



# SUPERANDO LOS RIESGOS DE UNA IA SIN ÉTICA NI REGULACIÓN

Grupo de Políticas Públicas y Regulación  
Colegio Oficial de Ingenieros de  
Telecomunicación, COIT



Colegio Oficial  
Ingenieros de  
Telecomunicación

## AUTORES

Eduardo Rodríguez Lorenzo  
José Antonio Portilla Figueras (Coordinador)  
Juan Santaella Vallejo  
Noelia Miranda Santos  
Sonia Castillo Triguero

## COORDINADOR DEL INFORME

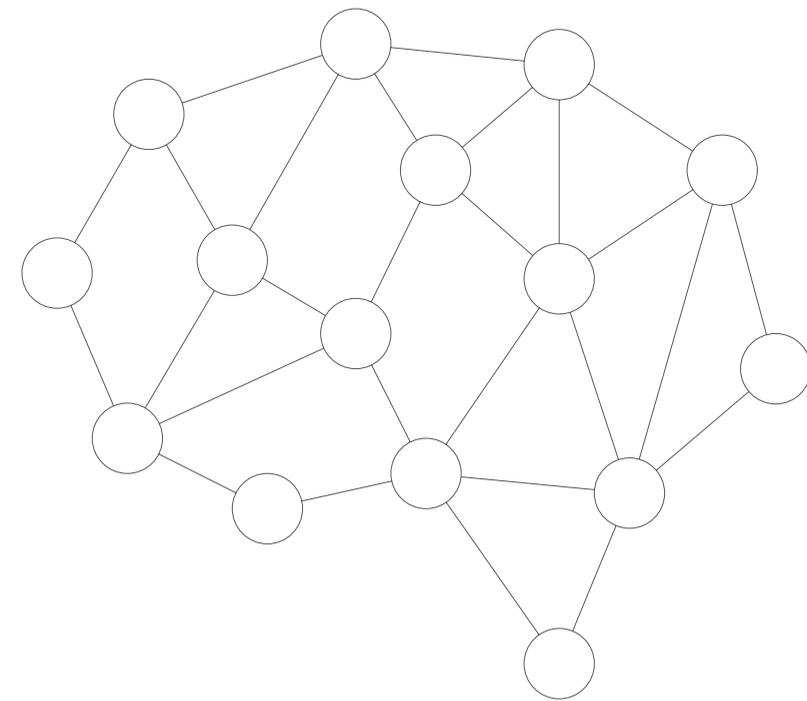
José Antonio Portilla Figueras

## Edita

Colegio Oficial de Ingenieros de Telecomunicación, COIT  
Calle Almagro, 2, 1ª Izqda, 28010 Madrid

1ª edición: enero 2020

Edición y diseño  
ICS Comunicación



# Índice

---

06	<b>INTRODUCCIÓN</b>
08	<b>CAPÍTULO 1</b> Técnicas y tecnologías de Inteligencia Artificial
18	<b>CAPÍTULO 2</b> Ejemplos de aplicación de la Inteligencia Artificial
26	<b>CAPÍTULO 3</b> Una reflexión sobre las implicaciones sociales y éticas de la IA
40	<b>CAPÍTULO 4</b> Estrategias Nacionales de IA
50	<b>CONCLUSIONES</b>
56	<b>REFERENCIAS</b>

---

# Introducción

Inteligencia Artificial (IA), uno de los términos más de moda actualmente. Sólo hace falta mirar una herramienta como *Google Trends* para observar que el interés por estas dos palabras en España no ha bajado de un ratio de 75/100 en el último año en España. Pero no sólo eso. Consultoras y empresas de talla mundial como Gartner, PWC, IBM, CISCO, Capgemini, Accenture o Ericsson, por citar algunas, establecen la inteligencia artificial como una de las tendencias tecnológicas en el futuro cercano, ya sea por sí misma o ya sea aplicada a uno de los múltiples (casi innumerables) ámbitos en los que puede ser utilizada.

Entonces ¿qué es la Inteligencia Artificial? Empecemos por el purismo más absoluto, cuando J. McCarthy, M. L. Minsky, N. Rochester y (el gran adorado por los “Telecos”) C. E. Shannon hacen su propuesta de proyecto de investigación sobre inteligencia artificial en el Dartmouth College (al que pertenecía McCarthy). En la propuesta<sup>1</sup> se indica, en su primer párrafo: “*An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves*”. Y casi en el último párrafo se expone: “*For the present purpose the artificial intelligence problem is taken to be that of making a machine behave in ways that would be called intelligent if a human were so behaving*.” Es decir, se trata de diseñar métodos, mecanismos, técnicas, algoritmos que permitan a los computadores comportarse ante una determinada situación de manera similar al comportamiento de una inteligencia humana.

¿Qué es para nosotros la inteligencia artificial? Para el común de los mortales es una caja negra en la que ya sea en modo local o conectada a Internet (esa caja se torna aún más negra en ocasiones) se producen una serie de operaciones. Estas operaciones se traducen en el entorno real a que sucedan cosas del tipo como que una vez que yo he visto un conjunto de series en la televisión, se me ofrezcan series que me pueden interesar, o que una vez que yo compro en un portal de internet una serie de artículos, se me ofrezcan en sitios web o redes sociales, teóricamente independientes productos similares o complementarios. También entendemos que es algo que maneja los coches autónomos y que, cada vez más, sale en películas y en las noticias de televisión, y que algunas de ellas pueden ser un poco catastrofistas.

Los breves ejemplos anteriores son una muestra de que la inteligencia artificial está presente en nuestras vidas, ya en el momento actual, y que va a estar mucho más presente, casi omnipresente, en el futuro. Y esto hace que la IA exceda el plano meramente técnico y aparezcan cuestiones económicas (todos podemos recordar el llamativo caso del algoritmo en el consejo de administración<sup>2</sup>), en los trabajos (según la OECD el 14% de los trabajos están en riesgo de automatización y el 32% de los trabajos se transformarán radicalmente<sup>3</sup>), en aspectos legales (la propiedad de los datos de partida con los que funcionan los algoritmos, más aún en el caso de datos de carácter personal), sociológicos, políticos e incluso filosóficos (las famosas tres leyes de Asimov, o la respuesta a la pregunta de qué haría un coche controlado por inteligencia artificial si para no atropellar a una mujer con un niño tiene que provocar un accidente donde va a morir el dueño del coche).

<sup>1</sup> <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>

<sup>2</sup> P. Pardo, “Un algoritmo, sentado en el consejo de un fondo chino”, *El mundo*, 9/06/2014, fecha de último acceso 1 de Junio de 2019

<sup>3</sup> OECD “The future of Work: OECD Employment Outlook 2019”.

Como vemos se nos abre un nuevo mundo de consecuencias todavía por determinar. El Colegio Oficial de Ingenieros de Telecomunicación es consciente de que los Ingenieros de Telecomunicación tienen un papel muy relevante que jugar en el desarrollo técnico de este nuevo mundo, pero además como entidad al servicio de la Sociedad entendemos que debemos arrojar un poco de luz sobre algunos de los aspectos anteriormente esbozados, e incluso plantear nuevas preguntas e inquietudes. Es por ello que desde el Grupo de Políticas Públicas y Regulación del COIT nos hemos planteado la elaboración de un informe sobre inteligencia artificial, que desde la visión de técnicos en la materia proporcione una visión de alto nivel sobre un amplio espectro de temas relacionados. De esta manera, en la primera sección daremos una introducción a las técnicas y tecnologías que mayor impacto están teniendo actualmente. Seguidamente abordaremos dónde se puede aplicar la IA (correcto, en todas partes). También abordaremos aspectos éticos y sociales, y finalmente haremos una revisión de las estrategias nacionales de inteligencia artificial.

En la elaboración del informe hemos solicitado (como ya se ha podido observar) la colaboración de profesionales de distintos sectores para que nos den una cierta visión lateral de la inteligencia artificial, con motivo expreso de crear inquietudes de las que puedan surgir líneas futuras de reflexión, o simplemente, tener la visión de otros profesionales en la materia.

## OPINIÓN

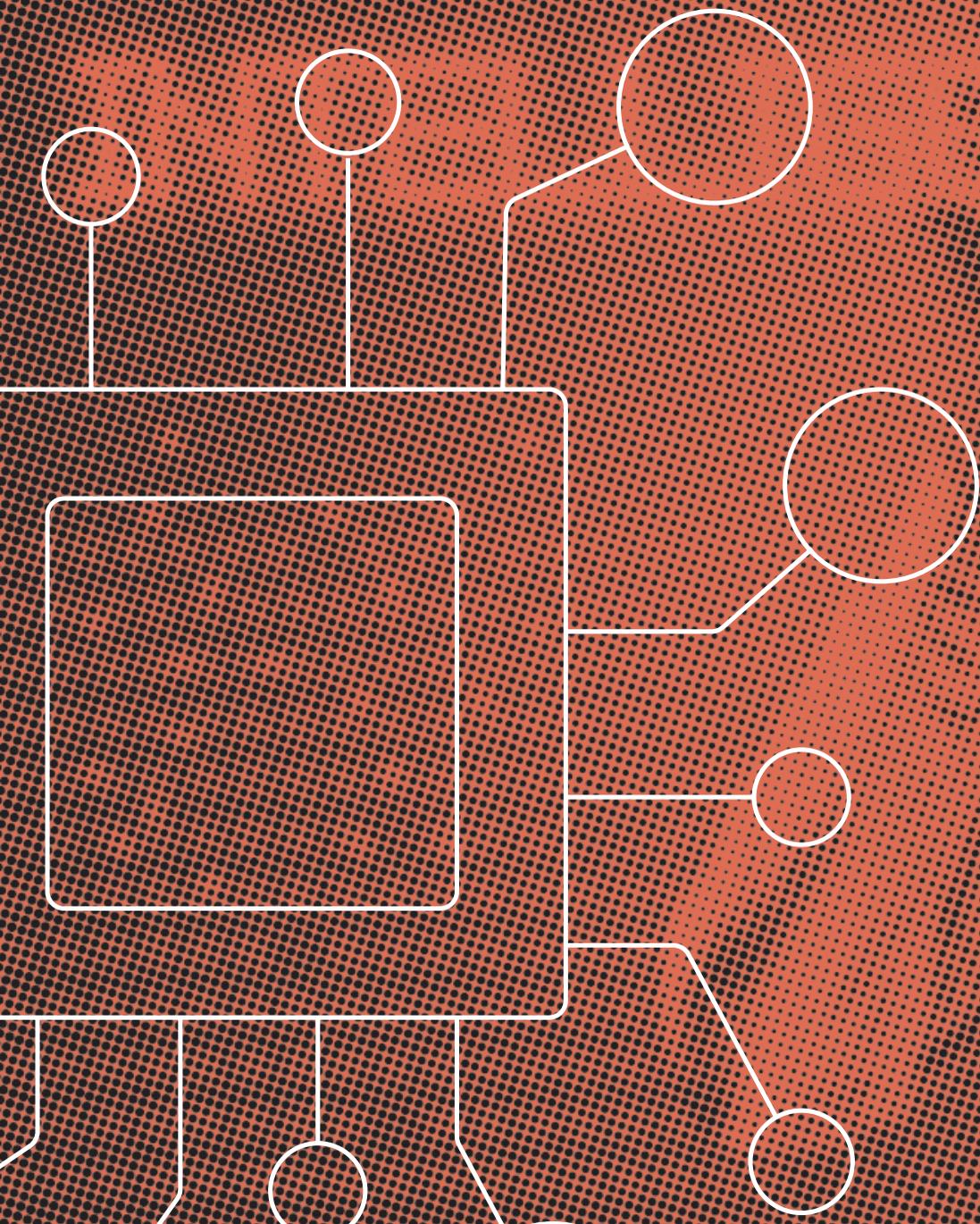


### Javier Torregrosa

GRADUADO EN PSICOLOGÍA Y MÁSTER EN METODOLOGÍA EN LAS CIENCIAS DEL COMPORTAMIENTO. UNIVERSIDAD AUTÓNOMA DE MADRID

El presente y, especialmente, **el futuro de las ciencias del comportamiento se encuentra ligado al desarrollo técnico que la inteligencia artificial** (entre otras técnicas computacionales) representa. Sus aplicaciones van desde la creación de modelos de ayuda a la toma de decisiones en seguridad e inteligencia, a modelos capaces de obtener patrones conductuales de individuos bajo una determinada patología psicológica, o la predicción de puntos calientes en infraestructuras y territorios. Especialmente, la psicología y la inteligencia artificial viven en la actualidad una situación de apoyo recíproco: mientras que **la inteligencia artificial aporta nuevas herramientas** para el estudio de fenómenos anteriormente complejos de enfocar desde el punto de vista conductual (tales como la creación de modelos de procesos cognitivos), **la psicología aporta nuevos enfoques, explicaciones y sustento teórico** a las conclusiones alcanzadas a través del uso de inteligencia artificial. Y, sin duda, esta relación continuará fortaleciéndose a medida que descubramos hasta qué punto ambas disciplinas pueden beneficiarse la una de la otra ante los nuevos (y viejos) retos que surjan durante su desarrollo.

TÉCNICAS Y  
TECNOLOGÍAS DE  
INTELIGENCIA  
ARTIFICIAL



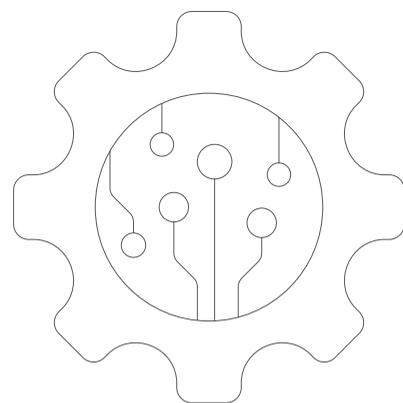
## ¿Por qué ahora?

Si tuviéramos que guiarnos por la prensa generalista o las actuaciones de políticas públicas parecería que la IA es un fenómeno que hubiera comenzado en los últimos dos o tres años, pero si se busca en la literatura vemos que muchas de estas técnicas, según la tecnología, empiezan en las décadas de los 60s o 70s del siglo pasado como, por ejemplo, con las redes neuronales o los algoritmos evolutivos<sup>4</sup>.

Entonces, ¿por qué nos encontramos con un sinfín de estrategias nacionales en IA, si estas tecnologías vienen de hace más de 50 años? La razón viene motivada porque ha habido una mejora en dos elementos básicos: disponibilidad de datos y capacidad de computación.

El incremento de la disponibilidad de datos viene ligado a diversos fenómenos, entre ellos la tan moderna transformación digital y la existencia de elementos ciberfísicos o sensorizados. Estas dos realidades han incrementado exponencialmente el número de datos en cualquier tipo de área o sector. Al disponer de gran cantidad de datos se puede realizar el entrenamiento (aprendizaje) de estos sistemas. Debemos tener en cuenta que estos sistemas resuelven problemas muy complejos, con muchas variables (en lenguaje técnico, dimensiones) y por lo tanto para encontrar la solución más adecuada en todo el espacio de soluciones posible se precisa una cantidad ingente de datos de partida desde los que la técnica de inteligencia artificial pueda “aprender”.

El otro factor es la potencia de cálculo. Para demostrar la evolución de la misma solo hay que recordar la “Ley de Moore”, que por ahora sigue vigente y que indica que aproximadamente cada dos años se duplica el número de transistores en un microprocesador. En 26 años el número de transistores en un chip se ha incrementado 3.200 veces (pensemos que el ordenador del Apolo 11 era mucho menos potente que cualquiera de nuestros Smartphone). Esta potencia de cálculo permite tener la capacidad de computación necesaria para que los sistemas de inteligencia artificial encuentren los patrones subyacentes en los datos de entrada y logren alcanzar la solución al problema.



## Más allá del hype

Como ya hemos introducido, el concepto no es nuevo, sino que se trata de una evolución en el mismo ámbito propulsada por la aparición de tecnologías que nos permiten su desarrollo. Sin embargo, no debemos pensar que es una evolución lineal, sino en una especie de espiral, dentro del mismo paradigma tecnológico donde el conjunto de técnicas que componen la IA son periódicamente revisadas y mejoradas. En los últimos 20 años, hemos dado varias vueltas en esa espiral, (aunque siempre avanzando y mejorando) y en cada una de ellas, hemos introducido un nuevo término, etiqueta o aplicación que ha sido el “hype” del momento.

Alrededor de 2005 se popularizó en la tecnología el término “smart”, y surgieron aunque fuera desde el marketing los “smart home”, “smart TV”, “smartphone”, .... Si se mira en Wikipedia<sup>5</sup> la primera entrada es de 2008 y lo identifican como “Sistemas que incorporan funciones de detección, actuación y control para describir y analizar una situación, y tomar decisiones basadas en los datos disponibles de forma predictiva o adaptativa, y de esta manera realizar acciones inteligentes. En la mayoría de los casos, la “inteligencia” del sistema puede atribuirse a un funcionamiento autónomo basado en el control de bucle cerrado, la eficiencia energética y las capacidades de red”. Avanzando en el tiempo, en 2010 encontramos “Internet of Things”, dispositivos con capacidad de comunicarse e interactuar con otros”, o en 2013 “Big Data”, donde se habla de análisis del comportamiento del usuario, extrayendo formulación de predicciones a través de los patrones observados. Todas estas definiciones se parece mucho a la que ofrece la Comisión Europea para IA en su comunicación<sup>6</sup>.

Queremos indicar que la Inteligencia Artificial es un conjunto de tecnologías, no es una única tecnología. Según quien hable de ella incluye más o menos. Por ejemplo, la Comisión Europea está tomando un enfoque muy amplio de lo que es IA, ya que consideran 5G, supercomputación y otras tecnologías. Si se observa la Estrategia Española de I+D+D en Inteligencia Artificial<sup>7</sup>, adquiere un sentido más estricto donde se consideran tecnologías como:

- > Big data y análisis de datos
- > Planificación y deducción automática
- > Sistemas inteligentes de teledetección
- > Aprendizaje automático
- > Procesamiento de flujo de datos
- > Sistemas inteligentes de tutorización
- > Cooperación hombre-máquina
- > Razonamiento, lógica y ontologías
- > Sistemas multiagente
- > Diseño de asistentes cognitivos
- > Robótica
- > Tecnologías, procesamiento y generación del lenguaje natural
- > Modelado basado en agentes inteligentes
- > Sistemas de recomendación
- > Procesamiento y aprendizaje de estructuras en imágenes
- > Optimización heurística
- > Sistemas inteligentes de predicción
- > Visión por computador
- > Sistemas empotrados

<sup>4</sup> J. Holland (1975). Adaptation in Natural and Artificial Systems, University of Michigan Press, Ann Arbor.

<sup>5</sup> [https://en.wikipedia.org/w/index.php?title=Smart\\_system&offset=20080930184831&action=history](https://en.wikipedia.org/w/index.php?title=Smart_system&offset=20080930184831&action=history)

<sup>6</sup> [https://ec.europa.eu/newsroom/dae/document.cfm?doc\\_id=56018](https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56018)

<sup>7</sup> <http://www.ciencia.gob.es/portal/site/MICINN/menuitem.26172fcf4eb029fa6ec7da6901432ea0/?vgnnextoid=70fcd-b77ec929610VgnVCM1000001d04140aRCRD>

## Alguno de los pilares tecnológicos

Hablar de todas estas tecnologías se sale fuera del alcance de este monográfico, pero sí queremos resaltar algunas de ellas mencionando sus principales fortalezas y su influencia en el desarrollo del paradigma tecnológico y que en el fondo engloban o se apoyan en las citadas.

### > Aprendizaje automático

Si algo se relaciona con la Inteligencia Artificial es esta rama de la técnica de aprendizaje automático<sup>8</sup> (*Machine Learning*), porque es una de las que ha permitido lograr sus éxitos más mediáticos, como cuando AlphaGo, perteneciente a Google DeepMind<sup>9</sup> derrotó al campeón del mundo de ajedrez.

La victoria de AlphaGo se basó en aprendizaje profundo (*Deep Learning*). Esta técnica es una particularización del aprendizaje automático donde se utilizan métodos que permiten a las máquinas “aprender” de forma automática un conjunto de datos proporcionados para generar un modelo que le permita tomar decisiones o hacer predicciones.

#### La mayoría de los algoritmos de aprendizaje automático se pueden dividir en dos categorías de “aprendizaje supervisado” y “aprendizaje no supervisado”

Los algoritmos de aprendizaje automático se pueden dividir en dos categorías de “aprendizaje supervisado” y “aprendizaje no supervisado”, dependiendo de si el conjunto de entrenamiento está supervisado (es decir, con una información del objetivo asociada con el objetivo) o no supervisado. Estas técnicas eliminan la necesidad de modelar y codificar el problema.

Las técnicas más populares, entre otras, son: regresiones lineales/lógica; árbol de decisión; máquinas de vectores de soporte; análisis de componentes principales; K-means; análisis de componentes independientes o clasificador Bayesiano.

Estas técnicas se basan en un conjunto de datos de entrenamiento y otros de test. A través de los datos de entrenamiento, el sistema va generando un modelo que se realimenta si la solución es la esperada o no. La mayoría de

### > Cooperación hombre-máquina

KUKA, uno de los principales fabricantes de robots industriales del mundo, dice en su página web<sup>10</sup>: “en la fábrica del futuro, no se distingue entre puestos de trabajo automatizados y puestos de trabajo manuales. Las personas y las máquinas cooperan perfectamente: **sin distinción, sin espacio de protección**”. Hay que tener en cuenta que esta es una particularización de este concepto, ya que la cooperación se puede entender desde un robot a ideas más futurísticas, como la conexión directa a una máquina como *neura-link* del empresario Elon Musk<sup>11</sup>.

La cooperación hombre-máquina incluye un conjunto de tecnologías como la percepción del ambiente. Un ejemplo de ello son los robots creados por Boston Dynamics.

## OPINIÓN



### Óscar Cordón

CATEDRÁTICO DE LA UNIVERSIDAD, UNIVERSIDAD DE GRANADA

La principal aplicación de la IA es crear sistemas y máquinas para automatizar tareas que requieran un comportamiento inteligente. Es necesario formalizar y programar los mecanismos del pensamiento humano y la conducta inteligente. Esto implica habilidades específicas como razonamiento, percepción, aprendizaje y uso del lenguaje que redundan en diversas tecnologías de IA. En razonamiento destacamos los sistemas basados en reglas, donde la lógica fuzzy permite modelar la incertidumbre y computar con palabras, y el razonamiento probabilístico con modelos como las redes Bayesianas. En aprendizaje automático se incluyen redes neuronales (como *deep learning*), *clustering*, árboles de decisión, reglas de asociación y aprendizaje evolutivo, entre otros. La percepción contempla la visión artificial con procesamiento y reconocimiento de imagen. El lenguaje integra el procesamiento y generación del lenguaje natural, así como la minería de textos. Es importante la vertiente global de la IA explicativa que persigue que los sistemas obtenidos puedan explicar el proceso de razonamiento seguido en sus decisiones.

<sup>8</sup> <https://www.deeplearningitalia.com/una-breve-mirada-al-aprendizaje-profundo-y-el-aprendizaje-automatico-2/>

<sup>9</sup> <https://www.nature.com/articles/nature16961>

<sup>10</sup> <https://www.kuka.com/es-es/tecnologia/C3%ADas/cooperacion/C3%B3n-hombre-robot>

<sup>11</sup> <https://www.neuralink.com>

## > Tecnologías, procesamiento y generación del lenguaje natural

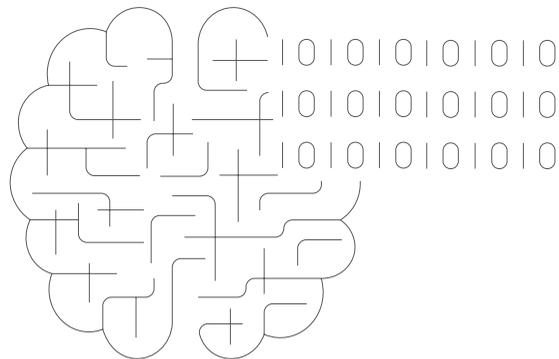
Actualmente, se han popularizado los asistentes de voz, como pueden ser Alexa, Google Home o Siri. Estos robots conversacionales son una muestra del avance que han sufrido en los últimos años las técnicas de procesamiento de lenguaje natural (PLN). Esta mejora viene motivada por pasar de sistemas basados en reglas a basarse principalmente en sistemas estadísticos.

### A nivel nacional existe desde el 2015 el Plan de Impulso de las Tecnologías del Lenguaje

El PLN tiene como elemento clave la creación de corpus lingüísticos que son la base para entrenar los sistemas, que se basan en un conjunto de tareas, como son, entre otras: Grammar induction, procesos de lematización; Part-of-speech ; Parsing, etc.

Las aplicaciones de PLN son múltiples: traducción máquina; reconocimiento de entidades (nombres); reconocimiento óptico de caracteres; análisis de sentimiento; text-to speech, etc.

Se debe destacar que a nivel nacional existe desde el 2015 el Plan de Impulso de las Tecnologías del Lenguaje<sup>12</sup>. Tiene como objetivo fomentar el desarrollo del procesamiento del lenguaje natural, la traducción automática y los sistemas conversacionales en lengua española y lenguas cooficiales, aprovechando la barrera natural que es el idioma para generar una industria nacional.



<sup>12</sup> <https://www.plantl.gob.es/Paginas/index.aspx>

## > Visión artificial

Es una rama de la inteligencia artificial que trata de que las computadoras puedan extraer información automáticamente a partir de imágenes. En un plano más profundo, la visión por computador constaría de varios procesos:

- > **Recogida de imágenes:** un sistema de cámaras conectadas al ordenador que tome fotografías del mundo real periódicamente o que haga un vídeo. Una ventaja añadida que suelen implementar las cámaras es que pueden registrar luz ultravioleta, infrarroja, rayos X, etc.
- > **Procesamiento de imágenes:** son tratadas y procesadas para convertirlas en nuevas imágenes con información complementaria o mejorada. Estas nuevas imágenes pueden jugar con espacios de color RGB o YCbCr, o hacer transformaciones basadas en niveles de intensidad, transformaciones lógicas, geométricas, binarizaciones, etc.
- > **Análisis de imágenes:** una vez procesadas las imágenes se procede a la extracción de la información. La mayoría de aplicaciones tratan de clasificar imágenes o detectar objetos dentro las mismas.

En los últimos años donde más se ha avanzado ha sido en los métodos de clasificación/detección. Esto es gracias a la aparición del Deep Learning, una rama del aprendizaje automático de las máquinas (Machine Learning) que se caracteriza por ser muy eficiente con datos matriciales, de ahí que se haya aplicado a imágenes.

Las áreas de aplicación son de lo más variadas y cada día aparece una nueva. Por ejemplo;

- > **Ingeniería Civil.** En este paper un dron graba un vídeo que manda a una computadora donde se procesa y realiza una detección de grietas y óxido mediante una red convolucional, válido para el cuidado de puentes y edificios<sup>13</sup>.
- > **Construcción.** En este trabajo se recogen los vídeos de las cámaras de seguridad de una obra y un computador hace un seguimiento (tracking) en tiempo real de los operarios. El ordenador es capaz de seguir a varios operarios a la vez<sup>14</sup>.
- > **Seguridad.** En este caso un ordenador detecta incendios a partir de las imágenes de vídeo de una cámara de seguridad. La detección la realiza una red convolucional<sup>15</sup>.
- > **Medicina.** En este trabajo se recogen numerosos ejemplos de aplicaciones de redes convolucionales en imágenes médicas. En estos se trata de detectar tumores, colonias de bacterias, roturas de órganos y huesos, hemorragias, etc.<sup>16</sup>

<sup>13</sup> Billie F. Spencer, Vedhus Hoskere, Yasutaka Narazaki, Advances in Computer Vision-Based Civil Infrastructure Inspection and Monitoring, Engineering, Volume 5, Issue 2, 2019,

<sup>14</sup> Adaptive computer vision-based 2D tracking of workers in complex environments, Automation in Construction, Volume 103, 2019,

<sup>15</sup> Hao Wu, Deyang Wu, Jinsong Zhao, An intelligent fire detection approach through cameras based on computer vision methods, Process Safety and Environmental Protection, Volume 127, 2019,

<sup>16</sup> Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. van der Laak, Bram van Ginneken, Clara I. Sánchez, A survey on deep learning in medical image analysis, Medical Image Analysis, Volume 42, 2017.

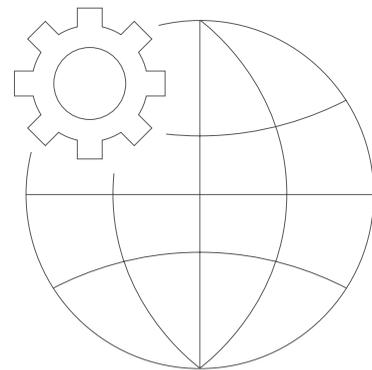
## > Sistemas empuotrados

Un sistema empuotrado es un sistema informático de uso específico que está encapsulado totalmente por el dispositivo que controla. Los sistemas empuotrados constituyen un sistema computacional fruto de la combinación de *hardware* y *software*. Esta combinación tiene como misión llevar a cabo una funcionalidad o un conjunto de funcionalidades determinadas. Se denominan empuotrados porque normalmente forman parte de un sistema completo o con funcionalidades más generales. Los sistemas empuotrados están optimizados para mejorar el tamaño, el coste, el consumo energético, la seguridad y el rendimiento del dispositivo. Se suelen implementar mediante FPGAs (Field-Programmable Gate Array ) o DSPs (Digital Signal Processor).

### Los sistemas empuotrados constituyen un sistema computacional fruto de la combinación de *hardware* y *software*

Aunque parezca mentira, nos encontramos completamente rodeados de sistemas empuotrados. Las cámaras digitales son un buen ejemplo de estos sistemas al permitir, por ejemplo, la detección automática de las caras de las personas. Nuestros coches están completamente llenos de sistemas empuotrados: el control de tracción, el detector de lluvia, los sistemas de seguridad activa, etc. Todos ellos están implementados por sistemas específicos y dedicados.

En los dos ejemplos anteriores hemos introducido ya algunos de las actividades que realizan los sistemas empuotrados, como la detección de imágenes, la clasificación de eventos o la predicción de movimiento. Estos son los problemas que ataca directamente la inteligencia artificial mediante mecanismos como SVMs, (Support Vector Machines), ELM's o Redes Neuronales. Sin embargo, la mayoría de nosotros asociamos estas técnicas a un coste computacional elevado que requiere importantes infraestructuras. El reto es, por tanto, la implementación de las técnicas de Inteligencia Artificial en dispositivos como DSP's o FPGAs<sup>17,18</sup>



<sup>17</sup> T. C. Yeam, N. Ismail, K. Mashiko and T. Matsuzaki, "FPGA implementation of extreme learning machine system for classification," TENCON 2017 - 2017 IEEE Region 10 Conference, Penang, 2017, pp. 1868-1873.

<sup>18</sup> Shuang Liang, Shouyi Yin, Leibo Liu, Wayne Luk, Shaojun Wei, FP-BNN: Binarized neural network on FPGA, Neuro-computing, Volume 275, 2018.

**La Inteligencia Artificial es un conjunto de tecnologías, no es una única tecnología**

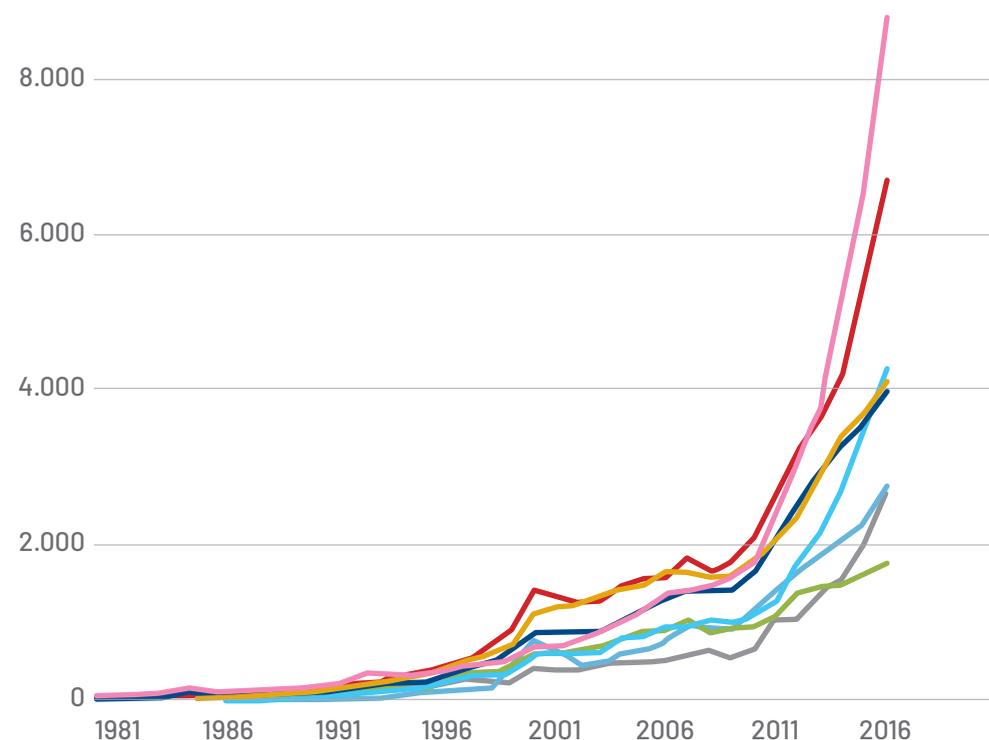


EJEMPLOS DE  
APLICACIÓN DE LA  
INTELIGENCIA  
ARTIFICIAL

Tendemos a pensar en la IA (Inteligencia Artificial) en tiempo futuro, especulando sobre cómo afectará este desarrollo tecnológico. Sin embargo, la IA y la robotización ya han empezado a transformar la vida cotidiana de las personas. Poco a poco, se ve cómo los sectores comienzan a utilizar la IA para revolucionar sus negocios y transformarse digitalmente después de haber pasado los últimos años poniendo en orden sus datos e identificando áreas donde esta tecnología puede ser de utilidad.

En un reciente estudio de la Organización Mundial de Propiedad Intelectual (WIPO, en sus siglas en inglés)<sup>19</sup>, se muestra que el número de patentes de este organismo relacionadas con técnicas de Inteligencia Artificial y que identifican ámbitos específicos de aplicación han experimentado un espectacular incremento desde 2011. En la figura siguiente se aprecia que las aplicaciones relacionadas con el transporte, las telecomunicaciones, la seguridad y las ciencias de la salud son las más numerosas.

● FIGURA 1. Familias de patentes relacionadas con la inteligencia artificial, según sus categorías de aplicación



Nota: Una patente puede referirse a más de una categoría

- Transporte ● Telecomunicaciones ● Seguridad ● Negocio
- Industria y Fabricación ● Dispositivos personales, computación e interacción hombre-máquina
- Gestión de documentos y publicación ● Ciencias de la vida y Medicina

<sup>19</sup> "WIPO Technology Trends 2019: Artificial Intelligence". World Intellectual Property Organization, 2019.

## > Un ejemplo de aplicación en el ámbito de las tecnologías de la salud

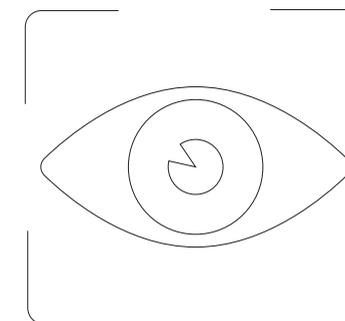
La aplicación de Inteligencia Artificial que describiremos a continuación surge como resultado, por un lado, de la investigación avanzada relativa a los efectos que sobre la actividad eléctrica de la piel (electrodermia o conductancia de la piel) produce la activación de zonas del hipocampo o de la amígdala del cerebro, y, por otro, del desarrollo de los algoritmos de aprendizaje máquina.

Una empresa italo-americana desarrolló un experimento clínico para recoger y validar señales biométricas de pacientes epilépticos, mediante wearables y aplicaciones de alerta asociadas, para compararlas con los informes de otros episodios de ataque epiléptico registrados anteriormente. De esta manera, se pudo obtener una información muy rica, en condiciones reales, sobre los casos en que fue posible detectar con antelación un ataque, los ataques desafortunadamente no detectados y las falsas alarmas.

Con la información recabada, y tras su etiquetado con precisión médica por un conjunto de neurólogos, se entrenó un algoritmo supervisado de tipo SVM (support vector machine) para que aprendiera a detectar que determinada actividad dérmica recogida en el smart watch podía revelar un inminente ataque epiléptico. Una vez entrenado el algoritmo SVM de detección de ataques, se pudo integrar en el wearable, en donde se mantendrá en continua ejecución para determinar si en los distintos eventos electrodérmicos acecha un episodio de ataque fatal. En caso de detectar un riesgo, se comunica con una aplicación que envía una alerta, junto con los datos registrados, para su atención y revisión por un médico.

Este equipamiento, ya en el mercado, permite que los enfermos epilépticos estén más protegidos frente a lo que se denomina muerte súbita en epilepsia (*sudden unexplained death in epilepsy -SUDEP-*), que provoca una muerte en el mundo cada 7-9 minutos.

> Más información: <https://www.empatica.com/en-eu/research/science/>



## > Un ejemplo de aplicación en el ámbito de la conservación de la naturaleza

WildTrack es un proyecto iniciado en 2004 por una veterinaria, Zoe Jewell, y un biólogo, Sky Alibhai, con el objetivo de obtener técnicas de identificación de huellas de animales salvajes, con la colaboración de rastreadores indígenas, para así permitir la monitorización de la vida salvaje de una forma no invasiva para los animales, y eficiente en coste.

Hasta 2018, su técnica de identificación de huellas (*footprint identification technique -FIT-*) se ha personalizado para 13 especies diferentes, a través de la colaboración de diferentes socios de proyecto en 19 países diferentes.

La técnica FIT comienza con la captación de imágenes digitales de huellas de animales salvajes mediante una cámara digital con GPS. Posteriormente, estas imágenes permiten componer un conjunto de datos, en donde cada imagen se convierte a una silueta geométrica, y tras esta transformación son aplicadas a algoritmos de aprendizaje profundo, desarrollados en diferentes proyectos con universidades de todo el mundo, para reconocer patrones de una forma similar a como lo hacen los rastreadores, pero a mayor escala y más rápidamente, para identificar cada especie, así como la edad y el sexo de cada individuo.

La precisión obtenida por WildTrack en la identificación de individuos de diferentes especies es superior al 95%.

> Más información: <https://wildtrack.org/>

## > Un ejemplo de aplicación en el ámbito del transporte ferroviario

“Tunnel Curiosity” es un proyecto de investigación, con fondos FEDER y CDTI, en el que participa un consorcio de seis empresas (Alstom, Adif, Ferrovial, Vias, Insitu y Tinámica), para el mantenimiento proactivo de infraestructuras ferroviarias. Se trata de un proyecto de innovación en la adopción de big data e Inteligencia Artificial en el ámbito ferroviario para la prolongación de la vida útil de los activos ferroviarios. El resultado es un vehículo automatizado y capacitado para la toma de datos de múltiples sensores en tiempo real, para posteriormente ser interpretados y analizados para monitorizar el estado y mantenimiento de túneles e infraestructuras críticas de la red de alta velocidad ferroviaria. Reduce en un 80% los costes de mantenimiento y mejora en un 35% la toma de decisiones.

> Más información: <https://tunnelcuriosity.es/>

## > Un ejemplo de aplicación en el ámbito de la seguridad

La empresa ForAllSecure ha creado un bot capaz de generar exploits totalmente automático. Su IA se denomina “Mayhem”. AEG (Generación Automática de Exploits) busca y encuentra errores de *software*. Una vez encontrado el bug, produce de manera autónoma la secuencia de control de la explotación; es decir, protege la vulnerabilidad. Este *software* tiene importantes aplicaciones para defensa. Por tanto, estamos ante una de las máquinas automáticas de *hacking* de sombrero blanco más impresionantes del mundo capaz de detectar, explotar y corregir las debilidades del *software* automáticamente, y a velocidades que ningún ser humano puede igualar.

> Más información: <https://spectrum.ieee.org/computing/software/mayhem-the-machine-that-finds-software-vulnerabilities-then-patches-them>

## OPINIÓN



### Rogelio Polanco

DIRECTOR DE CALIDAD DE ALHAMBRA EIDOS

En la sociedad del conocimiento en que vivimos, gestionar un volumen casi infinito de datos para obtener conclusiones que ayuden a la toma de decisiones de organizaciones de todo tipo se hace cada vez más complejo. En ocasiones, se trata de llegar a conclusiones objetivas, no contaminadas por las opiniones de las personas que participan. En otros muchos casos, la clave está en identificar qué datos aportan valor y cuales son sólo ruido y empobrecen las decisiones, e incluso llevan a conclusiones erróneas.

Teniendo en cuenta cómo las decisiones en organizaciones de todo sector y tamaño influyen en el futuro de la economía global, los países y las personas, los sistemas de IA ya están aportando capacidades analíticas muy superiores a las de la mente humana, sin apartarse de los criterios objetivos que definen las reglas de aprendizaje de entrenamientos máquina como *machine learning*.

Entraremos con ello en un nuevo terreno, el de integrar en nuestro día a día las conclusiones de estos sistemas, haciendo compatible la mejora en las capacidades analíticas y de toma de decisiones de la IA con los derechos y las libertades de las personas, ya sea en su calidad de clientes, de consumidores, de ciudadanos...

## > Un ejemplo de aplicación en el sector de las telecomunicaciones

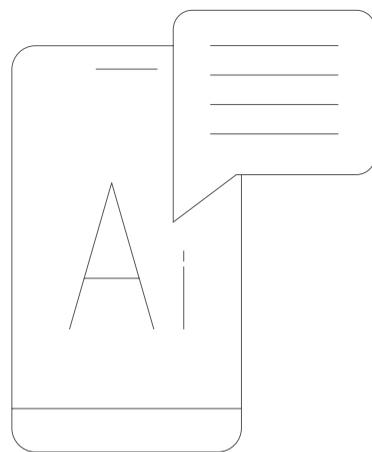
En los últimos tiempos los operadores de telecomunicación se han lanzado a una carrera por incorporar técnicas de Inteligencia Artificial en diversos ámbitos de su negocio, desde la optimización de la red, la segmentación de los clientes o la personalización de servicio. Aún más, la Inteligencia Artificial puede ser uno de los soportes que les permita expandir los límites de su negocio para competir en otras industrias verticales.

En esta tendencia se encuentra el operador coreano SK Telecom, que ya anunció en noviembre de 2018 el lanzamiento de su plan de convergencia 5G-AI, aplicando el símil orgánico de que la red 5G es el sistema circulatorio que comunica personas y objetos entre sí, mientras que la Inteligencia Artificial es el cerebro que analiza datos complejos y busca soluciones óptimas.

El lanzamiento del primer servicio comercial de este esquema convergente, en diciembre de 2018, se realizó en el entorno empresarial, aplicado a la verificación de la calidad de una línea de producción de un fabricante coreano, mediante técnicas de visión computerizada.

Pero no sólo SK Telecom utiliza la inteligencia artificial para diseñar un portfolio de servicios innovadores enfocados -por el momento- a la industria 4.0, sino que también la aplica con éxito en la gestión de su propia red 5G. Mediante estas técnicas, se puede predecir el momento y el lugar de las futuras situaciones de congestión de tráfico, para anticiparse y asignar dinámicamente más capacidad donde se necesite, ajustando automáticamente la directividad de las antenas y su área de propagación.

> Más información: <https://www.globalskt.com/home/info/2474>



**La intensidad de la investigación en IA a nivel mundial se aprecia en estadísticas como las de la Organización Mundial de Propiedad Intelectual: Desde 2011, el número de patentes de este organismo relacionadas con técnicas de IA ha crecido exponencialmente**



UNA REFLEXIÓN SOBRE  
LAS IMPLICACIONES  
SOCIALES Y ÉTICAS  
DE LA IA

## > Riesgo de fractura social: desempleo y brecha de cualificación

En los últimos años, estamos asistiendo a la extensión de la automatización inteligente a todos los ámbitos en los que se desarrollan no sólo la productividad, sino incluso la creatividad humana: desde la realización de diagnósticos médicos con una fiabilidad equivalente o incluso mayor que la de los propios especialistas<sup>20</sup> hasta la creación de arte en forma de música, videos o dibujos<sup>21</sup>. Y ello especialmente cuando el ritmo de crecimiento y sofisticación de las capacidades analíticas, predictivas y de resolución de problemas de los sistemas de inteligencia artificial los están haciendo cada vez más atractivos en términos de coste-beneficio para las empresas.

Existen no pocos estudios que advierten de que esta tendencia a la automatización inteligente es uno de los mayores riesgos, sino el principal, que afronta la sociedad con la generalización de la Inteligencia Artificial. Las predicciones de estos análisis son más o menos apocalípticas, pero todas apuntan a que en la cuarta Revolución Industrial se avecina una transformación del mercado de trabajo mayor que en las anteriores.

Varios estudios advierten del porcentaje de empleos que están en riesgo a medio plazo. Así, por ejemplo, un estudio de PwC estima en un 3% los puestos de trabajo en riesgo en los países de la OCDE en la siguiente década, que sube al 30% a mediados de la década de 2030<sup>22</sup>.

Otro informe de McKinsey sobre los 46 países que representan el 90% del PIB mundial estima que, dependiendo del escenario, entre un 3% y un 14% de sus trabajadores (entre 75 y 375 millones de trabajadores) tendrán que cambiar su ocupación en 2030. Recientemente, un estudio de la OCDE destaca que un 14% de los puestos de trabajo actuales en los países OCDE presenta una elevada probabilidad de automatización<sup>24</sup>.

<sup>20</sup> <https://healthcare-in-europe.com/en/news/artificial-intelligence-diagnoses-with-high-accuracy.html>  
<sup>21</sup> <https://www.newscientist.com/article/2193361-ai-can-diagnose-childhood-illnesses-better-than-some-doctors/>  
<sup>22</sup> <https://www.ibm.com/watson/advantage-reports/future-of-artificial-intelligence/ai-creativity.html>  
<sup>23</sup> "Will Robots Really Steal our Jobs?", PricewaterhouseCoopers, 2018. <https://www.pwc.co.uk/services/economics-policy/insights/the-impact-of-automation-on-jobs.html>  
<sup>24</sup> "Jobs Lost, Jobs Gained: Workforce Transitions in a Time of Automation", McKinsey Global Institute, 2017. <https://www.mckinsey.com/featured-insights/future-of-work/jobs-lost-jobs-gained-what-the-future-of-work-will-mean-for-jobs-skills-and-wages>  
<sup>25</sup> "OECD Employment Outlook 2019. The Future of Work", OCDE, 2019. <http://www.oecd.org/employment/future-of-work/>

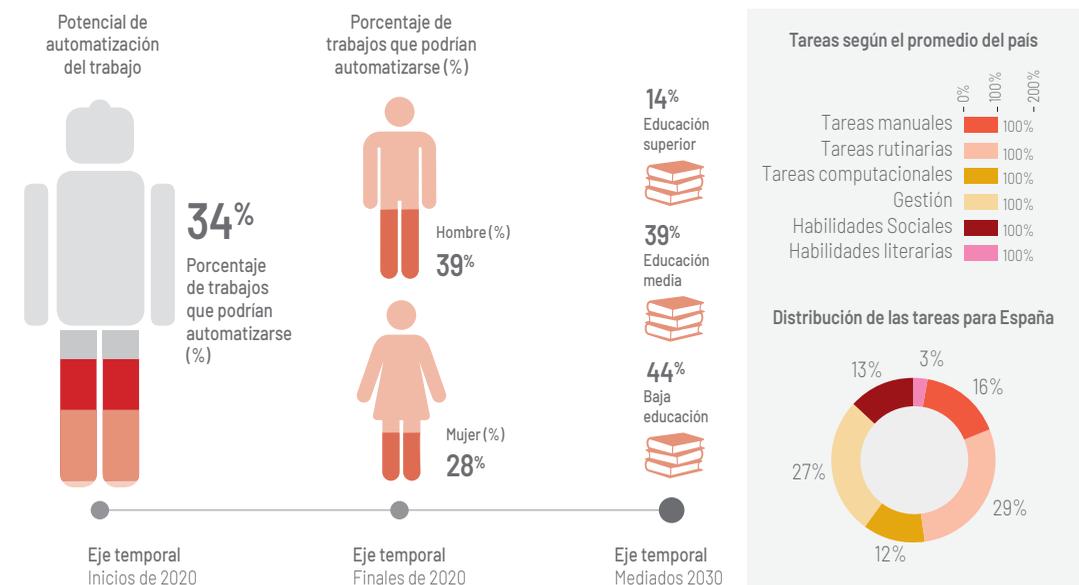
Un estudio de Frontiers<sup>25</sup> resume los tipos de trabajos más sujetos a la automatización:

- > Aquellos en los que se necesita un menor nivel de estudios (operarios de fábricas, servicios de hostelería, administrativos, etc.). Por el contrario, los puestos con requisitos de formación muy superiores son los menos expuestos al riesgo.
- > Aquellos puestos en donde no es necesario un grado de interacción con personas muy sofisticado y complejo (en estos casos, se entiende por grado de interacción elevado el que se produce cuando se realizan actividades como la supervisión y gestión de equipos, consultoría y asesoramiento especializado, asistencia y atención médica, formación, etc.).
- > Aquellas ocupaciones que requieren habilidades manuales.

Otros análisis sugieren otro enfoque más positivo. Así, en un estudio del MIT Sloan School of Management<sup>26</sup> se ha analizado el potencial impacto de las técnicas de *machine learning* en alrededor de 900 puestos de trabajo de la Oficina Federal de Estadísticas de Trabajo estadounidense, derivando que se verá afectada por igual toda la escala salarial que abarcan estos puestos, desde los CEO hasta los conductores de camión. El estudio concluye que la inteligencia artificial no desplazará en muchos casos un puesto de trabajo en su totalidad, sino que sustituirá muy probablemente tareas específicas, de forma que la automatización inteligente se centrará en aquello que mejor sabe hacer y se dejará a los humanos otras tareas para las que nuestra especie aún es más competente. Por tanto, según este estudio, nos queda por delante una importante y apasionante labor de reingeniería de procesos y reorganización de tareas para encontrar las sinergias adecuadas.

Figura 2. Resultados del informe de PricewaterhouseCoopers para España

### ESPAÑA > TODOS LOS SECTORES



<sup>25</sup> "The Impact of Artificial Intelligence on Work. An Evidence Review Prepared for the Royal Society and the British Academy", Frontiers Economics, 2018. <https://royalsociety.org/topics-policy/projects/ai-and-work/>  
<sup>26</sup> "What Can Machines Learn, and What Does It Mean for Occupations and the Economy?", Erik Brynjolfsson et al., 2018. Citado en: <https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-will-redesign-not-replace-work>

Un común denominador de todos ellos es que predicen que habrá una necesidad de nuevos empleos para abordar esta transformación de las actividades productivas, que incluso superará a los puestos de trabajo destruidos. De nuevo, como en otras muchas facetas de la Sociedad Digital que se han puesto de manifiesto en otros informes del Grupo de Políticas Públicas y Regulación<sup>27</sup>, la formación STEM se erige como un factor fundamental para el incremento de la competitividad y del desarrollo profesional y personal de nuestra sociedad.

## › La educación siempre será esencial

Entonces, en el marco de la educación, ¿los conocimientos que se imparten actualmente en los centros educativos (colegios, institutos, universidades,...) son útiles?, ¿los conocimientos y habilidades en los que estamos formando a nuestros niños y jóvenes tienen sentido o al menos parte de ellos?, ¿está el profesorado preparado para esta nueva irrupción digital?

Más allá de estas cuestiones, cada vez resulta menos extraño oír hablar de Inteligencia Artificial en el ámbito de la educación. El informe *'Artificial Intelligence and Life in 2030'*, impulsado por la Universidad de Stanford, destaca

la realidad virtual, la robótica educativa, los sistemas de tutoría inteligente y aprendizaje online o la analítica del aprendizaje como las tecnologías que, con toda probabilidad, ocuparán un lugar destacado en las aulas dentro de quince años. Los beneficios apuntan hacia la personalización del aprendizaje, la expansión del concepto de aula y una mayor y mejor interacción entre profesores y alumnos. Conclusiones similares alcanza también el documento de trabajo *'Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development'* [La Inteligencia Artificial en la educación: oportunidades y retos para el desarrollo sostenible], presentado por la UNESCO y ProFuturo durante la Semana del Aprendizaje Móvil 2019.

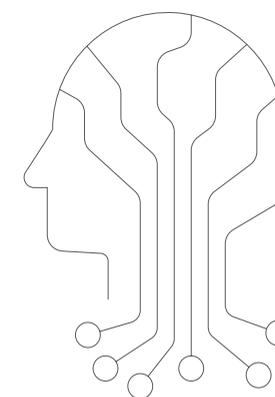
La Inteligencia Artificial en la educación: oportunidades y retos para el desarrollo sostenible], presentado por la UNESCO y ProFuturo durante la Semana del Aprendizaje Móvil 2019.

No obstante, en países en vía de desarrollo los retos aún son otros y están limitados por otras carencias, como la falta de una infraestructura básica y de recursos humanos de perfil alto en este campo.

A pesar de los indudables beneficios que potencialmente presenta el campo de la IA, es necesario afrontar los siguientes retos:

- › Garantizar un uso inclusivo y equitativo de la IA en la educación, ya que los países menos desarrollados se arriesgan a sufrir nuevas divisiones sociales y tecnológicas.
- › La científica de datos, Nuria Oliver, Ingeniera del Año 2018 por el COIT, destaca la necesidad de formar en el futuro a los alumnos en las siguientes disciplinas para ser competentes en pensamiento computacional: algoritmos, datos, redes, programación y *hardware*.

- › Preparar a los profesores para una educación impulsada por la IA, a la vez que se prepara la IA para que entienda la educación.
- › Desarrollar sistemas de datos inclusivos y de calidad. La calidad de los datos debe ser la principal inquietud. Es fundamental desarrollar capacidades estatales para mejorar la sistematización y recopilación de datos.
- › Resulta necesario conseguir que la investigación sobre la IA en educación sea significativa.
- › La IA plantea muchas inquietudes éticas sobre el acceso al sistema educativo, recomendaciones para estudiantes individuales, concentración de datos personales, responsabilidad, repercusión en el trabajo, privacidad de los datos y propiedad de los datos que se añaden a los algoritmos. Por tanto, la regulación de la IA en el ámbito de la educación requiere un debate público sobre ética, responsabilidad, transparencia y seguridad.



## OPINIÓN



### Miguel Camacho Collados

GABINETE DE LA SECRETARÍA DE ESTADO PARA EL AVANCE DIGITAL

En la próxima década la IA supondrá una revolución tan grande como la que supuso la llegada del ordenador, o muy posiblemente mayor. Gracias a ella, tenemos al alcance la posibilidad de aspirar a logros que jamás la humanidad pudo alcanzar antes.

Sin embargo, no hay que olvidar que la IA es un herramienta, y como tal no tiene moral intrínseca y puede ser mal utilizada. Por eso, tenemos que luchar para que la IA sea un instrumento de la democracia, que se use para lograr una sociedad más justa e igualitaria, que no deje a nadie atrás.

<sup>27</sup> *"Hacia la Sociedad Gigabit: Luces y Sombras"*, Grupo de Políticas Públicas y Regulación (COIT), 2018  
<https://www.coit.es/informes/hacia-la-sociedad-gigabit-luces-y-sombras-0>

## › Nuestros algoritmos deben ser mejores que nosotros

Actualmente, muchas decisiones relevantes, desde dónde invertir en bolsa hasta optar a un seguro de vida o ser seleccionado para un determinado puesto, se toman a partir de los análisis cada vez más sofisticados de un gran número de datos, tanto en cantidad como en diversidad, de diferentes acciones humanas, con el soporte de algoritmos derivados de técnicas de aprendizaje máquina.

¿Cómo podemos estar seguros de que estas decisiones asistidas por algoritmos son correctas? En un artículo de Forbes, un ingeniero jefe de IBM estimaba que un 94% de los proyectos de inteligencia artificial acaban en un cajón debido a problemas de falta de confianza o de transparencia<sup>28</sup>.

El laboratorio de medios del MIT (Media Lab) ha lanzado un proyecto, llamado con el inspirador nombre de *‘Liga de la Justicia Algorítmica’*, que está comprometido con la justicia, la responsabilidad y la transparencia de los sistemas codificados. El proyecto fue creado por la investigadora Joy Buolamwini, al descubrir que las tecnologías de reconocimiento facial no eran capaces de detectar su cara, aparentemente porque el diseño del reconocimiento no había tomado en cuenta datos para reconocer a las personas afrodescendientes.

Tenemos un serio riesgo de que se produzcan sesgos en las decisiones adoptadas que estén en abierto conflicto con nuestros principios de igualdad y no discriminación, que agraven el ya de por sí complicado problema de los prejuicios sociales de todo tipo y que condicionen negativamente las oportunidades y expectativas de vida de muchos ciudadanos.

Estos sesgos se deben fundamentalmente a tres aspectos:

- › Si se utilizan datos insuficientes, que no reflejan la diversidad humana ni la variedad de opiniones, para entrenar los sistemas.
- › Si existen condicionantes en los datos utilizados en el entrenamiento que llevan al sistema a tomar decisiones basadas en datos irrelevantes.
- › Si existe un error de código en el algoritmo resultante del entrenamiento que le lleva a adoptar decisiones erróneas.

Tenemos dos soluciones a estos problemas. La primera se basaría en la verificación posterior o auditoría de los algoritmos para determinar si existen evidencias de posibles sesgos en las decisiones asistidas por inteligencia artificial. La segunda sería la introducción de principios éticos en el propio código, buscando incorporar en el diseño toda nuestra larga tradición moral<sup>29</sup>, como luego veremos con más detalle.

En un informe del Servicio de Investigación del Parlamento Europeo de 2018<sup>30</sup>, se indica que no hay todavía procedimientos y modelos de certificación establecidos y estandarizados para

atender a los conflictos éticos de sesgo y falta de transparencia que se puedan producir del uso de algoritmos de decisión. También se destaca en este informe que obligar a dar información sobre las hipótesis de partida del algoritmo o la revelación del código del algoritmo en ciertos casos conflictivos exigiría una redefinición de las normas de propiedad intelectual actuales. En suma, dada la importancia de los desafíos que ya se están produciendo, el informe recomienda que la UE lidere el desarrollo de un marco ético general en el diseño, la implementación y el desarrollo de los algoritmos basados en inteligencia artificial. Y recientemente, se ha publicado por parte de este Servicio de Investigación un estudio con una propuesta de un marco público de gobernanza para la responsabilidad y transparencia de los sistemas algorítmicos<sup>31</sup>.

La posible pérdida de confianza pública abre el debate sobre la confiabilidad en la relación del ser humano y la Inteligencia Artificial. En las relaciones humanas, la confianza es clave para alcanzar un beneficio mutuo para las partes y no lo va a ser menos si en lugar de tratar con personas hay que interactuar con máquinas o *software* dirigidos por algoritmos.

Los expertos abogan por un enfoque de la inteligencia artificial centrada en el bienestar de los seres humanos, pero desde la confiabilidad, ya que es la única manera de que la sociedad se pueda beneficiar completamente de ella. La confiabilidad incluye la confianza en la tecnología, en las reglas, leyes y normas que rigen la inteligencia artificial, y en los modelos de negocio de los servicios, productos, y fabricantes de IA.

Siguiendo este enfoque, la Comisión Europea ha establecido en sus Principios Éticos la necesidad de un enfoque de la IA centrado en los seres humanos, de manera que el desarrollo y uso de la IA tenga como objetivo principal mejorar el bienestar de los seres humanos. De este modo los seres humanos confiarán en la IA y sólo podrán beneficiarse completamente si pueden confiar en la tecnología. La Comisión Europea indica que “la IA confiable tiene dos componentes: 1) debe respetar los derechos fundamentales, las leyes vigentes y los principios y valores esenciales, de manera que se garantice un fin ético, y 2) debe ser fiable y sólida técnicamente hablando, ya que un escaso dominio tecnológico puede provocar daños involuntarios, aunque las intenciones sean buenas.”

La Comisión Europea también ha destacado que es necesario “prestar especial atención a situaciones que impliquen grupos vulnerables como niños, personas con discapacidad o minorías”. Uno de los requerimientos esenciales es el diseño para todos, de manera que todos los ciudadanos puedan usar los productos o servicios, independientemente de su edad, estado de discapacidad o estado social. Otro requisito esencial es la no discriminación, referida a la variabilidad de los resultados de la IA entre individuos o grupos de personas en función de la explotación de las diferencias en sus características que pueden considerarse de manera intencional o no (como el origen étnico, el género, la orientación sexual o la edad), y que pueden afectar negativamente a tales individuos o grupos. Por último, el respeto a la privacidad es también muy importante. La privacidad y la protección de datos deben estar garantizadas en todas las etapas del ciclo de vida del sistema de IA, según las Directrices éticas mencionadas.

Otros requerimientos enunciados en las Directrices son la rendición de cuentas, la gobernanza de los datos, la gobernanza de la autonomía de la IA, el respeto por la autonomía humana, la robustez, la seguridad o la transparencia.

<sup>28</sup> “Managing The Ethics Of Algorithms”, *Forbes Insights*, 2019

<https://www.forbes.com/sites/insights-intelai/2019/03/27/managing-the-ethics-of-algorithms/#3c273ace3481>

<sup>29</sup> En este punto se abriría otro gran debate sobre qué moral introducir en el sistema, pero ello excede los objetivos de este informe

<sup>30</sup> “What if algorithms could abide by ethical principles?”, Mihalis Kritikos, Scientific Foresight Unit (STOA), European Parliamentary Research Service, 2018 [http://www.europarl.europa.eu/stoa/en/document/EPRS\\_ATA\(2018\)624267](http://www.europarl.europa.eu/stoa/en/document/EPRS_ATA(2018)624267)

<sup>31</sup> “A governance framework for algorithmic accountability and transparency”, Scientific Foresight Unit (STOA), European Parliamentary Research Service, 2019. [http://www.europarl.europa.eu/stoa/en/document/EPRS\\_STU\(2019\)624262](http://www.europarl.europa.eu/stoa/en/document/EPRS_STU(2019)624262)

<sup>32</sup> “Ethics guidelines for trustworthy AI”, High-Level Expert Group on Artificial Intelligence set up by the European Commission, 2019. <https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top>

## › El riesgo del desequilibrio del sistema de valores

En 1985, Judith Jarvis Thomson puso de manifiesto en la revista *The Yale Law Journal* (una publicación de la Facultad de Derecho de Yale) algunas preguntas<sup>33</sup> que llevaban rondando desde principios del siglo XX en el conocimiento humano. Esta reconocida filósofa moral anticiparía la complejidad que supondría para una Inteligencia Artificial encontrar la solución de un dilema humano complejo. Imaginemos que un tranvía<sup>34</sup> corre fuera de control por una vía. En su camino se hallan cinco personas atadas a la vía. Afortunadamente, es posible accionar un botón que encaminará al tranvía por una vía diferente; por desgracia, hay otra persona atada a ésta. ¿Debería pulsarse el botón?

La mayoría de los que consideran este problema creen que está permitido accionar el interruptor. La mayor parte cree que no sólo es una acción permitida, sino también la mejor opción moral en este caso, siendo la otra no hacer nada.

Ahora el mismo tranvía descontrolado se dirige hacia cinco personas. Una persona se sitúa en un puente sobre la vía y podría detener el paso del tren lanzando un gran peso delante del mismo. Mientras esto sucede, al lado del sujeto del puente sólo se halla un hombre muy gordo; de este modo, la única manera de parar el tren es empujar al hombre gordo desde el puente hacia la vía, acabando con su vida para salvar otras cinco. ¿Qué debe hacer el sujeto?

En este caso se encuentra una gran resistencia a decidir una participación activa. La mayoría, que inicialmente en el caso anterior aprobaba el sacrificio de uno en favor de los otros cinco no aprueba, en esta situación, lanzar al hombre gordo a la vía. Esto ha llevado a que se intente encontrar una diferencia moral relevante entre ambos casos.

**Judith Jarvis Thomson, reconocida filósofa moral, anticiparía la complejidad que supondría para una Inteligencia Artificial encontrar la solución de un dilema humano complejo**

en que se pretende la muerte de alguien para salvar a cinco, lo cual es malo, mientras que en el primer caso no existe tal intención.

Este dilema filosófico es un paso previo a la necesidad de plantearse un diseño de los sistemas de Inteligencia Artificial basado en algoritmos morales. La problemática planteada se

está materializando en la *Fundación OpenAI*, creada por Elon Musk y Sam Altman, en la que se invita a los principales líderes tecnológicos a identificar dilemas éticos y prejuicios. Su objetivo primordial es establecer unas reglas de juego basadas en un marco de comportamiento moral, donde la Inteligencia Artificial pueda desarrollarse en representación de la humanidad y no contra la humanidad. El primer paso es que los gobiernos tomen conciencia de la importancia de este asunto y lo regulen, y que ninguna empresa privada acumule demasiado poder a la hora de influir sobre la evolución futura de la Inteligencia Artificial.

De nuevo, consideramos que para ello es fundamental respetar el sistema de valores. Los valores son la base de las opiniones personales y colectivas sobre lo que es importante en la vida, influenciados por la cultura, la religión y las leyes. Son factores importantes en la toma de decisiones sobre asuntos sociales, ambientales y políticos, y sobre los mejores usos de nuestro tiempo o dinero.

## › Las amistades peligrosas de la Inteligencia Artificial

Seguro que todos recordamos el video de Barack Obama que presentó hace unos años la Universidad de Washington donde se le puede ver a un lado de la pantalla decir una cosa y, en el otro lado, con otra ropa y en otro lugar, decir exactamente lo mismo.

Es lo que se ha denominado *deep fakes* o vídeos manipulados, la temible evolución de las *fake news*. Se trata de una amenaza latente, dado lo complicado de diferenciar lo verdadero de lo falso. Sin embargo, la amenaza real está en el poder de manipulación que puede tener la difusión a gran escala de un vídeo similar, especialmente si algunos gobiernos o grupos malintencionados están dispuestos a aprovechar estas técnicas hasta sus últimas consecuencias. El CCN (Centro Criptológico Nacional) advierte de la doctrina Gerasimov<sup>35</sup>, un general ruso que, a raíz de las Primaveras árabes, señaló que “los métodos más utilizados en un conflicto están cada vez más relacionados con el uso de la información política y económica [...]”.

Precisamente, en España el Tribunal Constitucional ha hecho pública la sentencia contra el Artículo 58 bis de la Ley Orgánica del Régimen Electoral General (LOREG). Se trata de un dictamen que tumba la norma que permitía a los partidos políticos rastrear las opiniones políticas de los ciudadanos en Internet y redes sociales, con la consecuente posibilidad de manipulación del voto de la ciudadanía, ya que la norma les autorizaba a conocer las opiniones políticas de los ciudadanos y enviarles en consecuencia publicidad o propaganda dirigida.

Una de las principales consecuencias de esta desinformación es la pérdida de confianza pública en las democracias, en las instituciones y en los medios de comunicación. Los medios, precisamente, son un objetivo claro porque pueden servir de altavoz legítimo a una manipulación.

<sup>33</sup> Judith Jarvis Thomson, *The Trolley Problem*, 94 YALE L.J. (1985). <https://digitalcommons.law.yale.edu/yj/vol94/iss6/5>

<sup>34</sup> La primera versión de este dilema fue creada en 1967 por la filósofa Philippa Foot

<sup>35</sup> Desinformación en el Espacio. CCN-CERT BP/13. Febrero 2019

## > Asimetrías de información que derivan en asimetrías de poder

La Inteligencia Artificial puede ser uno de los avances más disruptivos de los próximos años, configurando el poder geoestratégico del futuro, al igual que la evolución social. Basada en grandes cantidades de datos y de computación, ¿qué puede ocurrir si el acceso es asimétrico y solo una minoría puede obtenerlo y utilizarlo dejando al resto como meros usuarios? Minimizar esta asimetría es una tarea esencial para garantizar que su impacto sea positivo para todos, no solo para unos pocos.

**Minimizar esta asimetría es una tarea esencial para garantizar que su impacto sea positivo para todos, no solo para unos pocos**

Stefaan Verhulst, co-fundador de The GovLab, aboga por definir estrategias de gobernanza colectiva que minimicen las asimetrías de poder creadas por la Inteligencia Artificial.

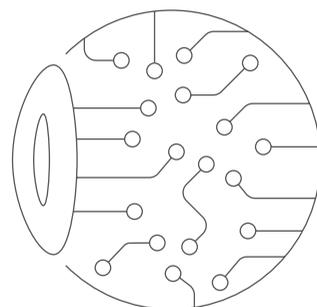
Las asimetrías de poder creadas por la Inteligencia Artificial pueden producirse entre los países con mayor desarrollo de IA y el resto,

entre empleador y empleado o entre empresas y consumidores, como destaca la Comisión Europea.

Las grandes potencias internacionales tratan de aprovechar su poder disruptivo principalmente en tres dimensiones de poder: la económica, la militar y la informativa, según un estudio del Belfer Center sobre *'Artificial Intelligence and National Security'*.

En el apartado anterior ya planteábamos la dimensión de poder que supone el control de la información, de lo que es real frente a lo manipulado, de lo que se publica frente a lo que se oculta. En el plano económico, de acuerdo con el estudio del McKinsey Global Institute *'AI, ¿the Last Digital Frontier?'*, los gigantes tecnológicos que lideran la digitalización son quienes más están invirtiendo en IA, seguidos de sectores como automoción, servicios financieros, energía y recursos, medios y entretenimiento o transporte y logística.

Y, por último, pero no por ello menos importante, en el ámbito militar se está produciendo una carrera entre las principales potencias del mundo para optar al liderazgo geoestratégico, geopolítico y geoeconómico, a través de líneas de investigación para introducir autonomía o semi-autonomía en robots, el uso de aviones de combate no tripulados, o sistemas armamentísticos autónomos (*lethal autonomous weapon systems*).



## ... ¿Y a largo plazo?

### > Reescribiendo a Asimov: las leyes de la IA

Desde hace unos años, la futura evolución de los sistemas de inteligencia artificial constituye un motivo de gran preocupación entre la comunidad científica. La cuestión fundamental que surge es si en un mundo dominado por la IA podrá la humanidad coexistir de manera digna.

En 2015, personalidades relevantes como Stephen Hawking y Elon Musk firmaron un manifiesto<sup>36</sup> apelando a la necesidad de que la investigación en este campo persiga no sólo hacer sistemas más expertos y capaces, sino también maximizar los beneficios de la Inteligencia Artificial para la sociedad. La llamada 'Carta de Copenhague', suscrita por prestigiosos científicos, invita a compartir una visión de progreso que trascienda al concepto de innovación recordando que el ser humano debe estar en el corazón de cualquier acción tecnológica.

En el mismo sentido se expresaron los manifiestos éticos difundidos por el Future of Humanity Institute, de la Universidad de Oxford, y por Future of Society, que cuenta con el apoyo de Harvard, que señalan los riesgos de las aplicaciones militares de la Inteligencia Artificial y abogan por una gobernanza global que brinde un marco ético.

En el estado actual de la tecnología, la Inteligencia Artificial se aplica a tareas muy concretas y es capaz de superar a los seres humanos en muchas de ellas, pero no puede resolver problemas o tareas de propósito general (esta fase se denomina *narrow / soft artificial intelligence*). Si el desarrollo tecnológico nos lleva a obtener sistemas que demuestren el mismo nivel de inteligencia que nuestra especie (la llamada Inteligencia Artificial general -*Artificial General Intelligence*-), y se utilizan recursivamente estos sistemas para mejorar las siguientes versiones, existe una corriente de opinión que considera muy probable que se alcance lo que los expertos denominan 'singularidad', a partir de la cual se genere una superinteligencia artificial que supere al ser humano, sin posibilidad de control. Este riesgo se agrava por el hecho de que el punto de 'singularidad' puede alcanzarse sin que seamos conscientes de ello.

Esto supone un reto ético mayúsculo, ya que estaría en juego la propia supervivencia de la especie humana en un futuro más o menos próximo; especialmente, si ambas inteligencias, la nuestra biológica y la cibernética, competimos por los mismos recursos del planeta o nuestras aspiraciones no están alineadas. Los manifiestos que hemos citado derivan fundamentalmente de la preocupación de que quizá, como especie, estemos lanzándonos a una carrera para crear una superinteligencia artificial sin haber ideado siquiera los mecanismos de control y escape, en caso de necesidad extrema, adecuados para evitar riesgos.

<sup>36</sup> "An Open Letter. Research Priorities for Robust and Beneficial Artificial Intelligence", 2015. <https://futureoflife.org/ai-open-letter>

# OPINIÓN



## Eduvigis Ortiz Moronta

STRATEGIC ALLIANCES LEADER SAS WIMLDS MADRID

La Inteligencia Artificial (IA) tiene innumerables usos que pueden hacer nuestra vida más fácil, así como ayudarnos a tomar decisiones de manera más rápida y eficiente. Cada vez más, la IA es parte de innumerables decisiones que tomamos en el día a día. No obstante, es importante tener en cuenta que, como toda creación humana responde a los parámetros y prejuicios introducidos por sus creadores y en los datos en los que basa su aprendizaje.

Está sobradamente demostrado que los algoritmos que se aplican y los resultados que se extraen de un conjunto de datos pueden excluir o perjudicar a ciertos grupos de personas respondiendo a los prejuicios de sus creadores. Esto es a lo que llamamos sesgos de la IA, que tiene mayor incidencia cuando en el equipo de creadores no han estado involucradas personas que aporten diversidad y distintas miradas en el desarrollo de la tecnología.

Debido a que la IA puede reflejar los prejuicios de las personas que la crean, a veces con sesgos inconscientes, es primordial que estos equipos se complementen tanto desde el punto de vista multidisciplinar como desde el experiencial. La diversidad en todos sus aspectos enriquece cualquier creación y evita que ocurran fallos tan llamativos como algunos que hemos conocido a través de los medios de comunicación, tales como:

1.-El organismo público que regula Wall Street tiene abierta una investigación contra Apple tras haber sido acusada de ofrecer límites de crédito más altos a hombres que a mujeres con su tarjeta de crédito Apple Card.

2.- La empresa de atención médica UnitedHealth Group también está siendo investigada debido a quejas sobre un algoritmo que favorecía a los pacientes blancos frente a otras razas.

Está claro que la IA se ha convertido en una de las principales fuentes de ventaja competitiva para las empresas y una palanca de innovación social, por lo que las mujeres deben formar parte de su desarrollo en los roles de creadoras, investigadoras, programadoras y usuarias finales. Dada la rápida transformación y digitalización de la economía y el mercado laboral, actualmente el 90% de los puestos de trabajo requieren capacidades digitales básicas. Las mujeres solo representan el 17% de las personas que estudian o trabajan en el campo de las tecnologías de la información en la Unión Europea y solo el 36% de los graduados en tecnología, a pesar de que las chicas tienen mejores resultados que los chicos en las competencias digitales<sup>37</sup>.

Esta es una de las principales razones por la que, desde la comunidad de 'Women In Machine Learning & Data Science' (WiMLDS), apostamos y luchamos por visibilizar la labor de las mujeres que trabajan en tecnología persiguiendo unos objetivos claros:

- 1.- Crear y apoyar referentes femeninos en el sector que den testimonio de la gran oportunidad que es formar parte de quiénes cambian el mundo a través de la tecnología.
- 2.- Promover vocación y curiosidad por conocer el mundo de la tecnología para todos los miembros de la comunidad WiMLDS, independientemente de su género, religión, raza, orientación sexual, generación o condición social.
- 3.- Ayudar a las organizaciones y empresas a identificar el talento femenino y conectarlo con nuestra comunidad para fomentar la creatividad y la innovación.

Gracias a iniciativas como esta, contaremos con más talento femenino creando tecnología y minimizaremos uno de los sesgos más presentes en nuestra sociedad, el de género. Si logramos que las mujeres sean creadoras de tecnología y de Inteligencia Artificial en la misma proporción en la que estamos representadas en la sociedad confío plenamente en que las soluciones que desarrollemos serán más innovadoras, más éticas, resolverán mejor nuevos problemas.

Cuando la IA es creada por equipos que reflejan el mundo en el que vivimos, nos acercamos más al mundo que aspiramos tener: más diverso, más inclusivo y más justo.

Es en este punto del debate en donde surge la reflexión de la necesidad de un marco de investigación y regulación que aborde materias de seguridad y control de los sistemas de Inteligencia Artificial, que además debería ser de ámbito mundial y con un enfoque multidisciplinar –que no sólo involucrara a los agentes de la investigación científica (en tecnología, pero también en humanidades), sino también a los agentes políticos, económicos o sociales–.

Algunas opiniones se muestran en contra de establecer regulación alguna<sup>38</sup>, puesto que consideran que puede conllevar refrenar la investigación en Inteligencia Artificial. Según estas opiniones, hacerlo en un momento como el actual, cuando las técnicas de IA se están haciendo imprescindibles en campos como la medicina y la atención sanitaria o la movilidad sostenible y segura (seguridad en el control aéreo, coches conectados, etc.), supondría unos costes humanos y económicos inasumibles y difícilmente defendibles ante la opinión pública.

Pero otras opiniones abogan por que ambas ideas son compatibles<sup>39</sup>; es decir, que los sistemas de Inteligencia Artificial que se diseñen a futuro ya deberían incorporar la seguridad y control 'by design' sin que esto suponga retroceso en sus capacidades futuras y en su efecto beneficioso para la sociedad. En este sentido, algunas ideas que se proponen implican una mayor investigación en mecanismos de verificación (para comprobar que el sistema satisface los requisitos formales), validación (para asegurar que el sistema, aunque cumple los requisitos formales, no tiene comportamientos no deseados), seguridad (para evitar su manipulación por parte de terceros no deseados) y control (para permitir tomar el control humano del sistema una vez que está operativo).

SUPERANDO  
LOS RIESGOS DE  
UNA IA SIN ÉTICA  
NI REGULACIÓN

Una reflexión  
sobre las  
implicaciones  
sociales y  
éticas de la IA

36

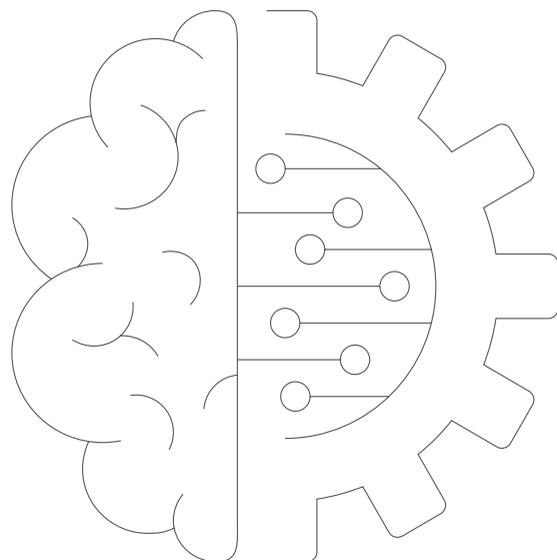
SUPERANDO  
LOS RIESGOS DE  
UNA IA SIN ÉTICA  
NI REGULACIÓN

Una reflexión  
sobre las  
implicaciones  
sociales y  
éticas de la IA

37

<sup>37</sup> Una Unión de la igualdad: Estrategia para la Igualdad de Género 2020-2025 COMISIÓN EUROPEA Bruselas, 5.3.2020 COM(2020)152 final

Otros estudios<sup>40</sup> recomiendan que todo sistema de Inteligencia Artificial que se diseñe debe incorporar un 'déficit intelectual'; es decir, la inhabilitación de fábrica de determinadas capacidades, para conseguir que nunca se alcance la superinteligencia, o incorporar una especie de talón de Aquiles que se active automáticamente cuando el sistema empiece a descontrolarse. El debate está muy lejos de haberse cerrado, y algunas voces autorizadas advierten de que no hay todavía las suficientes líneas de investigación en estos aspectos<sup>41</sup>, para la importancia capital que tienen en nuestro futuro como especie.



<sup>38</sup> "Should Artificial Intelligence Be Regulated?", Amitai Etzioni y Oren Etzioni, Issues in Science and Technology, 2017.

<https://issues.org/perspective-should-artificial-intelligence-be-regulated/>

<sup>39</sup> "Research Priorities for Robust and Beneficial Artificial Intelligence", Stuart Russell, Daniel Dewey y Max Tegmark, Future of Life Institute, 2015. [https://futureoflife.org/static/data/documents/research\\_priorities.pdf?x60419](https://futureoflife.org/static/data/documents/research_priorities.pdf?x60419)

<sup>40</sup> "Near-term Worries" and "Long-term Concerns" of the Artificial Intelligence "Revolution": An Analysis of Ethics and Ontology", Zhao Tingyang, Academia China de Ciencias Sociales, 2018. Traducción en inglés en: <https://docs.google.com/document/d/1b9n11KvMF6kj1NTwd-mP-lvOGFbe3bzqQPOPD-7BVRE/edit#heading=h.qmvhr4t12cuw>

<sup>41</sup> "Is superintelligence a threat for human decision-making?", entrevista a Jaan Tallinn, fundador de Skype y Kazaa, 2018. <https://e-estonia.com/goals-humanity-superintelligence/>

## Inteligencia Artificial: ¿Para bien o para mal?



### Cristina Sánchez Rebollo

CRISTINA SÁNCHEZ REBOLLO, SYSTEMS ENGINEER @ SAS

Uno de los principales retos a los que se enfrenta la integración de la IA es, como se comenta en este documento y ya adelanta el título, la falta de regulación y el control sobre un uso ético de esta nueva y potente herramienta. Una tecnología que delega la toma de decisiones en máquinas que 'aprenden', el largo debate sobre la destrucción de puestos de trabajo (y la creación de otros), así como la utilización como materia prima de datos incluyendo algunos de carácter personal hacen que exista cierta reticencia o temor hacia la integración de la inteligencia artificial en nuestro día a día.

Muchos de los argumentos que reclaman una regulación estricta y urgente hacen alusión al mal uso que se puede hacer con los datos personales si no hay leyes que regulen qué datos son más o menos vulnerables. En estos últimos meses hemos vivido una situación en la que el control de la población mediante geolocalización (el tipo de información cuyo acceso genera más reticencias) parecía ayudar a controlar la pandemia por COVID-19, ya ahí surgen varias dudas ¿es realmente necesario? ¿en qué situaciones se puede tener acceso a este tipo de datos? ¿quién puede estar capacitado para controlar y manipular esta información?

Por otro lado, el desconocimiento de las reglas que rigen las máquinas que han aprendido en base a estos datos suscita desconfianzas en cuanto al sesgo o incluso posible manipulación de estos datos. De sobra es conocido el algoritmo para la selección de personal de una gran empresa que aprendía en base a sus plantillas pasadas. Como la mayoría de su plantilla hace años eran hombres, la variable sexo descartaba muchos perfiles de mujeres directamente.

En general, todos estos criterios regulatorios se basan en dos aspectos: el miedo a que la toma de decisiones la haga una inteligencia desconocida y el uso malintencionado que se pueda hacer de la misma.

Sin embargo, en este artículo me gustaría hablar de la otra parte, la cara amable de este avance tecnológico. A este movimiento se le llama *'data for good'* y desde voluntarios individuales hasta empresas privadas, como SAS, están muy concienciadas en impulsar esta faceta de la inteligencia artificial con el convencimiento de que éste es el verdadero avance para la humanidad. Este movimiento anima a usar los datos para resolver problemas humanitarios relacionados con la pobreza, la salud, derechos humanos, educación y medio ambiente.

Con objetivos que van desde la detección y prevención de enfermedades hasta la protección de especies en peligro para poder reconstruirse después de desastres naturales, distintas organizaciones en todo el mundo están exprimiendo datos para poder llegar más allá mediante técnicas nuevas y creativas. Un par de casos en los que la aplicación de IA ha sido un éxito son:

#### Superar limitaciones humanas en la detección de cáncer

El cáncer colorrectal es el tercer cáncer más común en todo el mundo, y se extiende al hígado en aproximadamente la mitad de los pacientes. Amsterdam UMC uno de los centros académicos de investigación onco-

lógica y en particular el área que se especializa en cirugía hepática, establece que la mejor manera de tratar este tipo de cáncer es extirparlo. Sin embargo, algunos tumores son demasiado grandes para ser extirpados, y estos pacientes deben someterse a terapia sistémica como la quimioterapia para reducir los tumores.

Después de un período de tratamiento, los tumores se evalúan manualmente mediante una tomografía computarizada (TC). En ese momento, los profesionales médicos pueden ver si el tumor ha encogido o cambiado. La reacción de un tumor a la terapia sistémica determina si la cirugía es posible o si es necesario un régimen de quimioterapia diferente. Esta aproximación manual presenta algunos riesgos.

Esta evaluación manual de tumores es un proceso que requiere mucho tiempo de los radiólogos. Y para cada tomografía computarizada, por lo general sólo se miden los dos tumores más grandes, posiblemente dejando pistas vitales ocultas en los tumores restantes, si un paciente tiene más de dos. Además, la evaluación manual es propensa a la subjetividad, lo que conlleva a una respuesta distinta entre los radiólogos. Por otro lado, en esta revisión faltaría también la evaluación 3D y otros aspectos difíciles de detectar por el ojo humano. Y estos son problemas a los que habría que añadir el error humano. Desafortunadamente, debido a limitaciones humanas, ocurren errores radiológicos, y un diagnóstico erróneo puede someter a un paciente a riesgos potencialmente mortales, como cirugía innecesaria o quimioterapia.

Actualmente, Amsterdam UMC está transformando su evaluación de tumores con IA. Mediante técnicas de *deep learning* y *computer vision* han sido capaces de agilizar y perfeccionar las evaluaciones de la respuesta a la quimioterapia. El proyecto comenzó entrenando un modelo de *deep learning* con datos de 52 pacientes de cáncer. Cada píxel de 1.380 metástasis fue analizado y segmentado. Con esto el sistema aprendió cómo identificar instantáneamente las características del tumor y compartir información vital con los médicos. Los métodos que usaban previamente estaban limitados por lo que los médicos podían ver, pero los modelos actuales basados en IA proporcionan el volumen completo y una representación 3D de cada tumor, permitiendo a los médicos determinar con una base más precisa si la cirugía es viable o es preferible otra estrategia de tratamiento ([https://www.youtube.com/watch?v=R\\_Png784GzE](https://www.youtube.com/watch?v=R_Png784GzE)).

#### Inteligencia artificial ante la tragedia

Después de un devastador terremoto en Nepal en 2015, la Organización Internacional para las Migraciones (OIM) necesitaba proporcionar refugio a miles de familias desplazadas. Más de 45.000 familias ocuparon más de 200 campamentos de tiendas de campaña que se montaron en la lucha por un refugio seguro a medida que se acercaban las lluvias torrenciales de la temporada de monzones. La OIM necesitaba encontrar grandes cantidades de chapa metálica lo antes posible para comenzar a reconstruir casas.

Con la ayuda de la IA la OIM pudo acceder rápidamente a los datos del comercio mundial que las Naciones Unidas recopilan cada año. Décadas de datos comerciales de más de 200 países se analizaron en cuestión de minutos y la OIM tuvo respuestas sobre los principales productores y exportadores de metales fabricados de la zona. Gracias a esto se pudo agilizar la compra de los materiales. La recopilación y el análisis de datos fueron clave para dar a las personas una sensación de seguridad, estabilidad y esperanza (<https://blogs.sas.com/content/sascom/2015/07/08/disaster-relief-efforts-show-promise-of-analytics-and-seemingly-unrelated-data-sources/>).

Estos son sólo dos ejemplos reales de acciones en los que claramente la IA está a disponibilidad de la humanidad. Existen numerosos ejemplos con esta finalidad e incluso para protegerla, como son el sistema viogén contra la violencia de género o la detección de redes criminales a través de las redes sociales mediante teoría de grafos, análisis de texto y fuzzy logic (<https://www.hindawi.com/journals/complexity/2019/1238780/>).

Y aunque estos ejemplos se han englobado desde un punto de vista más organizacional es sorprendente la cantidad de iniciativas a nivel individual que han surgido en esta línea: proyectos final de carrera, proyectos individuales, grupos organizados sin ánimo de lucro, formaciones y ponencias gratuitas...

Todo esto evidencia que impulsar la Inteligencia Artificial es prioritario porque incluso en el contexto actual, en que existen reticencias fundamentadas especialmente por la falta de un sistema regulatorio, se ha conseguido hacer avances orientados al bien de todos y, teniendo la tecnología (que todavía tiene mucho por avanzar) y sobre todo la disponibilidad, no podemos dejar pasar este tren: ¡ahora es el momento!

SUPERANDO  
LOS RIESGOS DE  
UNA IA SIN ÉTICA  
NI REGULACIÓN

Una reflexión  
sobre las  
implicaciones  
sociales y  
éticas de la IA

40



La IA representará una gran oportunidad para el avance de la humanidad, siempre y cuando seamos capaces de identificar los retos éticos y sociales, a escala global, que conlleva y de diseñar respuestas eficaces para afrontarlos



ESTRATEGIAS  
NACIONALES DE LA  
INTELIGENCIA  
ARTIFICIAL

Durante décadas, el desarrollo de la IA se ha desarrollado en el ambiente académico y de investigación, siendo más recientemente cuando se han comenzado a observar resultados aplicables en el mundo de los negocios, con gran repercusión social y en los medios. Estas aplicaciones de la IA, algunas descritas en este informe, han devenido en negocios que han explotado vacíos regulatorios con los que grandes empresas han generado beneficios en un entorno incierto y con reglas poco definidas. No obstante, los beneficios potenciales de la IA para el desarrollo de la economía del conocimiento, las economías de escala, la competitividad y el bienestar social pasan por afrontar el reto definiendo unas reglas claras para todos los actores, no por la simple prohibición o regulación excesivamente estricta, lo que significaría un yugo a la innovación y a la futura competitividad de los países o bloques que opten por esta vía.

De esta forma, los países más industrializados han liderado, desde 2017, la puesta en marcha de políticas nacionales de Inteligencia Artificial, entendiendo por éstas las enfocadas a la maximización de los beneficios de la IA, mientras que minimizan sus riesgos y costes potenciales.

En esta parte del informe, haremos una recopilación de estos países pioneros en políticas IA, así como una revisión de las distintas categorías y tipos de iniciativas que existen, incluyendo ejemplos de implementación en esta fase temprana de implementación.

Pondremos más foco en nuestro marco de referencia próximo, la Unión Europea, así como en las incipientes iniciativas que se están desarrollando en España desde 2019.

Finalmente, se plantean una serie de recomendaciones del COIT para dotar a España de las herramientas necesarias para aprovechar esta revolución global.

## Líderes en Políticas nacionales de IA

Desde el año 2017, cinco países lideraron el desarrollo e implementación de políticas Inteligencia Artificial a nivel internacional. Estos son:

- > Canadá, primer país en lanzar una estrategia nacional de IA, lo que le ha llevado a capturar centro de I+D de empresas pioneras en este campo (Uber y Thomson Reuters, por ejemplo).
- > Japón, con su Estrategia de Tecnología IA que define un gran plan de industrialización 4.0 para el país, que desembocaría en lo que ha denominado Sociedad 5.0.
- > Singapur, con su iniciativa AI Singapore, enfocada en el desarrollo de talento para formar un ecosistema de investigación, tecnología e innovación alrededor de la IA.
- > China, con un agresivo plan que primero se propone igualar el nivel de desarrollo tecnológico de los líderes occidentales, con el objetivo de liderar globalmente para 2030.
- > Emiratos Árabes Unidos, siendo el primer país en crear un Ministerio de Inteligencia Artificial, enfocado en grandes programas que difícilmente podrían iniciarse de forma privada y poniendo a disposición grandes recursos a través de un acelerador de ideas.
- > Finlandia, que fue el primer país de la Unión Europea en publicar una estrategia de IA poniendo el foco en la eficiencia del sector público, la competitividad económica y la mejora del bienestar de los ciudadanos.

FIGURA 3. Estrategias Nacionales de IA



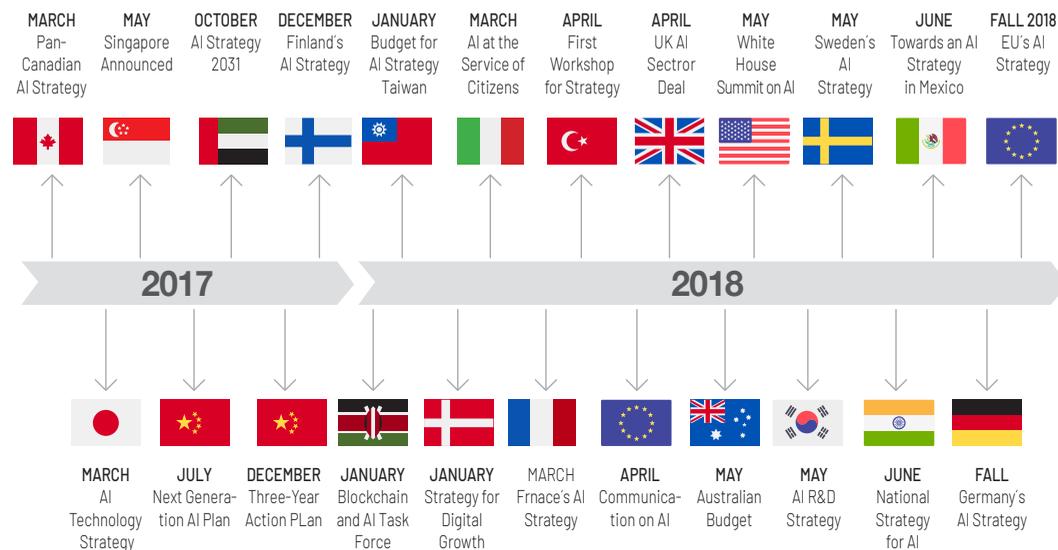
Es importante destacar que, desde el punto de vista de la estrategia nacional, el énfasis de estos países pioneros es tomar una posición de liderazgo, aprovechándose en cierta manera del *first-mover* advantage que ello supone, incluso desde el punto de vista de elevar el perfil del país y situarlo en el centro de la agenda mundial al respecto.

Ya en 2018, encontramos otros países y bloques que lanzaron iniciativas IA, en algunos casos siguiendo la estela de los seis pioneros ya mencionados:

- > Dinamarca
- > Francia
- > Reino Unido
- > Unión Europea
- > Corea del Sur
- > India

En el siguiente gráfico se ofrece una línea temporal muy interesante con los movimientos de los países y bloques que han desarrollado algún tipo de estrategia nacional de IA hasta el año 2018.

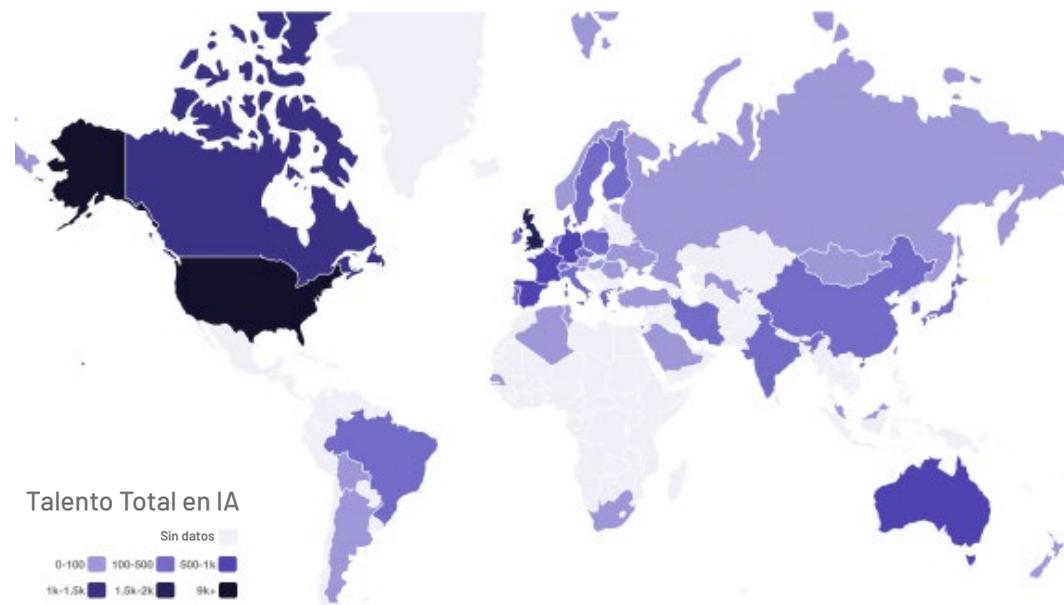
● FIGURA 4. Estrategias para la Inteligencia Artificial.



Fuente: <https://jfgagne.ai/talent/>

Más allá de la voluntad e iniciativa política de cada país, uno de los grandes habilitadores para el desarrollo de estrategias de IA es la formación, atracción y retención del talento IA.

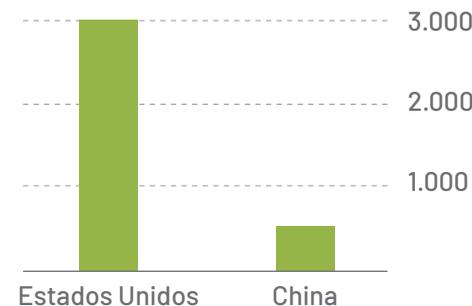
● FIGURA 5. Talento Global en IA en 2018, basado en perfiles de LinkedIn y número de conferencias relevantes.



Fuente: <https://jfgagne.ai/talent/>

Se estima que existen unas 22.000 personas con conocimientos académicos avanzados (hasta nivel Doctor), de las que el 40% se encuentran en EEUU, lo que explica la hegemonía de los últimos años por parte de empresas de IA de EEUU.

● FIGURA 6. Número de empresas de IA en 2016.



Fuente: Wuzhen Institute

Las estrategias nacionales de IA se han dotado de importantes presupuestos, así como de expertos de la industria y el sector público para dar cobertura a las iniciativas políticas. Cabe destacar:

- > UK, con iniciativas en educación, gestión de datos e inversiones privadas de más de 300€.
- > Francia, iniciando un fondo público de inversión para el apoyo a las startups con base IA.
- > China, con un programa de investigación de más de 2.000 millones de dólares para dar cabida a más de 400 empresas.

Desde el punto de vista político, los legisladores han de tener en cuenta además otros retos de gran calado social y con el potencial de crear tensiones y protestas, como por ejemplo:

- > La automatización del trabajo y su implicación en el tejido social, con predicciones que afirman que podría suponer una pérdida del 40% de los empleos para el año 2035: <https://www.axios.com/artificial-intelligence-automation-jobs-robots-fbca4208-ffe4-43c4-9bd3-8102b63cae-db.html>
- > La influencia de la IA en la formación de la opinión pública y la ingeniería social, lo que puede impactar en el funcionamiento de la democracia, y su posible manipulación, hasta puntos aún desconocidos.
- > Los aspectos éticos y morales asociados al avance y uso comercial de la tecnología, con ejemplos recientes y de alto impacto. A saber:
  - Cambridge Analytica: <https://www.wired.com/story/facebook-price-of-tech-utopia/>
  - Privacidad de los asistentes del hogar de Google, Amazon y Apple: <https://www.theverge.com/2018/5/9/17334658/google-ai-phone-call-assistant-duplex-ethical-social-implications>
  - La programación del coche autónomo, en relación con la seguridad de los pasajeros y su entorno: <http://moralmachine.mit.edu/>
  - La existencia de sesgos (bias) en los algoritmos y modelos.
  - El uso de tecnologías de reconocimiento facial y el papel de las empresas en ausencia de leyes y regulación al respecto, por ejemplo el caso de Microsoft y el uso de su tecnología en el Estado de California.
  - Provisión de infraestructuras habilitadoras, como el 5G y la fibra óptica.
  - La generación y gestión de los datos públicos generados por administraciones, ciudades y la "Internet de la Cosas".

# Áreas de actuación en política IA

La política de Inteligencia A de cada país tiene que guiarse por los objetivos estratégicos establecidos en sus estrategias, en base a las ventajas competitivas que ofrezcan su tejido productivo, empresarial y de I+D+i.

De entre todas las opciones e iniciativas políticas nacionales que pueden llevarse a cabo en el ámbito IA, se han establecido 10 categorías de actuaciones a tener en cuenta:

## 1. Investigación básica y aplicada

**Ejemplo de aplicación: Turing Institute del Reino Unido**

Se trata de una institución nacional para el desarrollo de la IA y la Ciencia de Datos, con un programa muy ambicioso alrededor de 3 objetivos:

1. La investigación avanzada y su aplicación a problemas reales.
2. Formar a los líderes del futuro.
3. Liderar la opinión pública.

## 2. Desarrollo, atracción y retención del talento

**Ejemplo de aplicación: CIFAR Chairs in AI Program de Canadá**

El Gobierno de Canadá provee de fondos a tres Centros de Excelencia para la investigación e innovación en IA, con el que se pretende atraer y retener al mejor talento disponible en la industria.

## 3. Educación STEM, continua y digital

**Ejemplo de aplicación: Technology Pact de Dinamarca**

Un pacto entre más de 80 organizaciones, públicas y privadas, para desarrollar las habilidades y conocimientos necesarios en la población danesa, para afrontar los trabajos de alto componente tecnológico y digital de hoy y del futuro. Se reconoce la importancia de estos conocimientos para mantener la posición competitiva del país, la importancia de la educación STEM, y que los colegios enseñen a los niños cómo programar y los fundamentos de la ciberseguridad. El objetivo es que para 2020, 150.000 jóvenes formen parte de este pacto, y que en 2030 al menos el 20% de los estudios superiores sea en un área STEM.

## 4. Industrialización de la IA

**Ejemplo de aplicación: Roadmap de Industrialización de Japón**

El objetivo de este plan es la materialización de una sociedad 'Super Smart', que Japón denomina Sociedad 5.0, para lo que se proponen derribar cinco muros:

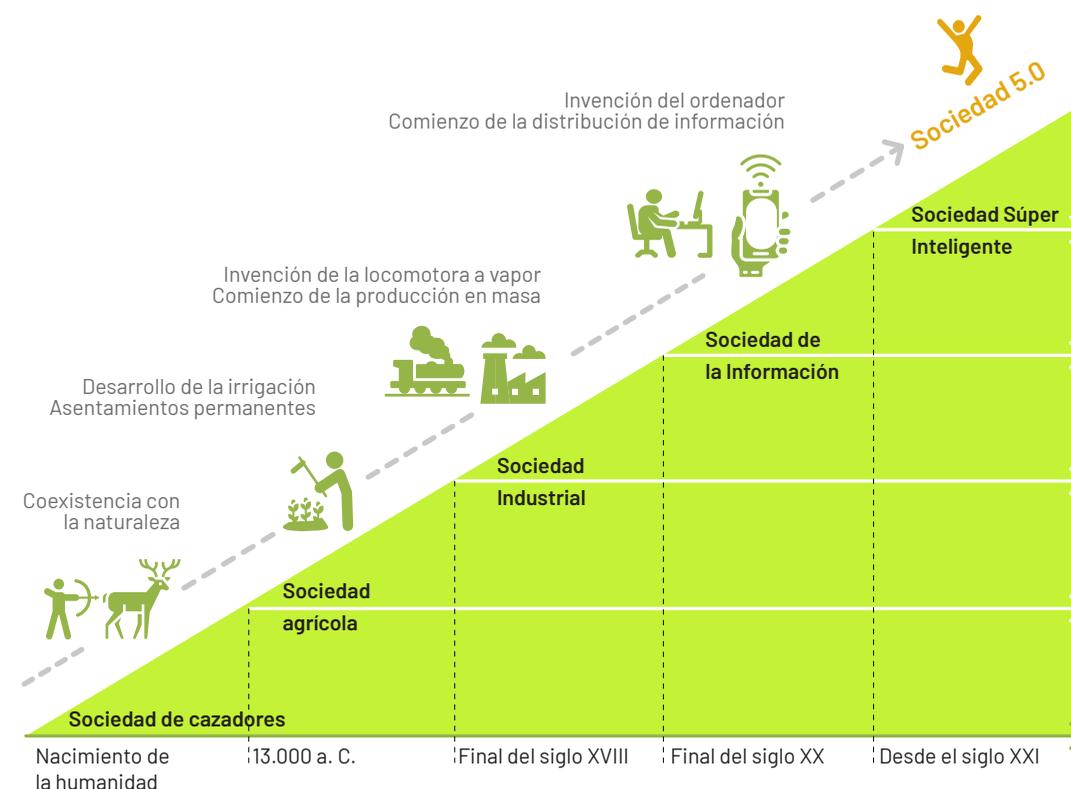
1. El muro del sector público.
2. El muro de la legislación.
3. El muro de la tecnología.
4. El muro de los recursos humanos y el talento.
5. El muro de la aceptación social.

## 5. Uso de IA en los gobiernos

**Ejemplo de aplicación: Ministerio de Inteligencia Artificial de Emiratos Árabes Unidos**

El primer Ministerio en el mundo dedicado a tal efecto, como ya comentamos anteriormente en relación a los pioneros en esta materia.

FIGURA 7. Evolución histórica de la humanidad como sociedad.



## 6. Infraestructura de datos

**Ejemplo de aplicación: Health Data Hub de Francia**

Como parte de la estrategia global de Francia en esta materia, uno de sus objetivos es liderar la adopción de la ciencia de datos e IA en el sector sanitario, mediante la creación de un gran banco de datos abiertos para investigadores y el público, que le sitúe como referente a nivel mundial.

## 7. Ética: sesgos, privacidad, seguridad

**Ejemplo de aplicación: AI Ethics Guidelines de la Unión Europea**

La Unión Europea nombró en 2018 un grupo de 52 expertos de todos los ámbitos para evaluar las oportunidades de la IA, pero también para proveer guías de actuación para la Unión Europea sobre justicia, seguridad, transparencia, el futuro del trabajo y el impacto de la IA en materia de derechos fundamentales, privacidad, datos personales, dignidad, protección de los consumidores y no discriminación. En Abril de 2019, dicho grupo de expertos ha publicado el informe: 'Ethics Guidelines for Trustworthy Artificial Intelligence (AI)'.

## 8. Regulación

**Ejemplo de aplicación: Ethics Commission on Automated and Connected Driving de Alemania**

Creación por parte del Ministerio de Transporte e Infraestructura Digital de una comisión ética para los coches autónomos y conectados, cuyo objetivo ha sido el desarrollo de guías éticas para la regulación de los coches autónomos y conectados. Los miembros

bros de esta comisión ética son filósofos, juristas, tecnólogos, sociólogos y además de miembros de las industrias de la automoción y del *software*. Otra nación que está afrontando el reto de la regulación es Singapur, que con su estándar nacional TR-68 para el desarrollo de los coches autónomos, está creando el marco idóneo para atraer talento y empresas que quieran depurar sus diseños en un entorno real.

## 9. Inclusión

### Ejemplo de aplicación: India -> AI For All de la India

India ha publicado unas líneas de actuación el eslogan #AIForAll. Este eslogan captura la esencia de su idea sobre la posición India respecto a la IA y la inclusión: adopción de tecnologías transformadoras que aseguren un crecimiento social e inclusivo de su sociedad. Además, plantean un papel de liderazgo replicando este modelo en otros países en vías de desarrollo. Según este manifiesto, los mayores beneficios sociales se obtendrán en cinco áreas:

- Salud: acceso universal y asequible.
- Agricultura: mejorar los ingresos de los agricultores con incrementos de la productividad y reducción del desperdicio.
- Educación: mejorar el acceso y la calidad de la educación.
- Ciudades inteligentes e infraestructuras: asegurar una conectividad eficiente en entorno urbanos.
- Movilidad inteligente y transporte: medios de transportes seguros e inteligentes, con mejores mecanismos para gestionar el tráfico y la congestión.

## 10. Política exterior

### Ejemplo de aplicación: Libro Blanco de la Estandarización de la IA de China

Publicación de un Libro Blanco de estandarización de Inteligencia Artificial, con grandes contribuciones del principal tejido industrial y académico chino, colocando a China como pionera en la definición de estándares mundiales en el uso de la tecnología, incluyendo aspectos éticos y regulatorios, dentro del comité técnico ISO/IEC JTC 1/SC 42. China es consciente de que sus algoritmos de IA tendrán efectos más allá de sus fronteras, y por eso su interés en tomar una función preponderante y de liderazgo en todo lo relacionado con los aspectos legislativos, sociales y éticos de la aplicación de la tecnología.

Esta clasificación y ejemplos demuestran que Gobiernos de toda índole y en todas las regiones están diseñando sus estrategias y regulaciones de IA, además de servir de guía para que otros Gobiernos evalúen su grado de madurez y estrategia en cada una de estas categorías.

Cabe destacar dos países por encima del resto en cuanto a su estrategia IA:

- > **Australia**, con la estrategia Australia 2030: Prosperity Through Innovation, que prioriza las inversiones en IA para la digitalización, e incluye una extensa selección de métricas para medir el éxito de las políticas IA en el país (página 110 del documento enlazado).
- > **China**, que presenta la estrategia nacional más avanzada de todos los países considerados, lo que encaja con su objetivo de convertirse en el líder absoluto de esta revolución tecnológica. Cabe destacar la clara orientación a objetivos claros:
  - Estar a la par de los líderes mundiales en 2020.
  - Llegar a ser líderes mundiales en algunas disciplinas de IA en 2025.
  - Ser el principal centro de innovación en IA en 2030.

● FIGURA 8. Línea temporal del marco estratégico y de implementación de la UE y España



Fuente: Estrategia Española en I+D+I en IA

## OPINIÓN



### Nuria Oliver

JEFA DE CIENCIA DE DATOS EN DATA-POP ALLIANCE

Dada la ubicuidad, transversalidad e invisibilidad de la Inteligencia Artificial, su gobernanza es un tema que está siendo debatido tanto a nivel nacional como supra-nacional. En este sentido he propuesto un marco de gobernanza que queda captado por el acrónimo FATEN en inglés, y que incluye:

**La F de fairness** o justicia, es decir sin discriminar. Las decisiones basadas en algoritmos pueden discriminar porque los datos utilizados para entrenar dichos algoritmos tengan sesgos que den lugar a decisiones discriminatorias; por el uso de un determinado algoritmo; o por el mal uso de ciertos modelos en diferentes contextos.

**La A es de autonomy** o autonomía, valor central en la ética occidental, según la cual cada persona debería tener la capacidad de decidir sus propios pensamientos y acciones, asegurando por tanto la libre elección, la libertad de pensamiento y acción. Sin embargo, hoy en día podemos construir modelos computacionales de nuestros deseos, necesidades, personalidad y comportamiento con la capacidad de influenciar nuestras decisiones y comportamientos de manera subliminal.

Por ello, deberíamos garantizar que los sistemas inteligentes autónomos adapten su toma de decisiones preservando la autonomía y dignidad humanas. Es importante destacar la importancia de que todos los

OPINIÓN  
Nuria Oliver

desarrolladores y profesionales que trabajen en el desarrollo de sistemas de Inteligencia Artificial que afecten o interaccionen con personas (por ejemplo, algoritmos para la toma de decisiones, sistemas de recomendación y personalización, chatbots...) se comporten de acuerdo con un claro Código de Conducta y de Ética definido por las organizaciones en las que trabajan. Como sabiamente dijo Disney "No es difícil tomar decisiones cuando tienes claro cuáles son tus valores".

**La A también es de *accountability***; es decir, con claridad respecto a la atribución de responsabilidad de las consecuencias de las decisiones algorítmicas.

**Y de *augmentation***, de manera que los sistemas de Inteligencia Artificial se utilicen para aumentar o complementar la inteligencia humana, no para reemplazarla.

**La T de *trust* o confianza**. La confianza es necesaria en las relaciones entre humanos e instituciones. Las empresas tecnológicas necesitan una relación de confianza con sus usuarios que cada vez delegan (delegamos) mas aspectos de nuestras vidas a los servicios digitales. Sin embargo, hoy en día la confianza hacia el sector tecnológico se esta perdiendo, en parte debido a escándalos recientes como el de Facebook/Cambridge Analytica y Huawei. Para que exista confianza, han de cumplirse tres condiciones: (1) la competencia para desempeñar con solvencia la tarea asignada; (2) la fiabilidad, es decir, la competencia sostenida en el tiempo; y (3) la honestidad y transparencia. Por ello, la T también es de:

***Transparency* o transparencia**. Un modelo es transparente si una persona puede observarlo y entenderlo con facilidad. Las decisiones algorítmicas pueden ser no transparentes (opacas) por tres motivos: intencionalmente para proteger la propiedad intelectual de los creadores de dichos algoritmos; por la falta de conocimiento por parte de los usuarios que les impida entender cómo funcionan los algoritmos y modelos computacionales construidos a partir de los datos; e intrínsecamente dado que ciertos métodos de aprendizaje por ordenador (por ejemplo, modelos de deep learning o aprendizaje profundo) son extremadamente complejos.

Asimismo, es imprescindible que los sistemas de Inteligencia Artificial sean transparentes no solo con relación a qué datos captan y analizan sobre el comportamiento humano y para qué propósitos sino también respecto en qué situaciones los humanos están interaccionando con sistemas artificiales (por ejemplo, chatbots) vs con otros humanos.

**La E de *education* o educación**, invirtiendo en educación a todos los niveles, desde la educación obligatoria a la ciudadanía en general, pasando por los profesionales y los decision makers y trabajadores del sector público;

**E de *bEneficence* o beneficencia**; es decir, maximizando el impacto positivo del uso de la Inteligencia Artificial, con sostenibilidad, diversidad, honestidad y veracidad y E de equality o igualdad, ya que desgraciadamente con la Cuarta Revolución Industrial se esta produciendo una polarización en la acumulación de la riqueza en el mundo. Deberíamos establecer los marcos regulatorios necesarios para fomentar la igualdad y solidaridad con el desarrollo de la IA.

**La N de *non-maleficence* o no maleficencia**, minimizando el impacto negativo que pueda derivarse del uso de las decisiones algorítmicas. Para ello, es importante aplicar un principio de prudencia, garantizar la seguridad, fiabilidad y reproducibilidad de los sistemas, preservando siempre la privacidad de las personas. En este último punto, Europa ha asumido cierto liderazgo a nivel mundial con la reciente entrada en vigor del Reglamento General de Protección de Datos o RGPD, que se suma a derechos fundamentales como los derechos a establecer y desarrollar relaciones con otros seres humanos, a la desconexión tecnológica y a estar libre de vigilancia. En este contexto, otros derechos que podríamos/deberíamos agregar incluyen el derecho a un contacto humano significativo —por ejemplo, en servicios de atención operados exclusivamente por chatbots— y el derecho a no ser medido, analizado, perfilado, orientado o influenciado subliminalmente mediante algoritmos.

Considero que será solamente cuando respetemos estos principios que seremos capaces de avanzar y conseguir un modelo de gobernanza democrática basado en los datos y la inteligencia artificial, por y para las personas.

SUPERANDO  
LOS RIESGOS DE  
UNA IA SIN ÉTICA  
NI REGULACIÓN

Estrategias  
Nacionales de  
la Inteligencia  
Artificial

52

**Más allá de la voluntad e iniciativa política de cada país, uno de los grandes habilitadores para el desarrollo de estrategias de IA es la formación, atracción y retención del talento IA**

**CONCLUSIONES**

## Conclusiones “No Tecnológicas”

El avance del conocimiento humano nos está llevando a fronteras inimaginables hace unas décadas en todos los ámbitos del conocimiento y del saber, pero todos ellos con un elemento común: la tecnología (o más bien las tecnologías). Una de ellas, quizás de las que tienen una componente más de fantasía, es la Inteligencia Artificial, puesto que se trata de la recreación en máquinas de un comportamiento esencialmente humano, la capacidad de pensar y razonar para hallar soluciones a problemas complejos. Desde la creación del “hombre de palo”, el autómatas que iba pidiendo limosna fabricado por Juanelo Turriano, relojero de la corte de Carlos I de España, el mito de la hija autómatas de Rene Descartes o el caminante de George Moore, el hombre ha intentado replicar su comportamiento y razón en las máquinas. Sin embargo, es ahora, en nuestros tiempos, y gracias principalmente a la enorme capacidad de computación, cuando las técnicas matemáticas y de ingeniería que rigen los métodos y modelos de inteligencia artificial, cuando se ha hecho posible.

En este informe hemos introducido algunas de las técnicas y tecnologías como el aprendizaje automático, el procesamiento y generación del lenguaje natural, la visión artificial o los sistemas empotrados que se están utilizando actualmente en diferentes sectores, y en todos ellos

con un alto impacto y rendimiento. Como hemos descrito, la utilización de la Inteligencia Artificial en las telecomunicaciones o la salud, pero también en muchos otros sectores, como el marketing, la agricultura o las energías renovables, está siendo cada vez más masiva

Pero todo esto, queridos lectores de este informe, lo conocemos ya, puesto que mucho se ha escrito (y queda por escribir) con respecto a las tecnologías y la aplicación por sectores de la Inteligencia Artificial. Es por ello que, en este informe, queremos darle un peso relevante a las implicaciones éticas y sociales, así como a las actuaciones que los diferentes gobiernos están teniendo, o planificando, sobre la Inteligencia Artificial

Es muy posible que, en un futuro no muy lejano, la biotecnología, por un lado, y la Inteligencia Artificial, por otro, conducirán al diseño intensivo de organismos vivos y de sistemas inteligentes, con mayor o menor grado de autonomía, destinados a mejorar nuestra existencia, pero que no están exentos de peligro por un uso inadecuado o por no haber valorado adecuadamente los riesgos que pueden derivarse de su uso.

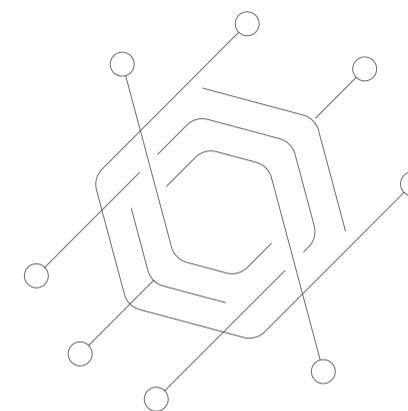
Muchas de las cuestiones éticas que hemos puesto de relevancia en este informe se vienen produciendo, desde los albores de la humanidad, fuera del ámbito de la Inteligencia Artificial (los prejuicios, la discriminación, la manipulación, etc.), pero es precisamente la aceleración creciente de las capacidades de estos sistemas, y su empleo recursivo para diseñar sistemas

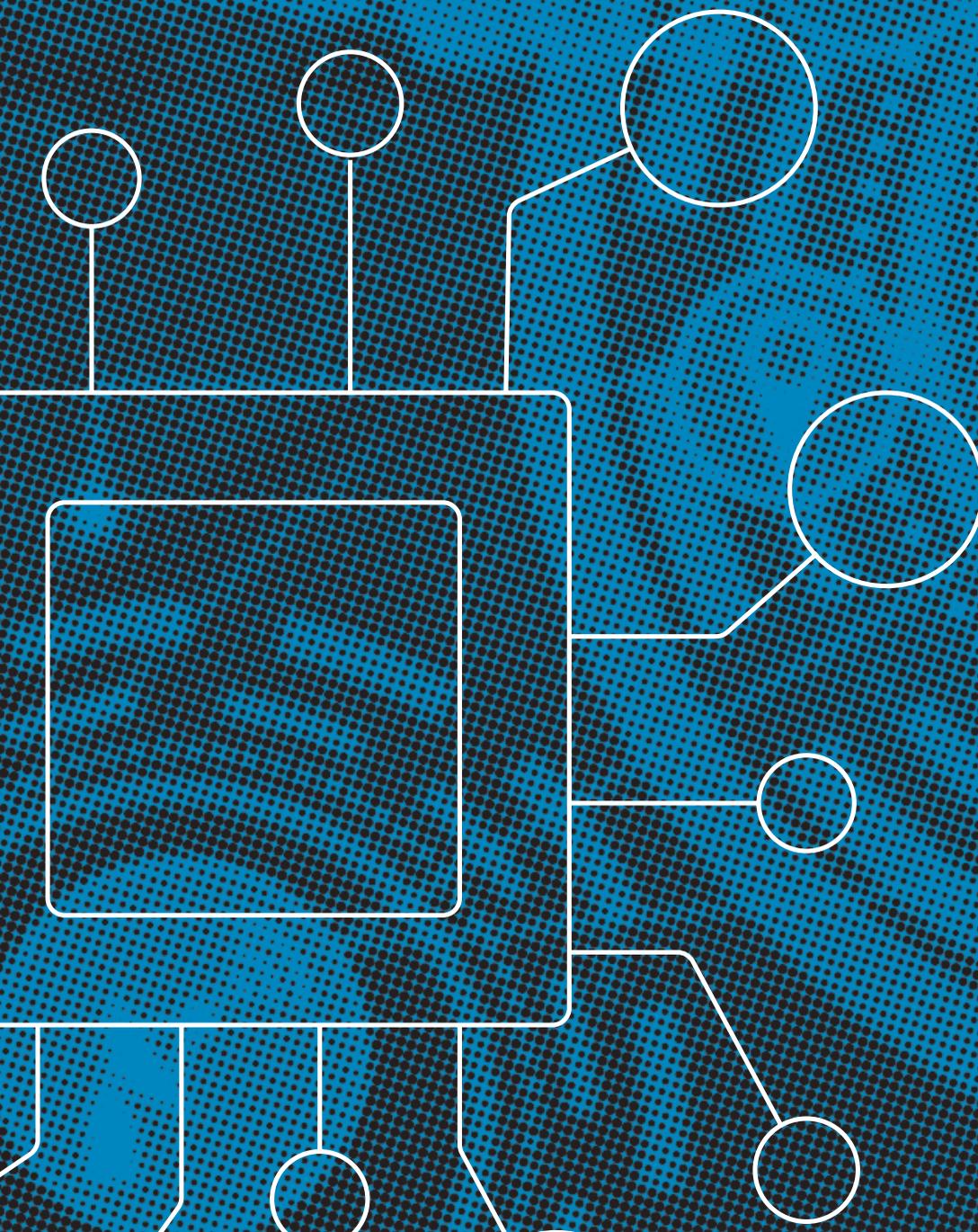
más potentes y capaces, lo que está permitiendo su generalización a todas las esferas de la actuación humana, y cuando hablamos de todas las esferas contamos las que buscan mejorar nuestra vida y la del planeta, pero también las que pertenecen al reverso tenebroso de nuestra especie.

Por tanto, queremos llamar la atención sobre el hecho de que no debe renunciarse a las posibilidades y bondades de la Inteligencia Artificial, pero ello debe hacerse desde la óptica de la ética y responsabilidad.

En suma, podríamos concluir lo siguiente:

- La inteligencia artificial puede constituir un pilar fundamental del humanismo si en el núcleo de su concepción y diseño tiene un lugar preeminente el desarrollo y bienestar de la humanidad, de su entorno y del planeta.
- Las asimetrías de información derivan en asimetrías de poder políticas, económicas y sociales. Garantizar la utilización equitativa e inclusiva de la inteligencia artificial es fundamental para evitar nuevas brechas tecnológicas, económicas y sociales, y más aún porque, en la actualidad todavía hay que superar obstáculos derivados de la falta de infraestructuras tecnológicas básicas.
- Las implicaciones morales y éticas de la Inteligencia Artificial conciernen a todos los agentes involucrados de una u otra manera en ella. En particular, los perfiles científicos y técnicos desarrolladores de la IA han de ser conscientes de estas nuevas implicaciones que conlleva su trabajo. La preocupación ética por la creación de nuevos tipos de inteligencia requiere de un exquisito criterio moral de las personas que diseñan estas nuevas formas de tecnología.
- En este nuevo contexto de avance tecnológico, la educación en todas sus facetas (STEM y humanidades), así como la formación continua, juegan un papel fundamental para asumir todos los beneficios de la Inteligencia Artificial y evitar los riesgos identificados.



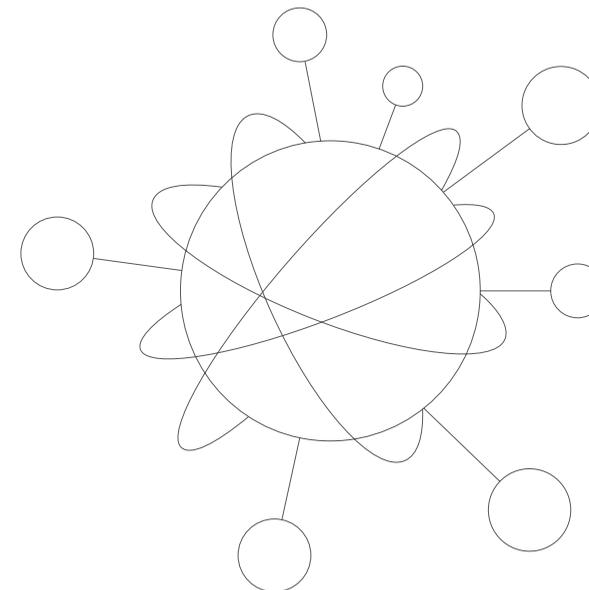


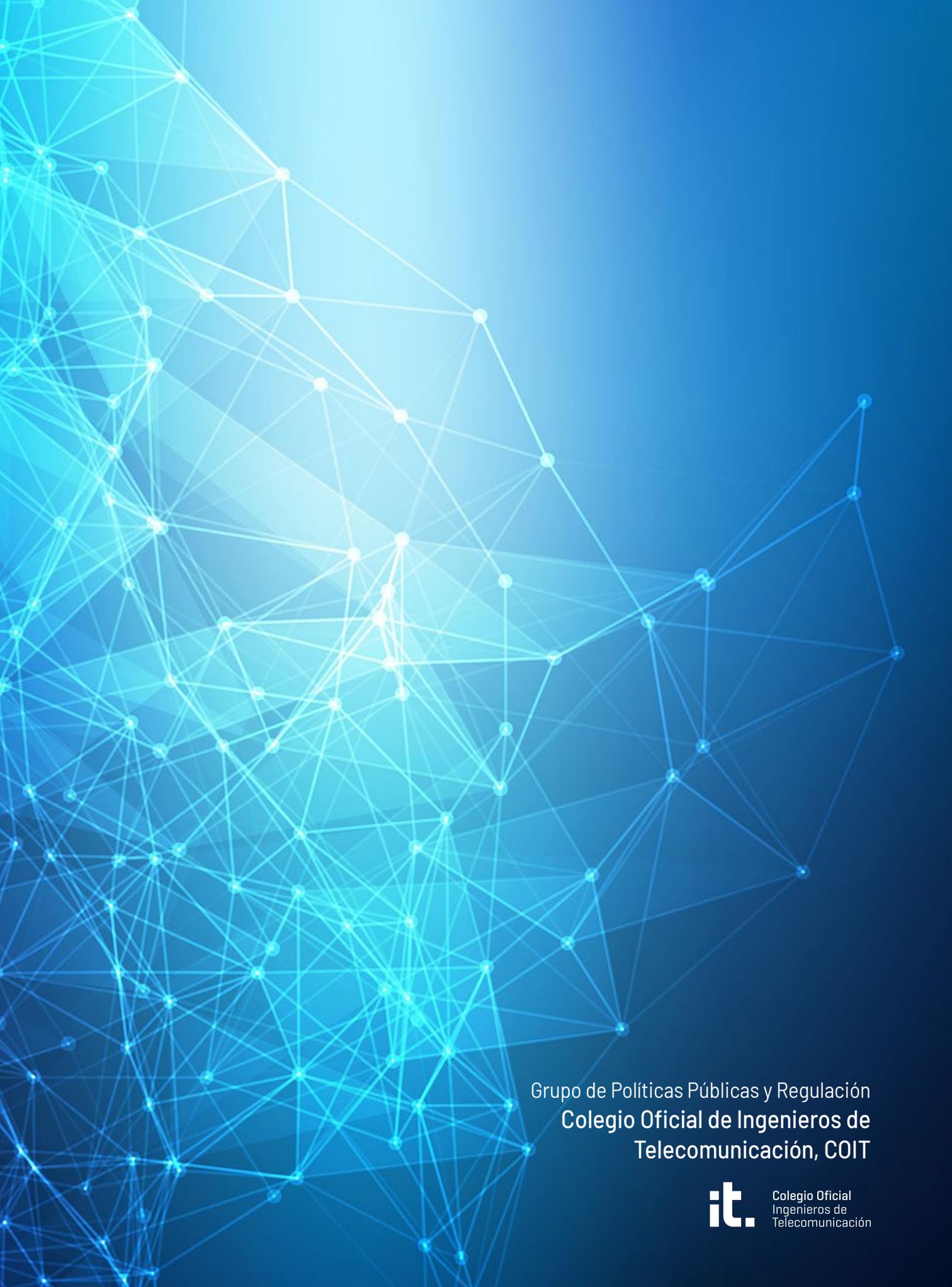
# REFERENCIAS

# Referencias

- > <https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd>
- > <https://medium.com/politics-ai/ai-policy-101-what-you-need-to-know-about-ai-policy-163a2bd68d65a>
- > <https://www.nedo.go.jp/content/100865202.pdf> -> Japan
- > <https://jfgagne.ai/talent/>
- > <https://www.elementai.com/>
- > <https://www.gov.uk/government/publications/artificial-intelligence-sector-deal/ai-sector-deal>
- > <https://brookfieldinstitute.ca/report/intro-to-ai-for-policymakers/> -> [https://brookfieldinstitute.ca/wp-content/uploads/2017/07/20/content\\_5211996.htm](https://brookfieldinstitute.ca/wp-content/uploads/2017/07/20/content_5211996.htm)
- > <https://www.technologyreview.com/s/610546/china-wants-to-shape-the-global-future-of-artificial-intelligence/> China
- > <https://www.technologyreview.com/s/609038/chinas-ai-awakening/> China
- > [http://www.gov.cn/zhengce/content/2017-07/20/content\\_5211996.htm](http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm)
- > How China is preparing for an AI-powered Future
- > AI Policy – China
- > EAU Ministerio de Inteligencia Artificial
- > <https://futurism.com/uae-minister-artificial-intelligence>
- > <https://futurism.com/dubai-announces-accelerator-to-solve-the-21st-century-s-most-pres-sing-challenges>
- > <https://dubaifutureaccelerators.com/en/>
- > <https://www.politico.eu/article/finland-one-percent-ai-artificial-intelligence-courses-learning-training/>
- > [http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEMrap\\_47\\_2017\\_verkko-julkaisu.pdf?sequence=1&isAllowed=y](http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEMrap_47_2017_verkko-julkaisu.pdf?sequence=1&isAllowed=y) -> Finlandia estrategia IA.
- > [https://www.turing.ac.uk/sites/default/files/2018-07/turing\\_pfs\\_digital-version.pdf](https://www.turing.ac.uk/sites/default/files/2018-07/turing_pfs_digital-version.pdf)
- > <https://www.elmundo.es/tecnologia/innovacion/2019/04/22/5cbda0cefc6c8314588b458d.html>
- > <https://www.cifar.ca/ai/pan-canadian-artificial-intelligence-strategy>
- > <https://eng.em.dk/news/2018/april/the-danish-government-launches-the-technology-pact/>
- > <https://www.teknologipagten.dk/>
- > <https://www.i-scoop.eu/industry-4-0-society-5-0/>
- > <https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>
- > <https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines>
- > AI For All: [http://www.niti.gov.in/writereaddata/files/document\\_publication/NationalStrategy-for-AI-Discussion-Paper.pdf](http://www.niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf)
- > [https://ec.europa.eu/commission/news/artificial-intelligence-2018-dec-07\\_en](https://ec.europa.eu/commission/news/artificial-intelligence-2018-dec-07_en)
- > Estrategia IA: <https://ec.europa.eu/digital-single-market/en/news/communication-artificial-intelligence-europe>
- > <https://ec.europa.eu/digital-single-market/en/artificial-intelligence>
- > [http://www.europarl.europa.eu/RegData/etudes/STUD/2018/626074/IPOL\\_STU\(2018\)626074\\_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/STUD/2018/626074/IPOL_STU(2018)626074_EN.pdf)
- > <https://ec.europa.eu/digital-single-market/en/news/factsheet-artificial-intelligence-europe>
- > <https://www.globalgovernmentforum.com/eu-launches-ai-strategy/>
- > <http://ai-europe.eu/artificial-intelligence-in-europe-ai4eu-project-is-launched/>
- > <https://www.ai4eu.eu/>
- > [http://europa.eu/rapid/press-release\\_IP-18-4043\\_en.htm](http://europa.eu/rapid/press-release_IP-18-4043_en.htm)

- > [https://ec.europa.eu/commission/sites/beta-political/files/budget-june2018-digital-trans-formation\\_en.pdf](https://ec.europa.eu/commission/sites/beta-political/files/budget-june2018-digital-trans-formation_en.pdf)
- > <https://www.zdnet.com/article/singapore-releases-guidelines-for-deployment-of-autono-mous-vehicles/>
- > <https://sbr.com.sg/transport-logistics/news/tr68-could-attract-foreign-autonomous-tech-nology-firms-singapore-fitch>
- > <http://www.ciencia.gob.es/portal/site/MICINN/menuitem.edc7f2029a2be-27d7010721001432ea0/?vgnextoid=4c7cffe4d2849610VgnVCM1000001d04140aRCRD>
- > [http://www.ciencia.gob.es/stfls/MICINN/Ciencia/Ficheros/Estrategia\\_Inteligencia\\_Artifi-cial\\_IDI.pdf](http://www.ciencia.gob.es/stfls/MICINN/Ciencia/Ficheros/Estrategia_Inteligencia_Artifi-cial_IDI.pdf)
- > <http://www.ciencia.gob.es/portal/site/MICINN/menuitem.26172fcf4eb029fa6ec-7da6901432ea0/?vgnextoid=70fcd77ec929610VgnVCM1000001d04140aRCRD>
- > [https://sevilla.abc.es/andalucia/granada/sevi-espana-presenta-granada-estrategia-pa-rra-inteligencia-artificial-etica-201903041536\\_noticia.html](https://sevilla.abc.es/andalucia/granada/sevi-espana-presenta-granada-estrategia-pa-rra-inteligencia-artificial-etica-201903041536_noticia.html)
- > [https://retina.elpais.com/retina/2018/12/14/innovacion/1544784100\\_806508.html](https://retina.elpais.com/retina/2018/12/14/innovacion/1544784100_806508.html)
- > <https://sustainabledevelopment.un.org/post2015/transformingourworld>





Grupo de Políticas Públicas y Regulación  
Colegio Oficial de Ingenieros de  
Telecomunicación, COIT

**it.** Colegio Oficial  
Ingenieros de  
Telecomunicación