas éticos que deriven de la IA vendrán de las posibles contradicciones sociales que a su vez derivaran de la programación que le hayan impuesto sus creadores.

De estos puntos se entiende que los usos maliciosos que se le pueda dar a la IA son los que pueden terminar convirtiéndose en dilemas éticos, punto que remarcó un estudio de un conjunto de autores de diversas universidades norteamericanas: *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*<sup>37</sup> siendo este resumido por Mario González Arencibia y Dagmaris Martínez Cardero en su artículo: *Dilemas éticos en el escenario de la inteligencia artificial*.

Estos destacan los posibles usos maliciosos que se le podrían dar a las IA señalando cuatro campos donde podría haber más repercusión: identidad, invasión de los entornos, intimidad y seguridad.

Por consiguiente, los usos maliciosos de la inteligencia artificial, podrían llegar a ser potencialidades perjudiciales para la sociedad. Los errores surgidos

---

<sup>37</sup> BRUNDAGE, M., AVIN, S., CLARK, J., TONER, H., ECKERSLEY, P., GARFINKEL, B., ... & AMODEI, D. *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*. (Primera ed.). Future of Humanity Institute, University of Oxford, Arizona State University, 2018. Citado por: Mario GONZÁLEZ ARENCIBIA y Dagmaris MARTÍNEZ CARDERO: *Dilemas éticos en el escenario de la inteligencia artificial*, Economía y Sociedad vol.25 n.57 Heredia Jan. /jun. 2020, pp 4 – 5. Disponible en: [https://www.scielo.sa.cr/scielo.php?pid=S2215-34032020000100093&script=sci\\_arttext#fn4](https://www.scielo.sa.cr/scielo.php?pid=S2215-34032020000100093&script=sci_arttext#fn4), [consulta:13/04/2023]

los desarrollos tecnológicos de la IA, posicionan en riesgo la dignidad y la vida humana, la libertad de expresión, los derechos fundamentales, así como en este escenario se pueden distinguir riesgos asociados a la creación de sesgos humanos.

Como sucede con la seguridad física y la ciberseguridad, los usos maliciosos también pueden surgir para proteger la soberanía nacional de los Estados y prevenir amenazas que afecten a países y regiones.

## **5. DIVERSIDAD, NO DISCRIMINACION Y EQUIDAD.**

A lo largo de todo el trabajo, sobre todo dentro del ámbito europeo, se ha venido hablando de la importancia de hacer realidad una IA fiable. Una IA fundada en el usuario y diseñada de forma que todas las personas sin distinción de género, capacidad o características puedan emplear los servicios que conforman la IA. Por consiguiente, uno de los pilares de ésta es el principio de no discriminación, ligado al ya mencionado principio de beneficencia.

Pero esto no solo es un principio ético, también lo es jurídico, ya que existe tal mandato jurídico expresado tanto en el ámbito europeo como en el nacional. Dentro del ámbito europeo se encuentra el Convenio Europeo de Derechos Humanos en su artículo 14 y el artículo 21 de la Carta de los Derechos Fundamentales de la UE. En cuanto al ámbito nacional hacer referencia al artículo 14 de la Constitución Española de 1978, donde se prohíbe cualquier tipo de discriminación por razón de religión, sexo, raza, nacimiento o cualquier otra condición o circunstancia personal o social.

### **5.1. ALGORITMOS, SESGOS Y DISCRIMINACIÓN.**

Antes de iniciar conviene definir lo que es el aprendizaje automático o *machine learning*.

El aprendizaje automático es una rama de la inteligencia artificial que se centra en el desarrollo de modelos y algoritmos que permiten a las computadoras, sin ser programadas explícitamente, aprender y mejorar su rendimiento en tareas específicas. Los sistemas de aprendizaje automático utilizan datos y experiencias previas para identificar patrones, realizar predicciones y tomar decisiones en lugar de seguir instrucciones precisas.

Los modelos se entrenan utilizando conjuntos de datos de entrenamiento, que son ejemplos del problema que se desea resolver. Estos datos incluyen características o atributos pertinentes, así como una etiqueta o el resultado esperado. El modelo ajusta sus parámetros internos de manera iterativa durante el entrenamiento para encontrar patrones y relaciones en los datos.

La gran ventaja que presenta es su capacidad de adaptación a nuevas situaciones basándose en el aprendizaje adquirido. Esta toma de decisiones automáticas se realiza mediante algoritmos. Pero el hecho de que estas IA mediante los datos de los que disponen puedan generalizar una respuesta a un hecho, puede desembocar en un sesgo.

Anteriormente ya se vio la diferencia entre sesgo y discriminación, ya que suelen ser dos conceptos que se suelen considerar semejante.

Por un lado, está el sesgo, definido como “*error sistemático en el que se puede incurrir cuando al hacer muestreos o ensayos se seleccionan o favorecen unas respuestas frente a otras.*”<sup>38</sup>. Generalmente es un término muy ligado a la estadística, pero cuando en el resultado obtenido influyen variables personales, deja de ser un sesgo para ser considerada una discriminación.

Existen un gran abanico de sesgos entre los que se pueden encontrar el sesgo histórico, el sesgo de representación o muestreo, pero el que realmente resulta de interés para el trabajo es el “sesgo algorítmico”.

En el sesgo algorítmico se debe tener en cuenta el diseño, el entrenamiento, así como las pruebas del propio algoritmo. Teniendo esto en

---

<sup>38</sup> REAL ACADEMIA ESPAÑOLA: Diccionario de la lengua española, 23.ª ed., [versión 23.6 en línea]. <<https://dle.rae.es>> [28-06-2023].

cuenta, el sesgo algorítmico se dará lugar cuando en el resultado que de un sistema IA hayan influido características personales de un colectivo o de una persona.

En el lado opuesto esta la discriminación, siendo esta “el acto de dar un trato desigual a una persona o colectividad por motivos raciales, religiosos, políticos, de sexo, de edad, de condición física o mental”<sup>39</sup>

La discriminación siempre ha estado relacionada con el contexto geográfico o cultural de un colectivo o persona, es lo que se llama “atributos sensibles”. Los más comunes son: la edad, el estatus social o socioeconómico, la nacionalidad, el género, la raza, la religión entre otros muchos más dependiendo del país.

Siendo este el caso, la discriminación se podría dividir en distintos tipos:

- Discriminación directa: cuando en la decisión de la IA han influido los atributos sensibles. Dentro de esta pueden encontrarse otros tipos de discriminaciones como, por ejemplo, la discriminación grupal, es decir, cuando se discrimina a una persona o a un colectivo por el simple hecho de pertenecer o ser de dicho grupo.
- Discriminación indirecta: para la toma de la decisión no se emplean de forma directa los atributos sensibles, sino otros atributos que están íntimamente relacionados con los atributos sensibles.
- Discriminación algorítmica: está relacionada con la discriminación directa, pero en este caso el sistema IA ante dos posibles resultados, optará por uno desechando el segundo por un atributo sensible que pueda poseer. Dentro de esta pueden existir dos tipos:
  - Discriminación algorítmica directa: al igual que en la discriminación directa, es necesario que en el algoritmo se introduzcan datos referentes a una persona o grupo que puedan tener un valor negativo o desfavorable.

---

<sup>39</sup> REAL ACADEMIA ESPAÑOLA: Diccionario de la lengua española, 23.ª ed., [versión 23.6 en línea]. <<https://dle.rae.es>> [28-06-2023].

- Discriminación algorítmica indirecta: en este caso se introducen datos que en un primer lugar pueden considerarse neutros, pero que pueden desembocar en efectos negativos para las personas o grupos afectados.

Por último, conviene matizar tres conceptos clave que serán relevantes en la sucesión del trabajo. Estos conceptos son: los datos personales, el tratamiento y la elaboración de perfiles.

Los datos personales será toda la información que se posea sobre una física, ya sea por ejemplo su nombre, datos de localización, datos económicos, culturales, sociales, número de identificación, etcétera.

El tratamiento se definirá como cualquier operación encaminada al análisis de los datos personales o el conjunto de estos.

Y por último está la elaboración de perfiles. Se trata de un tratamiento automatizado de los datos personales donde se evaluarán y se analizarán con el fin de predecir características o aspectos importantes de la persona para el correcto desempeño de la IA.

### **5.1.1. ¿Por qué discriminan los algoritmos? El diseño de los algoritmos y elaboración de perfiles**

El diseño de los algoritmos es un factor muy importante en la discriminación derivada del uso de los sistemas IA. Esto es así ya que los algoritmos emplean datos los cuales les han sido suministrados o que han sido recopilados de la sociedad. Y si en esa misma sociedad hay un arraigado componente de desigualdad, exclusión o rechazo a determinados colectivos por razón de sus atributos sensibles, el aprendizaje automático de la IA reproducirá dichos comportamientos en sus decisiones.

Además, se debe tener en cuenta que los algoritmos son programados por personas y en consecuencia puede darse el caso de que éstos plasmen sus prejuicios o sesgos en propio algoritmo. Resultando que los algoritmos estén

cargados de un cierto componente ideológico, lo que influirá en el tratamiento de los datos por parte de la IA. Ya que al final la muestra con la que trabaje el algoritmo sea poco repetitiva

En otro punto esta la elaboración de los perfiles, los cuales toman información, datos de la persona e incluso el comportamiento que tiene en la red con el fin de elaborar un perfil que luego será analizado por el algoritmo. Esta forma de actuar puede conllevar en una segmentación de la población, así como la vulneración de la privacidad de los datos.

Es por ello por lo que con el fin de evitar estos problemas presentados se busca auditar el desarrollo de estos algoritmos, surgiendo el concepto de “responsabilidad algorítmica”. Este concepto viene a significar que los creadores, desarrolladores o coordinadores de los sistemas IA no podrán eximirse de la responsabilidad por la violación de los derechos humanos y el principio de no discriminación por las decisiones autónomas que tomen las IA.

### **5.1.2. Probar la discriminación algorítmica**

Soriano Arnanz<sup>40</sup> expone que, en el caso de la discriminación algorítmica directa, incluso sin acceder al contenido del algoritmo, será bastante sencillo probar el supuesto de discriminación algorítmica directa, en el que la pertenencia al grupo explícitamente incluido determine de manera automática un resultado negativo. Sin embargo, una discriminación algorítmica directa que no afecte claramente a todas las personas del grupo protegido será mucho más complicada y solo se puede resolver accediendo al contenido del algoritmo, es decir, al código fuente.

---

<sup>40</sup> A. SORIANO ARNAZ, “Decisiones automatizadas y discriminación: aproximación y propuestas generales”, pp. 14 y ss. Citado por Mar MORENO REBATO: *Principios éticos y derecho. Autorregulación y/o regulación: Inteligencia artificial (umbrales éticos, derecho y administraciones públicas)*, Capítulo III Diversidad, no discriminación y equidad. Algoritmos y discriminación, Editorial Aranzadi, S.A.U, Pamplona, 2020, p. 58

Algo diferente será en el caso de la discriminación algorítmica indirecta donde al final, la totalidad del sistema será considerado como un criterio neutral que general el resultado discriminatorio.

### **5.1.3. Detección de la discriminación por algoritmos**

Al igual que los algoritmos pueden ser una fuente de discriminación, también pueden ser un fuerte aliado para combatirla.

Algunos ejemplos se pueden encontrar en Twitter y Facebook que emplean algoritmos para localizar y eliminar posibles contenidos discriminatorios, fraudes o discursos de odio en sus plataformas, es decir, el algoritmo se emplea como herramienta para filtrar los contenidos. Pero estas no dejan de ser normas no jurídicas que no tiene relevancia más allá de las plataformas donde se aplican.

Lo cierto es que hoy en día las únicas herramientas jurídicas que se pueden llegar a emplear con este fin son los procedimientos de certificación, evaluación de impacto y auditoria de los algoritmos.

## **6. LA RESPONSABILIDAD Y LA IA**

La responsabilidad al igual que la no discriminación es otro de los pilares que construyen la IA confiable, por lo que es conveniente estudiar el impacto que puede tener estas nuevas tecnologías IA respecto la responsabilidad.

### **6.1 CONCEPTO DE RESPONSABILIDAD.**

Estudiar el concepto de responsabilidad podría llevar mucho tiempo y demasiados matices por lo que para el caso que se estudia bastaría con

descomponer la palabra en dos “respons—abilidad”<sup>41</sup> y con ello se podría definir como “la capacidad de dar respuesta frente a determinadas circunstancias, o bien como la capacidad de responder por determinadas acciones”.

Llegados este punto hay que diferenciar entre la responsabilidad legal y la responsabilidad moral. Cuando una persona comete ya sea un delito o por su conducta provoca un perjuicio que debe ser subsanado, se estaría hablando de responsabilidad legal ya que ese daño tiene que ser reparado.

Pero el enfoque no se puede quedar solo en la reparación de los daños, tiene que ir más allá, a la prevención y a la obligación de prevenirlos. Es la llamada responsabilidad moral. De aquí nace por tanto una de las primeras acepciones de responsabilidad, la de rendir cuentas.

Pero con la llegada de las nuevas tecnologías y hoy en día, con los sistemas IA, este enfoque de responsabilidad legal, moral y de redención de cuentas puede llegar a suponer un sobreesfuerzo, dado que la relación agente/acción/efecto se extiende y se prolonga en el tiempo. Será más complicado medir el alcance de los impactos, así como sus consecuencias. Esto obligará a establecer una nueva dimensión de responsabilidad, al igual que una gobernanza de la tecnología.

## **6.2. LA RESPONSABILIDAD Y SU RELACIÓN CON LA IA.**

Cualquier decisión o acción que haya tomado una IA o un ser humano siguiendo las directrices de una IA tendrán un contenido responsable que podrá repercutir en cuestiones o consecuencias de naturaleza ética.

Aquí se pueden abrir varios caminos, como por ejemplo quien es el responsable, de que es responsable, frente a quien y en base a que.

Respecto quién es responsable. Este punto resulta sencillo cuando solo existe un agente individual al que se le puede atribuir la responsabilidad. Pero la

---

<sup>41</sup> Mónica VILLAS OLMEDA y Javier CAMACHO IBÁÑEZ: *Manual de Ética Aplicada En Inteligencia Artificial*, Capítulo IV Responsabilidad. Madrid: Anaya, 2022, p. 104

situación se torna compleja cuando se introducen más agentes. Es por ello por lo que conviene establecer dos criterios fundamentales para poder atribuir la responsabilidad:<sup>42</sup>

- Vincular la responsabilidad moral al control moral. Esto quiere decir que la responsabilidad deberá recaer sobre el sujeto moral.
- Y comprender que la responsabilidad no desaparece, aunque sea compartida.

Siendo el caso de los sistemas IA el análisis se puede complicar más, ya que la acción se puede delegar en el algoritmo, pero la responsabilidad es diferente y habría que analizar la responsabilidad del programador y de la persona que emplea el sistema IA.

En cuanto de lo que se es responsable hay que acudir a los tres enfoques del razonamiento ético<sup>43</sup>: la ética de la virtud, la ética del deber y la ética utilitarista.

- Desde el punto de vista de la ética de la virtud es la propia persona la responsable de sí misma, de desarrollarse para alcanzar su mejor versión.
- La ética del deber dice, por otro lado, que cada uno somos responsable de nuestros actos, teniendo que aceptar tanto una obra bien hecha como la responsabilidad derivada de una mala. Somos responsables de lo que somos capaces de hacer.
- Finalmente, la ética utilitarista expone que se es responsable de las consecuencias de los actos que uno emprenda.

Este último enfoque es realmente importante de cara a los sistemas IA dada la importancia que adquieren las circunstancias y las consecuencias que en un principio pueden ser previsibles.

Establecer frente a quién se es responsable puede ser algo más complicado si se está en el ámbito de los sistemas IA. Esto es debido a la

---

<sup>42</sup> *Ibidem*, p. 109

<sup>43</sup> *Ibidem*, p. 111

dificultad que puede suponer identificar todas las personas, colectivos que han sido afectados.

Finalmente, en base a qué se es responsable, hay que tener en cuenta que muchos de los preceptos éticos se acaban consolidando en cuestiones legales, esto no quiere decir que sean lo mismo. Pero si entramos en el ámbito de los sistemas IA, la mayoría de los principios que se consideran convenientes para su correcta llevanza suelen venir dados desde los principios éticos y acaban siendo regulados.

Pero ahora puede surgir otra pregunta ¿cómo tomar una decisión o llevar a cabo una acción de manera responsable?

Aquí es donde entran en juego los principios éticos que se han planteado en puntos anteriores.

En primer lugar, el principio de beneficencia, donde se presume que los sistemas IA se basarán en la utilidad, persiguiendo siempre ofrecer el servicio por el que ha sido desarrollado sin desembocar en un perjuicio para el usuario.

En segundo lugar, el principio de autonomía se centra en el respeto a la capacidad que tienen los terceros que han resultado afectados por las decisiones tomadas por el agente.

En tercer y último lugar está el principio de justicia, que ayuda a enfocar y matizar determinadas situaciones que, bajo una visión utilitarista, pueden resultar correctas, pero que sin embargo para un determinado colectivo han podido resultar perjudiciales.

### **6.3 LA GOBERNANZA DE LA IA**

La gobernanza ética se podría definir como *“el conjunto de procesos, procedimientos, cultura y valores que aseguran los más altos estándares de comportamiento, durante todo el ciclo de vida del producto”*<sup>44</sup>. Es por ello un

---

<sup>44</sup> *Ibídem*, p. 124

punto clave para asegurar una IA confiable y que transmita confianza a los ciudadanos y los usuarios.

De su definición se extrae que es un proceso que afecta a todas las fases de desarrollo de los sistemas IA, incluso a las personas encargadas, a las herramientas, procesos, etc. Llegando a comprender aspectos del este como la seguridad, verificabilidad, evaluación de riesgos, protección, requisitos de auditoría interna y externa.

En definitiva, la gobernanza es la encargada de establecer las políticas, procedimientos, controles y evaluaciones.

Por último, se pueden ofrecer algunas recomendaciones con el fin de garantizar una correcta gobernanza ética de los sistemas IA:<sup>45</sup>

- La organización debe establecer unos valores y principios fundamentales que se reflejen en las políticas y procedimientos de la organización. Estos valores y principios deben superar la ley y no limitarse a cumplir estrictamente con las regulaciones actuales. Esto requiere un proceso continuo de reflexión y la creación de un grupo de expertos que se reúne periódicamente para lograrlo.
- Llevar a cabo evaluaciones periódicas de los peligros y el impacto de los sistemas de IA en todas las facetas de la vida de las personas, desde la recopilación de datos hasta la implementación de resultados. Hay que definir claramente los riesgos y los beneficios cuando el uso de datos puede tener un impacto potencial en las personas. Los riesgos deben ser justificados y proporcionales a los beneficios.
- La dirección buscará capacitar y formar a los desarrolladores, así como a los colaboradores que trabajen con los sistemas IA.
- Asimismo la organización deberá ser transparente acerca de los valores éticos que rigen los sistemas IA que desarrollen.

---

<sup>45</sup> *Ibidem*, p. 125

## **7. PRIVACIDAD.**

### **7.1 CONCEPTO DE PRIVACIDAD**

El derecho a la privacidad es un derecho recogido en el artículo 12 de la Declaración Universal de los Derechos Humanos siendo este inherente a las personas e intransferible. Este derecho garantiza que cada persona pueda mantener su vida privada alejada del alcance del resto garantizando de esta forma un espacio de dignidad para el individuo.

Dentro del ámbito tecnológico y de los sistemas IA, este derecho se manifestará en la protección de los datos personales, sobre todo los que sirvan para identificar a un individuo. Estos datos personales pueden ser los nombres y apellidos, la dirección de correo electrónico, la dirección de protocolo de internet, el domicilio, etc.

Pero estos datos personales también pueden conformar la identidad del individuo, en este caso vendrán a definir la identidad digital. Es decir, la identidad digital será la definida por una serie de datos e interacciones en el ámbito digital que con el paso del tiempo habrá conformado una representación de la identidad propia del individuo. En este punto es donde se puede apreciar la conexión entre identidad y privacidad.

La pregunta que puede asaltar en este punto es, siendo el ámbito digital donde están estos datos, ¿a quién le pertenecen estos datos?

Partiendo de la base de que los datos conforman la identidad digital del individuo, éstos no pueden ser sometidos a posesión o propiedad, pero si a los derechos humanos que protegen dicha identidad. Es en este punto donde pueden surgir dificultades para responder a la pregunta, ya que la separación del interés por los datos como propiedad, al interés en la intimidad personal, es algo difuso.

Teniendo esto en cuenta, el interés va a estar en la protección del flujo de datos, es decir, que, pese al continuo uso que se le dé a los datos personales y a la información, no se menoscabe la privacidad del individuo.

## **7.2. GESTIÓN DE LA PRIVACIDAD Y DE LOS DATOS.**

Para hablar de la gestión de los datos por parte de los sistemas IA habrá que acudir al Reglamento general de protección de datos 679/2016.

En el ámbito de las IA destaca el artículo 22, siendo en este artículo donde se establecen algunos puntos que pueden ser aplicables a los sistemas IA referidos sobre todo al tratamiento automatizado de datos.

Principalmente se reconoce en el punto 1 del artículo 22 el derecho de los individuos a no ser objeto de decisiones que hayan sido tomadas siguiendo un tratamiento automatizado de los datos, por ejemplo, la elaboración de perfiles, y que estas tengan efectos jurídicos.

A su vez, el punto 4 del mismo artículo, donde se prohíbe el tratamiento de datos personales que pueden ser considerados sensibles, son los siguientes:

- origen étnico o racial.
- las opiniones políticas.
- las convicciones religiosas o filosóficas, o la afiliación sindical.
- el tratamiento de datos genéticos, datos biométricos dirigidos a identificar de manera unívoca a una persona física.
- datos relativos a la salud o datos relativos a la vida sexual o las orientaciones sexuales de una persona física.

En este caso cabría hacer una matización, ya que se permitirá el tratamiento de estos datos cuando:

- el individuo haya otorgado su consentimiento.
- haya expuesto sus datos de forma pública.
- cuando la ley permita su tratamiento por razones de interés público
- cuando la ley garantice las medidas de protección idóneas para el tratamiento de datos sensibles.

En definitiva, para evitar el tratamiento de datos de alto riesgo que puedan suponer una vulneración de los derechos y libertades de los individuos se emplea la denominada “seudonimización”, definida en el artículo 4.5 del Reglamento como *“el tratamiento de datos personales de manera tal que ya no puedan atribuirse a un interesado sin utilizar información adicional, siempre que dicha información adicional figure por separado y esté sujeta a medidas técnicas y organizativas destinadas a garantizar que los datos personales no se atribuyan a una persona física identificada o identificable”*.

Este mecanismo está enfocado en lo ya mencionado anteriormente, en la protección del flujo de datos.

## 8. CONCLUSIONES

Es algo indudable que los múltiples usos y aplicaciones que se puede dar a la IA representan un gran potencial, ya sea para el desarrollo social, así como para el avance de la industria y la tecnología. Sin embargo, es necesario que los Estados sean los encargados de definir un código ético sobre la IA reflejando en el mismo cuestiones éticas sobre la forma en que se está empleando la tecnología, y cómo está afectando la sociedad.

Algunos de estos posibles efectos los expone el grupo independiente de expertos de alto nivel sobre inteligencia artificial en su informe llamado “Directrices éticas para una IA fiable” dando algunos ejemplos de cuales son algunas de las oportunidades o precauciones que se pueden plantear con la IA. Entre las oportunidades que aprecian los expertos están los avances en salud y bienestar social, en la lucha contra el cambio climático y sobre todo en la educación. Pero las precauciones pueden resultar más alarmantes, por ejemplo, que se usen los sistemas IA como sistemas de armas, o que se empleen para vulnerar los derechos fundamentales de los ciudadanos.

Es por ello por lo que este gran avance de las tecnologías ha hecho evidente la necesidad de evolucionar en el ámbito jurídico y ético para de esta forma promover acciones que estén en línea con los valores sociales y el bienestar de las personas. Esto tiene un impacto en el diseño y la distribución de los resultados, así como en la construcción y gestión del ecosistema digital.

Es fundamental abordar los desafíos actuales mediante una práctica social basada en la responsabilidad individual y colectiva. En lugar de priorizar los posibles fines maliciosos que pueden perjudicar el bienestar social, la ciencia de los datos debe reconsiderarse para proteger los intereses humanos. Siendo necesario que se inicien políticas gubernamentales, respuestas institucionales y enfoques humanísticos para conseguir dicho objetivo.

De esta forma se reduda en la necesidad de que este código de ética sobre la IA tome como base la defensa y garantía de los derechos fundamentales, la autonomía del ser humano, la dignidad y la supervisión, transparencia y explicabilidad de la IA.

Pero no es suficiente con asentar las bases éticas y jurídicas, los profesionales del derecho también deben avanzar de la mano con estas nuevas tecnologías basadas en los sistemas IA. Éstos pueden suponer un gran avance dentro de los servicios legales ayudando a resolver situaciones con mayor agilidad, contribuyendo de esta forma a una justicia rápida y eficiente.

## 9. BIBLIOGRAFIA

CASTILLO, Enrique, GUTIÉRREZ, José Manuel y HADI, Ali S: *Sistemas expertos y modelos de redes probabilísticas*, Academia de Ingeniería, 1997.

Comisión Europea (2018): *Artificial intelligence: Commission kicks off work on marrying cutting-edge technology and ethical standards*. Comunicado de prensa. Disponible en:

[https://ec.europa.eu/commission/presscorner/detail/es/IP\\_18\\_1381](https://ec.europa.eu/commission/presscorner/detail/es/IP_18_1381),

[consulta: 02/04/2023]

Comisión Europea (2020): *Libro Blanco sobre la inteligencia artificial - un enfoque europeo orientado a la excelencia y la confianza*, pp. 1-31. Disponible en:

<https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52020DC0065>, [consulta:05/04/2023]

Comisión Europea. (s. f): *Regulatory Framework for Artificial Intelligence*. Disponible en:

<https://digital-strategy.ec.europa.eu/es/policies/regulatory-framework-ai>, [consulta:05/04/2023]

Comité Económico y Social Europeo: *Inteligencia artificial: las consecuencias de la inteligencia artificial para el mercado único (digital), la producción, el consumo, el empleo y la sociedad*, Diario Oficial de la Unión Europea C 288/1, 31 de agosto de 2017, pp. 1-9

DELGADO, Miguel: *La inteligencia artificial. Realidad de un mito moderno*, Granada, Servicio de Publicaciones de la Universidad de Granada, 1996.

DÍAZ, María José, HERRERA TRIGUERO, Francisco, CORDÓN GARCÍA, Oscar: *Una visión actual de la inteligencia artificial: recorrido histórico, datos y aprendizaje confiabilidad y datos*, en PERALTA, Alfonso, SALVADOR TORRES, Leopoldo y HERRERA, Francisco (ed.): *El derecho y la Inteligencia*

*Artificial*, Título II Aplicaciones técnicas de la inteligencia artificial, Universidad de Granada, Granada, 2022, pp. 51-80.

GONZÁLEZ ARENCIBIA, Mario y MARTÍNEZ CARDERO, Dagmaris: *Dilemas éticos en el escenario de la inteligencia artificial*, Economía y Sociedad vol.25 n.57 Heredia Jan. /jun. 2020, pp 1–11. Disponible en:

[https://www.scielo.sa.cr/scielo.php?pid=S2215-34032020000100093&script=sci\\_arttext#fn4](https://www.scielo.sa.cr/scielo.php?pid=S2215-34032020000100093&script=sci_arttext#fn4), [consulta:13/04/2023]

Grupo Europeo de Ética en Ciencia y Nuevas Tecnologías: *Statement on Artificial Intelligence, Robotics and “Autonomous” Systems*. Comisión Europea, Dirección General de Investigación e Innovación, 2018.

HERNÁNDEZ GIMÉNEZ, María: “Inteligencia artificial y derecho penal”, *Actualidad Jurídica Iberoamericana Nº 10 bis, junio 2019*, pp. 794-843

HINTZE, A (2016): “Understanding the four types of AI, from reactive robots to self-aware beings”. *The Conversation*, 14. Disponible en:

<https://theconversation.com/understanding-the-four-types-of-ai-from-reactive-robots-to-self-aware-beings-67616>, [consulta: 30/03/2023]

HUESO, Lorenzo Cotino: “Ética en el diseño para el desarrollo de una inteligencia artificial, robótica y big data confiables y su utilidad desde el derecho”. *Revista catalana de dret públic*, 2019, vol. 58, pp 30-48.

KVALNES, Ø: *Moral Reasoning at Work. Rethinking Ethics in Organizations*. (Second ed.). Capítulo 2 Dilemas Morales. Department of Leadership and Organizational Development, Oslo, Noruega, p 11 Disponible en: <https://doi.org/10.1007/978-3-030-15191-1> , [consulta:13/04/2023]

Ministerio de Economía y Transformación Digital. (s. f.): *Carta de los Derechos Digitales*. Disponible en:

[https://portal.mineco.gob.es/RecursosArticulo/mineco/ministerio/participacion\\_publica/audiencia/ficheros/SEDIACartaDerechosDigitales.pdf](https://portal.mineco.gob.es/RecursosArticulo/mineco/ministerio/participacion_publica/audiencia/ficheros/SEDIACartaDerechosDigitales.pdf), [consulta:10/04/2023]

Ministerio de asuntos económicos y transformación digital: *Inteligencia artificial*. Disponible en: <https://portal.mineco.gob.es/es-es/ministerio/areas-prioritarias/Paginas/inteligencia-artificial.aspx>, [consulta:05/04/2023]

Ministerio de asuntos económicos y transformación digital: *Plan Estratégico de Inteligencia Artificial*, Madrid, 2020.

MORALES MORENO, África María: “Inteligencia artificial y derecho penal: primeras aproximaciones”, *Revista jurídica de Castilla y León* (53), pp. 177-202

MORENO REBATO, Mar: *Inteligencia artificial (umbrales éticos, derecho y administraciones públicas)*, Aranzadi, Pamplona, 2020.

OCDE (s. f): *Artificial Intelligence (AI) Principles*. OECD.AI. Disponible en: <https://oecd.ai/en/ai-principles>, [consulta: 30/03/2023]

OCDE (s. f): *Recommendation of the Council on Artificial Intelligence*, OECD/LEGAL/0449. Disponible en: <https://oecd.ai/en/assets/files/OECD-LEGAL-0449-en.pdf>, [consulta: 30/03/2023]

Parlamento Europeo (Bruselas). ¿Qué es la inteligencia artificial y como se usa? [en línea]: en Portada / Sociedad: Noticias Parlamento Europeo. Creado el 8 de septiembre de 2020, actualizado el 23 de marzo de 2021, [Consulta: 28 mar. 2023]. Disponible en:

<https://www.europarl.europa.eu/news/es/headlines/society/20200827STO85804/que-es-la-inteligencia-artificial-y-como-se-usa>, [consulta:28/03/2023]

Parlamento Europeo: *Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión sobre normas de Derecho civil sobre robótica (2015/2103(INL))*, 2017.

PEGUERA POCH, Miquel: *En búsqueda de un marco normativo para la inteligencia artificial*, en Agustín Cerrillo y Martínez y Miguel Peguera Poch (eds.): *Retos jurídicos de la inteligencia artificial*, Título I Impacto, ética y marco legal, Aranzadi, Pamplona, 2020, pp. 41-56

Propuesta de Reglamento del Parlamento Europeo y del Consejo estableciendo normas armonizadas sobre inteligencia artificial (Ley de Inteligencia Artificial) y modificando ciertos actos legislativos de la Unión. Bruselas, 21-4-2021. Disponible en:

<https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=celex%3A52021PC0206>, [consulta:05/04/2023]

REAL ACADEMIA ESPAÑOLA: *Diccionario de la lengua española*, 23.<sup>a</sup> ed., [versión 23.6 en línea]. Disponible en: <https://dle.rae.es>, [28/03/2023].

Resolución del Parlamento Europeo, de 20 de octubre de 2020, con recomendaciones destinadas a la Comisión sobre un marco de los aspectos éticos de la inteligencia artificial, la robótica y las tecnologías conexas 2020/2012INL. Disponible en:

[https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275\\_ES.html](https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_ES.html), [consulta:05/04/2023]

SOLAR CAYÓN, José Ignacio (2020): “La inteligencia artificial jurídica: nuevas herramientas y perspectivas metodológicas para el jurista”, *Revus. Journal for Constitutional Theory and Philosophy of Law/Revija za ustavno teorijo in filozofijo prava*, (41). Disponible en:

<https://journals.openedition.org/revus/6547>, [consulta 28/03/2023]

Supervisor Europeo de Protección de Datos: *Dictamen 4/2015. Hacia una nueva ética digital. Datos, dignidad y tecnología*, 2015.

Telefónica (2018): *Principios para la Inteligencia Artificial*. Disponible en: <https://www.telefonica.com/es/wp-content/uploads/sites/4/2021/08/principios-ai-esp-2018.pdf>, [consulta:10/04/2023]

UNESCO. (2021): *Ética de la inteligencia artificial*. Oficina de la UNESCO en Montevideo. Disponible en:

[https://es.unesco.org/fieldoffice/montevideo/EticaInteligenciaArtificial#:~:t\\_ext=La%20Recomendaci%C3%B3n%20proh%C3%ADbe%20expl%C3%ADcita](https://es.unesco.org/fieldoffice/montevideo/EticaInteligenciaArtificial#:~:t_ext=La%20Recomendaci%C3%B3n%20proh%C3%ADbe%20expl%C3%ADcita)

[mente%20el,se%20utilizan%20de%20forma%20generalizada,](#)

[consulta:

30/03/2023]

VILLAS OLMEDA, Mónica y CAMACHO IBÁÑEZ, Javier: *Manual de Ética Aplicada En Inteligencia Artificial*. Madrid: Anaya, 2022.